COPING WITH CLIMATE CHANGE: A GENOMIC PERSPECTIVE ON THERMAL ADAPTATION

EDITED BY: Margarida Matos, Pedro Simões, Inês Fragata, Ana Sofia Quina, Torsten Nygaard Kristensen and Mauro Santos

PUBLISHED IN: Frontiers in Genetics and Frontiers in Ecology and Evolution

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88966-494-8
DOI 10.3389/978-2-88966-494-8

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

COPING WITH CLIMATE CHANGE: A GENOMIC PERSPECTIVE ON THERMAL ADAPTATION

Topic Editors:
Margarida Matos, University of Lisbon, Portugal
Pedro Simões, University of Lisbon, Portugal
Inês Fragata, University of Lisbon, Portugal
Ana Sofia Quina, University of Lisbon, Portugal
Torsten Nygaard Kristensen, Aalborg University, Denmark
Mauro Santos, Universitat Autònoma de Barcelona, Spain

Citation: Matos, M., Simões, P., Fragata, I., Quina, A. S., Kristensen, T. N., Santos, M., eds. (2021). Coping with Climate Change: A Genomic Perspective on Thermal Adaptation. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-494-8

Table of Contents

frontiers in Genetics

EDITORIAL
published: 13 January 2021
doi: 10.3389/fgene.2020.619441

Editorial: Coping With Climate Change: A Genomic Perspective on Thermal Adaptation

Margarida Matos[1], Pedro Simões[1], Inês Fragata[1], Ana Sofia Quina[2], Torsten Nygaard Kristensen[3] and Mauro Santos[4]*

[1] cE3c – Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal, [2] CESAM, Centre for Environmental and Marine Studies, Universidade de Aveiro and Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal, [3] Section of Biology and Environmental Science, Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark, [4] Departament de Genètica i de Microbiologia, Grup de Genòmica, Bioinformàtica i Biologia Evolutiva (GBBE), Universitat Autònoma de Barcelona, Barcelona, Spain

Keywords: candidate genes for thermal adaptation, genomic basis of evolutionary potential, genome-wide association studies, transcriptomics, thermal plasticity

Editorial on the Research Topic

Coping With Climate Change: A Genomic Perspective on Thermal Adaptation

Current human-induced climate warming poses a threat to many organisms (Somero, 2012; Buckley and Huey, 2016; Walsh et al., 2019). Species respond to climate change in different ways, from plasticity, evolutionary adaptation, and dispersal, to extinction (Holt, 1990; Parmesan, 2006). In ectotherms, the upper thermal limits have limited plasticity compared to lower thermal limits (Gunderson and Stillman, 2015). Additionally, physiological tolerance to critically high temperatures (when performance drops to zero) may be genetically constrained (Araújo et al., 2013; Hoffmann et al., 2013). Consequently, behavioral thermoregulation can be an important mechanism in buffering exposure to extreme temperatures (Sunday et al., 2014). Predicting how species may adapt to new thermal conditions requires robust ways of evaluating their underlying evolutionary and plastic potentials. Given that intraspecific differentiation to upper critical thermal limits is commonly observed (e.g., Herrando-Pérez et al., 2019, 2020), selection for tolerance to high temperatures may be occurring, although it is not clear how. This calls for a deeper understanding of the underlying genetic basis of thermal adaptation (Porcelli et al., 2015).

In recent decades there has been a boom in studies that use genome-wide sequencing (Savolainen et al., 2013; Ellegren, 2014). Next-generation sequencing techniques, when applied to experimental thermal evolution have contributed to understanding these genomic responses to changing thermal conditions (e.g., Porcelli et al., 2015; Mallard et al., 2018). The combination of genome-wide screenings with more classical approaches could pave the way for an integrative understanding of how populations cope with climate change. This Research Topic aims to: (1) expand knowledge on the genomic basis of thermal adaptation; (2) assess whether and how genetic and genomic diversity can lead to common or different adaptive routes; and (3) discuss ways to improve the contribution of different studies to community-level knowledge.

Logan and Cox suggest that there is moderate heritability for upper thermal tolerance and, hence, the potential for heat tolerance to evolve. However, this may also be constrained by unfavorable genetic correlations with other thermal performance traits. They also suggest that the plastic response of the transcriptome depends on the magnitude of thermal shifts. Rodrigues and Beldade also study genomics and transcriptomics of plasticity, which are usually assumed to have the potential to enhance thermal adaptation. However, more research is needed because high

OPEN ACCESS

Edited and reviewed by:
Samuel A. Cushman, United States Forest Service (USDA), United States

****Correspondence:***
Margarida Matos
mmmatos@fc.ul.pt

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *20 October 2020*
Accepted: *15 December 2020*
Published: *13 January 2021*

Citation:
Matos M, Simões P, Fragata I, Quina AS, Kristensen TN and Santos M (2021) Editorial: Coping With Climate Change: A Genomic Perspective on Thermal Adaptation. Front. Genet. 11:619441. doi: 10.3389/fgene.2020.619441

phenotypic plasticity may be maladaptive in warmer and more variable future climates (Kreyling et al., 2019).

Using a genome-wide association study and transcriptomic profiling in lines from the *Drosophila* Genetic Reference Panel (DGRP, Mackay et al., 2012), Lecheta et al. found that heat tolerance is less variable than cold tolerance, with ∼50% more identified genes affecting the latter. These results reinforce that heat tolerance is more constrained than cold tolerance (Araújo et al., 2013). However, all lines came from one single population, which could produce biased estimates of heat tolerance (Herrando-Pérez et al., 2020), and further research in other populations is needed. The Research Topic also includes genome-wide approaches and gene expression analyses. Sørensen et al. studied *D. simulans* populations after 20 generations of experimental evolution under predictable or unpredictable thermal fluctuations. The strongest response involved unpredictable fluctuations, and the genes under selection were distinct from the genes that are important for the adaptive plastic response under predictive thermal fluctuations. Other studies have shown that constant and fluctuating temperatures induce different plastic and evolutionary responses (Botero et al., 2015; Dey et al., 2016), and more studies like this are important in uncovering the complexities of thermal evolution.

A potential limitation of studies on the evolution of heat tolerance is that most ignore the possible negative impacts of sublethal temperatures on oogenesis and spermatogenesis. This could lead to a higher vulnerability to climate warming in many organisms than is currently thought (David et al., 2005; Walsh et al., 2019). Using the DGRP, Zwoinska et al. found that males are more affected than females when flies were exposed at high sublethal temperatures. At the same time, they did not find additive genetic variance for reproductive performance at these temperatures. Similar results were obtained for the egg-to-adult viability assessed at different temperatures in *D. melanogaster* (Kristensen et al., 2015). However, in Zwoinska et al. the power to map the genetic variants of relatively small effects may be reduced due to the low line number. Thus more studies, using several populations or more DGRP lines are needed.

Few studies have demonstrated that adaptive evolution is occurring as a consequence of climate change. Latitudinal and long-term trends in the frequency of inversions in *Drosophila subobscura* are remarkably consistent worldwide and highly correlated with environmental temperature, respond to seasonal changes and frequency shifts shortly after a heatwave (Balanyà et al., 2006; Rezende et al., 2010; Rodríguez-Trelles et al., 2013). Karageorgiou et al. focused on the breakpoints of a particular inversion that shows cyclic seasonal changes and speculate that this might be partly due to antagonistic pleiotropic effects on reproduction and immunity resulting from a position effect affecting the expression of functional genes located at the breakpoints.

An important ingredient in a warming world is the ecological and evolutionary implications of parasite-host dynamics and prevalence. Mazzucco et al. focussed on the endosymbiont *Wolbachia* infecting *D. melanogaster* lines evolved in cold and hot environments, and found that these dynamics cannot be straightforwardly linked to temperature, making it difficult to predict the impact of climate change.

A goal of genome-wide analyses is to detect and understand the signatures left by natural selection on the genome. Cortés et al. summarize some of the tools available to reveal the genetic consequences of climate change, but there are some shortcomings to linking fitness relevant genes with environmental factors; e.g., those related to data reporting as highlighted by Waldvogel et al. This is important, as genotype-environment associations can be a key ingredient in forecasting the response of natural populations to climatic variation.

Relevant insights into the genetic basis of thermal evolution can come from studying the adaptations of organisms that live in extreme natural environments. Using a metagenomics approach, Alcorta et al. shed light on the genomic features and taxonomy of thermophilic cyanobacteria living in hot springs. These features included genome reduction, changes in GC content, coding density, and size of biosynthetic gene clusters.

The contributions to this Research Topic add to our understanding of thermal adaptation and its multifactorial nature, and highlight the challenges that are still ahead of us in striving for a deeper understanding of adaptation to expected higher and more variable future temperatures. As shown in this Research Topic, increased knowledge should be brought about by complementary approaches comprising different levels of biological organization and their interaction, using a variety of methodologies and study organisms.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication. All authors contributed to the writing of the editorial.

FUNDING

MM and PS research received funding by Fundação para a Ciência e a Tecnologia (FCT) for the projects PTDC/BIA-EVL/28298/2017 and cE3c Unit UIDB/BIA/00329/2020. PS and AQ are funded by FCT through a contract foreseen in ns. 4–6 of article 23 of DL 57/2016, changed by Law 57/2017. IF was funded by a Junior Researcher contract from FCT (CEECIND/02616/2018). TK was supported by grants from the Danish Council for Independent Research (DFF-8021-00014B). MS was funded by grants CGL2017-89160P from the Ministerio de Economía y Competitividad (Spain; co-financed with the European Union FEDER funds), and 2017SGR 1379 from Generalitat de Catalunya.

ACKNOWLEDGMENTS

We are grateful to all contributors to this Research Topic.

REFERENCES

Araújo, M. B., Ferri-Yáñez, F., Bozinovic, F., Marquet, P. A., Valladares, F., and Chown, S. L. (2013). Heat freezes niche evolution. *Ecol. Lett*. 16, 1206–1219. doi: 10.1111/ele.12155

Balanyà, J., Oller, J. M., Huey, R. B., Gilchrist, G. W., Raymond, M. O., and Serra, L. (2006). Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science* 313, 1773–1775. doi: 10.1126/science.1131002

Botero, C. A., Weissing, F. J., Wright, J., and Rubenstein, D. R. (2015). Evolutionary tipping points in the capacity to adapt to environmental change. *Proc. Natl. Acad. Sci. U.S.A*. 112, 184–189. doi: 10.1073/pnas.1408589111

Buckley, L. B., and Huey, R. B. (2016). How extreme temperatures impact organisms and the evolution of their thermal tolerance. *Integr. Comp. Biol*. 56, 98–109. doi: 10.1093/icb/icw004

David, J. R., Araripe, L. O., Chakir, M., Legout, H., Lemos, B., Pétavy, G., et al. (2005). Male sterility at extreme temperatures: a significant but neglected phenomenon for understanding *Drosophila* climatic adaptations. *J. Evol. Biol*. 18, 838–846. doi: 10.1111/j.1420-9101.2005.00914.x

Dey, S., Proulx, S. R., and Teotónio, H. (2016). Adaptation to temporally fluctuating environments by the evolution of maternal effects. *PLoS Biol*. 14:e1002388. doi: 10.1371/journal.pbio.1002388

Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends Ecol. Evol*. 29, 51–63. doi: 10.1016/j.tree.2013.09.008

Gunderson, A. R., and Stillman, J. H. (2015). Plasticity in thermal tolerance has limited potential to buffer ectotherms from global warming. *Proc. R. Soc. B* 282:20150401. doi: 10.1098/rspb.2015.0401

Herrando-Pérez, S., Ferri-Yáñez, F., Monasterio, C., Beukema, W., Gomes, V., Belliure, J., et al. (2019). Intraspecific variation in lizard heat tolerance alters estimates of climate impact. *J. Anim. Ecol*. 88, 247–257. doi: 10.1111/1365-2656.12914

Herrando-Pérez, S., Monasterio, C., Beukema, W., Gomes, V., Ferri-Yáñez, F., Vieites, D. R., et al. (2020). Heat tolerance is more variable than cold tolerance across species of Iberian lizards after controlling for intraspecific variation. *Funct. Ecol*. 34, 631–645. doi: 10.1111/1365-2435.13507

Hoffmann, A. A., Chown, S. L., and Clusella-Trullas, S. (2013). Upper thermal limits in terrestrial ectotherms: how constrained are they? *Funct. Ecol*. 27, 934–949. doi: 10.1111/j.1365-2435.2012.02036.x

Holt, R. D. (1990). The microevolutionary consequences of climate change. *Trends Ecol. Evol*. 5, 311–315. doi: 10.1016/0169-5347(90)90088-U

Kreyling, J., Puechmaille, S. J., Malyshev, A. V., and Valladares, F. (2019). Phenotypic plasticity closely linked to climate at origin and resulting in increased mortality under warming and frost stress in a common grass. *Ecol. Evol*. 9, 1344–1352. doi: 10.1002/ece3.4848

Kristensen, T. N., Overgaard, J., Lassen, J., Hoffmann, A. A., and Sgrò, C. (2015). Low evolutionary potential for egg-to-adult viability in *Drosophila melanogaster* at high temperatures. *Evolution* 69, 803–814. doi: 10.1111/evo.12617

Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., et al. (2012). The *Drosophila melanogaster* genetic reference panel. *Nature* 482, 173–178. doi: 10.1038/nature10811

Mallard, F., Nolte, V., Tobler, R., Kapun, M., and Schlötterer, C. (2018). A simple genetic basis of adaptation to a novel thermal environment results in complex metabolic rewiring in *Drosophila*. *Genome Biol*. 19:119. doi: 10.1186/s13059-018-1503-4

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst*. 37: 637–669. doi: 10.1146/annurev.ecolsys.37.091305.110100

Porcelli, D., Butlin, R. K., Gaston, K. J., Joly, D., and Snook, R. R. (2015). The environmental genomics of metazoan thermal adaptation. *Heredity* 114, 502–514. doi: 10.1038/hdy.2014.119

Rezende, E. L., Balanyà, J., Rodríguez-Trelles, F., Rego, C., Fragata, I., Matos, M., et al. (2010). Climate change and chromosomal inversions in *Drosophila subobscura*. *Clim. Res*. 43, 103–114. doi: 10.3354/cr00869

Rodríguez-Trelles, F.,Tarrío, R., and Santos, M. (2013). Genome-wide evolutionary response to a heat wave in *Drosophila*. *Biol. Lett*. 9:20130228. doi: 10.1098/ rsbl.2013.0228

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nat. Rev. Genet*. 14, 807–820. doi: 10.1038/nrg3522

Somero, G. N. (2012). The physiology of global change: linking patterns to mechanisms. *Annu. Rev. Mar. Sci*. 4, 39–61. doi: 10.1146/annurev-marine-120710-100935

Sunday, J. M., Bates, A. E., Kearney, M. R., Colwell, R. K., Dulvy, N. K., Longino, J. T., et al. (2014). Thermal-safety margins and the necessity of thermoregulatory behavior across latitude and elevation. *Proc. Natl. Acad. Sci. U.S.A*. 111, 5610–5615. doi: 10.1073/pnas.1316145111

Walsh, B. S., Parratt, S. R., Hoffmann, A. A., Atkinson, D., Snook, R. R., Bretman, A., et al. (2019). The impact of climate change on fertility. *Trends Ecol. Evol*. 34, 249–259. doi: 10.1016/j.tree.2018.12.002

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 May 2020
doi: 10.3389/fgene.2020.00482

Long-Term Dynamics Among *Wolbachia* Strains During Thermal Adaptation of Their *Drosophila melanogaster* Hosts

*Rupert Mazzucco[1], Viola Nolte[1], Thapasya Vijayan[1,2] and Christian Schlötterer[1]**

[1] *Institut für Populationsgenetik, Veterinärmedizinische Universität Wien, Wien, Austria,* [2] *Vienna Graduate School of Population Genetics, Vienna, Austria*

Climate change is a major evolutionary force triggering thermal adaptation in a broad range of species. While the consequences of global warming are being studied for an increasing number of species, limited attention has been given to the evolutionary dynamics of endosymbionts in response to climate change. Here, we address this question by studying the dynamics of *Wolbachia*, a well-studied endosymbiont of *Drosophila melanogaster*. *D. melanogaster* populations infected with 13 different *Wolbachia* strains were exposed to novel hot and cold laboratory environments for up to 180 generations. The short-term dynamics suggested a temperature-related fitness difference resulting in the increase of clade V strains in the cold environment only. Our long-term analysis now uncovers that clade V dominates in all replicates after generation 60 irrespective of temperature treatment. We propose that adaptation of the *Drosophila* host to either temperature or *Drosophila* C virus (DCV) infection are the cause of the replicated, temporally non-concordant *Wolbachia* dynamics. Our study provides an interesting case demonstrating that even simple, well-controlled experiments can result in complex, but repeatable evolutionary dynamics, thus providing a cautionary note on too simple interpretations on the impact of climate change.

Keywords: experimental evolution, evolve-and-resequence, strain frequency, copy number, microecology, microbe–host interaction, DCV resistance

OPEN ACCESS

Edited by:
Margarida Matos,
University of Lisbon, Portugal

Reviewed by:
Perran Ross,
The University of Melbourne, Australia
Pieter Arnold,
The Australian National University, Australia

***Correspondence:**
Christian Schlötterer
Christian.Schloetterer@vetmeduni.ac.at

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *15 November 2019*
Accepted: *17 April 2020*
Published: *14 May 2020*

Citation:
Mazzucco R, Nolte V, Vijayan T and Schlötterer C (2020) Long-Term Dynamics Among Wolbachia Strains During Thermal Adaptation of Their Drosophila melanogaster Hosts. Front. Genet. 11:482. doi: 10.3389/fgene.2020.00482

INTRODUCTION

The global change in climate imposes strong pressure on many species to deal with increasing temperatures (Thuiller et al., 2005; Cheung et al., 2009; Klausmeyer and Shaw, 2009)—using either mitigation strategies (e.g., shifts in range or activity periods; Davis, 2001; Menzel et al., 2006; Chen et al., 2011), or through genetic changes (e.g., thermal adaptation; Calosi et al., 2008; Marshall et al., 2010; Somero, 2010; Hoffmann and Sgrò, 2011). In the presence of endosymbiotic bacteria, adaptation to temperature could occur by genetic changes either in the host or bacteria—even co-evolutionary processes between both of them could contribute to thermal adaptation.

Wolbachia are intracellular α-Proteobacteria found in many insect and other arthropod species (Baldo et al., 2006; Mateos et al., 2006; Werren et al., 2008), infecting about two thirds of all insects (Hilgenboecker et al., 2008; Miller, 2013). They are predominantly transmitted through the female germline and often confer fitness advantages; e.g., virus protection (Hedges et al., 2008;

Chrostek et al., 2013; Faria et al., 2018), learning ability (Bi et al., 2018), increased fecundity (Fry et al., 2004), resistance to heat stress (Gruntenko et al., 2017), and influence longevity (Maistrenko et al., 2016). On the other hand, *Wolbachia* frequently also imposes considerable costs on its host through the reduction in effective population size by male-killing (Hurst et al., 1999), feminization of genetic males (Rigaut, 1997) and cytoplasmic incompatibility (Bourtzis et al., 1996; Hoffman and Turelli, 1997). In addition to fitness effects of *Wolbachia* on its host, the fitness of the infected host and the probability of vertical transmission also affect the fitness of *Wolbachia*. Among the factors contributing to these fitness components are temperature (Jia et al., 2009; Bordenstein and Bordenstein, 2011), bacterial density in the host (Breeuwer and Werren, 1993; Bourtzis et al., 1996; Noda et al., 2001), and the genetic background of the host (Olsen et al., 2001; Reynolds and Hoffmann, 2002; Fry et al., 2004).

Multiple strains—sometimes several supergroups—of *Wolbachia* may compete within a host population (Dean et al., 2003; Mouton et al., 2003). While coinfection and thus competition within single hosts has been described (Fleury et al., 2000; Hiroki et al., 2004; Ant and Sinkins, 2018), competition mainly occurs between hosts. The relative fitness of multiple *Wolbachia* strains can be measured by the spread of the fitter strain(s) in sexual populations. A particularly interesting question is how the fitness of different *Wolbachia* strains is affected by the environment.

A pioneering study used experimental evolution to study temperature adaptation by exposing a replicated polymorphic *Drosophila melanogaster* population infected by multiple *Wolbachia* strains to two different temperature regimes (Versace et al., 2014). The dynamics of *Wolbachia* infection were monitored by clade-specific SNPs in Pool-Seq data (Schlötterer et al., 2014) from up to four replicates at multiple time points in hot and cold temperature regimes. The striking result was that in the cold temperature regime *Wolbachia* from a single clade (V) very rapidly predominated. Even in hot-evolved replicates that were shifted to the cold temperature regime the same *Wolbachia* clade V dominated. This consistent association of clade V with cold temperatures was considered strong support for environmentally triggered fitness differences between *Wolbachia* strains. Here, we extend the previous work by characterizing the *Wolbachia* dynamics on the level of individual strains rather than clades and our analyses cover substantially more generations in more replicates.

MATERIALS AND METHODS

Drosophila melanogaster Population and Culture Conditions

We reanalyze an evolve-and-resequence experiment (Turner et al., 2011) for which allele frequency changes in *D. melanogaster* (Orozco-terWengel et al., 2012; Tobler et al., 2014; Franssen et al., 2015), and *Wolbachia* strain turnover during the first 50 generations were reported (Versace et al., 2014); detailed descriptions of the experimental setup can be found there. Briefly, 10 replicate populations each with approximately 1000 individuals were created from 113 *D. melanogaster* isofemale lines collected in Portugal and were subsequently kept in two different temperature regimes: five replicates in a hot environment fluctuating between 18 and 28°C, and five replicates in a cold environment fluctuating between 10 and 20°C. Of the 113 isofemale lines, 47 were known to carry *Wolbachia*. The 10 replicates at generation 0 are considered as the base population.

The previous datasets included only time points from the early phase of the experimental evolution cages. Here, we extend the analyses to advanced phases of the experiment by including additional time points and replicates for both the hot and the cold evolved populations: while Orozco-terWengel et al. (2012) and Versace et al. (2014) analyzed up to three replicates in the hot evolved populations until generation F37, we now include data for up to five replicates at multiple earlier and later time points until generation F180 in the hot environment. For the cold environment, Versace et al. (2014) analyzed four replicates in generation F15. Here, we add the fifth replicate and multiple time points up to generation F100 (**Supplementary Datasheet S2**).

Sequencing and Postprocessing

Single females of each of the 47 isofemale lines infected with *Wolbachia* were sequenced individually (2 × 100 bp; ∼10–30× autosome coverage). The infection status of the isofemale lines had been previously determined using the protocol described below (section "Confirmation of *Wolbachia* Infection Status"). Pools of flies (Kofler et al., 2011; Schlötterer et al., 2014) were sequenced at different time points over the course of the experiment, including three replicates of the base population (∼500 flies per generation and replicate; paired-end; ∼30× autosome coverage; various read lengths, library preparation protocols, providers, and sequencing platforms following the development of Illumina sequencing over a decade; **Supplementary Datasheet S2**).

Reads were trimmed with ReadTools v1.2.1 (Gómez-Sánchez and Schlötterer, 2017; parameters: –disable5pTrim –mottQualityThreshold 20 –minReadLength 34); mapped with novoalign v3.08 (Novocraft, 2018; parameters: -r RANDOM) and bwa v0.7.17 (Li, 2013; parameters: mem) using our standard DistMap pipeline (Pandey and Schlötterer, 2013) against the combined reference genome of *D. melanogaster* v6.03 (Thurmond et al., 2019), *w*Mel (AE017196.1), and common gut bacteria (Petkau et al., 2016; *Acetobacter pasteurianus*, AP011170.1; *Lactobacillus brevis*, CP000416.1; *Lactobacillus plantarum*, CP013753.1); filtered for quality and overlap with the *w*Mel genome or the mtDNA genome with samtools v1.9 (Li et al., 2009; parameters: -f0x02 -q 5); and had duplicates removed with picard v2.12.1-SNAPSHOT (The Broad Institute, 2018; parameters: MarkDuplicates REMOVE_DUPLICATES = true).

Variant Calling and Marker Sites

Variants in the 47 sequenced individuals were called using freebayes v1.2.0 (Garrison and Marth, 2012; parameters: -p2 –pooled-discrete) using the alignments of novoalign and bwa jointly to account for the mapper-specific influence of insert

size differences on SNP calling (Kofler et al., 2016). Among the resulting variants, we selected SNPs that only occurred in a true subset of the 47 samples, i.e., allow to discern among strains, and met minimum coverage and quality criteria using bcftools v1.9 [Li, 2011; parameters: -i 'TYPE = "SNP" & INFO/DP < 3*mean(DP) & NS = 47 & NUMALT = 1 & QUAL > 40 & MQM > 50 & MQM/MQMR > 4/5 & MQM/MQMR < 5/4 & RPL/RPR > 1/3 & RPL/RPR < 3 & SAF/SAR > 1/3 & SAF/SAR < 3 & SRF/SRR > 1/3 & SRF/SRR < 3' -e 'FORMAT/GT! = "hom,"' where "mean(DP)" in the first condition is the mean depth of all 47 samples over all sites calculated beforehand, and separately for *Wolbachia* and mtDNA contigs], leaving us with 197 high-quality marker SNPs to discern among the *Wolbachia* strains present in the 47 infected founder lines (**Supplementary Datasheet S3**: markers_wmel.vcf.gz) and 29 marker SNPs to discern among the mitochondrial clades (**Supplementary Datasheet S3**: markers_mtDNA.vcf.gz). Not all SNPs are equally informative, being shared by two or more strains.

Strain and Clade Identification

Based on the polymorphic sites (markers), we distinguished 13 *Wolbachia* strains (**Figure 1**), 10 of which have private SNPs. Comparing the markers to the strain-specific SNPs identified previously (Versace et al., 2014), we identify the same overall clade structure with some additional, previously unresolved, fine structure. Accordingly, we continue using the same naming convention. This assignment is fully consistent for *Wolbachia* and mitochondria strains to clades based on the SNPs provided in Richardson et al. (2012), where the clade structure was originally established.

Estimation of Strain Frequencies

SNP frequencies in all Pool-Seq samples at the previously detected marker sites were called with freebayes (parameters: -F 0.01 -C 2 –pooled-continuous). Given markers for n strains and a Pool-Seq sample with reference allele frequencies b at m marker sites, we estimate the n-vector of corresponding strain frequencies x by minimizing $w|Ax - b|$ subject to constraints $0 < x_j < 1$ and $\sum x_j = 1$ with the conjugate-gradient method, in which A is an $m \times n$-matrix with columns containing 1 where the corresponding strain is marked by the reference allele and 0 otherwise, b an m-vector of called reference allele frequencies, and w an m-vector containing the coverage depth at the marker sites, serving as weights to the minimization procedure.

Ideally, only SNPs private to a strain (i.e., that have a multiplicity of 1) should be used to estimate the frequency of a given strain as the median over the frequencies of the strain-specific SNPs. However, private SNPs are not available for all strains. We thus use the most informative subset of marker sites large enough to differentiate among all strains, obtained by iteratively removing the marker sites with the highest multiplicity until the Shannon entropy (Shannon, 1948) per matrix row is maximized. This retains 75 of the 197 SNPs that differentiate *Wolbachia* strains, 120 of the 180 SNPs that differentiate clades, and all 155 SNPs that differentiate superclades, as well as all 21 mtDNA SNPs that differentiate mtDNA superclades. Since the median is the central point that minimizes the mean absolute deviation, our procedure is equivalent to median estimation when used with only private SNPs and equal weights (Stepniak, 2016). Code is provided in **Supplementary Datasheet S4**.

Confirmation of *Wolbachia* Infection Status

While Versace et al. (2014) demonstrated that all flies were infected with *Wolbachia* within less than 37 generations in either temperature regime, the infection could be lost at later generations. We therefore confirmed the infection status at the final generation (F100 in the cold, F180 in the hot environment) via PCR of at least 30 individual male flies per replicate. We extracted DNA using a salting-out procedure (Miller et al., 1988). To determine the *Wolbachia* infection status, we performed PCR using primers wsp81F (5′-TGGTCCAATAAGTGATGAAGAAAC-3′) and wsp691R (5′-AAAAATTAAACGCTACTCCA-3′) (Braig et al., 1998) resulting in a 630 bp fragment of the *Wolbachia wsp* gene. To rule out that the absence of a *wsp* PCR fragment was due to low quality DNA or suboptimal PCR conditions, we chose primers LV125-F (5′-GAGTCGGTTTCCCACAAAG-3′) and LV125-R (5′-GAGCACATCTACGAGTTTCC-3′) to amplify in parallel a 349 bp fragment of *D. melanogaster* DNA in the same PCR reaction. PCR amplifications were performed in 20 µl reaction volumes using 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 10 pmol of each primer, 0.4 U FIREPol *Taq* Polymerase in buffer B (Solis Biodyne, Tartu, Estonia), and ca. 10 ng genomic DNA. PCRs were run under the following conditions: 3 min at 94°C for initial denaturation followed by 32 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 50 s, and a final extension step of 72°C for 7 min. In a few samples with weak or absent amplification of the *wsp* PCR fragment, an additional PCR with *Wolbachia*-specific primers (wMel-clades_fw: 5′-CACTTTTCTGCTGCTGTTATAC-3′, wMel-clades_rv: 5′-AGAGGGTATTTATGGTAGCAAG-3′) was used with the same conditions to verify the presence or absence of *Wolbachia*.

Copy Number Estimation

We estimated *Wolbachia* and mitochondrial copy numbers from the coverage depth of the *Wolbachia*, or mitochondria, genome relative to the coverage depth of the *Drosophila* autosomes to account for read depth heterogeneity among libraries. Since the low GC content results in a systematic underestimation of read coverage, we corrected for GC bias by GC matching: all positions in the reference genome are assigned an effective GC content, defined as the average GC content of a DNA fragment that covers this position, and calculated as a weighted count of GC bases around the focal position, with weights constructed from the estimated read length and insert size distributions. Positions are then binned by GC content. The copy number is obtained as a weighted mean over GC bins of the relative coverage depths on the target contigs (*w*Mel, mtDNA) and normalization contigs (all *Drosophila* autosomes), with weights

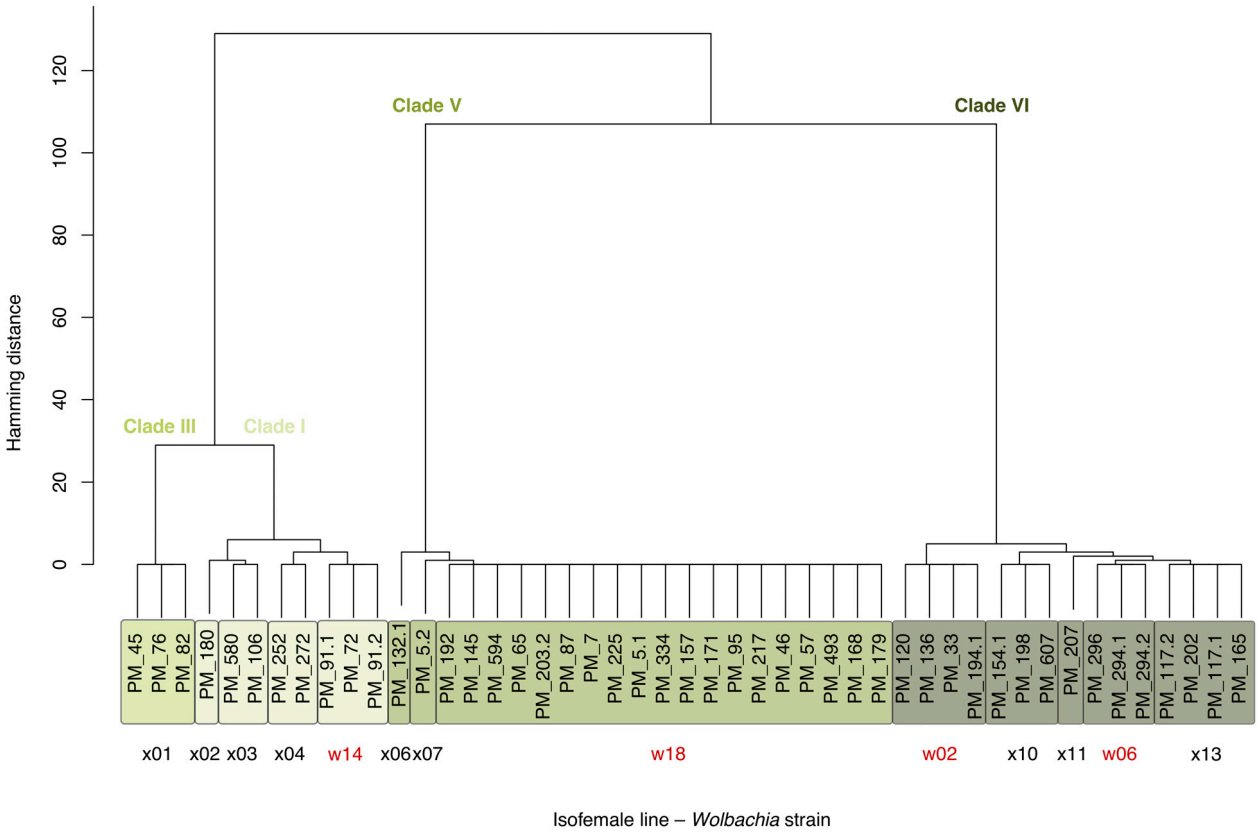


FIGURE 1 | 47 Isofemale lines of the *D. melanogaster* base population carry 13 *Wolbachia* strains representing four clades. The tree shows the isofemale lines clustered by Hamming distance h over 197 high-quality marker SNPs on the *w*Mel genome. The three clusters at $h = 29$ correspond to the *Wolbachia* superclades discussed in Versace et al. (2014), where also strains w02, w06, w14, and w18 (red) were introduced. The four clusters at $h = 6$ (color labels) correspond to the *Wolbachia* clades as defined in Richardson et al. (2012). Clade VI strains resemble the variant also known as *w*MelCS.

$mn/(m+n)$ accounting for the number of positions m on the target contig and n on the normalization contigs within each GC bin (**Supplementary Datasheet S4**). Copy numbers are given as copies per host cell.

RESULTS

Based on 197 informative polymorphisms, we distinguished 13 distinct *Wolbachia* strains from 47 *Wolbachia*-infected isofemale lines. These 13 strains cluster into three major groups and can be assigned to clades I, III, V, and VI as defined by Richardson et al. (2012). No strain belonging to clade II or clade IV was identified (**Figure 1**).

Previous studies reported variation in copy number between different *Wolbachia* strains (Min and Benzer, 1997; Ijichi et al., 2002; Salzberg et al., 2005; Chrostek et al., 2013). Grouping the 47 samples into four clades confirmed pronounced differences. The highest copy number was seen in clade VI and the lowest in clade III. The copy numbers of clade V and VI, in particular, are clearly differentiated (**Figure 2**).

The relative frequency of the *Wolbachia* strains is obtained from Pool-Seq data by taking advantage of the strain/clade specific SNPs. Nevertheless, this method is not informative about the fraction of flies being infected. Versace et al. (2014) tested individual flies and found that at generation 37, all flies tested were infected. We confirm the infection status after long-term evolution (generations 100 in the cold and 180 in the hot environment). With at least 30 sampled flies from each replicate, we conclude that the infection status did not change between generation 37 and generations 100 or 180 (**Supplementary Table S2**).

Consistent with the results of Versace et al. (2014) we find that in the cold environment, clade V very quickly replaces the other *Wolbachia* clades (**Figure 3**, upper row). In the hot environment, a different pattern is observed. In all five replicates, clade VI is predominant during the first generations, as already noted in Versace et al. (2014). Starting around generation 70, however, the same clade V that dominates in the cold environment replaces the other *Wolbachia* genotypes (**Figure 3**, lower row, **Supplementary Figure S1**, and **Supplementary Table S3**). Owing to the delayed response, the anti-correlation between clade V and clade VI is a bit weaker in the hot environment (**Supplementary Figure S2**). This long-term behavior differs from expectations based on the results of Versace et al. (2014), who only studied the dynamics in the hot environment until generation 37.

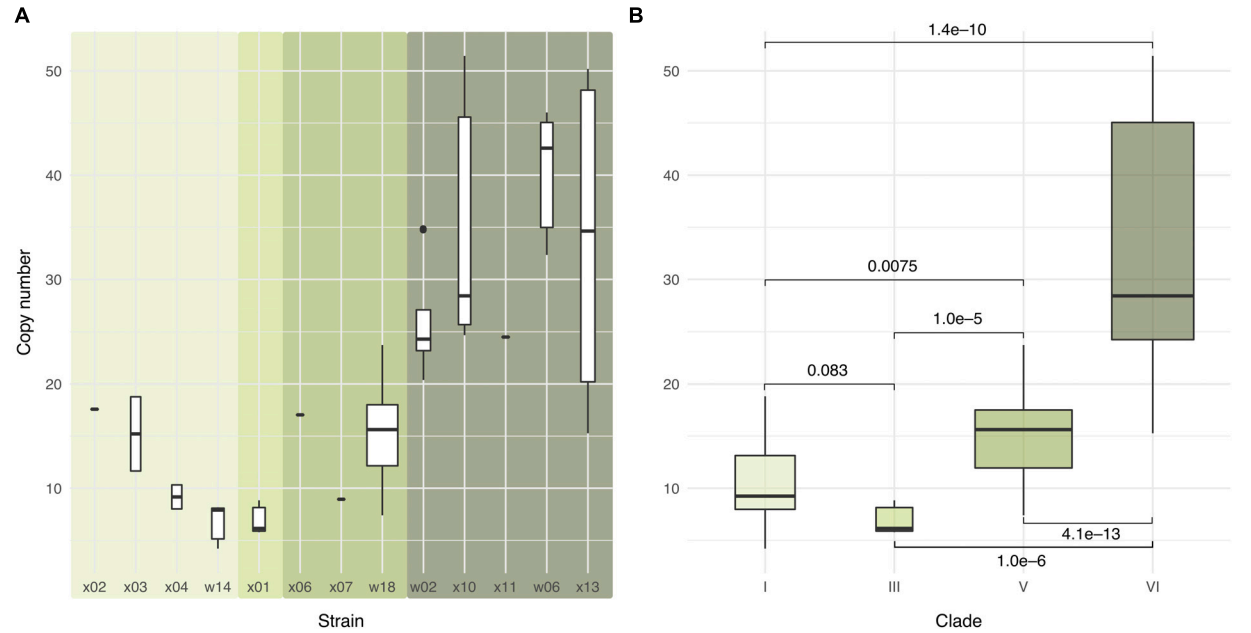


FIGURE 2 | *Wolbachia* copy numbers differ substantially among strains. Box plots in the upper row show *Wolbachia* copy numbers in representative individuals from the 47 isofemale lines **(A)** by strain and **(B)** by clade, with boxes indicating the two middle quartiles of the distribution and whiskers extending to the largest or smallest value no further than 1.5 times the inter-quartile range from the hinge. Background shading in **(A)** indicates the clades depicted in **(B)**. Brackets labels in **(B)** show the *p*-values from a two-sided Wilcox test for the null hypothesis of the same mean copy number (**Supplementary Table S1**, also for strains).

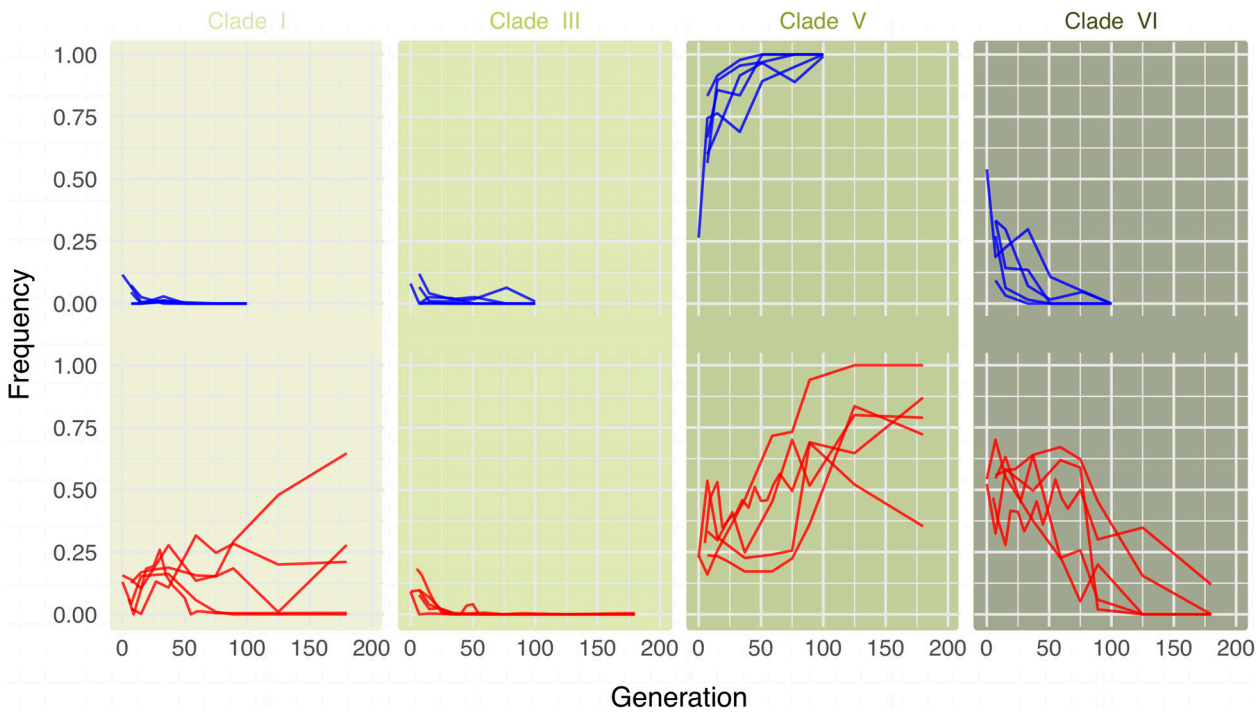


FIGURE 3 | Long-term dynamics reveal that the same *Wolbachia* clades are successful in both hot and cold environments. Lines show clade frequencies over time for five replicates each in either the cold (upper row; blue) or the hot (lower row; red) environment, based on the most informative 120 of the 180 marker SNPs differentiating among clades. Clade V is consistently more successful in the long term in either environment.

The dynamics of *Wolbachia* strain turnover are also reflected in the mean *Wolbachia* coverage at the different time points. In the hot environment, we first observe an increase in coverage, which reflects mainly the increasing infection frequency. After generation 37, the copy number drops, reflecting the taking over of low-copy-number strains (**Figure 4** and **Supplementary Figure S3**). In the cold, the infection frequency also increases, but unlike the hot environment, this does not result in a higher coverage, because the low-copy number strains predominate already at the early generations. Rather, we notice first a drop in coverage, followed by a recovery of the coverage as the entire population becomes infected with *Wolbachia* until generation 33 (Versace et al., 2014).

Given the high consistency of the phylogenetic relationship of mtDNA and *Wolbachia* seen in 290 *melanogaster* lines (Richardson et al., 2012), we expected that the *Wolbachia* dynamics are mirrored by the mtDNA dynamics. A direct comparison is, however, complicated by the lower number of SNPs in the mtDNA, resulting in a lower resolution and more noise. Therefore, we compared the dynamics on the level of super clades, as defined by Versace et al. (2014), which combines clade I, II, and III. Consistent with our expectation, we find an excellent overall correlation between *Wolbachia* and mitochondria (**Supplementary Figure S4**). We attribute the minor deviations to difficulties with an unambiguous clade assignment, rather than biological differences (**Figure 5**).

Like *Wolbachia* copy number, we also evaluated whether mtDNA copy numbers change during the experiment. Unlike *Wolbachia*, the mtDNA copy number is very stable. This observation is fully consistent with the very similar copy numbers in all strains analyzed (**Figure 6**).

DISCUSSION

Compared to Versace et al. (2014), this study covers three advancements. First, we increased the number of replicates and show that the results of Versace were robust. Second, we provide a full SNP catalog of all *Wolbachia* strains in the experiment. This analysis showed that multiple different strains contribute to the clade specific dynamics previously described. Hence, *Wolbachia* strains belonging to the same clade are behaving similarly in their evolutionary response. Third, we increase the number of generations by more than fourfold. While the long-term dynamics in the cold environment do not change, we notice an interesting difference in the hot environment. While during the first generations the turnover of *Wolbachia* genotypes is rather modest, at later generations the same clade V that predominates in the cold cage outcompetes all other ones in the hot cage.

The temporal inconsistency of the evolutionary response cannot be explained by stochastic changes, as it is observed in all five replicates—albeit with different dynamics. The dynamics of the *Wolbachia* strains is not consistent with temperature being the only factor determining the frequency of the *Wolbachia* strains in the evolving replicates. Because the temperature regime is not changing over time, a consistent trend would have been expected. We observe, however, that the early phase differs from later time points. In the following, we will discuss some scenarios, which may explain the repeatable pattern of temporal heterogeneous *Wolbachia* dynamics.

The first hypothesis is that some uncontrolled environmental variables have changed during the experiment. This may include slight modifications in the food, due to different suppliers or modification in the maintenance protocol. As evolved flies are more fecund than ancestral ones from the base population (e.g., Barghi et al., 2019), egg laying time was reduced and larval density may also have changed.

The second hypothesis assumes that adaptation of the host affected the dynamics of the different *Wolbachia* strains, as has been shown for other stocks maintained in the laboratory (Correa and Ballard, 2014). Given that our experiment was designed to study the impact of temperature, adaptation of *Drosophila* to the new temperature regime may explain the dynamics. In a similar experiment, *Drosophila simulans* has been shown to have phenotypically converged at generation 60 (Barghi et al., 2019). Thus, it may be possible that different *Wolbachia* strains may be favored before and after the flies reaching trait optimum.

The third hypothesis is motivated by the observation that the evolved populations sometimes showed symptoms that are typical hallmarks of *Drosophila* C virus (DCV) infection (black, elongated, dying larvae and pupae; Ashburner and Roote, 2007). We propose that the dynamics may relate to the impact of *Wolbachia* copy number on host fitness in the presence of the DCV. A high *Wolbachia* copy number has been shown to be favorable in DCV infected flies when clade VI was compared to clades I, II, and III (Chrostek et al., 2013). Consistent with this, in a DCV-challenged population infected with clades I, II, III, and V, *Wolbachia* of the clade with the higher copy number (V) increased relative to clade I/III, but in the control population, no change was observed (Faria et al., 2016). A particularly interesting feature of our evolving populations is that, for the first time, two high-copy clades, V and VI, can be directly compared against each other. Because in our experiment clade VI has the highest copy number (**Figure 2**), this *Wolbachia* strain should provide the highest protection against DCV, but it is outcompeted by clade V. We attribute this apparent discrepancy to the fitness costs caused by high *Wolbachia* copy numbers (Fleury et al., 2000; Fry et al., 2004; Zhukova and Kiseleva, 2012; Chrostek and Teixeira, 2015; Martinez et al., 2015).

Another explanation for the increase of clade V is that if the *Drosophila* host responds to this DCV challenge by developing resistance, the advantage of the high copy *Wolbachia* may be diminished and another *Wolbachia* strain with lower copy number may take over. We addressed this hypothesis and analyzed the dynamics of two sequence variants, which confer DCV resistance in *Drosophila* (Martins et al., 2014) as an indicator of the resistance level of the *Drosophila* host. The resistance allele of the *pastrel* locus occurs at very low frequencies only and does not respond during the experiment (**Supplementary Figure S5**). The second resistance allele increases in some replicates, but not in all. While it is possible that other DCV resistance loci contribute, we do not have strong evidence for the *Drosophila* host developing DCV resistance during the experiment. This implies that if the evolving

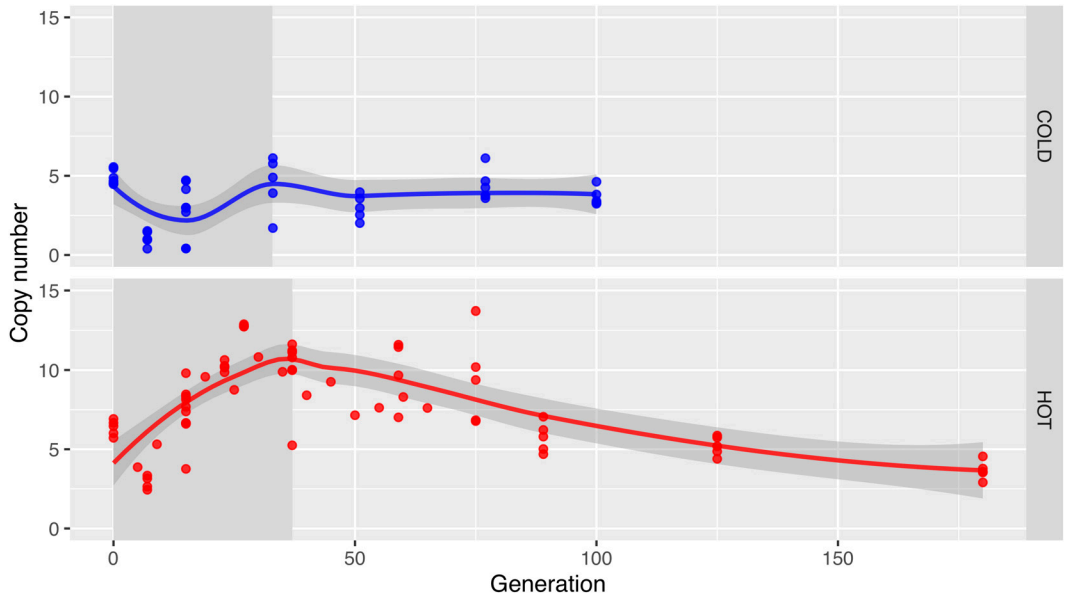


FIGURE 4 | In the long term, *Wolbachia* copy numbers decrease in the hot environment. Dots show the *Wolbachia* copy numbers in replicates estimated with GC matching from the relative coverage depths in cold (upper row; blue) and hot (lower row; red) environments, the lines visualize the main trend obtained by Loess-smoothing. Only a fraction of the population was infected before generation 33 in the cold environment or generation 37 in the hot environment (gray background shading).

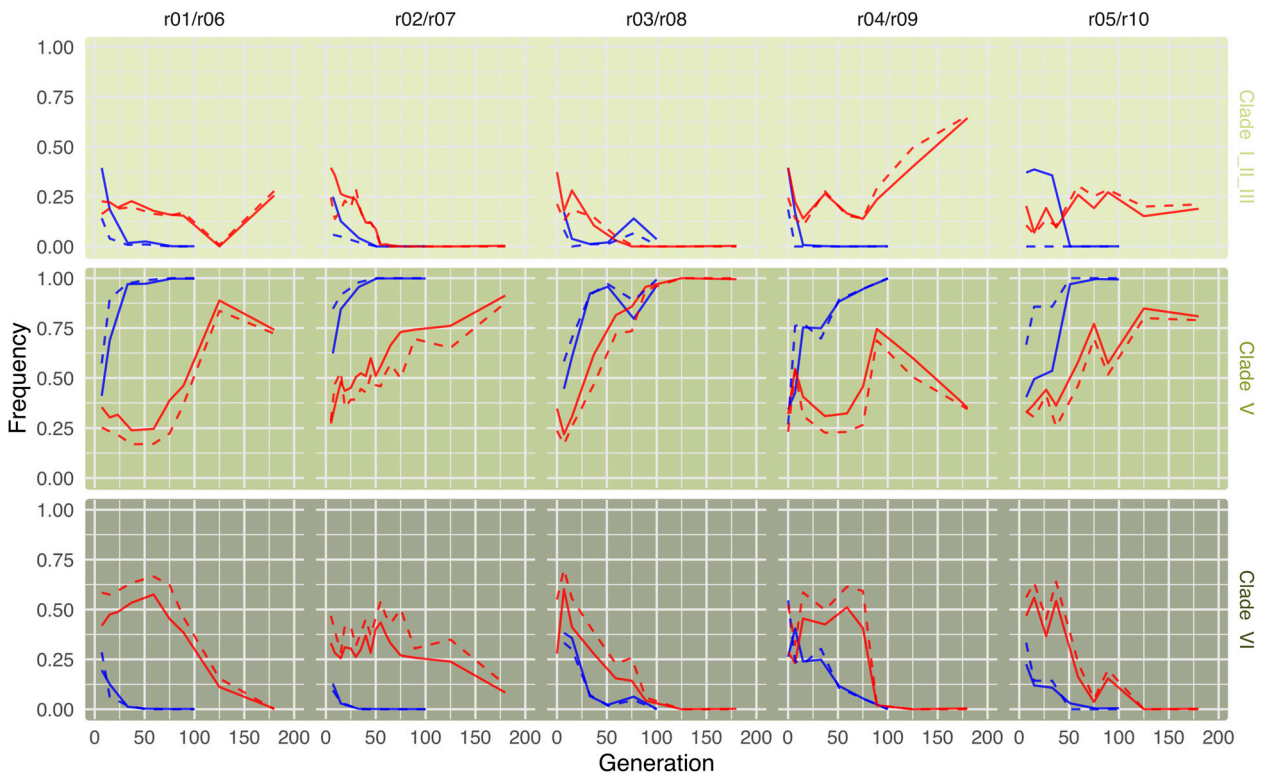


FIGURE 5 | Mitochondrial dynamics on the superclade level are consistent with *Wolbachia* dynamics. Lines indicated superclade (as defined by Versace et al., 2014) frequencies for both mitochondria (solid lines) and *Wolbachia* (dashed lines) in cold (blue) and hot (red) environments. While frequencies are based on 155 marker SNPs differentiating among the three superclades, mitochondria frequencies are based on only 21 differentiating marker SNPs (owing to their smaller genome). The grouping of two replicates from different temperature treatments in a panel is random. All replicates evolved independently.

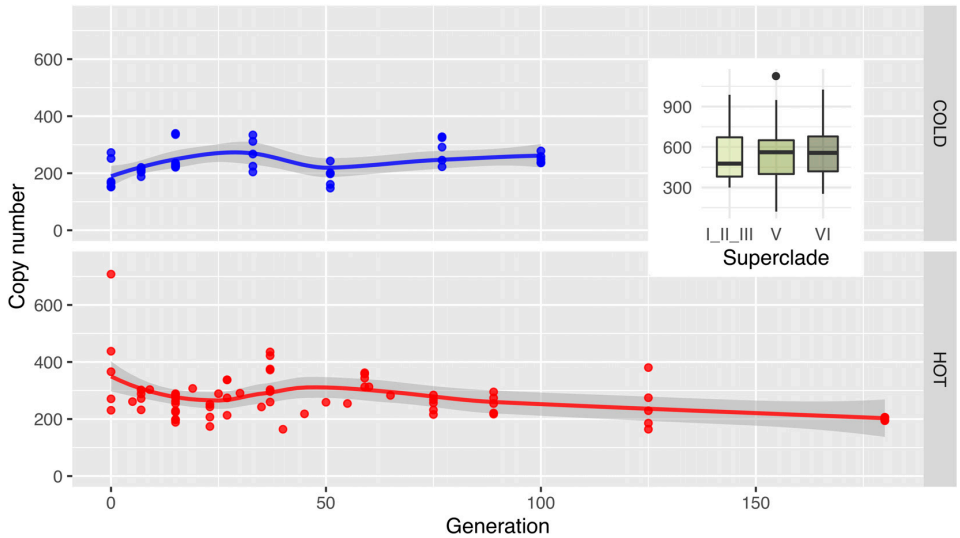


FIGURE 6 | Mitochondrial copy numbers remain constant. Dots show the mitochondrial copy number for all replicates in cold (upper row; blue) and hot (lower row; red) environments, while the lines visualize the main trend via Loess-smoothing. Copy numbers were estimated with GC matching from the relative coverage depths. Boxplots in the insert show the copy numbers of mitochondria in the isofemale lines on superclade level (compare **Figure 2**; *p*-values from a Wilcox test in **Supplementary Table S1**).

populations were challenged by DCV and developed strategies against DCV, this has been mainly achieved by changing the *Wolbachia* strain composition—a hypothesis that could be experimentally tested in future studies.

Finally, as a fourth hypothesis, *Wolbachia* may have adapted to their new environment by the acquisition of new mutations. We consider this highly unlikely because in all five replicates the same three *Wolbachia* strains increased in frequency in the hot environment at the later generations. This would require that all three, highly similar, strains independently acquired new mutations providing a fitness advantage. Furthermore, one would need to find additional explanations, such as epistasis, for the observation that only a single clade increases in frequency at the later generations. Finally, we did not detect new mutations in these strains that could explain the increase in fitness (data not shown).

Independent of the actual cause for the changes in *Wolbachia* dynamics, our study demonstrated that long-term experimental evolution may uncover evolutionary dynamics that remain unnoticed in short-term experiments. Particularly interesting would be further work to illuminate the influence of the host genotype on the observed *Wolbachia* dynamics.

As the hot environment was found to have short- and long-term dynamics, our experiments also highlight the difficulty in making predictions about the impact of temperature changes, thus providing a cautionary note on too simple interpretations on the impact of climate change.

DATA AVAILABILITY STATEMENT

All short-read data used in this study are available from the European Nucleotide Archive (PRJEB37761), but have in parts also been made available earlier (see **Supplementary Datasheet S2** for details). Variant data are included as **Supplementary Datasheet S3**.

AUTHOR CONTRIBUTIONS

CS designed the study. VN conducted the experiments. TV performed a subset of the PCR measurements. RM performed the analysis. RM and CS wrote the manuscript. All authors contributed to manuscript revision and approved the submitted version.

FUNDING

This work was supported by the Austrian Science Funds (FWF, P27630 and W1225) and the European Research Council (ERC, ArchAdapt). Open-access funding was provided by the Austrian Science Fund.

ACKNOWLEDGMENTS

Illumina sequencing for a subset of the data was performed at the VBCF NGS Unit (www.viennabiocenter.org/facilities).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00482/full#supplementary-material

REFERENCES

Ant, T. H., and Sinkins, S. P. (2018). A *Wolbachia* triple-strain infection generates self-incompatibility in *Aedes albopictus* and transmission instability in *Aedes aegypti*. *Parasit. Vectors* 11:295. doi: 10.1186/s13071-018-2870-0

Ashburner, M., and Roote, J. (2007). Maintenance of a *Drosophila* laboratory: general procedures. *Cold Spring Harb. Protoc.* 2007:pdb.ip35. doi: 10.1101/pdb.ip35

Baldo, L., Dunning Hotopp, J. C., Jolley, K. A., Bordenstein, S. R., Biber, S. A., Choudhury, R. R., et al. (2006). Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl. Environ. Microbiol.* 72, 7098–7110. doi: 10.1128/AEM.00731-06

Barghi, N., Tobler, R., Nolte, V., Jakšiæ, A. M., Mallard, F., Otte, K. A., et al. (2019). Genetic redundancy fuels polygenic adaptation in *Drosophila*. *PLoS Biol.* 17:e3000128. doi: 10.1371/journal.pbio.3000128

Bi, J., Zheng, Y., Wang, R.-F., Ai, H., Haynes, P. R., Brownlie, J. C., et al. (2018). *Wolbachia* infection may improve learning and memory capacity of *Drosophila* by altering host gene expression through microRNA. *Insect Biochem. Mol. Biol.* 106, 47–54. doi: 10.1016/j.ibmb.2018.11.007

Bordenstein, S. R., and Bordenstein, S. R. (2011). Temperature affects the tripartite interactions between bacteriophage WO, *Wolbachia*, and cytoplasmic incompatibility. *PLoS One* 6:e29106. doi: 10.1371/journal.pone.0029106

Bourtzis, K., Nirgianaki, A., Markakis, G., and Savakis, C. (1996). *Wolbachia* infection and cytoplasmic incompatibility in *Drosophila* species. *Genetics* 144, 1063–1073.

Braig, H. R., Zhou, W., Dobson, S. L., and O'Neill, S. L. (1998). Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. *J. Bacteriol.* 180, 2373–2378.

Breeuwer, J. A. J., and Werren, J. H. (1993). Cytoplasmic incompatibility and bacterial density in *Nasonia vitripennis*. *Genetics* 135, 565–574.

Calosi, P., Bilton, D. T., and Spicer, J. I. (2008). Thermal tolerance, acclimatory capacity and vulnerability to global climate change. *Biol. Lett.* 4, 99–102. doi: 10.1098/rsbl.2007.0408

Chen, I.-C., Hill, J. K., Ohlemuller, R., Roy, D. B., and Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science* 333, 1024–1026. doi: 10.1126/science.1206432

Cheung, W. W. L., Lam, V. W. Y., Sarmiento, J. L., Kearney, K., Watson, R., and Pauly, D. (2009). Projecting global marine biodiversity impacts under climate change scenarios. *Fish Fish.* 10, 235–251. doi: 10.1111/j.1467-2979.2008.00315.x

Chrostek, E., Marialva, M. S. P., Esteves, S. S., Weinert, L. A., Martinez, J., Jiggins, F. M., et al. (2013). *Wolbachia* variants induce differential protection to viruses in *Drosophila melanogaster*: a phenotypic and phylogenomic analysis. *PLoS Genet.* 9:e1003896. doi: 10.1371/journal.pgen.1003896

Chrostek, E., and Teixeira, L. (2015). Mutualism breakdown by amplification of *Wolbachia* genes. *PLoS Biol.* 13:e1002065. doi: 10.1371/journal.pbio.1002065

Correa, C. C., and Ballard, J. W. O. (2014). What can symbiont titres tell us about co-evolution of *Wolbachia* and their host? *J. Invertebr. Pathol.* 118, 20–27. doi: 10.1016/j.jip.2014.02.009

Davis, M. B. (2001). Range shifts and adaptive responses to quaternary climate change. *Science* 292, 673–679. doi: 10.1126/science.292.5517.673

Dean, M. D., Ballard, K. J., Glass, A., and Ballard, J. W. O. (2003). Influence of two *Wolbachia* strains on population structure of East African *Drosophila simulans*. *Genetics* 165, 1959–1969.

Faria, V. G., Martins, N. E., Magalhães, S., Paulo, T. F., Nolte, V., Schlötterer, C., et al. (2016). Drosophila adaptation to viral infection through defensive symbiont evolution. *PLoS Genet.* 12:e1006297. doi: 10.1371/journal.pgen.1006297

Faria, V. G., Martins, N. E., Schlötterer, C., and Sucena, É (2018). Readapting to DCV infection without *Wolbachia*: frequency changes of *Drosophila* antiviral alleles can replace endosymbiont protection. *Genome Biol. Evol.* 10, 1783–1791. doi: 10.1093/gbe/evy137

Fleury, F., Vavre, F., Ris, N., Fouillet, P., and Boulétreau, M. (2000). Physiological cost induced by the maternally-transmitted endosymbiont *Wolbachia* in the *Drosophila* parasitoid *Leptopilina heterotoma*. *Parasitology* 121, 493–500. doi: 10.1017/S0031182099006599

Franssen, S. U., Nolte, V., Tobler, R., and Schlötterer, C. (2015). Patterns of linkage disequilibrium and long range hitchhiking in evolving experimental *Drosophila melanogaster* populations. *Mol. Biol. Evol.* 32, 495–509. doi: 10.1093/molbev/msu320

Fry, A. J., Palmer, M. R., and Rand, D. M. (2004). Variable fitness effects of *Wolbachia* infection in *Drosophila melanogaster*. *Heredity* 93, 379–389. doi: 10.1038/sj.hdy.6800514

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *ArXiv* [Preprint]. arXiv:1207.3907.

Gómez-Sánchez, D., and Schlötterer, C. (2017). *ReadTools*: a universal toolkit for handling sequence data from different sequencing platforms. *Mol. Ecol. Resour.* 18, 676–680. doi: 10.1111/1755-0998.12741

Gruntenko, N. Å, Ilinsky, Y. Y., Adonyeva, N. V., Burdina, E. V., Bykov, R. A., Menshanov, P. N., et al. (2017). Various *Wolbachia* genotypes differently influence host *Drosophila* dopamine metabolism and survival under heat stress conditions. *BMC Evol. Biol.* 17:252. doi: 10.1186/s12862-017-1104-y

Hedges, L. M., Brownlie, J. C., O'neill, S. L., and Johnson, K. N. (2008). *Wolbachia* and virus protection in insects. *Science* 322, 702–702.

Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A., and Werren, J. H. (2008). How many species are infected with *Wolbachia*? – a statistical analysis of current data: *Wolbachia* infection rates. *FEMS Microbiol. Lett.* 281, 215–220. doi: 10.1111/j.1574-6968.2008.01110.x

Hiroki, M., Tagami, Y., Miura, K., and Kato, Y. (2004). Multiple infection with *Wolbachia* inducing different reproductive manipulations in the butterfly *Eurema hecabe*. *Proc. R. Soc. Lond. B Biol. Sci.* 271, 1751–1755. doi: 10.1098/rspb.2004.2769

Hoffman, A. A., and Turelli, M. (1997). "Cytoplasmic incompatibility in insects," in *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction*, eds S. L. O'Neill, A. A. Hoffmann, and J. H. Werren (Oxford: Oxford University Press).

Hoffmann, A. A., and Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature* 470, 479–485. doi: 10.1038/nature09670

Hurst, G. D. D., Jiggins, F. M., Schulenburg, J. H. G., von der, Bertrand, D., West, S. A., Goriacheva, I. I., et al. (1999). Male-killing *Wolbachia* in two species of insect. *Proc. R. Soc. Lond. B Biol. Sci.* 266, 735–740. doi: 10.1098/rspb.1999.0698

Ijichi, N., Kondo, N., Matsumoto, R., Shimada, M., Ishikawa, H., and Fukatsu, T. (2002). Internal spatiotemporal population dynamics of infection with three *Wolbachia* strains in the Adzuki bean beetle, *Callosobruchus chinensis* (Coleoptera: Bruchidae). *Appl. Environ. Microbiol.* 68, 4074–4080. doi: 10.1128/AEM.68.8.4074-4080.2002

Jia, F.-X., Yang, M.-S., Yang, W.-J., and Wang, J.-J. (2009). Influence of continuous high temperature conditions on *Wolbachia* infection frequency and the fitness of *Liposcelis tricolor* (Psocoptera: Liposcelididae). *Environ. Entomol.* 38, 1365–1372. doi: 10.1603/022.038.0503

Klausmeyer, K. R., and Shaw, M. R. (2009). Climate change, habitat loss, protected areas and the climate adaptation potential of species in mediterranean ecosystems worldwide. *PLoS One* 4:e6392. doi: 10.1371/journal.pone.0006392

Kofler, R., Langmüller, A. M., Nouhaud, P., Otte, K. A., and Schlötterer, C. (2016). Suitability of different mapping algorithms for genome-wide polymorphism scans with pool-seq data. *G3 (Bethesda)* 6, 3507–3515. doi: 10.1534/g3.116.034488

Kofler, R., Orozco-terWengel, P., De Maio, N., Pandey, R. V., Nolte, V., Futschik, A., et al. (2011). PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS One* 6:e15925. doi: 10.1371/journal.pone.0015925

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv* [Preprint]. arXiv:1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Maistrenko, O. M., Serga, S. V., Vaiserman, A. M., and Kozeretska, I. A. (2016). Longevity-modulating effects of symbiosis: insights from *Drosophila–Wolbachia* interaction. *Biogerontology* 17, 785–803. doi: 10.1007/s10522-016-9653-9

Marshall, D. J., McQuaid, C. D., and Williams, G. A. (2010). Non-climatic thermal adaptation: implications for species' responses to climate warming. *Biol. Lett.* 6, 669–673. doi: 10.1098/rsbl.2010.0233

Martinez, J., Ok, S., Smith, S., Snoeck, K., Day, J. P., and Jiggins, F. M. (2015). Should symbionts be nice or selfish? Antiviral effects of *Wolbachia* are costly but reproductive parasitism is not. *PLoS Pathog.* 11:e1005021. doi: 10.1371/journal.ppat.1005021

Martins, N. E., Faria, V. G., Nolte, V., Schlotterer, C., Teixeira, L., Sucena, E., et al. (2014). Host adaptation to viruses relies on few genes with different cross-resistance properties. *Proc. Natl. Acad. Sci. U.S.A.* 111, 5938–5943. doi: 10.1073/pnas.1400378111

Mateos, M., Castrezana, S. J., Nankivell, B. J., Estes, A. M., Markow, T. A., and Moran, N. A. (2006). Heritable endosymbionts of *Drosophila*. *Genetics* 174, 363–376. doi: 10.1534/genetics.106.058818

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., et al. (2006). European phenological response to climate change matches the warming pattern. *Glob. Change Biol.* 12, 1969–1976. doi: 10.1111/j.1365-2486.2006.01193.x

Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16, 1215–1215. doi: 10.1093/nar/16.3.1215

Miller, W. J. (2013). Bugs in transition: the dynamic world of *Wolbachia* in insects. *PLoS Genet.* 9:e1004069. doi: 10.1371/journal.pgen.1004069

Min, K.-T., and Benzer, S. (1997). *Wolbachia*, normally a symbiont of *Drosophila*, can be virulent, causing degeneration and early death. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10792–10796. doi: 10.1073/pnas.94.20.10792

Mouton, L., Henri, H., Bouletreau, M., and Vavre, F. (2003). Strain-specific regulation of intracellular *Wolbachia* density in multiply infected insects. *Mol. Ecol.* 12, 3459–3465. doi: 10.1046/j.1365-294X.2003.02015.x

Noda, H., Koizumi, Y., Zhang, Q., and Deng, K. (2001). Infection density of *Wolbachia* and incompatibility level in two planthopper species, *Laodelphax striatellus* and *Sogatella furcifera*. *Insect Biochem. Mol. Biol.* 31, 727–737. doi: 10.1016/S0965-1748(00)00180-6

Novocraft (2018). *Novoalign. Powerful Tool Designed for Mapping of Short Reads Onto a Reference Genome*. Available online at: http://www.novocraft.com/products/novoalign/ (accessed June 17, 2019).

Olsen, K., Reynolds, K. T., and Hoffmann, A. A. (2001). A field cage test of the effects of the endosymbiont *Wolbachia* on *Drosophila melanogaster*. *Heredity* 86, 731–737.

Orozco-terWengel, P., Kapun, M., Nolte, V., Kofler, R., Flatt, T., and Schlötterer, C. (2012). Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. *Mol. Ecol.* 21, 4931–4941. doi: 10.1111/j.1365-294X.2012.05673.x

Pandey, R. V., and Schlötterer, C. (2013). DistMap: a toolkit for distributed short read mapping on a hadoop cluster. *PLoS One* 8:e72614. doi: 10.1371/journal.pone.0072614

Petkau, K., Fast, D., Duggal, A., and Foley, E. (2016). Comparative evaluation of the genomes of three common *Drosophila* -associated bacteria. *Biol. Open* 5, 1305–1316. doi: 10.1242/bio.017673

Reynolds, K. T., and Hoffmann, A. A. (2002). Male age, host effects and the weak expression or non-expression of cytoplasmic incompatibility in *Drosophila* strains infected by maternally transmitted *Wolbachia*. *Genet. Res.* 80, 79–87. doi: 10.1017/S0016672302005827

Richardson, M. F., Weinert, L. A., Welch, J. J., Linheiro, R. S., Magwire, M. M., Jiggins, F. M., et al. (2012). Population genomics of the *Wolbachia* endosymbiont in *Drosophila melanogaster*. *PLoS Genet.* 8:e1003129. doi: 10.1371/journal.pgen.1003129

Rigaut, T. (1997). "Inherited microorganisms and sex determination of arthropod hosts," in *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction*, eds S. L. O'Neill, A. A. Hoffmann, and J. H. Werren (Oxford: Oxford University Press), 81–108.

Salzberg, S. L., Hotopp, J. C. D., Delcher, A. L., Pop, M., Smith, D. R., Eisen, M. B., et al. (2005). Serendipitous discovery of *Wolbachia* genomes in multiple *Drosophila* species. *Genome Biol.* 6:R23.

Schlötterer, C., Tobler, R., Kofler, R., and Nolte, V. (2014). Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. *Nat. Rev. Genet.* 15, 749–763. doi: 10.1038/nrg3803

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Somero, G. N. (2010). The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine "winners" and "losers." *J. Exp. Biol.* 213, 912–920. doi: 10.1242/jeb.037473

Stepniak, C. (2016). Interesting properties of variation characteristics. *Aust. J. Stat.* 39:341. doi: 10.17713/ajs.v39i4.254

The Broad Institute (2018). *Picard. Picard*. Available online at: http://broadinstitute.github.io/picard/ (accessed June 17, 2019).

Thuiller, W., Lavorel, S., Araujo, M. B., Sykes, M. T., and Prentice, I. C. (2005). Climate change threats to plant diversity in Europe. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8245–8250. doi: 10.1073/pnas.0409902102

Thurmond, J., Goodman, J. L., Strelets, V. B., Attrill, H., Gramates, L. S., Marygold, S. J., et al. (2019). FlyBase 2.0: the next generation. *Nucleic Acids Res.* 47, D759–D765. doi: 10.1093/nar/gky1003

Tobler, R., Franssen, S. U., Kofler, R., Orozco-terWengel, P., Nolte, V., Hermisson, J., et al. (2014). Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. *Mol. Biol. Evol.* 31, 364–375. doi: 10.1093/molbev/mst205

Turner, T. L., Stewart, A. D., Fields, A. T., Rice, W. R., and Tarone, A. M. (2011). Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet.* 7:e1001336. doi: 10.1371/journal.pgen.1001336

Versace, E., Nolte, V., Pandey, R. V., Tobler, R., and Schlötterer, C. (2014). Experimental evolution reveals habitat-specific fitness dynamics among *Wolbachia* clades in *Drosophila melanogaster*. *Mol. Ecol.* 23, 802–814. doi: 10.1111/mec.12643

Werren, J. H., Baldo, L., and Clark, M. E. (2008). *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* 6, 741–751. doi: 10.1038/nrmicro1969

Zhukova, M. V., and Kiseleva, E. (2012). The virulent *Wolbachia* strain *w*MelPop increases the frequency of apoptosis in the female germline cells of *Drosophila melanogaster*. *BMC Microbiol.* 12:S15. doi: 10.1186/1471-2180-12-S1-S15

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 03 June 2020
doi: 10.3389/fgene.2020.00573

Phenotypic Responses to and Genetic Architecture of Sterility Following Exposure to Sub-Lethal Temperature During Development

Martyna K. Zwoinska[1], Leonor R. Rodrigues[1], Jon Slate[2] and Rhonda R. Snook[1]**

[1] Department of Zoology, Stockholm University, Stockholm, Sweden, [2] Department of Animal and Plant Sciences, University of Sheffield, Sheffield, United Kingdom

OPEN ACCESS

Edited by:
Mauro Santos, Autonomous University of Barcelona, Spain

Reviewed by:
Francesc Mestres, University of Barcelona, Spain
Amitabh Joshi, Jawaharlal Nehru Centre for Advanced Scientific Research, India

***Correspondence:**
Martyna K. Zwoinska
zwoinska@gmail.com;
martyna.zwoinska@zoologi.su.se
Rhonda R. Snook
rhonda.snook@zoologi.su.se

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *04 March 2020*
Accepted: *11 May 2020*
Published: *03 June 2020*

Citation:
Zwoinska MK, Rodrigues LR, Slate J and Snook RR (2020) Phenotypic Responses to and Genetic Architecture of Sterility Following Exposure to Sub-Lethal Temperature During Development. Front. Genet. 11:573. doi: 10.3389/fgene.2020.00573

Thermal tolerance range, based on temperatures that result in incapacitating effects, influences species' distributions and has been used to predict species' response to increasing temperature. Reproductive performance may also be negatively affected at less extreme temperatures, but such sublethal heat-induced sterility has been relatively ignored in studies addressing the potential effects of, and ability of species' to respond to, predicted climate warming. The few studies examining the link between increased temperature and reproductive performance typically focus on adults, although effects can vary between life history stages. Here we assessed how sublethal heat stress during development impacted subsequent adult fertility and its plasticity, both of which can provide the raw material for evolutionary responses to increased temperature. We quantified phenotypic and genetic variation in fertility of *Drosophila melanogaster* reared at standardized densities in three temperatures (25, 27, and 29°C) from a set of lines of the *Drosophila* Genetic Reference Panel (DGRP). We found little phenotypic variation at the two lower temperatures with more variation at the highest temperature and for plasticity. Males were more affected than females. Despite reasonably large broad-sense heritabilities, a genome-wide association study found little evidence for additive genetic variance and no genetic variants were robustly linked with reproductive performance at specific temperatures or for phenotypic plasticity. We compared results on heat-induced male sterility with other DGRP results on relevant fitness traits measured after abiotic stress and found an association between male susceptibility to sterility and male lifespan reduction following oxidative stress. Our results suggest that sublethal stress during development has profound negative consequences on male adult reproduction, but despite phenotypic variation in a population for this response, there is limited evolutionary potential, either through adaptation to a specific developmental temperature or plasticity in response to developmental heat-induced sterility.

Keywords: climate change, heat stress, thermal fertility limits, heat-induced male sterility, *Drosophila* Genetic Reference Panel, *Drosophila melanogaster*, GWAS, phenotypic plasticity

INTRODUCTION

An increase in mean temperatures and temperature variation associated with ongoing climate change threatens biodiversity (Pachauri et al., 2015). Ectotherms play a critical role in ecosystem functioning (Weisser and Siemann, 2008) and can be particularly vulnerable to the effects of climate change because their physiology and biochemistry depend directly upon ambient temperatures (Hochachka and Somero, 2002; Deutsch et al., 2008). Climate change risk assessments are frequently based on quantification of thermal parameters (Deutsch et al., 2008; Sinclair et al., 2016; Kellermann and van Heerwaarden, 2019), such as thermal tolerance (e.g., either critical tolerance or lethal temperatures, such as lower temperatures (CTmin) and higher temperatures (CTmax), representing a species lower and upper operational temperature), and thermal performance curves, e.g., reaction norms in which individuals are exposed to different temperatures until performance fails at CTmin and CTmax. These parameters are associated with latitudinal species' range distributions (Addo-Bediako et al., 2000; Kellermann et al., 2012; Overgaard et al., 2014). For ectotherms, thermal performance is skewed such that performance drops sharply at increasing, but not decreasing, temperatures. Upper critical thermal limits of terrestrial ectotherms show considerably less geographical variation than lower limits (Addo-Bediako et al., 2000; Deutsch et al., 2008; Kellermann et al., 2012) and many species are thought to operate close to their upper performance limits (Huey et al., 2012; van Heerwaarden et al., 2016). Although some *Drosophila* species have latitudinal clines of CTmax (Castañeda et al., 2015; O'Brien et al., 2017), suggesting the ability to locally adapt to varying temperatures, most evidence on the evolutionary potential for increasing heat tolerance (e.g., shifting critical thermal maximum) suggests limited genetic variability to respond to selection (Castañeda et al., 2019; Kellermann and van Heerwaarden, 2019). Phenotypic plasticity of thermal tolerance parameters, such as CTmax, may be critical to species persistence but many species appear to have a small capacity to shift CTmax via phenotypic plasticity (Sørensen et al., 2016; Kellermann and Sgrò, 2018). Because the capacity for adaptation to climate warming will depend on the underlying genetic architecture and the extent to which adaptation and plasticity contributes to responses to climate warming, these patterns indicate much concern about the consequences of a warming climate on ectotherm species' distributions and persistence.

Thermal performance measures used in these analyses commonly are based on performance proxies of survival, such as when respiration or movement stop, or death occurs (Kellermann and van Heerwaarden, 2019; Walsh et al., 2019b). However, reproductive performance, such as fertility, can be negatively affected by temperatures that are neither incapacitating nor lethal (Jørgensen et al., 2006; Austin et al., 2013; Kjærsgaard et al., 2013; Manenti et al., 2014; Kingsolver et al., 2015; Porcelli et al., 2017; Sales et al., 2018; Saxon et al., 2018). This raises concerns over whether predictions for species' responses to increased temperature based on critical thermal limits alone may be too conservative (Walsh et al., 2019b). Complementary to studies of critical thermal limits, knowledge of thermal fertility limits, the reproductive equivalent of critical thermal limits, is necessary to assess the extent of the problem. Yet, few studies have systematically determined either the upper temperatures at which reproduction fails or described the thermal fertility reaction norm within a population. Likewise, to our knowledge, there have been no studies determining the underlying genetic architecture of this response.

While most studies predicting species' response to climate change incorporate only data from the adult stage, thermal sensitivity may vary across different life cycle stages (Kingsolver et al., 2011; Sinclair et al., 2016; Moghadam et al., 2019). Sublethal but stressful temperatures experienced during juvenile development of *Drosophila subobscura* resulted in fertility loss whereas keeping adults at the same temperature had no negative fertility effect. Likewise, in *Drosophila melanogaster,* the effect of brief high temperature exposure on survival varied across life history stages with adaptive hardening (i.e., previously briefly exposed to high temperatures) more pronounced at juvenile stages (Moghadam et al., 2019). This result suggests that *D. melanogaster* juveniles exhibit higher plasticity in response to temperatures than adults, who can rely to a larger degree on behavioral responses.

Male reproductive performance is thought to be affected by temperature to a greater degree than female reproductive performance because spermatogenesis, which in many insects starts during the juvenile period (Nijhout, 1998), is more thermally sensitive than oogenesis (Setchell, 1998; David et al., 2005; Hansen, 2009). Heat stress experienced during development can render males either temporarily or permanently sterile (Chakir et al., 2002; Araripe et al., 2004; Rohmer, 2004; Vollmer et al., 2004; David et al., 2005; Jørgensen et al., 2006; David, 2008; Pedersen et al., 2011; Porcelli et al., 2017). Even when changes to the male reproductive system are reversible, heat stress can have serious negative consequences for short-lived organisms such as many insects (Sinclair and Roberts, 2005). However, few studies have directly addressed sex-specific thermal sensitivity of reproductive performance, particularly following developmental heat stress (Walsh et al., 2019a).

Here we aim to characterize phenotypic and genetic variation in developmentally heat induced sterility and its plasticity. We used genome-sequenced lines from the *Drosophila* Genetic Reference Panel (DGRP), a set of inbred *D. melanogaster* lines (Mackay et al., 2012; Huang et al., 2014), exposing juveniles to three different temperatures and measuring subsequent fertility in the adult stage. We quantified phenotypic variation and examined the correlation with traits measured on different abiotic stressors that have been published using the panel. We determined the genetic architecture of the reproductive traits and performed a genome-wide association study (GWAS) to identify trait-associated genetic variants. We concentrated on males but, for a smaller subset of lines, we also provide phenotypic data on female reproductive performance.

MATERIALS AND METHODS

Fly Stocks and Maintenance

We used isogenic, genome-sequenced lines from the DGRP, initiated from a natural population from Raleigh, North Carolina that underwent 20 generations of full-sib mating (Mackay et al., 2012; Huang et al., 2014). Climate in Raleigh is humid subtropical, characterized by hot and humid summers with average high temperatures reaching ∼32°C Weather-us.com (2020)[1]. We quantified male fertility responses from 127 DGRP lines. Following determination of high and low performing male lines, we then quantified female fertility responses from 40 lines. We standardized the female (or males for female fertility responses) used as mates across our experiments, using a wild-type Canton Special (CS) strain (gift from Dick Nässel, originally obtained from the Bloomington Stock Center). CS and DGRP stock flies were maintained in standard culture vials using cornmeal medium (10 L: 9 L water, 720 *g* cornmeal, 162 *g* dried yeast, 90 *g* soya flour, 720 *g* malt extract, 360 *g* molasses, 72 *g* agar, 36 mL propionic acid, and 225 mL of 10% Nipagin) at 12−h light:12−h dark cycle at 25°C. Mates for experimental individuals were similarly reared at 25°C throughout development, under controlled density conditions (100 eggs/vial), subsequently collected as virgins under CO_2 anesthesia, transferred to vials in groups of about 20, stored at 25°C, and were 3–6 day old when used for experiments.

Responses of Fertility to Developmental Thermal Stress

Males

For focal experimental males of each DGRP line, we standardized egg number by placing 2-week old adult flies onto egg laying media (6 *g* agar, 57.5 *g* bread syrup, and 360 mL of water, seeded with 100 μL yeast paste upon drying) for 2–3 days prior to egg collection at 25°C. In the morning on the day of egg collections, flies were transferred onto fresh egg laying media for 2–4 h, eggs collected onto mesh screen (Snook et al., 1994), and groups of ca. 50 eggs were counted and transferred into vials filled with cornmeal medium. Replicate vials were made for each DGRP, with subsets of vials placed into each control temperature incubator (Panasonic MIR-154) set to 25, 27, or 29°C, 12−h light: 12−h dark cycle. Virgin males from each line were collected under CO_2 anesthesia and transferred into individual vials then stored at 25°C. The day following eclosion, a single virgin control female was added to each vial. Experimental pairs were allowed to interact for 3 days, then removed from vials. Reproductive performance was scored 2 days later as a binomial trait; fertile males produced at least one larva and sterile males did not. We had to run the experiment in blocks. This was due to the large number of DGRP lines assayed (n = 127) and because for each line we simultaneously tested fertility in response to three different temperatures, while strictly controlling egg density in vials generating experimental individuals. We ran 13 blocks, each consisting of 7–11 DGRP lines, with most blocks having 10 lines. To account for block variation, we assessed fertility of CS males at 25°C and 29°C in each block. The mean number of individual males/line/temperature was 24.8 (median 28). See **Supplementary Table S1** for details of the DGRP lines used, number of flies/line/temperature, and trait values.

Females

To determine the consequences of developmental heat stress on female fertility we used a subset (n = 40) of tested male lines. This subset represented the lines performing well across all temperatures ("high lines" – 19 lines) and lines performing poorly as temperature increases ("low lines" – 21 lines). Performance category was based on a variety of considerations, but firstly on males' phenotypic response at 29°C and the slope of response (based on random intercepts and slope model described in the below section "Genome-wide association response variables: temperature-specific reproductive performance and plasticity"), then on other considerations such as the number of replicates contributing to the values, and whether we could obtain sufficient number of individuals before the experiment. We ran the female experiment three times, each consisting of 12–15 DGRP lines averaging 27.39 (median 28) females/line/temperature and CS. See **Supplementary Table S1** for details of the DGRP lines used, number of flies/line/temperature, and trait values.

Statistical Analyses

Average Phenotypic Responses

We categorized the effects of each temperature on reproductive performance as binomial; for each mating pair, the reproductive response was either a success (at least one larva was produced) or a failure (no larva produced), taking into account the number of males tested in each line (function *cbind*, in R; see **Supplementary Table S1** for data). For each line then we get a proportion of males assayed that are fertile. To address the use of this conservative estimate for temperature-induced impacts on fertility, we fitted binomial mixed-effect models with a logit link function using the *lme4* package in R (Bates et al., 2015). The model for male dataset included temperature as a fixed factor and block and DGRP line as random factors [*cbind (Reproduced, Did not reproduce)* ∼ *Temperature* + *(1| Block)* + *(1| DGRP line)*]. The model for female data set had an additional fixed factor, *Line status*, which indicated whether the line was classified as high or low performing based on the reproductive performance of males [*cbind (Reproduced, Did not reproduce)* ∼ *Temperature* × *Line status* + *(1| Block)* + *(1| DGRP line)*]. In these models we included CS to account for variation in each block but model fits with CS included did not perform better than model fits without CS included. This is likely because there was little variation across temperatures in CS performance. Thus, to assess the extent to which DGRP responses were repeatable (and therefore potentially impacted by variation between blocks), we analyzed male data from two subsequent experiments we have run only on the high and low lines, with the experimental design exactly the same as here. That is, we have measured male fertility, under the same conditions, in 40 lines, three different times. We used the *corrplot R* package to obtain a matrix of Pearson correlations and the *Hmisc* package to calculate the p-values of

[1] https://www.weather-us.com/en/north-carolina-usa/raleigh-climate?c,mm,mb,km#temperature

the correlation between fertility measures for DGRP lines across the three different sampling periods. We found consistent results at 27°C and especially at 29°C, with correlations across these experiments at the higher temperature ranging from 0.73 to 0.88; low repeatability at 25°C is likely a consequence of little variation across lines (**Supplementary Table S2**). This analysis confirms responses are repeatable, particularly at the highest temperature which is also the most phenotypically variable across the lines. Experimental block was included as a random factor to help account for the non-independence of observations within a single experimental unit (Harrison et al., 2018). *Wolbachia* and inversion status of the DGRP lines used were fitted as fixed factors but inclusion of all of them caused convergence issues. Fitting them individually returned no significant effect.

Models for each temperature treatment were run separately using block and DGRP line as random factors [*cbind (Reproduced, Did not reproduce)* ∼ *(1| Block)* + *(1| DGRP line)*]. The variance components of line and block, along with the residual variance, assumed to be $\pi^2/3$ (Nakagawa and Schielzeth, 2010), of each of these binomial models was used to calculate broad-sense heritabilities, and such that $H^2 = V_{\text{line}}/(V_{\text{line}} + V_{\text{block}} + V_{\text{resid}})$. We note that although the variance explained by line is assumed to be some form of genetic variance (i.e., additive genetic variance, dominance genetic variance, and epistasis or gene-by-environmental interaction), it is not possible to partition the line variance between these different genetic components. We also note that part of the variance explained by block is likely to include some form of genetic variance but it is not possible to partition the part of block variance that would contribute to genetic components of line variance. Thus, the H^2 estimate is likely conservative. We used the package *lsmeans* (Lenth, 2016) to obtain least squares means for each temperature treatment and converted them from the logit scale to obtain predicted probabilities of reproducing at a given temperature.

Genome-Wide Association Response Variables: Temperature-Specific Reproductive Performance and Plasticity

We initially intended to perform a mixed-model GWAS with binomial response variables using the package GMMAT (Chen et al., 2016; Chen et al., 2019) but sensible results at 25°C and 27°C were not produced, likely because all lines performed almost equally well at these temperatures. Thus, we used line-specific intercepts, as a measurement of temperature-specific reproductive performance, at each temperature and line-specific slopes, as a measure of phenotypic plasticity of reproductive performance. In the random slopes model used to extract line-specific intercepts and slopes, temperature was added as fixed factor, while block and DGRP line were treated as random factors [*cbind (Reproduced, Did not reproduce)* ∼ *Temperature* + *(1| Block)* + *(Temperature| DGRP line)*]. Temperature was a continuous variable in this model and was centered at each temperature treatment (25, 27, or 29°C) to extract treatment-specific intercepts. The random factor of the DGRP line *(Temperature| DGRP line)* allowed for the effect of temperature to vary between the lines and provided the line-specific slope estimate. We extracted the model intercepts and slopes with the *coef* command from the binomial random slope model with a logit link function fitted using the *lme4* package (Bates et al., 2015). The model terms for each line's slope and intercept were continuous, making them more tractable for GWAS than if they were binomial variables (see **Supplementary Table S1** for estimates, **Supplementary Figure S1** for temperature-specific and slope values, and **Supplementary Figure S2** for frequency distribution of values).

SNP Filtering and Quality Control

Quality control of the genomic data was performed in Plink v1.9 (Purcell et al., 2007; Chang et al., 2015). We set the minor allele frequency threshold (MAF) to be at least 5% and we filtered out all variants that were missing in more than 10% of lines (*–geno* in Plink). The rate of genotype missingness (*–mind* in Plink) for each line was set to be not more than 15%, which ensured the retention of all 127 phenotyped lines. A total of 1,465,358 variants were retained after quality control.

SNP-Based Heritability

We estimated the proportion of variance for male phenotypes explained by all genetic variants, the SNP-based heritability (sensu Yang et al., 2010), using the GREML approach implemented in the GCTA software (Yang et al., 2011). GREML uses a genetic relatedness matrix (GRM) to perform a marker-based animal model to measure the proportion of variation explained by additive genetic effects. Here the GRM was created using autosomal markers only (1,230,417 variants). Male traits were phenotypic plasticity of fertility in response to developmental heat stress (slope) and fertility at each temperature (intercept for each temperature).

Genome-Wide Association Analyses

For association tests, we used GMMAT, implemented in R (Chen et al., 2016, 2019). First, we fitted linear mixed models to adjusted phenotype data obtained from the mixed models described above. GWAS phenotypes were the model intercepts (male fertility at each temperature) and slopes (the plasticity of male fertility) with family set to *gaussian* and the link function set to *identity*. *Wolbachia* status of each DGRP line and 5 major inversions present in the DGRP panel [*In(2L)t*, *In(2R)NS*, *In(3R)P*, *In(3R)K*, and *In(3R)Mo*] were included as fixed factors. To account for cryptic genetic relatedness, we fitted a centered GRM, created using the GEMMA package, as a random factor in our model. Only the autosomal markers were used to create this matrix. GMMAT performs the GWAS by adding each SNP, in turn, to the model to test for associations between genotype and phenotype. Because some of the genetic markers will be in linkage disequilibrium with one another, we estimated the effective number of tests (M_e) in the GWAS, using the Genetic Type 1 error calculator (Li et al., 2012). M_e was 722,833 which means that a genome-wide significance threshold at $P < 0.05$ requires a SNP to be nominally significant at $P = 6.92 \times 10^{-8}$.

Correlations With Other DGRP Datasets

We used the *corrplot R* package to obtain and plot a matrix of Pearson correlations and the *Hmisc* package to calculate the *p*-values of each correlation between fertility measures for males and other stress-related traits measured in the DGRP. Because phenotypic variation in fertility was highest among the lines at 29°C, and because the random slope and intercept model corrects for block effects, we compared the intercept of fertility at 29°C against traits associated with other abiotic stressors: chill coma resistance and starvation resistance (Mackay et al., 2012), desiccation tolerance (Rajpurohit et al., 2018), CTmax (Rolandi et al., 2018), and two measures of oxidative stress based on two different oxidative stress-inducing agents, paraquat and menadione sodium bisulfite (MSB) resistance (Weber et al., 2012). Traits were analyzed based on male trait value/line for all comparisons.

RESULTS

Phenotypic Response

Males

Reproductive performance was significantly negatively affected as developmental temperature increased, with 25°C as the least affected, 27°C intermediate, and 29°C the most affected (**Figure 1A** and **Table 1**). Impact of different temperatures resulted in probabilities of reproducing, derived from least square means, of 0.98 at 25°C, 0.97 at 27°C, and 0.55 at 29°C. Block effect explained about 0.26 of variance of the model (σ^2_{Block} divided by $\sigma^2_{\text{Block}} + \sigma^2_{\text{Line}} + \sigma^2_{\text{Residual}}$) and DGRP line explained about 0.20 of variance (σ^2_{Line} divided by $\sigma^2_{\text{Block}} + \sigma^2_{\text{Line}} + \sigma^2_{\text{Residual}}$, **Table 1**). Recall, however, that among high and low performing lines, repeatability of results was ca. 80% at 29°C, suggesting block variance is at least partially a consequence of biological variation in the lines tested in any given block. Broad-sense heritability, determined by variance explained by the DGRP line, differed between temperature treatments but was highest at 29°C where there was the most phenotypic variability among the lines (**Table 2**).

Intercept values, used for GWAS, resulted in similar probabilities of reproducing, based on binomial data; 1 (0.995) at 25°C, 0.94 at 27°C, and 0.59 at 29°C (values on a logit scale based on intercept means of 5.36, 2.87, and 0.38 at 25, 27, and 29°C, respectively; **Supplementary Table S3**). The slope of reproductive performance declines by the proportion of 0.22 across the treatments (value on a logit scale was –1.25; **Supplementary Table S3**). Note that the model for 25°C returned warnings about convergence failure, but generated estimate outputs.

Females

Temperature also significantly affected female reproductive performance, although unlike males, females reproductive performance did not differ between 25°C and 27°C (**Figure 1B** and **Table 3**). Reproductive performance of males from the same line (Line status) was not a significant predictor in the female model ($z = -1.40$, $P = 0.16$) indicating little or no association between male and female fertility in response to thermal stress. Probabilities of reproducing at different temperatures, derived from least square means, were 0.95 for 25°C, 0.93 for 27°C, 0.90 for 29°C, substantially higher at 29°C than for males. Broad-sense heritabilities returned singular fits except for 27°C treatment, estimated to be 0.01.

SNP-Based Heritability

SNP-based heritability analyses in GCTA revealed narrow-sense heritability and additive genetic variance of 0 for all traits analyzed (**Table 2**), although it should be noted that the standard errors around these estimates were quite large, meaning that the possibility of some genetic variance in these traits cannot be ruled out.

Genome-Wide Association Analyses

Standard practice for GWAS analyses in the DGRP panel (Mackay et al., 2012) is to use a nominal *p* value of $P < 1 \times 10^{-5}$ threshold for reporting significant SNPs (indicated as a red line in **Figure 2**). The number of variants meeting this threshold was: two at 25°C, 21 at 27°C, 10 at 29°C, and 13 for plasticity (**Figure 2**). Three variants overlapped between the 29°C and slope analysis (see **Supplementary Table S5** for list of nominally significant variants). Importantly, with 1,465,358 variants analyzed and a *p* value threshold of 1×10^{-5} one would expect ∼15 significantly associated variants by chance alone. Quantile-quantile plots (**Supplementary Figure S3**) further illustrated no enrichment of associations exceeding the $P < 1 \times 10^{-5}$ threshold. No variant passed a more stringent significance threshold, for instance one based on an M_e of 722,833 which is equivalent to a *p* value of 6.92×10^{-8} or -log10(*p*) ∼ 7.16 (indicated as a black line in **Figure 2**). The lowest *p* values were in the range of 1.16×10^{-6}, corresponding to -log10(*p*) ∼ 5.94. Thus there is no statistical support for the claim that any of the variants that passed the $p < 10^{-5}$ threshold represents a true positive finding. It is therefore unsurprising that GO enrichment analysis (Ashburner et al., 2000; The Gene Ontology Consortium, 2019) revealed no significant enrichment for any of the measured phenotypes (not shown).

Comparison With Other DGRP Datasets

There was a significant positive correlation between MSB resistance (survival time) and fertility at 29°C but all other comparisons between male fertility during developmental heat stress across lines and other traits responsive to abiotic stressors in the DGRP were not significant (**Table 4**).

DISCUSSION

Understanding the consequences of increasing temperature on sex-specific fertility effects, and the evolutionary and plastic responses of natural populations to thermal challenges, will help improve predictions for species' persistence. In this study, we determined the impact on adult fertility of sublethal heat stress following developmental exposure to three different temperatures, determined the thermal reaction norm, and assessed genetic architecture of measured traits in a mapped

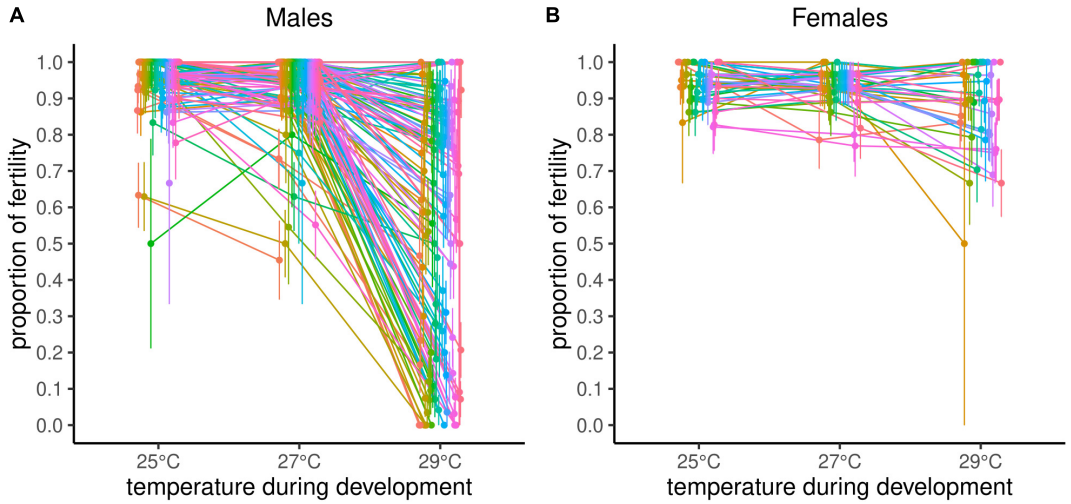


FIGURE 1 | Natural genetic variation in fertility of males **(A)** and females **(B)** at three different developmental temperatures.

TABLE 1 | Generalized mixed-effect model on male reproductive performance (binomial; number of successfully reproducing and unsuccessful females for high and low performing male lines) following developmental heat stress at 25, 27, and 29°C.

	Estimate	Std. Error	*Z* value	*P* value
Fixed effects				
Intercept (Temperature 25°C)	4.16	0.33	12.48	<0.001
Temperature (27°C)	–0.63	0.13	–5.035	<0.001
Temperature (29°C)	–3.96	0.12	–33.058	<0.001
	Parameter	**Variance**	**Std. Dev.**	
Random effects				
Block	Intercept	1.61	1.077	
Line	Intercept	1.21	1.10	
	Estimate	**Std. Error**	***Z* value**	***P* value**
***Post hoc* contrasts**				
29°C–27°C	–3.33	0.10	–32.39	<0.001
Residual variance of the model: 3.29				
Residual deviance: 594.842 on 367 degrees of freedom (ratio: 1.621)				

Post hoc contrast between 27°C and 29°C provided as this is not directly tested in the main model.

population of *D. melanogaster*. We found males were affected to a larger degree than females by higher developmental temperatures, and the difference was particularly striking at the highest temperature. The average male fertility for each DGRP line, for both 29°C and for the slope of fertility across all temperatures, was correlated with previous DGRP results on male survival after oxidative stress (Weber et al., 2012). Despite significant phenotypic variance in male thermal fertility limits at 29°C and in the slope of response across all temperatures, we found little evidence of heritable genetic variation for these reproductive traits. The number of genetic variants significantly associated with the traits analyzed at the nominal p value threshold of 10^{-5} did not exceed what would be expected by chance alone. We discuss our results in light of what may be driving the sterility effect, the genetic architecture of fitness-related traits in light of previous DGRP results, and the impact that temperature-induced sterility may have on population persistence.

Sex-specific thermal sensitivity was observed. Female fertility was not assessed in all DGRP lines that male reproductive performance was measured in as early results suggested females were not as affected. However, we found there was no effect of whether male reproductive performance was either relatively insensitive (high performing), or sensitive (low performing) to developmental temperature on female fertility, indicating male and female reproductive performance in response to developmental heat-stress is unlinked. While both sexes showed increased between-line variation after developing at 29°C, male reproductive performance was affected to a much larger degree, both with respect to estimates of thermal fertility limit and its phenotypic plasticity (slope of the reaction norm), than female reproductive performance. Similar sex-specific results were recently reported following adult heat stress in the red flour beetle *Tribolium castaneum* (Sales et al., 2018).

We speculate the larger male effect is due to the high thermal sensitivity of spermatogenesis, which in many insects starts during development. We previously found that *D. subobscura* males had reduced sperm motility after experiencing sublethal heat stress during development (Porcelli et al., 2017). Heat wave exposure in adult males in *T. castaneum* caused reduced sperm production and sperm viability (Sales et al., 2018). Thus, sublethal heat stress has effects on sperm quantity and quality (Snook, 2005). Intriguingly, we found that DGRP lines that were less sensitive to heat-induced sterility lived longer following exposure to MSB, an oxidative stress-inducing chemical agent. Oxidative stress is strongly linked with the production of reactive oxygen species (ROS) and is considered to be a main cause of male infertility, causing damage

TABLE 2 | Heritability measurements for male fertility following developmental heat stress at 25, 27, and 29°C.

H^2 Estimate	25°C	27°C	29°C	Slope
σ^2 for DGRP line	0.74	0.84	2.20	
σ^2 total	4.03	4.13	5.49	
H^2	0.18	0.20	0.40	
SNP-based heritability	**25°C variance ± SE**	**27°C variance ± SE**	**29°C variance ± SE**	**Slope variance ± SE**
Source				
V_g	0.00 ± 0.58	0.00 ± 0.12	0.00 ± 0.67	0.00 ± 0.11
V_e	2.33 ± 1.20	0.41 ± 0.24	1.83 ± 1.35	0.40 ± 0.23
V_p	2.33 ± 0.65	0.41 ± 0.13	1.83 ± 0.70	0.40 ± 0.12
V_g/V_p	0.00 ± 0.25	0.00 ± 0.28	0.00 ± 0.37	0.00 ± 0.28

*Broad-sense heritabilities (H^2) calculated from between-line variances for each temperature derived from the mixed effect model (see **Table 1**) and SNP-based heritability estimates on the intercepts of each temperature and slope obtained using the GCTA software and the GREML approach.*

TABLE 3 | Generalized mixed-effect model on female reproductive performance (binomial; number of successfully reproducing and unsuccessful females for high and low performing male lines) following developmental heat stress at 25, 27, and 29°C.

	Estimate	Std. Error	*Z* value	*P* value
Fixed effects				
Intercept (Temperature 25°C, High)	3.02	0.23	13.28	<0.001
Temperature (27°C)	–0.19	0.17	–1.12	0.26
Temperature (29°C)	–0.64	0.16	–4.01	<0.001
Line status (Low)	–0.32	0.23	–1.40	0.16
	Parameter	**Variance**	**Std. Dev.**	
Random effects				
Block	Intercept	0.32	0.56	
Line	Intercept	0.04	0.19	
	Estimate	**Std. Error**	***Z* value**	***P* value**
***Post hoc* contrasts**				
29°C–27°C	–0.45	0.15	–2.93	0.003
Residual variance of the model: 3.29				
Residual deviance: 117.983 on 114 degrees of freedom (ratio: 1.035)				

Post hoc contrast between 27°C and 29°C provided as this is not directly tested in the main model.

to sperm membranes that impairs sperm-egg interactions, reduces ejaculate quality, including sperm velocity, and can cause sperm DNA damage that also negatively impacts fertilization (Mora et al., 2017). High temperature increases metabolism and therefore increases ROS production (dos Hamilton et al., 2016) and GWAS indicates an association between SNPs in antioxidant genes and male infertility (Yu and Huang, 2015). Future work will assess directly the relationship between developmental heat-induced sterility, consequences to sperm quantity and quality, and the relationship with ROS in *Drosophila*.

CTmax values have been used to assess species consequences to future warming. Previous work on the DGRP has measured adult CTmax which ranged from ca. 40.1°C to 41.5°C (Rolandi et al., 2018). We found no correlation between developmentally heat-induced sterility in this study and adult CTmax (Rolandi et al., 2018). This may be because there were only 21 lines that overlapped between studies and CTmax of those lines did not vary substantially. Regardless, DGRP CTmax temperatures are substantially higher than temperatures that cause lowered male fertility. Our original experimental design included a 31°C temperature treatment to more completely describe fertility limits in this population but, at this higher temperature, substantial juvenile mortality was observed [matching previous descriptions of other *D. melanogaster* populations from temperate collections; (David et al., 2005)]. The comparison between these studies suggest that developmental heat stress, relative to adult heat stress, could have considerable negative impacts on population persistence. This, however, depends on whether future temperatures are expected to go above 29°C for extended periods of time during juvenile development and from which they cannot escape. Rolandi et al. (2018) compared historic climatic records (1980–2005) from Raleigh North Carolina, where the DGRP originated, and found only 10 days above the adult CTmax, but future climate projections (2045–2070) based on the RCP6.0 emissions scenario predicted an increase to 243 days of extreme high temperatures above CTmax. Together, these results suggest that heat-induced sterility during development occurs at temperatures substantially lower than adult CTmax (i.e., those temperatures used to project species response to climate change), and that future temperature regimes are likely to frequently reach temperatures that result in developmentally-induced sterility. Estimates of fertility here were based on binomial quantification, in which producing one larva would count as a male being fertile. This is a conservative estimate of the consequences of developmental heat stress on adult reproduction. Our impression after assaying ca. 10000 males in this study, and based on an experiment we are currently conducting, is that progeny number is substantially reduced at 29°C, even for lines characterized as being fertile. Thus, it is likely that the potential consequences of exposure to sublethal heat stress during development on adult fertility we document here is conservative.

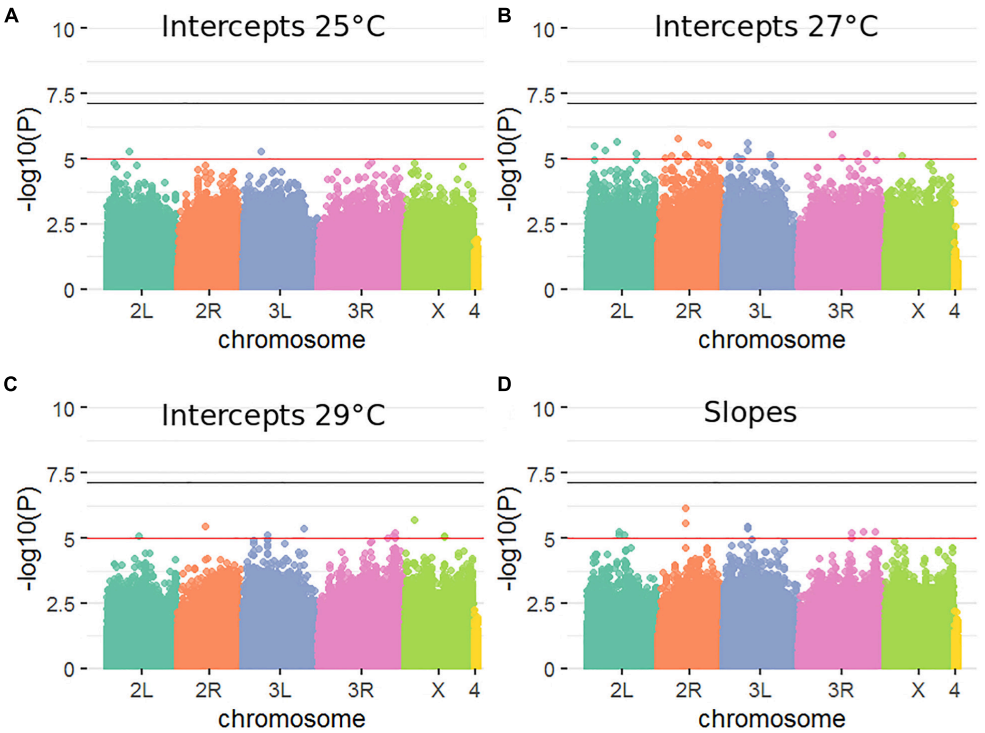


FIGURE 2 | Genetic variants influencing male fertility at three different temperatures and the slope of response. **(A–D)**. Manhattan plots corresponding to four GWAS analyses performed. Horizontal lines are *p*-value = 1×10^{-5} (red) and *p*-value = 6.92×10^{-8} (black), where 1×10^{-5} corresponds to -log(10) of 5 and 6.92×10^{-8} corresponds to -log(10) of 7.16.

TABLE 4 | Phenotypic co-variation between male fertility at 29°C (intercept extracted from the random slopes model, corrected for block effect) and other fitness traits following abiotic stress using Pearson-correlation coefficient.

Environmentally relevant trait	Correlation coefficient	*P*-value	Nr lines
Chill coma recovery	0.132	0.199	96
Starvation resistance	0.084	0.403	102
Desiccation tolerance	0.080	0.413	108
CTmax	− 0.187	0.417	21
Paraquat resistance	0.184	0.064	102
MSB resistance	0.210	0.034	102

The number of lines (nr lines) that overlap between this fertility study and the other variables ranged from 21 to 108. All comparisons were made using data from only males. NB: similar results were found using the fertility slope. Data was obtained from each reference's Supplementary Material: Mackay et al., 2012, Table S20 (Chill Coma Recovery, Starvation Resistance); Rajpurohit et al., 2018, Supplementary Table S3 (Desiccation Tolerance); Rolandi et al., 2018, Supplementary Table S1 (CTmax); and Weber et al., 2012, Supplementary Table S1 (Paraquat and MSB Resistance).

The extent to which the population can respond to temperature selection is critical to determine as this will impact population persistence. Estimates of additive genetic variance and heritability for temperature-specific fertility effects and its phenotypic plasticity do not give cause for optimism. We found little to no additive genetic variance or heritability using SNP-based animal models and no significant SNPs were detected via GWAS. The DGRP can only be used to reliably detect genetic variants of moderate effects, and the mapping power of the panel is considered to be low because of a relatively small number of lines (Turner et al., 2013; Long et al., 2014; Mackay and Huang, 2018). There appears to be epistatic interactions impacting the genetic architecture of quantitative traits in the DGRP (Huang et al., 2012; Shorter et al., 2015); epistatic interactions are not detectable in the analyses we performed as they are only designed to identify additive genetic variation (Yang et al., 2011; Evans et al., 2018). However, the larger estimate of broad-sense heritability compared to narrow-sense heritability suggests that some non-additive genetic variance, possibly due to epistasis, is present. The influence of epistasis on trait expression in these lines has been suggested for several traits (Huang et al., 2012; Shorter et al., 2015; Huang and Mackay, 2016).

Heritability estimates are sensitive to environmental conditions (Hoffmann and Parsons 191). Low heritability estimates for thermal performance traits has been suggested to be a function of the intensity and duration of the thermal treatment (Castañeda et al., 2019). In some studies, increasing the length of the thermal assay lowers heritability, perhaps because additional stress factors (e.g., resource depletion, cellular damage, and dessication resistance), arising under chronic but not acute stress, increase environmental variance (Mitchell and Hoffmann, 2010; Castañeda et al., 2019). However, we

find increased broad-sense heritability at higher temperatures with no correlation between heat-induced sterility and other environmental stress factors such as desiccation resistance that may contribute to environmental variance. Previous reviews have described examples of heritability being greatest in stressful conditions (Hoffmann and Parsons, 1991; Hoffmann and Merilä, 1999; Charmantier and Garant, 2005) and in our experiment genetic variation is revealed at the most stressful temperature of 29°C.

Heritability estimates are also impacted by how close the trait is to fitness. We assayed fertility *per se*, a trait intimately related to fitness. Other life history traits closely linked to fitness exhibit lower narrow-sense heritabilities than morphological or physiological traits (Mousseau and Roff, 1987). While low heritabilities can result from high levels of residual genetic variance, rather than low levels of additive genetic variance *per se* (Houle, 1992), numerous studies have found very low levels of additive genetic variance for fitness, and/or fertility (Kruuk et al., 2000; Teplitsky et al., 2009; McFarlane et al., 2014; Sztepanacz and Blows, 2015; Noble et al., 2017). This includes data on *D. melanogaster* outbred and inbred populations (Hughes, 1995a,b; Snoke and Promislow, 2003). Low to zero additive genetic variation, but high dominance genetic variance, for fitness-linked traits has been found in *D. serrata* (Sztepanacz and Blows, 2015), and a study using a *C. elegans* mapping panel of recombinant inbred lines found estimates of the heritability of fertility to not be significantly different from 0. Instead, around 40% of variance in fertility was explained by epistasis (Noble et al., 2017). The study concluded that numerous small-effect epistatic interactions explained non-additive genetic variation in fitness-related traits in this population (Noble et al., 2017), similar to findings on the genetic architecture of quantitative traits in the DGRP (Huang et al., 2012; Mackay and Huang, 2018).

Genome-wide association study analyses did not identify any SNPs that were genome-wide significant for heat-induced sterility and Q-Q plots did not reveal an excess of nominally significant SNPs at lower thresholds. Many DGRP GWAS papers show evidence for a modest excess of loci with p values below the 1×10^{-5} threshold, suggesting an enrichment of true positive associations (Mackay and Huang, 2018). However, comparisons between genetic variants discovered using the DGRP and other mapping panels or populations rarely reveal overlapping loci (Huang et al., 2012; Swarup et al., 2013; Morozova et al., 2015; Najarro et al., 2015, 2017; Shorter et al., 2015; Carbone et al., 2016; Rajpurohit et al., 2018; Everman et al., 2019). While we report any loci significant at $P < 1 \times 10^{-5}$ in **Supplementary Table S5**, we place a caveat that many or perhaps all of these associations are likely to be false positives.

In summary, we showed that male fertility was less thermally tolerant than female fertility, and that males exhibited within-population variation in the response of fertility to sublethal heat stress during development and in phenotypic plasticity of this response. Lines in which males were susceptible to heat-induced sterility were also more susceptible to oxidative stress and oxidative stress has known negative consequences on sperm quantity and quality. Despite a moderate broad-sense heritability at 29°C, we found no evidence of additive genetic variation although some non-additive genetic variation may be present. Likewise, we observed no genetic variants that could be robustly associated with either temperature-specific fertility consequences, even at the most stressful temperature tested, or its plasticity. Future climate scenarios predict increased likelihood for temperatures that could result in at least portions of the population becoming sterile, at temperatures well below those resulting in reduced performance associated with survival, and our current measure of the impact of developmentally-induced sterility is conservative. Therefore, the impact of thermal fertility limits on population persistence under future climate scenarios will need to be considered to help predict responses to increased temperatures.

DATA AVAILABILITY STATEMENT

All datasets generated and analyzed for this study are included in **Supplementary Table S1**. These data are also published on Figshare: https://su.figshare.com/articles/Zwoinska_et_al_2020_Heat-induced_sterility_data/12248576/1.

AUTHOR CONTRIBUTIONS

RS conceived the experiments. MZ and LR collected the data. The data was analyzed by ZW, JS, and LR. MZ and RS wrote the manuscript with contributions from LR and JS. All authors agreed on the final version of the manuscript.

FUNDING

This study was supported by awards to RS, including a Carl Tryggers Stiftelse (CTS 18:359) postdoctoral scholarship (LR), and funding from Stockholm University Faculty of Science for postdoctoral research (RRS to ZW). ZW also was funded by a Lars Hiertas Minne Stiftelsen grant (FO2018-0190) and a Kungl Vetenskaps Akademien grant (BS2018-0028). JS was funded by a Leverhulme Trust/Royal Society Senior Research Fellowship (SRF/R1/180116).

ACKNOWLEDGMENTS

The authors would like to thank Mirre Simons (University of Sheffield) for the DGRP lines and to Dick Nässel (Stockholm University) for the Canton-S stock.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00573/full#supplementary-material

REFERENCES

Addo-Bediako, A., Chown, S. L., and Gaston, K. J. (2000). Thermal tolerance, climatic variability and latitude. *Proc. R. Soc. B Biol. Scie. B* 267, 739–745. doi: 10.1098/rspb.2000.1065

Araripe, L. O., Klaczko, L. B., Moreteau, B., and David, J. R. (2004). Male sterility thresholds in a tropical cosmopolitan drosophilid. *Zaprionus indianus*. *J. Ther. Biol.* 29, 73–80. doi: 10.1016/j.jtherbio.2003.11.006

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Austin, C. J., Moehring, A. J., and Plaistow, S. (2013). Optimal temperature range of a plastic species. *Drosophila simulans*. *J. Anim. Ecol.* 82, 663–672. doi: 10.1111/1365-2656.12041

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Carbone, M. A., Yamamoto, A., Huang, W., Lyman, R. A., Meadors, T. B., Yamamoto, R., et al. (2016). Genetic architecture of natural variation in visual senescence in *Drosophila*. *Proc. Natl. Acad. Scie. U.S.A.* 113, E6620–E6629.

Castañeda, L. E., Rezende, E. L., and Santos, M. (2015). Heat tolerance in *Drosophila subobscura* along a latitudinal gradient: contrasting patterns between plastic and genetic responses. *Evolution* 69, 2721–2734. doi: 10.1111/evo.12757

Castañeda, L. E., Romero-Soriano, V., Mesas, A., Roff, D. A., and Santos, M. (2019). Evolutionary potential of thermal preference and heat tolerance in *Drosophila subobscura*. *J. Evol. Biol.* 32, 818–824. doi: 10.1111/jeb.13483

Chakir, M., Chafik, A., Moreteau, B., Gibert, P., and David, J. R. (2002). Male sterility thermal thresholds in *Drosophila*: *D. simulans* appears more cold-adapted than its sibling *D. melanogaster*. *Genetica* 114, 195–205.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7. doi: 10.1186/s13742-015-0047-8

Charmantier, A., and Garant, D. (2005). Environmental quality and evolutionary potential: lessons from wild populations. *Proc. R. Soc. B Biol. Sci.* 272, 1415–1425. doi: 10.1098/rspb.2005.3117

Chen, H., Huffmann, J. E., Brody, J. A., Wang, C., Lee, S., Li, Z., et al. (2019). Efficient variant set mixed model association tests for continuous and binary traits in large-scale whole-genome sequencing studies. *Am. J. Hum. Genet.* 104, 260–274. doi: 10.1016/j.ajhg.2018.12.012

Chen, H., Wang, C., Conomos, M. P., Stilp, A. M., Li, Z., Sofer, T., et al. (2016). Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. *Am. J. Hum. Genet.* 98, 653–666. doi: 10.1016/j.ajhg.2016.02.012

David, J. R. (2008). Male sterility at high and low temperatures in *Drosophila*. *J. Soc. Biol.* 202, 113–118. doi: 10.1051/jbio:2008014

David, J. R., Araripe, L. O., Chakir, M., Legout, H., Lemos, B., Petavy, G., et al. (2005). Male sterility at extreme temperatures: a significant but neglected phenomenon for understanding *Drosophila* climatic adaptations. *J. Evol. Biol.* 18, 838–846. doi: 10.1111/j.1420-9101.2005.00914.x

Deutsch, C. A., Tewksbury, J. J., Huey, R. B., Sheldon, K. S., Ghalambor, C. K., Haak, D. C., et al. (2008). Impacts of climate warming on terrestrial ectotherms across latitude. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6668–6672. doi: 10.1073/pnas.0709472105

dos Hamilton, T. R. S., Mendes, C. M., de Castro, L. S., de Assis, P. M., Siqueira, A. F. P., de Delgado, J. C., et al. (2016). Evaluation of lasting effects of heat stress on sperm profile and oxidative status of ram semen and epididymal sperm. *Oxid Med. Cell. Longevity* 2016, 1–12. doi: 10.1155/2016/1687657

Evans, L. M., Rasool, T., Vrieze, S. I., Abecasis, G. R., Das, S., Gazal, S., et al. (2018). Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat. Genet.* 50, 737–745. doi: 10.1038/s41588-018-0108-x

Everman, E. R., McNeil, C. L., Hackett, J. L., Bain, C. L., and Macdonald, S. J. (2019). Dissection of complex, fitness-related traits in multiple *Drosophila* mapping populations offers insight into the genetic control of stress resistance. *Genetics* 211, 1449–1467. doi: 10.1534/genetics.119.301930

Hansen, P. J. (2009). Effects of heat stress on mammalian reproduction. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 364, 3341–3350. doi: 10.1098/rstb.2009.0131

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., et al. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ* 6:e4794. doi: 10.7717/peerj.4794

Hochachka, P. W., and Somero, G. N. (2002). *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. Oxford: Oxford University Press.

Hoffmann, A. A., and Merilä, J. (1999). Heritable variation and evolution under favourable and unfavourable conditions. *Trends Ecol. Evol.* 14, 96–101. doi: 10.1016/s0169-5347(99)01595-5

Hoffmann, A. A., and Parsons, P. A. (1991). *Evolutionary Genetics and Environmental Stress*. Oxford: Oxford University Press.

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics* 130, 195–204.

Huang, W., and Mackay, T. F. C. (2016). The genetic architecture of quantitative traits cannot be inferred from variance component analysis. *PLoS Genet.* 12:e1006421. doi: 10.1371/journal.pgen.1006421

Huang, W., Massouras, A., Inoue, Y., Peiffer, J., Ràmia, M., Tarone, A. M., et al. (2014). Natural variation in genome architecture among 205 *Drosophila melanogaster* genetic reference panel lines. *Genome Res.* 24, 1193–1208. doi: 10.1101/gr.171546.113

Huang, W., Richards, S., Carbone, M. A., Zhu, D., Anholt, R. R. H., Ayroles, J. F., et al. (2012). Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. *Proc. Natl. Acad. Sci. U.S.A.* 109, 15553–15559. doi: 10.1073/pnas.1213423109

Huey, R. B., Kearney, M. R., Krockenberger, A., Holtum, J. A., Jess, M., and Williams, S. E. (2012). Predicting organismal vulnerability to climate warming: roles of behaviour, physiology and adaptation. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 367, 1665–1679. doi: 10.1098/rstb.2012.0005

Hughes, K. A. (1995a). The evolutionary genetics of male life-history characters in *Drosophila melanogaster*. *Evolution* 49, 521–537. doi: 10.1111/j.1558-5646.1995.tb02284.x

Hughes, K. A. (1995b). The inbreeding decline and average dominance of genes affecting male life-history characters in *Drosophila melanogaster*. *Genet. Res. Camb.* 65, 41–52. doi: 10.1017/s0016672300032997

Jørgensen, K. T., Sørensen, J. G., and Bundgaard, J. (2006). Heat tolerance and the effect of mild heat stress on reproductive characters in *Drosophila buzzatii* males. *J. Ther. Biol.* 31, 280–286. doi: 10.1016/j.jtherbio.2005.11.026

Kellermann, V., Overgaard, J., Hoffmann, A. A., Fløjgaard, C., Svenning, J.-C., and Loeschcke, V. (2012). Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16228–16233. doi: 10.1073/pnas.1207553109

Kellermann, V., and Sgrò, C. M. (2018). Evidence for lower plasticity in CTmax at warmer developmental temperatures. *J. Evol. Biol.* 31, 1300–1312. doi: 10.1111/jeb.13303

Kellermann, V., and van Heerwaarden, B. (2019). Terrestrial insects and climate change: adaptive responses in key traits. *Physiol. Entomol.* 44, 99–115. doi: 10.1111/phen.12282

Kingsolver, J. G., Higgins, J. K., and Augustine, K. E. (2015). Fluctuating temperatures and ectotherm growth: distinguishing non-linear and time-dependent effects. *J. Exp. Biol.* 218, 2218–2225. doi: 10.1242/jeb.120733

Kingsolver, J. G., Woods, H. A., Buckley, L. B., Potter, K. A., MacLean, H. J., and Higgins, J. K. (2011). Complex life cycles and the responses of insects to climate change. *Integr. Comp. Biol.* 51, 719–732. doi: 10.1093/icb/icr015

Kjærsgaard, A., Pertoldi, C., Loeschcke, V., and Blanckenhorn, W. U. (2013). The effect of fluctuating temperatures during development on fitness-related traits of *Scatophaga stercoraria* (*Diptera*: *Scathophagidae*). *Environ. Entomol.* 42, 1069–1078. doi: 10.1603/EN13074

Kruuk, L. E. B., Clutton-Brock, T. H., Slate, J., Pemberton, J. M., Brotherstone, S., and Guinness, F. E. (2000). Heritability of fitness in a wild mammal population. *Proc. Natl. Acad. Sci. U.S.A.* 97, 698–703. doi: 10.1073/pnas.97.2.698

Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *J. Stat. Softw.* 69, 1–33.

Li, M. X., Yeung, J. M. Y., Cherny, S. S., and Sham, P. C. (2012). Evaluating the effective numbers of independent tests and significant *p*-value thresholds in

commercial genotyping arrays and public imputation reference datasets. *Hum. Genet.* 131, 747–756. doi: 10.1007/s00439-011-1118-2

Long, A. D., Macdonald, S. J., and King, E. G. (2014). Dissecting complex traits using the *Drosophila* synthetic population resource. *Trends Genet.* 30, 488–495. doi: 10.1016/j.tig.2014.07.009

Mackay, T. F., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., et al. (2012). The *Drosophila melanogaster* genetic reference panel. *Nature* 482, 173–178. doi: 10.1038/nature10811

Mackay, T. F. C., and Huang, W. (2018). Charting the genotype-phenotype map: lessons from the *Drosophila melanogaster* genetic reference panel. *Wiley Interdiscip. Rev. Dev. Biol.* 7:10.1002/wdev.289. doi: 10.1002/wdev.289

Manenti, T., Sørensen, J. G., Moghadam, N. N., and Loeschcke, V. (2014). Predictability rather than amplitude of temperature fluctuations determines stress resistance in a natural population of *Drosophila simulans*. *J. Evol. Biol.* 27, 2113–2122. doi: 10.1111/jeb.12463

McFarlane, S. E., Gorrell, J. C., Humphries, M. M., Buton, S., and McAdam, A. G. (2014). Very low levels of direct additive genetic variance in fitness and fitness components in a red squirrel population. *Ecol. Evol.* 4, 1729–1738. doi: 10.1002/ece3.982

Mitchell, K. A., and Hoffmann, A. A. (2010). Thermal ramping rate influences evolutionary potential and species differences for upper thermal limits in *Drosophila*. *Funct. Ecol.* 24, 694–700. doi: 10.1111/j.1365-2435.2009.01666.x

Moghadam, N. N., Ketola, T., Pertoldi, C., Bahrndorff, S., and Kristensen, T. N. (2019). Heat hardening capacity in *Drosophila melanogaster* is life stage-specific and juveniles show the highest plasticity. *Biol. Lett.* 15:20180628. doi: 10.1098/rsbl.2018.0628

Mora, A. R., Firth, A., Blareau, S., Vallat, A., and Helfenstein, F. (2017). Oxidative stress affects sperm performance and ejaculate redox status in subordinate house sparrows. *J. Exp. Biol.* 220, 2577–2588. doi: 10.1242/jeb.154799

Morozova, T. V., Huang, W., Pray, V. A., Whitham, T., Anholt, R. R. H., and Mackay, T. F. C. (2015). Polymorphisms in early neurodevelopmental genes affect natural variation in alcohol sensitivity in adult *Drosophila*. *BMC Genomics* 16:865. doi: 10.1186/s12864-015-2064-5

Mousseau, T. A., and Roff, D. A. (1987). Natural selection and the heritability of fitness components. *Heredity* 59, 181–197. doi: 10.1038/hdy.1987.113

Nakagawa, S., and Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol. Rev. Camb. Philos. Soc.* 85, 935–956. doi: 10.1111/j.1469-185X.2010.00141.x

Najarro, M. A., Hackett, J. L., and Macdonald, S. J. (2017). Loci contributing to boric acid toxicity in two reference populations of *Drosophila melanogaster*. *G3 Genes Genomes Genet.* 7, 1631–1641. doi: 10.1534/g3.117.041418

Najarro, M. A., Hackett, J. L., Smith, B. R., Highfill, C. A., King, E. G., Long, A. D., et al. (2015). Identifying loci contributing to natural variation in xenobiotic resistance in *Drosophila*. *PLoS Genet.* 11:e1005663. doi: 10.1371/journal.pgen.1005663

Nijhout, H. F. (1998). *Insect Hormones*.Princeton, NJ: Princeton University Press.

Noble, L. M., Chelo, I., Guzella, T., Afonso, B., Riccardi, D. D., Ammerman, P., et al. (2017). Polygenicity and epistasis underlie fitness-proximal traits in the *Caenorhabditis elegans* multiparental experimental evolution (CeMEE) panel. *Genetics* 207, 1663–1685. doi: 10.1534/genetics.117.300406

O'Brien, E. K., Higgie, M., Reynolds, A., Hoffmann, A. A., and Bridle, J. R. (2017). Testing for local adaptation and evolutionary potential along altitudinal gradients in rainforest *Drosophila*: beyond laboratory estimates. *Global Change Biol.* 23, 1847–1860. doi: 10.1111/gcb.13553

Overgaard, J., Kearney, M. R., and Hoffmann, A. A. (2014). Sensitivity to thermal extremes in Australian *Drosophila* implies similar impacts of climate change on the distribution of widespread and tropical species. *Global Change Biol.* 20, 1738–1750. doi: 10.1111/gcb.12521

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., and Christ, R. (2015). *Climate Change 2014: Synthesis Report [Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva: IPCC.

Pedersen, L. D., Bundgaard, J., Pedersen, A. R., and Bijlsma, R. (2011). The effects of inbreeding and heat stress on male sterility in *Drosophila melanogaster*. *Biol. J. Linnean Soc.* 104, 432–442. doi: 10.1111/j.1095-8312.2011.01725.x

Porcelli, D., Gaston, K. J., Butlin, R. K., and Snook, R. R. (2017). Local adaptation of reproductive performance during thermal stress. *J. Evol. Biol.* 30, 422–429. doi: 10.1111/jeb.13018

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Rajpurohit, S., Gefen, E., Bergland, A. O., Petrov, D. A., Gibbs, A. G., and Schmidt, P. S. (2018). Spatiotemporal dynamics and genome-wide association analysis of desiccation tolerance in *Drosophila melanogaster*. *Mol. Ecol.* 27, 3525–3540. doi: 10.1111/mec.14814

Rohmer, C. (2004). Heat induced male sterility in *Drosophila melanogaster*: adaptive genetic variations among geographic populations and role of the Y chromosome. *J. Exp. Biol.* 207, 2735–2743. doi: 10.1242/jeb.01087

Rolandi, C., Lighton, J. R. B., de la Vega, G. J., Schilman, P. E., and Mensch, J. (2018). Genetic variation for tolerance to high temperatures in a population of *Drosophila melanogaster*. *Ecol. Evol.* 8, 10374–10383. doi: 10.1002/ece3.4409

Sales, K., Vasudeva, R., Dickinson, M. E., Godwin, J. L., Lumley, A. J., Michalczyk, Ł, et al. (2018). Experimental heatwaves compromise sperm function and cause transgenerational damage in a model insect. *Nat. Commun.* 9:4771. doi: 10.1038/s41467-018-07273-z

Saxon, A. D., O'Brien, E. K., and Bridle, J. R. (2018). Temperature fluctuations during development reduce male fitness and may limit adaptive potential in tropical rainforest *Drosophila*. *J. Evol. Biol.* 31, 405–415. doi: 10.1111/jeb.13231

Setchell, B. P. (1998). The parkes lecture. Heat and the testis. *J. Reproduct. Fertil.* 114, 179–194. doi: 10.1530/jrf.0.1140179

Shorter, J., Couch, C., Huang, W., Carbone, M. A., Peiffer, J., Anholt, R. R. H., et al. (2015). Genetic architecture of natural variation in *Drosophila melanogaster* aggressive behavior. *Proc. Natl. Acad. Sci. U.S.A.* 112, E3555–E3563. doi: 10.1073/pnas.1510104112

Sinclair, B. J., Marshall, K. E., Sewell, M. A., Levesque, D. L., Willett, C. S., Slotsbo, S., et al. (2016). Can we predict ectotherm responses to climate change using thermal performance curves and body temperatures? *Ecol. Lett.* 19, 1372–1385. doi: 10.1111/ele.12686

Sinclair, B. J., and Roberts, S. P. (2005). Acclimation, shock and hardening in the cold. *J. Ther. Biol.* 8, 557–562. doi: 10.1016/j.jtherbio.2005.07.002

Snoke, M. S., and Promislow, D. E. (2003). Quantitative genetic tests of recent senescence theory: age-specific mortality and male fertility in *Drosophila melanogaster*. *Heredity* 91, 546–556. doi: 10.1038/sj.hdy.6800353

Snook, R. R. (2005). Sperm in competition: not playing by the numbers. *Trends Ecol. Evol.* 20, 46–53. doi: 10.1016/j.tree.2004.10.011

Snook, R. R., Markow, T. A., and Karr, T. L. (1994) Functional nonequivalence of sperm in *Drosophila pseudoobscura*. *Proc. Natl. Acad. Scie. U.S.A.* 91, 11222–11226 doi: 10.1073/pnas.91.23.11222

Sørensen, J. G., Kristensen, T. N., and Overgaard, J. (2016). Evolutionary and ecological patterns of thermal acclimation capacity in *Drosophila*: is it important for keeping up with climate change? *Curr. Opin. Insect. Sci.* 17, 98–104. doi: 10.1016/j.cois.2016.08.003

Swarup, S., Huang, W., Mackay, T. F. C., and Anholt, R. R. H. (2013). Analysis of natural variation reveals neurogenetic networks for *Drosophila* olfactory behavior. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1017–1022. doi: 10.1073/pnas.1220168110

Sztepanacz, J. L., and Blows, M. W. (2015). Dominance genetic variance for traits under directional selection in *Drosophila serrata*. *Genetics* 200, 371–384. doi: 10.1534/genetics.115.175489

Teplitsky, C., Mills, J. A., Yarrall, J. W., and Merilä, J. (2009). Heritability of fitness components in a wild bird population. *Evolution* 63, 716–726. doi: 10.1111/j.1558-5646.2008.00581.x

The Gene Ontology Consortium (2019). The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47, D330–D338. doi: 10.1093/nar/gky1055

Turner, T. L., Miller, P. M., and Cochrane, V. A. (2013). Combining genome-wide methods to investigate the genetic complexity of courtship song variation in *Drosophila melanogaster*. *Mol. Biol. Evol.* 30, 2113–2120. doi: 10.1093/molbev/mst111

van Heerwaarden, B., Kellermann, V., and Sgro, C. M. (2016). Limited scope for plasticity to increase upper thermal limits. *Funct. Ecol.* 30, 1947–1956. doi: 10.1111/1365-2435.12687

Vollmer, J. H., Sarup, P., Kaersgaard, C. W., Dahlgaard, J., and Loeschcke, V. (2004). Heat and cold-induced male sterility in *Drosophila buzzatii*: genetic variation among populations for the duration of sterility. *Heredity* 92, 257–262. doi: 10.1038/sj.hdy.6800405

Walsh, B. S., Parratt, S. R., Atkinson, D., Snook, R. R., Bretman, A., and Price, T. A. R. (2019a). Integrated approaches to studying male and female thermal fertility limits. *Trends Ecol. Evol.* 34, 492–493. doi: 10.1016/j.tree.2019.03.005

Walsh, B. S., Parratt, S. R., Hoffmann, A. A., Atkinson, D., Snook, R. R., Bretman, A., et al. (2019b). The impact of climate change on fertility. *Trends Ecol. Evol.* 34, 249–259. doi: 10.1016/j.tree.2018.12.002

Weather-us.com (2020). *Climate Raleigh, USA*. Avaliable at: https://www.weather-us.com/en/north-carolina-usa/raleigh-climate?c,mm,mb,km#temperature (accessed April 29, 2020).

Weber, A. L., Khan, G. F., Magwire, M. M., Tabor, C. L., Mackay, T. F. C., and Anholt, R. R. H. (2012). Genome-wide association analysis of oxidative stress resistance in *Drosophila melanogaster*. *PLoS One* 7:e034745. doi: 10.1371/journal.pone.0034745

Weisser, W. W., and Siemann, E. (2008). *Insects and Ecosystem Function*. Berlin: Springer.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569. doi: 10.1038/ng.608

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi: 10.1016/j.ajhg.2010.11.011

Yu, B., and Huang, Z. (2015). Variations in antioxidant genes and male infertility. *BioMed Res. Int.* 2015, 1–10. doi: 10.1155/2015/513196

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 23 June 2020
doi: 10.3389/fgene.2020.00658

Integrating GWAS and Transcriptomics to Identify the Molecular Underpinnings of Thermal Stress Responses in *Drosophila melanogaster*

Melise C. Lecheta[1], David N. Awde[1], Thomas S. O'Leary[2], Laura N. Unfried[1], Nicholas A. Jacobs[1], Miles H. Whitlock[1], Eleanor McCabe[1], Beck Powers[2], Katie Bora[2], James S. Waters[3], Heather J. Axen[4], Seth Frietze[5], Brent L. Lockwood[2], Nicholas M. Teets[1] and Sara H. Cahan[2]*

[1] Department of Entomology, University of Kentucky, Lexington, KY, United States, [2] Department of Biology, University of Vermont, Burlington, VT, United States, [3] Department of Biology, Providence College, Providence, RI, United States, [4] Department of Biology and Biomedical Sciences, Salve Regina College, Providence, RI, United States, [5] Department of Biomedical and Health Sciences, University of Vermont, Burlington, VT, United States

OPEN ACCESS

Edited by:
Ana Sofia Quina, Universidade de Lisboa, Portugal

Reviewed by:
Mary Anna Carbone, North Carolina State University, United States
Jesper Givskov Sørensen, Aarhus University, Denmark

***Correspondence:**
Melise C. Lecheta
melise.lecheta@uky.edu

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *06 April 2020*
Accepted: *29 May 2020*
Published: *23 June 2020*

Citation:
Lecheta MC, Awde DN, O'Leary TS, Unfried LN, Jacobs NA, Whitlock MH, McCabe E, Powers B, Bora K, Waters JS, Axen HJ, Frietze S, Lockwood BL, Teets NM and Cahan SH (2020) Integrating GWAS and Transcriptomics to Identify the Molecular Underpinnings of Thermal Stress Responses in Drosophila melanogaster. Front. Genet. 11:658. doi: 10.3389/fgene.2020.00658

Thermal tolerance of an organism depends on both the ability to dynamically adjust to a thermal stress and preparatory developmental processes that enhance thermal resistance. However, the extent to which standing genetic variation in thermal tolerance alleles influence dynamic stress responses vs. preparatory processes is unknown. Here, using the model species *Drosophila melanogaster,* we used a combination of Genome Wide Association mapping (GWAS) and transcriptomic profiling to characterize whether genes associated with thermal tolerance are primarily involved in dynamic stress responses or preparatory processes that influence physiological condition at the time of thermal stress. To test our hypotheses, we measured the critical thermal minimum (CT_{min}) and critical thermal maximum (CT_{max}) of 100 lines of the *Drosophila* Genetic Reference Panel (DGRP) and used GWAS to identify loci that explain variation in thermal limits. We observed greater variation in lower thermal limits, with CT_{min} ranging from 1.81 to 8.60°C, while CT_{max} ranged from 38.74 to 40.64°C. We identified 151 and 99 distinct genes associated with CT_{min} and CT_{max}, respectively, and there was strong support that these genes are involved in both dynamic responses to thermal stress and preparatory processes that increase thermal resistance. Many of the genes identified by GWAS were involved in the direct transcriptional response to thermal stress (72/151 for cold; 59/99 for heat), and overall GWAS candidates were more likely to be differentially expressed than other genes. Further, several GWAS candidates were regulatory genes that may participate in the regulation of stress responses, and gene ontologies related to development and morphogenesis were enriched, suggesting many of these genes influence thermal tolerance through effects on development and physiological status. Overall, our results suggest that thermal tolerance alleles can

influence both dynamic plastic responses to thermal stress and preparatory processes that improve thermal resistance. These results also have utility for directly comparing GWAS and transcriptomic approaches for identifying candidate genes associated with thermal tolerance.

Keywords: thermal limit, CT_{min}, CT_{max}, heat shock, cold shock, genomics, transcriptomics

INTRODUCTION

Temperature directly affects performance, survival, fitness, and consequently, the geographic distribution of organisms (Angilletta, 2009; Dowd et al., 2015). Ectotherms are particularly vulnerable to changes in temperature, and these organisms have evolved a suite of adaptations to cope with thermal variability. An ectotherm's thermal tolerance is determined by both fixed genetic factors and plastic changes in behavior, morphology, physiology, and gene expression. Genetic variation in thermal tolerance is well-documented (e.g., Sørensen et al., 2001; McMillan et al., 2005; Rako et al., 2007) and can occur through changes in basal stress tolerance and/or changes in the ability to quickly respond to thermal challenges (Ayrinhac et al., 2004). These heritable differences within populations permit evolutionary shifts in thermal response as selection acts (Hoffmann et al., 2003), and adaptive differences in thermal tolerance across latitudinal gradients and thermal environments are common (Hoffmann et al., 2002; Fallis et al., 2014). Specifically, populations from higher latitudes often are more tolerant of low temperatures than populations from lower latitudes, and the same pattern is also seen for heat stress, where populations that extend to lower latitudes often have improved survival at high temperatures (e.g., Calosi et al., 2010; but see Castañeda et al., 2015). Thus, thermal tolerance is a trait that is both highly plastic and highly adaptable, and understanding the genetic basis of thermal tolerance is critical for predicting future responses to environmental change.

Genome Wide Association Studies (GWAS) and other quantitative genetic approaches have characterized the genetic architecture of thermal tolerance and identified genes that regulate temperature-dependent traits (e.g., Morgan and Mackay, 2006; Rako et al., 2007; Svetec et al., 2011; Rohde et al., 2016). A series of recent studies with the *Drosophila* Genetic Reference Panel (DGRP; Mackay et al., 2012) have identified a number of candidate loci associated with thermal tolerance. Rolandi et al. (2018) found 12 SNPs associated with variation in critical thermal maximum (CT_{max}), and most of these SNPs were located within intronic regions, suggesting that variation in the heat stress response could be mediated by regulatory changes in gene expression or splicing. For cold, distinct but related traits often have non-overlapping genetic architectures, suggesting these traits have the capacity to evolve independently. For example, two plastic responses to cold, rapid cold hardening and developmental cold acclimation, have non-overlapping SNPs associated with them, although the genes associated with these traits share some functional similarities (Gerken et al., 2015). Similarly, Teets and Hahn (2018) found minimal overlap in genes associated with cold shock response and chill coma recovery, and Freda et al. (2017) found no overlap in genes associated with adult and larval cold hardiness. The candidate genes identified in Teets and Hahn (2018) were functionally tested with RNAi, and knockdown of most genes affected cold tolerance, indicating that GWAS is a robust method for identifying genes with functional roles in thermal tolerance. Taken together, the various GWAS studies of thermal traits indicate that the thermal stress response is a highly polygenic trait, but additional studies linking these polymorphisms to their functional consequences are needed to clarify their role in thermal tolerance.

One way the genetic makeup of an organism influences thermal tolerance is by modifying gene expression changes in response to temperature change (Stucki et al., 2017). Transcriptional responses to thermal variability have been described at various levels, including whole transcriptomic studies of specific stress treatments (e.g., Qin et al., 2005; Sørensen et al., 2005, 2016; Teets et al., 2012), targeted experiments for specific candidate genes (e.g., *Frost* in Goto, 2001; Sinclair et al., 2007; Zhu et al., 2017), and comparisons of transcriptomic responses to thermal stressors both among (e.g., in damselflies; Lancaster et al., 2016) and within populations (e.g., Telonis-Scott et al., 2009). A consistent theme from these studies is that changing temperatures can cause substantial changes in gene expression. For example, in *D. melanogaster*, acclimation that enhanced the cold response led to nearly one third of the transcriptome being differentially regulated (with around 60% of these genes being downregulated; MacMillan et al., 2016). This cold acclimation included upregulation of genes already known to have an association with stress and temperature responses, such as *Frost* and many genes encoding for heat shock proteins. Similar sets of genes are also upregulated following brief cold shock in *D. melanogaster* and the flesh fly *Sarcophaga bullata* (Qin et al., 2005; Teets et al., 2012), indicating that anticipatory acclimation responses share some mechanisms with dynamic responses that occur during and after stress. For heat stress, most genes that are differentially expressed following short-term heat hardening (Sørensen et al., 2005) and heat shock (Telonis-Scott et al., 2013) in *D. melanogaster* are downregulated, with the exception of heat shock proteins, which are generally upregulated. However, despite the large body of literature on transcriptional responses to thermal stress, additional work is needed to clarify the functional consequences of these transcriptomic changes and determine how segregating variation in thermal tolerance relates to these transcriptional mechanisms.

Thermal tolerance is a combination of dynamic plastic changes that occur during and after a stress event (i.e., processes that actively counter, repair or minimize the consequences of damage) and preparative processes that enhance stress resistance

(i.e., processes that prevent damage; Roy and Kirchner, 2000; Wos and Willi, 2015). Plastic processes that occur during and after thermal stress largely involve production of stress proteins (e.g., heat shock proteins), often at the expense of other biological processes (Feder and Hofmann, 1999). Preparative processes that enhance thermal resistance include production of protective osmolytes (e.g., cryoprotectants; Yancey, 2005; Storey and Storey, 2012), changes in membrane composition and cell structure that permit membrane function at extreme temperature (Sinensky, 1974; Koštál, 2010), and anticipatory production of stress proteins during dormancy and/or thermal acclimation (Manjunatha et al., 2010; Colinet et al., 2013; MacMillan et al., 2016). Thus, an allelic variant may contribute to basal tolerance to extreme temperature by altering either of these two components: enhancing the plastic ability to adjust to stress by participating in or regulating the dynamic temperature response, or by better preparing the organism for that stress. However, for genes associated with variation in thermal tolerance, whether these genes primarily affect dynamic plastic processes or preparative processes is unclear.

Here, we used a combination of GWAS and RNA-seq to address the extent to which genes associated with thermal tolerance variation are involved in the dynamic stress response and preparative responses. We measured critical thermal minimum (CT_{min}) and CT_{max} (Schou et al., 2017) in 100 lines from the DGRP, and the resulting phenotypic data were used in conjunction with genome-wide polymorphism data to identify genes associated with variation in thermal limits. These candidate genes were then compared to differentially expressed genes identified via transcriptomic assays of a single genotype exposed to heat and cold shock treatments to identify their roles in the stress response. Three non-mutually exclusive hypotheses were considered. To identify candidates involved in dynamic stress responses, we tested the following two specific hypotheses: (H1) Genes associated with thermal tolerance are part of the dynamic response, and are directly up- or downregulated during thermal stress; (H2) Genes associated with thermal tolerance are transcription factors and regulatory genes that regulate the dynamic transcriptional response to thermal stress. In support of H1, we predict that GWAS candidates will be more likely to be up- or downregulated in response to thermal stress, and these candidates will include genes directly activated during the stress response (e.g., heat shock proteins) as well as genes downregulated because they are incompatible with stressful temperatures (e.g., certain metabolic processes and reproduction). To identify genes involved in preparative processes that enhance thermal stress resistance, we tested the following hypothesis: (H3) Genes associated with thermal tolerance are involved in preparatory developmental and physiological processes that influence the condition of the organism at the time of thermal stress. Here we predict that specific GWAS candidates will be involved in the stress resistance processes discussed above (e.g., cell membrane remodeling, osmolyte production, etc.), but these candidates will not necessarily be part of the dynamic stress response. Our results indicate that genes associated with the thermal response have diverse functional roles that contribute to thermal tolerance in all three of these ways. There is considerable overlap between the genes associated with quantitative variation in thermal tolerance and those that are differentially expressed in response to thermal stress, and our GWAS analysis indicated an abundance of genes involved in developmental processes and cell morphogenesis that may have a role in enhancing stress resistance. Testing these hypotheses will advance our understanding of the functional consequences of genes polymorphisms associated with thermal tolerance. Furthermore, these results also have utility for directly comparing two commonly used methods for identifying and characterizing candidate genes associated with thermal tolerance.

MATERIALS AND METHODS

Insect Rearing

The *Drosophila* Genetic Reference Panel (DGRP) was established from a natural population in Raleigh, North Carolina, and isofemale lines were isogenized with 20 generations of full-sib mating (Mackay et al., 2012). DGRP lines were obtained from the Bloomington Drosophila Stock Center, maintained at 25°C on a 12:12 light–dark cycle, and fed a standard cornmeal/soy flour diet consisting of 0.58% agar, 1.73% yeast, 7.31% cornmeal, 1.00% soy flour, 0.13% Tegosept (w/v), 7.69% light corn syrup, and 0.48% propionic acid (v/v) in H_2O. To generate flies for CT_{min} and CT_{max} assays, 15 females and 10 males were added to vials containing food and dry active yeast and were allowed to mate and lay eggs for 4 days. Restricting the number of adults in each vial and limiting the time to lay eggs prevented vials from becoming overcrowded, as extremely high larval densities can impact thermal tolerance (Sørensen and Loeschcke, 2001). Ten days after removing the parental adults, adults of the resulting progeny were removed and held for 24 h to ensure that all flies had an opportunity to mate. After 24 h, males and females were sorted and placed into separate vials in groups of 20. Flies were held in the vials for 3–4 days prior to measuring CT_{min}. For CT_{max}, flies were held for 2–3 days, and 24 h prior to the experiment, flies were lightly anesthetized with CO_2 and individually transferred to small screw-top vials with food. All flies were between 4 and 9 days old at the time of the experiment.

Seven-day-old *D. melanogaster* Canton-S female flies were used to characterize gene expression responses after a cold or heat shock. Only females were used to minimize confounding variation due to sex differences in gene expression and to include gene expression responses associated with protection of egg production, which is strongly related to fitness and may be expected to be under selection in nature. We selected the Canton-S background for these experiments to (1) address the extent to which GWAS candidates predict gene expression in a standard, well-characterized genetic background, to increase the generalizability of our results, and (2) provide candidate genes for future functional experiments, as most mutant and transgenic strains are in the Canton-S background. To generate flies of known age for RNA-seq assays, all adults were removed from mixed-sex stock vials that had been maintained at

approximately 50 flies per vial and all newly eclosed adults were sampled daily. Same-day cohorts were maintained in mixed-sex vials at a density of ∼30 flies/vial on a nutrient-rich medium, consisting of 0.88% agar, 8.33% yeast, 10% cornmeal, 0.33% Tegosept (w/v), 4.66% molasses, and 0.66% propionic acid (v/v) in dH_2O (Buchanan et al., 2018), at 25°C on a 12:12 light–dark cycle. Cohorts were transferred to fresh food vials after 4 days.

Phenotypic Assays

To measure CT_{min} and CT_{max}, we used a dynamic ramping approach in which flies were gradually cooled or heated until motor function was lost. To assess CT_{min}, we used a vertical jacketed column (modified from Huey et al., 1992) connected to a temperature-controlled fluid bath, and the temperature was monitored inside the column with a type T thermocouple (**Supplementary Figure 1A**). For detailed assembly instructions for the jacketed column, see Awde et al. (2020). For each line, ∼20 males and ∼20 females were combined in the column and submitted to the following thermal program: 25°C for 5 min, 25°C to 10°C at 0.5°C min^{-1}, 10°C for 2 min, then 10°C to −10°C at 0.25°C min^{-1}. The ramping rates are in line with other studies of CT_{min} and were designed to maximize throughput and prevent cold hardening during the procedure (e.g., Sinclair et al., 2015). At 10°C we began collecting flies as they reached their CT_{min} and fell through the column into collection vials containing 70% ethanol. New vials were placed under the column at 0.25°C intervals as the temperature decreased. Flies were typically at the top or on the walls of the column at the beginning of a trial (since they are negatively geotropic), and any flies that remained at the bottom of the column were discarded once the temperature reached 10°C. Flies from each vial were then sexed and counted, and the CT_{min} was recorded as the maximum temperature for a given interval (e.g., flies collected between 10°C and 9.75°C had a CT_{min} of 10°C). CT_{min} for each line was estimated by averaging the CT_{min} of all flies tested across two independent cohorts. Due to variation in line productivity, escaping flies, and discarded flies, the total number of flies measured per line ranged from 15 to 44 for males (median = 28 and mode = 28) and from 9 to 38 for females (median = 26 and mode = 26).

CT_{max} was assessed using the same apparatus as CT_{min}, except the jacketed column was arranged horizontally and flies were contained individually in 2 ml screw-top vials to prevent them from voluntarily walking out of the column as temperature increased (**Supplementary Figure 1B**). For each line, ∼18 males and ∼18 females were individually placed in vials attached to a wooden dowel (**Supplementary Figure 1B**). The wooden dowel with the vials was placed inside the column and submitted to the following ramping program: 25°C for 5 min, 25°C to 35°C at 0.5°C min^{-1}, then 35°C to 45°C at 0.25°C min^{-1}. Flies were checked for movement after the temperature reached 35°C by flicking the wooden dowel every 0.2°C. The CT_{max} of flies was recorded when flies were motionless and no longer responded to stimulus. As with CT_{min}, CT_{max} for each line was estimated by averaging the CT_{max} of all flies tested across two independent cohorts. Due to variation in line productivity and escaping flies, the total number of flies per line ranged from 23 to 53 for males (median = 33 and mode = 33) and from 24 to 51 for females (median = 33 and mode = 32).

We also tested the extent to which CT_{min} and CT_{max} were correlated with other life-history parameters and other measures of thermal tolerance using previously collected phenotype data for the DGRP. We obtained data for lifespan and fecundity from Durham et al. (2014), *Wolbachia* infection status and chill coma recovery time from Mackay et al. (2012), rapid cold hardening and chronic and acute survival from cold from Gerken et al. (2015), CT_{min} from Ørsted et al. (2018), cumulative cold tolerance from Teets and Hahn (2018), heat knockdown from Rohde et al. (2016), CT_{max} from Rolandi et al. (2018), and cold and heat hardness from Freda et al. (2019). We used Pearson correlations (*cor.test*) to test for linear correlation between these measures in R (version 3.6.1; R Core Team, 2019).

Heritability and Genome Wide Association Study (GWAS)

Broad sense heritability (H^2), defined as the proportion of the total phenotypic variation that is due to all genetic factors, was estimated as $H^2 = \sigma^2_L/(\sigma^2_L + \sigma^2_\varepsilon)$, where σ^2_L is among-line and σ^2_ε is within-line variance components (Mackay and Huang, 2018). Variance components were estimated using a linear mixed model and treating line as a random effect, with the *lme4* package (Bates et al., 2015) in R.

Genome wide associative mapping was used to identify genetic polymorphisms associated with CT_{min} and CT_{max} using the GWAS platform available on the DGRP website[1] (Mackay et al., 2012). This analysis associates the phenotypic variation of DGRP lines with single nucleotide polymorphisms (SNPs), insertions, deletions, and multiple nucleotide polymorphisms (MNPs). One of the lines tested (208) was removed from the GWAS analysis by the DGRP server, thus the GWAS analysis included 99 lines. Variants with p-value < 1E-4 (using the average mixed p-value of the two sexes) were considered significant and were annotated to genes using FlyBase annotation v5.57. The average mixed p-value of the two sexes for GWAS analysis was chosen because both CT_{min} and CT_{max} were significantly correlated across sexes (see the section "Results"). To identify transcriptional regulators of thermal stress in support of H2 (see the section "Introduction"), we compared GWAS candidates to annotated transcription factors in FlyBase annotation v5.57.

As an alternative to GWAS on individual variants, we also conducted gene-based GWAS to test the aggregated effect of a set of SNPs (e.g., SNPs within a gene) on CT_{min} and CT_{max} phenotypes. Gene-based p-values were calculated by contrasting the observed T value to an empirical distribution generated from resampling under the null hypothesis with permutations using PLINK (version 1.9; Purcell et al., 2007; Liu et al., 2010). We controlled for confounding genetic relatedness between the DGRP lines used in this study, and used the Tracy-Windom test in the AssocTests package (Wang et al., 2015)

[1] http://dgrp2.gnets.ncsu.edu

to evaluate eigenvalues from 20 principal components (PCs) of genotype. We retained the first eight PCs as covariates in the PLINK model as described above (Patterson et al., 2006). As inversions and *Wolbachia* infection status can also influence the phenotypes of the DGRP lines, we used the adjusted phenotypes for these factors outputted from the DGRP2 website. Variants with MAF $\geq$ 5% and a genotype rate of 20% were used as well as the FlyBase v5.57 gene annotations. A total of 1,939,313 variants were tested over 8,954 and 8,270 genes with at least one significant variant for CT_{min} and CT_{max}, respectively. Genome-wide significance was determined by controlling for FDR using the q value method (Storey and Tibshirani, 2003).

RNA-Sequencing and Differential Gene Expression

To characterize gene expression responses to thermal shock, three replicates of three females each were exposed to cold or heat shock conditions by placing flies in sealed 15 × 150 mm glass test tubes and submerging in a circulating water bath programmed to cool or heat at a rate of 0.25°C min^{-1} until the temperature reached 4°C or 37°C, respectively. Flies were held at the final temperature for 5 min and then collected into pre-filled bead homogenization tubes (Benchmark Scientific) under CO_2 anesthesia, immediately snap-frozen in liquid nitrogen and held at −80°C. Control flies were similarly handled but remained at 25°C until collection and flash-freezing. Whole flies were homogenized using a Bullet Blender Bead Homogenizer (Next Advance) in 300 μL TRIzol Reagent (Life Technologies) followed by purification with a Direct-zol RNA MicroPrep Kit (Zymo Research #R2060) according to the manufacturer's instructions. DNA was removed using DNAse I on the column followed by two washes with RNA Wash Buffer. Total RNA was eluted with 15 μL DNAse/RNAse-free water, quantified using a Qubit 2.0 Fluorometer (Thermo Fisher), and RNA quality was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies). RNA-seq libraries were generated from 1 ug of rRNA-depleted total RNA using the NuQuant Universal RNA-Seq Library Preparation Kit (Nugen #M01506) according to the manufacturer's protocol with 12 cycles of PCR. A total of nine libraries were pooled and sequenced on a single lane of an Illumina HiSeq1500 single-read flow cell. The sequence quality of the resulting raw Illumina reads was assessed using FastQC (version 0.11.4) and reads were aligned to the *D. melanogaster* reference genome (Release 6) using STAR aligner (Dobin et al., 2013). Genes were quantified using featureCounts (part of the Rsubread package, version 2.0.0, Liao et al., 2019) against the DM6 build. Differential expression was performed using the DESeq2 package (version 1.24.0; Love et al., 2014) in R. All data have been deposited into the NCBI SRA database with accession Bioproject PRJNA612361.

Pathway Enrichment Analysis

Overrepresentation analysis (ORA) of significantly differentially expressed genes (Benjamini–Hochberg corrected p-value < 0.01, fold-change > 2) and the GWAS genes (p-value < 1E-4) was performed using WebGestalt (Wang et al., 2017; minimum five genes per category, maximum 2,000 genes) and a false discovery rate cut-off of 0.05. The results of the overrepresentation analysis were the primary means by which we identified GWAS candidates in support of H3 (see the section "Introduction").

Integration of Differential Gene Expression With GWAS

Genes associated with SNPs identified using GWAS were matched with corresponding genes in the expression data set. For cases where multiple genes were associated with a single SNP, each gene was included. The distribution of log-fold changes of the expression of all genes in the heat shock and cold shock versus control was compared to the fold-change distribution of the genes significantly associated with the corresponding thermal performance limit with Kolmogorov–Smirnov tests implemented in R. These analyses were the primary means by which we identified inducible genes in support of H1 (see the section "Introduction").

RESULTS

Genetic Variation in Thermal Tolerance

In this study we measured thermal limits (CT_{min} and CT_{max}) in a subset of lines from the DGRP. CT_{min} values across the DRGP lines varied considerably more than CT_{max} (**Figures 1A,B**). CT_{min} ranged from 0.81 to 8.55°C for males and from 2.29 to 8.64°C for females, while CT_{max} ranged from 38.63 to 40.72°C for males and from 38.51 to 40.80°C for females (**Supplementary Table 1**). The sex of the flies did not affect CT_{min} (p = 0.39), but it did affect CT_{max} (p < 0.001), with the males being slightly more heat tolerant than the females. However, the effect size was small (effect size: 0.13°C). The interaction of sex and line was also significant for both CT_{min} and CT_{max} (p < 0.01 and p < 0.001, respectively). Within each phenotype, values were significantly correlated across sexes (CT_{min}, r = 0.85, p-value < 0.001; CT_{max}, r = 0.52, p-value < 0.001; **Supplementary Figures 2A,B**). The sex-averaged CT_{min} and CT_{max} values were not significantly correlated across lines (r = 0.06, p-value = 0.54; **Figure 1C**).

We tested for trade-offs associated with thermal tolerance by comparing CT_{min} and CT_{max} with previously collected lifespan and fecundity data (Durham et al., 2014); however, we found no evidence of trade-offs among these traits, as neither thermal tolerance measurement was correlated with longevity or fecundity (**Supplementary Table 2**). DGRP lines have variable infection status by *Wolbachia pipientis*, a ubiquitous endosymbiont in insects that can significantly modulate physiology (Werren, 1997). Within the lines studied, we found no evidence that *Wolbachia* infection impacts CT_{min} or CT_{max} (p = 0.07 and p = 0.19, respectively). We also performed correlations between our data and other measures of thermal tolerance, and we found no correlation among our results and chill coma recovery time (Mackay et al., 2012), measures of thermal plasticity (i.e., rapid cold hardening and survival from cold; Gerken et al., 2015), cumulative

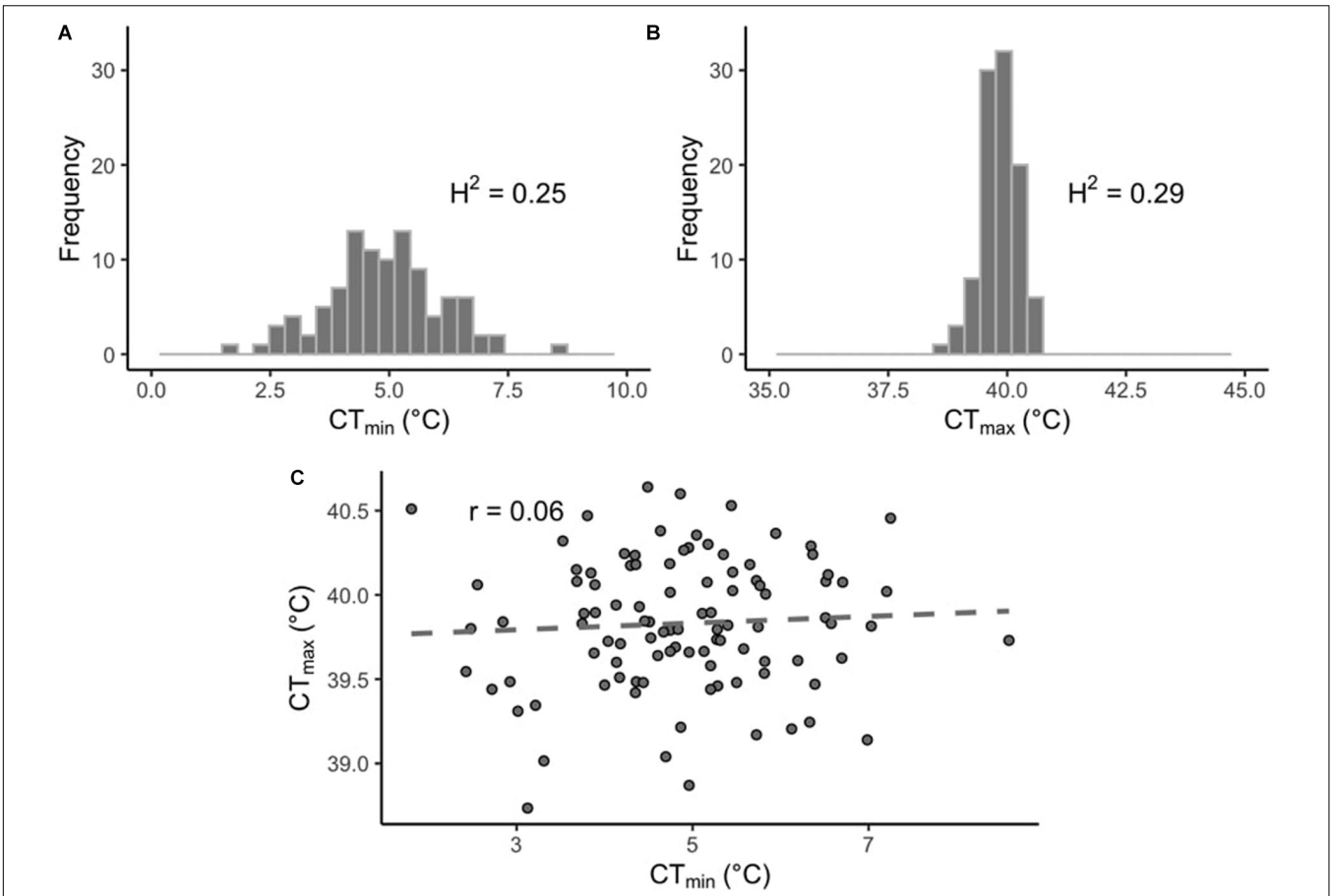


FIGURE 1 | Distribution of thermal limits in the DGRP. Histograms of mean phenotypes across 100 DGRP lines for **(A)** CT_{min} and **(B)** CT_{max}. **(C)** Correlation between line-means of CT_{min} and CT_{max}.

cold tolerance (Teets and Hahn, 2018), and heat knockdown time (Rohde et al., 2016; **Supplementary Table 2**). Ørsted et al. (2018) measured CT_{min} in males reared at 26°C (using a ramping method of 0.1°C min^{-1}), and we found significant correlations among their CT_{min} values with our CT_{min} values for males and females (r = 0.60, p-value < 0.01 and r = 0.61, p-value < 0.01, respectively; **Supplementary Table 2**). As in our study, Rolandi et al. (2018) also measured CT_{max} using a ramping method (0.25°C min^{-1}). We found significant correlations between their CT_{max} values for males with our CT_{max} values for males and females (r = 0.59, p-value < 0.01 and r = 0.55, p-value < 0.01, respectively; **Supplementary Table 2**).

Genetic Architecture of Thermal Limits

The estimated broad sense heritability (H^2) for CT_{max} was 0.29 and for CT_{min} was 0.25.

Using available genomic data for the DGRP, we identified genetic polymorphisms associated with CT_{min} and CT_{max}. For the 99 lines measured, more than 1.9 million variants were analyzed (mostly SNPs), and we found ∼550 unique allelic variants associated with these traits (p-value threshold of 1E-4; **Supplementary Table 3**). We identified 348 allelic variants (319 SNPs, 15 deletions, 13 insertions, and 1 MNP) significantly associated with CT_{min}, and 193 allelic variants (173 SNPs, 9 deletions, 8 insertions, and 3 MNPs) with CT_{max} (**Table 1**). Polymorphisms associated with CT_{min} and CT_{max} were identified on all chromosomes (**Figures 2A,B**). CT_{min} and CT_{max} did not share any allelic variants (**Figure 2C**). Among these allelic variants, 262 mapped to 151 unique genes for CT_{min}, and 169 mapped to 99 unique genes for CT_{max}. Three genes (*iab8*, *Btk29A* and *Sp1*) were common between both traits (**Figure 2D**). From all the genes associated with the allelic variants, 8% of the CT_{min} and

TABLE 1 | Overlap between the CT_{min} and CT_{max} SNPs and expression data.

GWAS	Relaxed SNPs	Strict SNPs	Unique genes	DEGs	TFs	DE TFs
CT_{min}	348	53	151	72	12	3
CT_{max}	193	21	99	59	9	9

Counts of Relaxed SNPs (p-value < 1E-4), Strict SNPs (p-value < 1E-5), Unique genes (based on relaxed SNPs), differentially expressed genes (DEGs; FDR < 0.01), transcription factors (TFs), and differentially expressed transcription factors (DE TFs).

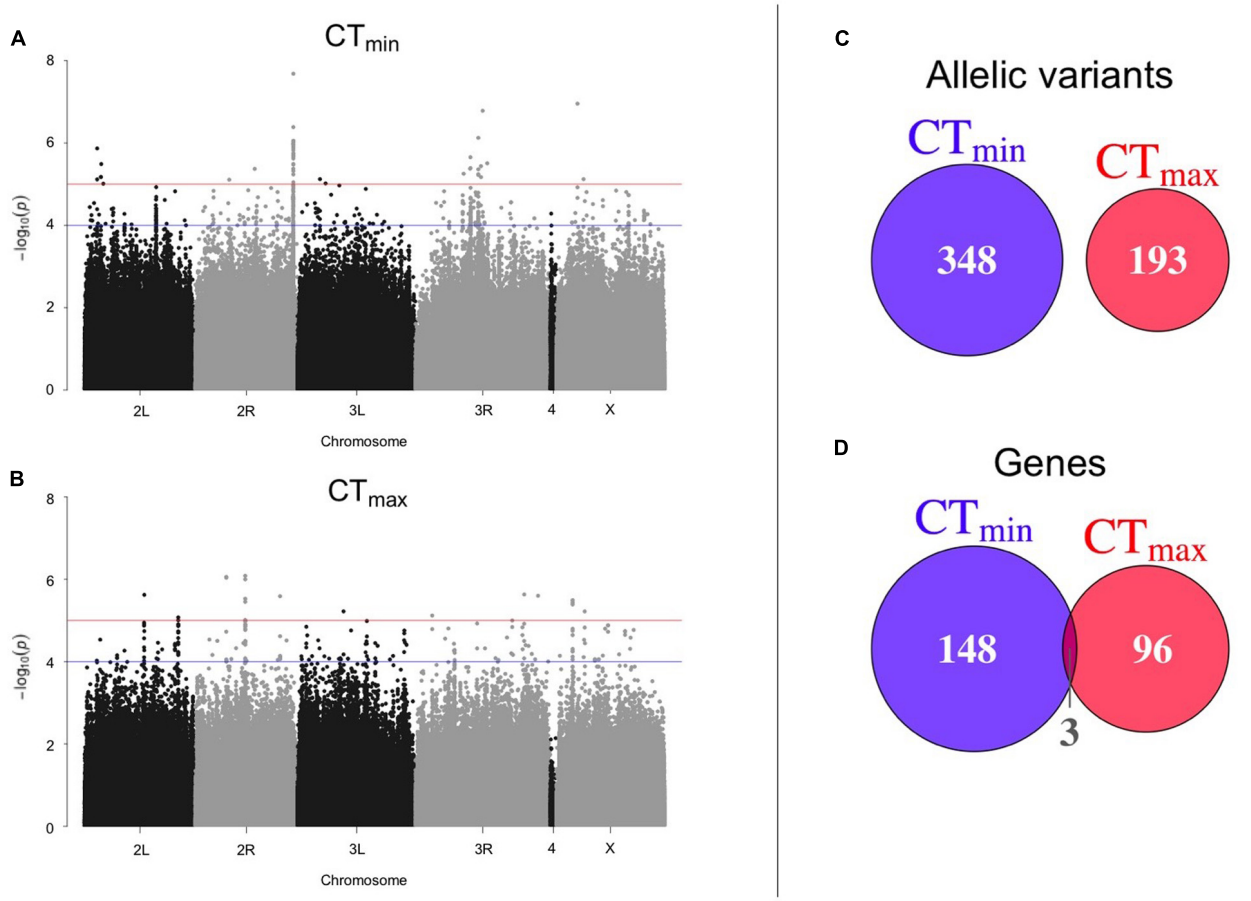


FIGURE 2 | Results of GWAS to identify polymorphisms associated with thermal tolerance. Manhattan plots of the results from the GWAS for **(A)** CT_{min} and **(B)** CT_{max}. The blue line corresponds to *p*-value < 1E-4 and the red line corresponds to *p*-value < 1E-5. The overlap of **(C)** allelic variants and **(D)** unique genes between CT_{min} and CT_{max} is also shown.

9% of the CT_{max} genes encode transcription factors (**Table 1**), including one of the genes common to both traits (*Sp1*). Most of the allelic variants significantly associated with both traits were located in introns (55% for CT_{min} and 71% for CT_{max}; **Supplementary Table 4**). The distribution of the direction of effect sizes differed between SNPs that underlined CT_{min} vs. SNPs that underlined CT_{max} (**Figure 3A**; Kolmogorov–Smirnov test, $D = 0.87$, *p*-value < 1E-9), such that the CT_{min}-associated alleles that were most common in the DGRP population caused individuals to have higher CT_{min} (i.e., worse cold tolerance; **Figure 3B**), whereas the CT_{max}-associated alleles that were most common in the population overwhelmingly caused individuals to have higher CT_{max} (i.e., better heat tolerance; **Figure 3B**). Additionally, there was a negative exponential relationship between effect size and minor allele frequency for both CT_{min}- and CT_{max}-associated SNPs (**Figure 3B**).

There is some evidence that low abundance alleles can be underpowered in the DGRP (Ivanov et al., 2015), so as an alternative to the variant-based GWAS performed above, we also conducted gene-based GWAS. However, in the gene-based GWAS analysis, almost no genes were detected as significant using the *p*-value threshold of 1E-4 (three for CT_{min} and one for CT_{max}; **Supplementary Table 5**). Thus, for the remainder of the paper, we will focus on the results of the variant-based GWAS described above.

Differential Gene Expression Following Acute Thermal Exposures

RNA-seq and differential expression analysis were used to determine the gene expression responses of the Canton-S strain of *D. melanogaster* under ramped cold shock and heat shock conditions. In total, 15,844 genes were expressed across all treatment groups. Principal component analysis (PCA) of expressed genes showed clustering of replicates for each condition (**Supplementary Figure 3**). Pairwise comparison of cold shock and heat shock to controls revealed a large number of significantly differentially expressed genes (*p*-adj. < 0.05, fold-change > 2; **Figure 4**). Among the differentially expressed genes, more were downregulated (5,126 in cold shock and 6,241 in heat shock) and fewer were upregulated relative to controls (1,826 in cold shock and 2,314 in heat shock). The direction of

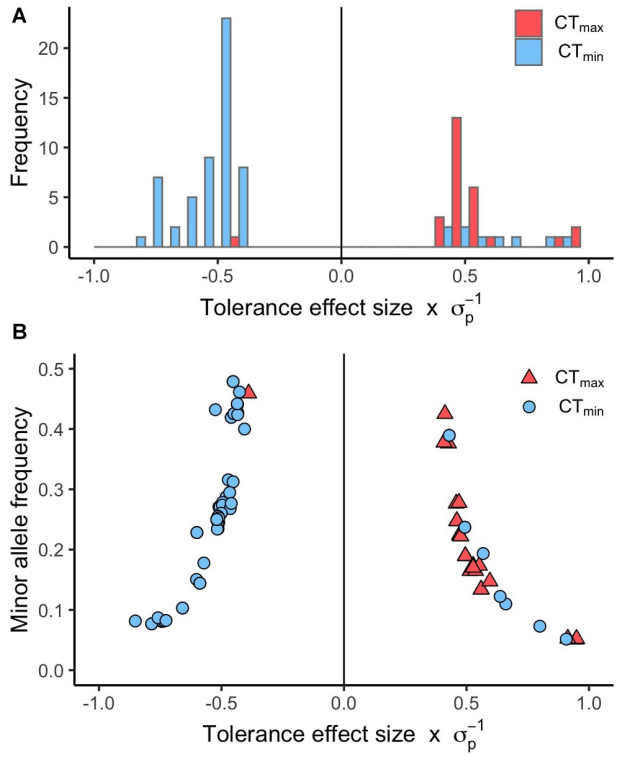


FIGURE 3 | Distribution of effect sizes for SNPs associated with CT_{min} and CT_{max}. **(A)** Histogram of the tolerance effect size scaled by the phenotypic standard deviation (σ_p) for the same allelic variants. **(B)** Scatter plot of the minor allele frequency versus the tolerance effect size scaled by the phenotypic standard deviation (σ_p) for the allelic variants associated with CT_{min} and CT_{max} (p-value < 1E-5).

regulation for most differentially expressed genes was consistent across treatments, with 6,081 genes changing similarly in both magnitude and direction in response to cold and heat shock conditions (**Figure 4**).

Integration of Transcriptomics and GWAS

Of the 151 unique genes found for CT_{min}, 72 (47.7%) were differentially expressed under cold shock. Of the 99 unique genes associated with CT_{max}, 59 (59.6%) were differentially expressed under heat shock. The distribution of log-fold change expression of heat shock versus control for the GWAS candidates associated with CT_{max} significantly differed from the distribution of all other genes under heat shock (**Figure 4**; p-value < 0.01). For the genes associated with CT_{min}, there was a similar trend, but the log-fold change distribution of those genes only marginally differed from the background expression of all other genes (**Figure 4**; p-value = 0.062).

Overrepresentation Analysis

Among the GWAS candidates, we identified six overrepresented GO biological process categories for CT_{min} and 17 categories for CT_{max} that met the FDR cut-off of 0.05 (**Figure 5**). For both traits we found many enriched GO terms related to development, differentiation, and morphogenesis. Among the differentially expressed genes from the RNA-seq experiments, we identified 99 enriched GO biological process categories for cold shock and 113 enriched categories for heat shock (**Supplementary Table 6**). Many of these categories were also related to development and differentiation. There was overlap between the enriched categories for GWAS and gene expression data. Five of the six categories enriched among genes associated with CT_{min} were also significantly enriched among genes differentially expressed under cold shock, and 13 of the 17 categories identified for CT_{max} were also enriched under heat shock (**Figure 5**).

DISCUSSION

Here, we characterized the genetic architecture of thermal tolerance and identified candidate genes that contribute to both dynamic responses to thermal stress and preparative processes that enhance stress resistance. Our study suggests that GWAS candidates are involved in both dynamic stress responses and preparative processes that influence the condition of the insect at the time of thermal stress. Together, our results indicate diverse functions for genes involved in thermal tolerance and allow us to generate new hypotheses for the genetic basis of thermal tolerance. Below, we discuss the genetic architecture of thermal tolerance in general, followed by a discussion of our three specific hypotheses to test the relative contribution of dynamic and preparative processes in shaping thermal tolerance.

Genetic Architecture of Thermal Tolerance

While several studies have separately assessed the genetic basis of cold and heat tolerance, here we measured both CT_{min} and CT_{max} across 100 lines of the DGRP. We found variation in both measures, although CT_{min} varied considerably more than CT_{max} (**Figure 1**). This pattern of variation in upper and lower thermal limits is also seen across species and populations with distinct geographic ranges, for both latitudinal and altitudinal gradients (e.g., Gaston and Chown, 1999; Addo-Bediako et al., 2000; Chown, 2001; Hoffmann et al., 2002; Nyamukondiwa et al., 2011; Sunday et al., 2011, 2019; Kellermann et al., 2012a,b). These patterns of variation in thermal limits, both within and across species, likely reflect stronger latitudinal and interannual variation in winter conditions relative to summer conditions (Williams et al., 2015). In addition, our results are consistent with previous studies indicating that upper and lower limits have distinct underlying mechanisms (e.g., Chown, 2001; Nyamukondiwa et al., 2011), as we found no phenotypic correlation between CT_{min} and CT_{max} across lines.

The variation in both CT_{min} and CT_{max} had a strong genetic component. Broad sense heritability was high for both CT_{max} (H^2 = 0.29) and CT_{min} (H^2 = 0.25), which is consistent with previous heritability estimates for thermal responses in the DGRP. Our heritability estimate for CT_{max} was higher than a previous estimate for a smaller subset of the DGRP population (H^2 = 0.14, Rolandi et al., 2018). For CT_{min}, heritability was

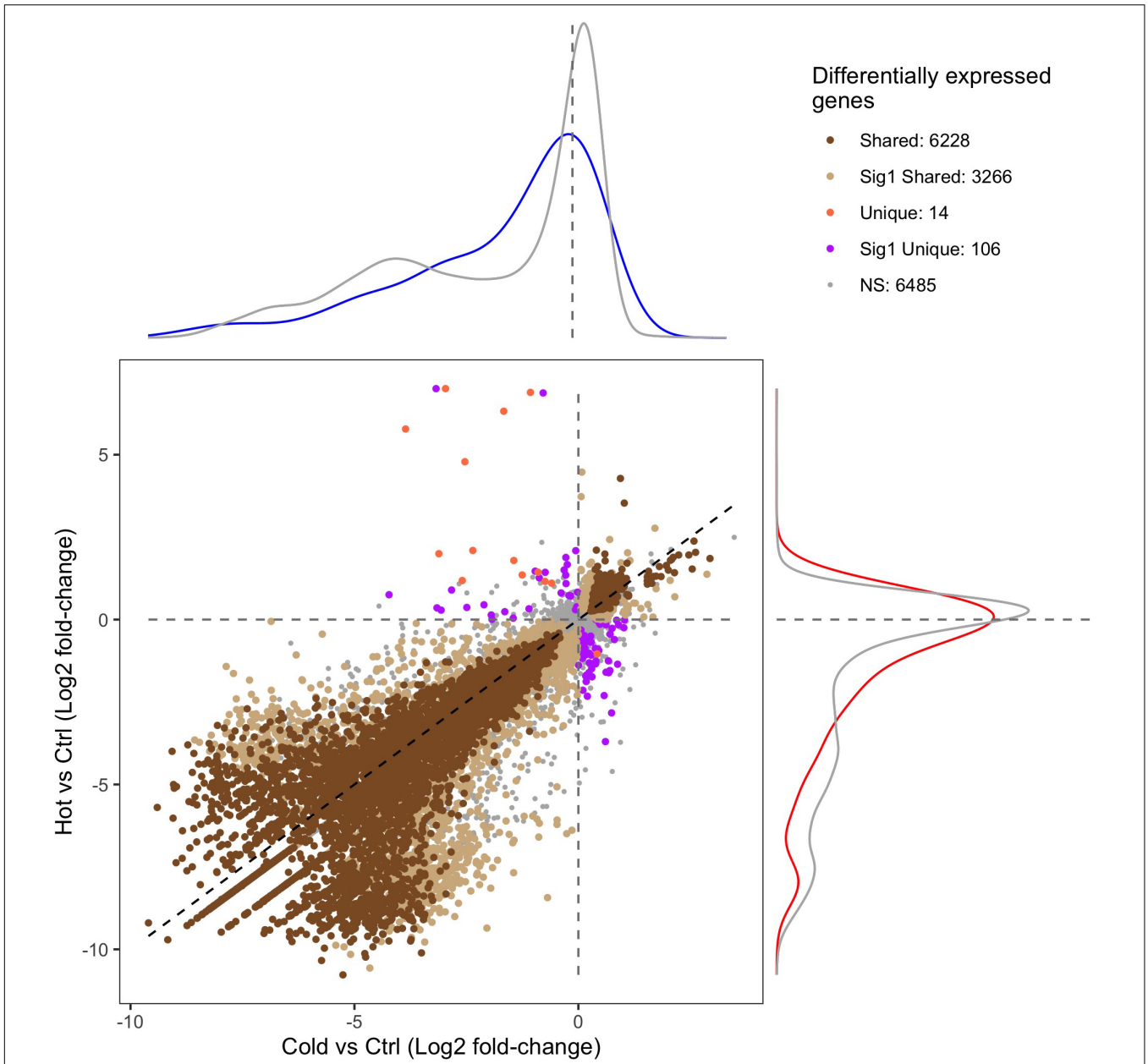


FIGURE 4 | Scatter plot of the gene expression $\log_2$ fold-change of heat-shock (37°C) and cold-shock (4°C) relative to the 25°C control. "Shared" DEGs indicate a significant expression relative fold-change in the same direction in both heat and cold shock. "Sig1 Shared" genes indicate that the relative fold-change is the same under heat shock and cold shock, but only significantly differentially expressed in one condition. "Unique" DEGs indicate that the genes were differentially expressed in each condition but in opposite directions. "Sig1 Unique" genes indicate that the genes are expressed in opposite directions relative to the control, but only significantly so in one condition. "NS" indicates genes that are not differentially expressed in both heat shock and cold shock. The density plots surrounding the figure indicate the density of the expression of genes that are associated with CT_{min} (top; blue) or CT_{max} (right; red) relative to the $\log_2$ fold-change expression of all other genes (gray).

within the range observed by Gerken et al. (2015) for acute and chronic cold exposure (H^2 = 0.15 and 0.44, respectively). The strong heritability for both traits suggests high evolutionary capacity for thermal tolerance in the mid-latitude population from which the DGRP was derived.

While variation in CT_{min} and CT_{max} was explained by distinct allelic variants, some variants mapped to the same genes, suggesting that some genes can affect both heat and cold tolerance. Of the 247 total unique genes, three genes were common to both traits: *iab-8* (a non-coding regulatory RNA), *Btk29A* (a tyrosine kinase involved in cellularization and morphogenesis), and *Sp1* (a zinc finger transcription factor involved in ventral thoracic appendage specification, leg growth and in the development of type-II neuroblasts). At the gene level,

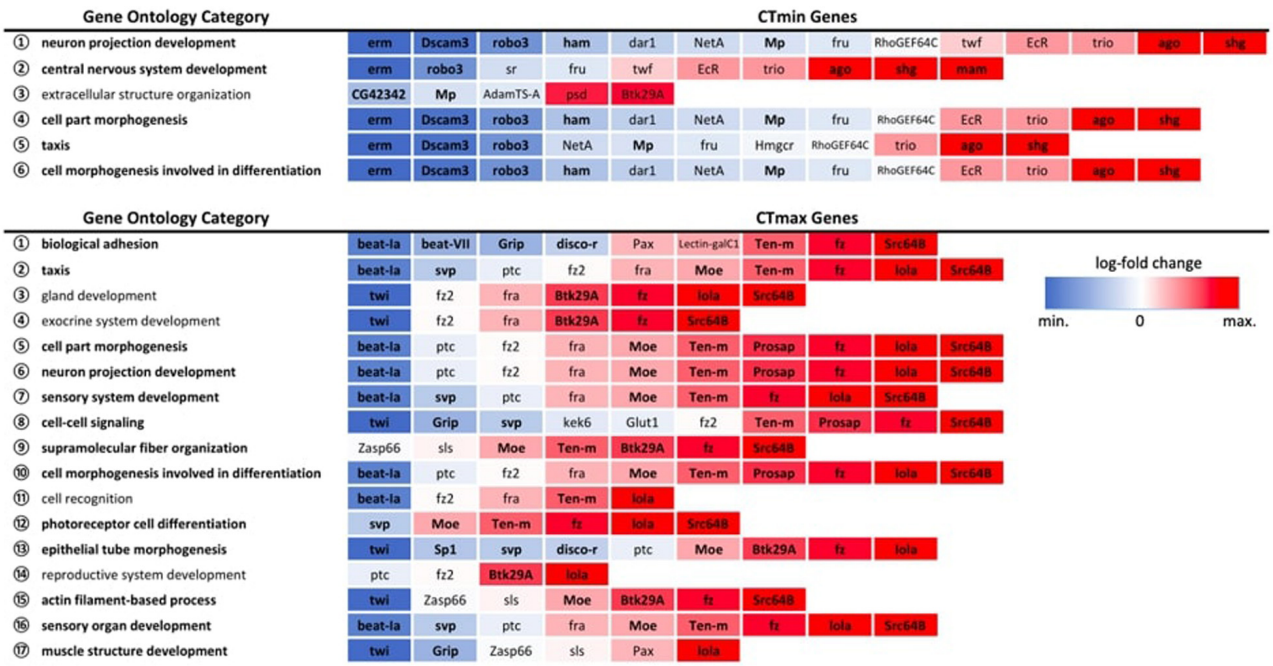


FIGURE 5 | Expression patterns of GWAS genes within significantly overrepresented GO biological process categories. All expression patterns are expressed relative to the 25°C control. Bolded categories indicate those also over-represented in the full DEG dataset for the corresponding temperature extreme. Bolded genes indicate significant differential gene expression at an FDR < 0.01.

there were also some candidate genes in common between our study and previous work. The gene CG42673 was associated with CT_{max} in this study and also with chill coma recovery time in Mackay et al. (2012). CG42673 is a putative nitric-oxide synthase binding protein, and while nitric oxide has not been linked to thermal tolerance in insects, nitric oxide is an important mediator of both heat and cold tolerance in plants (Parankusam et al., 2017; Costa-Broseta et al., 2018). In both our study and in Ørsted et al. (2018), *Mur89F* was associated with CT_{min}, and this gene is involved in chitin metabolic process and extracellular matrix.

Several patterns in our data suggest that thermal tolerance, especially heat tolerance, is an important component of organismal fitness in nature. Alleles that enhance heat tolerance (i.e., raise CT_{max}) were more common in the DGRP population than alleles that impair heat tolerance (**Figure 3A**). However, the pattern is opposite for cold tolerance; most alleles that improve cold tolerance (i.e., lower CT_{min}) were relatively infrequent in the population (**Figure 3A**). These results were counter to our expectations, since there is generally stronger latitudinal variation in cold tolerance than heat tolerance (see above). However, intraspecific latitudinal clines for heat tolerance have been observed in *D. melanogaster* (Hoffmann et al., 2002; Sgrò et al., 2010), indicating that there is selection for heat tolerance in this species. In the case of the DGRP, which originates in mid-latitude North Carolina, selection for cold tolerance may be lower than for heat tolerance, which could explain the relative rarity of alleles that improve cold tolerance. Alternatively, alleles that improve cold tolerance may have negative effects on other fitness-related traits, which would prevent these alleles from increasing in frequency in the population. Further, some polymorphisms in *D. melanogaster* oscillate in allele frequency across seasons (Bergland et al., 2014), so depending on when the DGRP was collected (presumably summer, although exact details are not provided in Mackay et al., 2012), cold tolerance alleles might be less common in the panel. Previously, a similar pattern was shown for oxidative stress resistance in the DGRP – i.e., most alleles that improve oxidative stress response are present at low frequency in the population (Weber et al., 2012). Finally, our study did not address plasticity in thermal tolerance, which may be an important means of response to cold challenges. Thus, the rarity of alleles improving basal cold tolerance may not be relevant in this population if the population possesses alleles for thermal plasticity. However, regardless of the reason for these results, these allele frequency patterns between heat and cold tolerance are not random and suggest that different forces may be affecting patterns of standing genetic variation for the two traits despite their similar overall heritabilities.

We also observed a negative exponential relationship between effect sizes and minor allele frequencies for SNPs that underlie both CT_{min} and CT_{max} (**Figure 3B**). Overwhelmingly, the SNPs that have the greatest effect sizes in both directions are present at the lowest frequencies in the population. Many other studies have reported this same pattern for genetic polymorphisms that underlie a wide range of different traits in the DGRP, including oxidative stress resistance, startle response, starvation resistance, chill coma recovery time, and position

effect variegation (Mackay et al., 2012; Weber et al., 2012; Kelsey and Clark, 2017). Overall, this pattern suggests that large-effect alleles that underlie CT_{min} or CT_{max} have undergone selection, either to increase the frequency of large-effect alleles in the population by positive selection (Barton and Keightley, 2002), thus driving the alternative allele to low frequency, or to remove large-effect alleles from the population by purifying selection (Keightley and Lynch, 2003).

Some of the genes underlying variation in CT_{min} and CT_{max} in the DGRP are also likely important for temperature adaptation in natural populations. Previous work by Fabian et al. (2012) and Bergland et al. (2016) used F_{ST} outlier analyses to identify loci across the genome that are likely to be undergoing adaptive divergence among populations of *D. melanogaster* that inhabit different thermal environments in eastern North America and eastern Australia. Among the 109 candidate genes that showed convergent patterns of clinal latitudinal differentiation in North America and Australia (Fabian et al., 2012; Bergland et al., 2016), seven of these genes were associated with CT_{max} and one was associated with CT_{min} in the DGRP. The clinal genes associated with CT_{max} were *beat-VII*, *dpr8*, *CG33970*, *CG42322*, *Tsp66E*, *A2bp1*, and *Moe*, and the sole clinal gene associated with CT_{min} was *fru*. Most of these clinal CT_{max} genes also changed expression following heat stress: three genes were downregulated (*beat-VII*, *dpr8*, and *CG42322*), two genes were upregulated (*Tsp66E* and *Moe*), and two genes (*CG33970* and *A2bp1*) were not differentially expressed (**Supplementary Table 7**). These genes are involved in a myriad of cellular and developmental processes, but a general theme is the potential role of neuronal processes in the thermal adaptation of heat tolerance (see below for additional discussion on the nervous system). It is also interesting to note the potential role of regulatory genes in thermal adaptation. Specifically, the CT_{max} clinal gene *A2bp1* is an RNA-binding Fox protein that regulates transcription and mRNA translation (Usha and Shashidhara, 2010; Carreira-Rosario et al., 2016), and the sole CT_{min} clinal gene *fru* is a zinc finger transcription factor known to be a regulator of transcriptional activity of many genes across various tissues (Sato and Yamamoto, 2020). While these genes did not dynamically respond to heat or cold stress (**Supplementary Table 7**), they may be important for setting up the developmental and/or cellular contexts in which stress responses operate.

H1: Genes Associated With Thermal Tolerance Are Involved in Dynamic Stress Responses

Thermal tolerance is shaped by a combination of preparative processes that improve stress resistance and dynamic changes in gene expression and activity that occur during and after stress. Dynamic changes in gene expression are well-established responses to thermal stress, and here we observed sweeping changes in gene expression in response to temperature change. Ramping at 0.25°C min^{-1} toward both temperature extremes elicited transcriptome-wide gene expression changes, with 43.9% and 54.0% of detected genes differentially expressed under cold and heat shock, respectively. These values are substantially larger than those reported in other studies, which range from minimal transcriptional response to ∼15% of the transcriptome, depending on the methodology used (Zhou et al., 2012; Telonis-Scott et al., 2013; von Heckel et al., 2016; Königer and Grath, 2018).

By pairing GWAS and RNA-seq using similar ramping methodologies, we can assess the extent to which GWAS candidates are directly involved in dynamic stress responses. GWAS-associated CT_{max} genes were significantly more temperature-responsive than the transcriptome at large and CT_{min} genes were marginally so (**Figure 4**), suggesting that genes associated with thermal tolerance are involved in the dynamic response to thermal stress. This congruence was also mirrored in the intersection of overrepresented GO terms in the two datasets, with five of six categories overrepresented in the CT_{min} GWAS also enriched among genes of cold shock response, and 13 of 17 CT_{max} categories shared with the heat shock response (**Figure 5**).

Despite a lower total number of genes identified that underlie CT_{max}, the total set of overrepresented biological process categories was more diverse than for CT_{min} and included cell signaling, muscle structure and function, and the sensory system (**Figure 5**). This may indicate a stronger role of active defensive responses in setting CT_{max}. At the individual gene level, the majority of top GWAS hits for CT_{max} were thermally responsive (**Table 1**) and were similarly diverse in function, including genes involved in neuropeptide signaling, mRNA processing, and protein dephosphorylation and catabolism. In contrast to the cold response, which included a mix of downregulated and upregulated genes, the majority of CT_{max} GWAS hits found within thermally responsive categories were upregulated under heat ramp conditions, suggesting that they are involved in active protection from or in response to heat damage (**Figure 4**). Thus, although the magnitude of phenotypic variation in CT_{max} is substantially lower than that of CT_{min}, standing genetic variation may be mediated via a wider range of defensive mechanisms, each of small effect.

Interestingly, the GWAS analysis did not indicate a role for the genes most commonly associated with thermal tolerance in experimental work. Much of the early literature on thermal limits focused on the effects of copy number and regulatory control of the heat shock protein (*hsp*) genes (Welte et al., 1993; Feder et al., 1996), and natural selection may also affect *hsp* allele frequencies (e.g., *hsp70*; Bettencourt et al., 2002). However, neither *hsp* genes nor their regulatory factors (e.g., hsf-1) were identified from the GWAS analysis as causal drivers of variation for either heat or cold limits within the DGRP. These *hsp* genes and their regulatory factors (hsf) did increase in expression in response to both cold and heat shock, so while these canonical stress genes had clear roles in dynamic stress responses, polymorphisms in these genes were not associated with thermal limits in the DGRP. For cold tolerance, *Frost* (*Fst*) is commonly upregulated in response to cold acclimation and during recovery from cold stress (Goto, 2001; Qin et al., 2005; Sinclair et al., 2007; Colinet et al., 2010), and it is also located within a QTL for chill coma

recovery time (Morgan and Mackay, 2006). In our study, *Frost* was not associated with variation for thermal limits. Further, it was not differentially expressed following cold shock but was upregulated after the heat shock. The lack of upregulation during cold stress is likely due to flies being sampled at 4°C, as *Frost* expression typically only increases during recovery from cold stress (Bing et al., 2012). The unexpected upregulation during a heat ramp suggests that *Frost* may be involved in heat stress, in addition to its role for cold and desiccation stress (Sinclair et al., 2007). Likewise, *Starvin* (*stv*), a poorly studied gene that is strongly upregulated during recovery from cold stress (Colinet and Hoffmann, 2010), was not associated with thermal tolerance but was upregulated after heat shock in this study.

H2: Genes Associated With Thermal Tolerance Have Regulatory Functions

Changes in gene transcription are one of the primary cellular responses to cold and heat stress in this and other studies (Gasch et al., 2000; Leemans et al., 2000; Gracey, 2007; Lockwood et al., 2010; Brown et al., 2014; Sørensen et al., 2016). Moreover, there is a direct connection between whole-organism stress tolerance and the ability to transcriptionally respond to stress, as organisms with limited transcriptional stress responses have lower survival following exposure to stress (Welte et al., 1993; Feder et al., 1996; Hofmann et al., 2000). Therefore, we expected to find candidate genes for thermal tolerance that have gene regulatory functions, such as transcription factors. Polymorphisms in regulatory genes or genomic regions may modify transcriptional responses to thermal stress, and thereby confer phenotypic differences in whole-organism thermal tolerance (Zatsepina et al., 2001; Bettencourt et al., 2002; Lerman et al., 2003). We report evidence for this potential regulatory effect among the SNPs that underlie both heat and cold tolerance, but consistent with the stronger pattern of upregulation in the dynamic response to heat, our results suggest a larger role of transcription factors in driving genetically based variation in CT_{max} than in CT_{min}.

Overall, there were nine CT_{max}-associated transcription factor genes (**Table 1**), and all were differentially expressed in response to heat stress (**Supplementary Table 7**). Indeed, the top four SNPs that were associated with CT_{max} (lowest *p*-value; **Supplementary Table 3**) were in two genes that encode transcription factors, *Oaz* and *lola*. *Oaz* encodes a transcription factor known to be involved in spiracle development (Krattinger et al., 2007), and thus may mediate developmental mechanisms that impact heat tolerance, especially since spiracles facilitate gas exchange and failings of aerobic respiration may set upper thermal limits (Dahlhoff and Somero, 1993; Pörtner, 2002). *Oaz* may also be involved in regulating the dynamic transcriptional response to heat stress, as it was the most strongly downregulated transcription factor following heat stress (**Supplementary Table 7**). *lola* is involved in a diverse array of cellular and developmental processes (Thurmond et al., 2019), and is represented among several GO categories in **Figure 5**. All four of the top CT_{max} SNPs lie in introns of the coding sequences of *Oaz* or *lola*, suggesting that these polymorphisms influence gene regulation (Bicknell et al., 2012). Indeed, in the case of *lola* both mutations lie in a region that is a putative transcription factor binding site. While previous work also showed that these two genes respond to heat stress in *D. melanogaster* (Brown et al., 2014), to our knowledge no previous studies have identified a functional role for these genes in heat tolerance. Another notable CT_{max}-associated SNP lies in two overlapping genes that encode the transcription factors HmgD and HmgZ; the CT_{max} SNP lies in the 5′ UTR intron of *HmgZ* and in the putative upstream regulatory region of *HmgD*. These genes encode proteins that belong to the family of high mobility group box transcription factors that are known to facilitate gene transcription by promoting DNA structural flexibility via chromatin remodeling (Štros, 2010), and both of these genes were differentially expressed following heat stress (**Supplementary Table 7**). Interestingly, high mobility group proteins have previously been reported to show expression patterns that track environmental temperature in killifish, *Austrofundulus limnaeus* (Podrabsky and Somero, 2004). Thus, the regulation of gene transcription may be a key aspect of heat tolerance in *D. melanogaster*.

The genetic architecture of cold tolerance also included genetic variation in transcription factor genes, but most of the CT_{min}-associated transcription factor genes did not change expression following cold stress, suggesting a qualitative difference in the role of transcription factors in heat vs. cold tolerance. While there were 12 transcription factor genes with significant associations with CT_{min} (**Table 1**), only three of these genes changed expression following cold stress (**Supplementary Table 7**). Importantly, one of the top SNPs in association with CT_{min} lies in the gene *blistered* (*bs*) (**Supplementary Table 3**). *bs* encodes a transcription factor known to be involved in a variety of developmental processes, including wing morphogenesis (Dworkin and Gibson, 2006), tracheal development (Affolter et al., 1994), and neural system development (Donlea et al., 2009; Thran et al., 2013). Similar to *lola*, the thermal tolerance SNP in *bs* lies in an intron with a transcription factor binding site; however, in contrast to *lola* and the other CT_{max}-associated transcription factor genes, *bs* did not change expression in response to cold shock (**Supplementary Table 7**).

One of the three genes associated with both CT_{min} and CT_{max}, the long non-coding RNA (lncRNA) *iab-8*, was downregulated in both cold and heat shock. LncRNAs have been implicated in a range of biological processes and are emerging as key regulators of gene expression at transcriptional and post-transcriptional levels (Li et al., 2019). *In vivo* studies of lncRNAs revealed that dysregulated expression of lncRNAs in *Drosophila* may result in poor stress resistance (Lakhotia et al., 2012). The *iab-8* lncRNA, expressed in the embryonic abdominal segment 8, represses the expression of the *abd-A* gene in the posterior central nervous system (Li et al., 2019). The *abd-A* gene is linked with neural system development (Bello et al., 2003; Cenci and Gould, 2005), and as discussed below, the nervous system likely plays an important role in thermal tolerance.

H3: Genes Involved in Thermal Tolerance Affect the Developmental and Structural Context

Thermal tolerance occurs within a developmental and structural context, making physiological systems more or less resistant to temperature challenges (González-Tokman et al., 2020). Thus, genes associated with thermal tolerance may do so by altering the physiological condition of the organism at the time of thermal stress. Because these genes alter baseline preparedness prior to application of cold or heat, the expression of these genes may not directly respond to temperature changes. Moreover, we would expect their biological function to be concentrated in processes underlying thermal stability of physiological functions, such as the central and peripheral nervous system, cell membrane composition, and proteome stability (Cossins and Prosser, 1978; Gu and Hilser, 2009; Cooper et al., 2012; Fields et al., 2015; MacMillan et al., 2015a; Willot et al., 2017).

Our results suggest that although segregating variation in both heat and cold tolerance is likely to include some structural effects, preparative processes that enhance thermal stress resistance appear to play a stronger role for CT_{min}. Fully half of the CT_{min} GWAS genes did not change significantly in expression in response to cold exposure, regardless of the significance cutoff used (**Table 1**). These included a cluster of functionally related genes (*dar1, fru, NetA, RhoGEF64C, trio, twf)* involved in nervous system development, which was reflected in overrepresentation of the nervous system and cell morphogenesis and differentiation GO categories in the CT_{min} GWAS gene set. Neuronal failure operationally defines both CT_{min} and CT_{max} (Andersen et al., 2018; Andersen and Overgaard, 2019; Jørgensen et al., 2019), and dynamic stabilization of the neuromuscular circuit under temperature stress is a likely mechanism for altering thermal limits. Indeed, previous investigation of the genetic architecture of cold hardiness and electrophysiological analyses of the rapid hardening response both suggest an important role for stabilization of ion channels and cytoskeletal structures supporting the synapse and neuromuscular junction (Klose and Robertson, 2004; Robertson and Money, 2012; Freda et al., 2017). Aside from genes involved in neural morphogenesis, the GO term extracellular structure organization was overrepresented among cold tolerance genes, and this was the only overrepresented category that did not overlap with the differential expression categories. The glial-derived extracellular matrix is integrally involved in development, stabilization and plasticity of neuronal synapses, and is involved in promoting cell survival, facilitating repair, and maintaining synaptic current amplitude under stress conditions (Dityatev et al., 2010; Faissner et al., 2010; Wang et al., 2018). Together, these results suggest that stabilization of nervous function is an important component of cold tolerance, which is consistent with recent physiological literature (reviewed by Overgaard and MacMillan, 2017).

For CT_{max}, more of the GWAS candidate genes were differentially expressed, especially when considering the strongest candidates (**Supplementary Table 7**). Relatively few of the non-responsive genes were found within overrepresented categories, with a range of only 0–3 included in each of the 17 enriched CT_{max} GO categories (**Figure 5**). The few genes that consistently appeared in overrepresented categories, *fra, fz2* and *ptc*, are also functionally associated with the nervous system, including axon and dendrite guidance and synapse organization. Spreading depolarization of the central nervous system (triggered by failure to maintain ion gradients between the intra- and extracellular compartments) is linked with heat tolerance across *Drosophila* species (Jørgensen et al., 2019), indicating that neuronal failure is an important component of heat tolerance in addition to cold tolerance. An additional set of three genes, *Pax, sls* and *Zasp66*, co-occurred in several associated categories of muscle structure and development (**Figure 5**). Depolarization of muscles has been associated with lower thermal limits (e.g., MacMillan et al., 2014, 2015b) but to our knowledge has not been linked to upper limits.

CONCLUSION

Several studies have separately assessed the genetic architecture and plastic transcriptional responses to thermal stress, but the extent to which genes associated with thermal tolerance are involved in preparative and dynamic stress responses has not been assessed. Here, we show that genes associated with variation in thermal tolerance, identified via GWAS, included differentially expressed genes directly involved in the dynamic stress response, as well as a number of transcription factors that likely regulate these processes. However, while GWAS candidate genes were more likely to be differentially expressed, genes commonly associated with thermal stress, such as heat shock proteins (*hsp*), were not identified among GWAS genes. These core stress genes tend to be highly conserved, so it is likely that these genes have little genetic variation, especially in functional regions. In addition, consistent with previous studies (e.g., Sørensen et al., 2005; Telonis-Scott et al., 2013; MacMillan et al., 2016), most of the differentially expressed genes were downregulated for both hot and cold stresses. This result suggests an important role for shutting down certain biological processes during stress, and future studies should address these processes that are incompatible with stress tolerance, in addition to the well-studied protective pathways that are activated by stress.

A noteworthy finding of our study is a prominent role for genes involved in preparatory physiological processes that influence the condition of the organism at the time of thermal stress. While we did not observe an abundance of genes involved in processes commonly associated with preparation for thermal stress, such as cell membrane, circadian function, and immune response (Sinensky, 1974; Koštál, 2010; Teets and Hahn, 2018), we found a strong representation of nervous system processes for both CT_{min} and CT_{max} GWAS genes. Many of these genes have defined roles in development or morphogenesis, suggesting that developmental processes can influence thermal tolerance later in life, or that some of these genes are co-opted for thermal tolerance later in life. Future validation with other tools, such as RNA interference and transgenic overexpression, can clarify the precise role of these genes in thermal tolerance.

While all three of our hypotheses were supported for both cold and heat tolerance, genes associated with upper thermal

limits tended to be more involved in dynamic stress responses than those associated with cold tolerance. Furthermore, some genes associated with thermal tolerance appear to play multiple roles, highlighting that our three hypotheses are not mutually exclusive and that some genes likely have pleiotropic roles to shape thermal tolerance. Both upper and lower thermal limits had a strong genetic component, but the genetic signatures for these traits were largely distinct, with no overlapping SNPs and only three overlapping genes. Thus, heat and cold tolerance involve distinct molecular processes and can likely independently involve in response to changing environmental conditions.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**. Sequencing data have been deposited into the NCBI SRA database with accession Bioproject PRJNA612361.

AUTHOR CONTRIBUTIONS

NT, BL, SC, JW, and HA acquired the financial support for the project leading to this publication. NT, BL, and SC conceived and designed the study. ML, DA, LU, NJ, MW, EM, BP, and KB performed the experiments and data collection. ML, DA, TO'L, SF, BL, NT, and SC analyzed the data. ML, DA, TO'L, SF, BL, NT, and SC wrote the manuscript. All authors approved the final version of the manuscript.

FUNDING

This work was supported by the National Science Foundation grant OIA-1826689 to NT, BL, SF, JW, HA, and SC, Hatch Project 1010996 from the USDA National Institute of Food and Agriculture to NT. ML was partially supported by the National Science Foundation I/UCRC, The Center for Arthropod Management Technologies, under grant IIP-1821936, and by industry partners to NT.

ACKNOWLEDGMENTS

We thank Mark Garcia for help constructing and maintaining the apparatuses used for CT_{min} and CT_{max} measurements. We also thank Alison Gerken for providing their DGRP data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00658/full#supplementary-material

REFERENCES

Addo-Bediako, A., Chown, S. L., and Gaston, K. J. (2000). Thermal tolerance, climatic variability and latitude. *Proc. R. Soc. B* 267, 739–745. doi: 10.1098/rspb.2000.1065

Affolter, M., Montagne, J., Walldorf, U., Groppe, J., Kloter, U., LaRosa, M., et al. (1994). The Drosophila SRF homolog is expressed in a subset of tracheal cells and maps within a genomic region required for tracheal development. *Development* 120, 743–753.

Andersen, M. K., Jensen, N. J. S., Robertson, R. M., and Overgaard, J. (2018). Central nervous system shutdown underlies acute cold tolerance in tropical and temperate Drosophila species. *J. Exp. Biol.* 221:jeb179598. doi: 10.1242/jeb.179598

Andersen, M. K., and Overgaard, J. (2019). The central nervous system and muscular system play different roles for chill coma onset and recovery in insects. *Comp. Biochem. Physiol. Part A* 233, 10–16. doi: 10.1016/j.cbpa.2019.03.015

Angilletta, M. (2009). *Thermal Adaptation: A Theoretical and Empirical Synthesis*. Oxford: Oxford University Press.

Awde, D. N., Fowler, T. E., Pérez-Gálvez, F., Garcia, M. J., and Teets, N. M. (2020). High throughput assays of critical thermal limits in insects. *J. Vis. Exp.* e61186. doi: 10.3791/61186

Ayrinhac, A., Debat, V., Gibert, P., Kister, A.-G., Legout, H., Moreteau, B., et al. (2004). Cold adaptation in geographical populations of *Drosophila melanogaster*: phenotypic plasticity is more important than genetic variability. *Funct. Ecol.* 18, 700–706. doi: 10.1111/j.0269-8463.2004.00904.x

Barton, N. H., and Keightley, P. D. (2002). Understanding quantitative genetic variation. *Nat. Rev. Genet.* 3, 11–21. doi: 10.1038/nrg700

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *J. Stat. Softw.* 67:48. doi: 10.18637/jss.v067.i01

Bello, B. C., Hirth, F., and Gould, A. P. (2003). A Pulse of the Drosophila Hox protein abdominal-a schedules the end of neural proliferation via neuroblast apoptosis. *Neuron* 37, 209–219. doi: 10.1016/S0896-6273(02)01181-9

Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S., and Petrov, D. A. (2014). Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in drosophila. *PLoS Genetics* 10:e1004775. doi: 10.1371/journal.pgen.1004775

Bergland, A. O., Tobler, R., González, J., Schmidt, P., and Petrov, D. (2016). Secondary contact and local adaptation contribute to genome-wide patterns of clinal variation in *Drosophila melanogaster*. *Mol. Ecol.* 25, 1157–1174. doi: 10.1111/mec.13455

Bettencourt, B. R., Kim, I., Hoffmann, A. A., and Feder, M. E. (2002). Response to natural and laboratory selection at the Drosophila Hsp70 genes. *Evolution* 56, 1796–1801. doi: 10.1111/j.0014-3820.2002.tb00193.x

Bicknell, A. A., Cenik, C., Chua, H. N., Roth, F. P., and Moore, M. J. (2012). Introns in UTRs: Why we should stop ignoring them. *Bioessays* 34, 1025–1034. doi: 10.1002/bies.201200073

Bing, X., Zhang, J., and Sinclair, B. J. (2012). A comparison of Frost expression among species and life stages of Drosophila. *Insect Mol. Biol.* 21, 31–39. doi: 10.1111/j.1365-2583.2011.01108.x

Brown, J. B., Boley, N., Eisman, R., May, G. E., Stoiber, M. H., Duff, M. O., et al. (2014). Diversity and dynamics of the Drosophila transcriptome. *Nature* 512, 393–399. doi: 10.1038/nature12962

Buchanan, J. L., Meiklejohn, C. D., and Montooth, K. L. (2018). Mitochondrial dysfunction and infection generate immunity-fecundity tradeoffs in Drosophila. *Integr. Comp. Biol.* 58, 591–603. doi: 10.1093/icb/icy078

Calosi, P., Bilton, D. T., Spicer, J. I., Votier, S. C., and Atfield, A. (2010). What determines a species' geographical range? Thermal biology and latitudinal range size relationships in European diving beetles (Coleoptera: Dytiscidae). *J. Anim. Ecol.* 79, 194–204. doi: 10.1111/j.1365-2656.2009.01611.x

Carreira-Rosario, A., Bhargava, V., Hillebrand, J., Kollipara, R. K., Ramaswami, M., and Buszczak, M. (2016). Repression of Pumilio protein expression by Rbfox1 promotes germ cell differentiation. *Dev. Cell* 36, 562–571. doi: 10.1016/j.devcel.2016.02.010

Castañeda, L. E., Rezende, E. L., and Santos, M. (2015). Heat tolerance in Drosophila subobscura along a latitudinal gradient: contrasting patterns

between plastic and genetic responses. *Evolution* 69, 2721–2734. doi: 10.1111/evo.12757

Cenci, C., and Gould, A. P. (2005). Drosophila Grainyhead specifies late programmes of neural proliferation by regulating the mitotic activity and Hox-dependent apoptosis of neuroblasts. *Development* 132, 3835–3845. doi: 10.1242/dev.01932

Chown, S. L. (2001). Physiological variation in insects: hierarchical levels and implications. *J. Insect Physiol.* 47, 649–660. doi: 10.1016/S0022-1910(00)00163-3

Colinet, H., Fai Lee, S., and Hoffmann, A. (2010). Functional characterization of the frost gene in *Drosophila melanogaster*: importance for recovery from Chill Coma. *PLoS One* 5:e10925. doi: 10.1371/journal.pone.0010925

Colinet, H., and Hoffmann, A. (2010). Gene and protein expression of Drosophila Starvin during cold stress and recovery from chill coma. *Insect Biochem. Mol. Biol.* 40, 425–428. doi: 10.1016/j.ibmb.2010.03.002

Colinet, H., Overgaard, J., Com, E., and Sørensen, J. G. (2013). Proteomic profiling of thermal acclimation in *Drosophila melanogaster*. *Insect Biochem. Mol. Biol.* 43, 352–365. doi: 10.1016/j.ibmb.2013.01.006

Cooper, B. S., Hammad, L. A., Fisher, N. P., Karty, J. A., and Montooth, K. L. (2012). In a variable thermal environment selection favor greater plasticity of cell membranes in *Drosophila melanogaster*. *Evolution* 66, 1976–1984. doi: 10.1111/j.1558-5646.2011.01566.x

Cossins, A. R., and Prosser, C. L. (1978). Evolutionary adaptation of membranes to temperature. *Proc. Natl. Acad. Sci. U.S.A.* 75, 2040–2043. doi: 10.1073/pnas.75.4.2040

Costa-Broseta, Á, Perea-Resa, C., Castillo, M.-C., Ruíz, M. F., Salinas, J., and León, J. (2018). Nitric oxide controls constitutive freezing tolerance in Arabidopsis by attenuating the levels of osmoprotectants, stress-related hormones and anthocyanins. *Sci. Rep.* 8:9268. doi: 10.1038/s41598-018-27668-8

Dahlhoff, E. A., and Somero, G. N. (1993). Effects of temperature on mitochondria from Abalone (genus Haliotis): adaptive plasticity and its limits. *J. Exp. Biol.* 185, 151–168.

Dityatev, A., Schachner, M., and Sonderegger, P. (2010). The dual role of the extracellular matrix in synaptic plasticity and homeostasis. *Nat. Rev. Neurosci.* 11, 735–746. doi: 10.1038/nrn2898

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Donlea, J. M., Ramanan, N., and Shaw, P. J. (2009). Use-dependent plasticity in clock neurons regulates sleep need in Drosophila. *Science* 324, 105–108. doi: 10.1126/science.1166657

Dowd, W. W., King, F. A., and Denny, M. W. (2015). Thermal variation, thermal extremes and the physiological performance of individuals. *J. Exp. Biol.* 218, 1956–1967. doi: 10.1242/jeb.114926

Durham, M. F., Magwire, M. M., Stone, E. A., and Leips, J. (2014). Genome-wide analysis in Drosophila reveals age-specific effects of SNPs on fitness traits. *Nat. Commun.* 5:4338. doi: 10.1038/ncomms5338

Dworkin, I., and Gibson, G. (2006). Epidermal growth factor receptor and transforming growth factor-beta signaling contributes to variation for wing shape in *Drosophila melanogaster*. *Genetics* 173, 1417–1431. doi: 10.1534/genetics.105.053868

Fabian, D. K., Kapun, M., Nolte, V., Kofler, R., Schmidt, P. S., Schlötterer, C., et al. (2012). Genome-wide patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North America. *Mol. Ecol.* 21, 4748–4769. doi: 10.1111/j.1365-294X.2012.05731.x

Faissner, A., Pyka, M., Geissler, M., Sobik, T., Frischknecht, R., Gundelfinger, E. D., et al. (2010). Contributions of astrocytes to synapse formation and maturation - Potential functions of the perisynaptic extracellular matrix. *Brain Res. Rev.* 63, 26–38. doi: 10.1016/j.brainresrev.2010.01.001

Fallis, L. C., Fanara, J. J., and Morgan, T. J. (2014). Developmental thermal plasticity among *Drosophila melanogaster* populations. *J. Evol. Biol.* 27, 557–564. doi: 10.1111/jeb.12321

Feder, M. E., Cartaño, N. V., Milos, L., Krebs, R. A., and Lindquist, S. L. (1996). Effect of engineering Hsp70 copy number on Hsp70 expression and tolerance of ecologically relevant heat shock in larvae and pupae of *Drosophila melanogaster*. *J. Exp. Biol.* 199, 1837–1844.

Feder, M. E., and Hofmann, G. E. (1999). Heat-Shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu. Rev. Physiol.* 61, 243–282. doi: 10.1146/annurev.physiol.61.1.243

Fields, P. A., Dong, Y., Meng, X., and Somero, G. N. (2015). Adaptations of protein structure and function to temperature: there is more than one way to 'skin a cat'. *J. Exp. Biol.* 218, 1801–1811. doi: 10.1242/jeb.114298

Freda, P. J., Alex, J. T., Morgan, T. J., and Ragland, G. J. (2017). Genetic decoupling of thermal hardiness across metamorphosis in *Drosophila melanogaster*. *Integr. Comp. Biol.* 57, 999–1009. doi: 10.1093/icb/icx102

Freda, P. J., Ali, Z. M., Heter, N., Ragland, G. J., and Morgan, T. J. (2019). Stage-specific genotype-by-environment interactions for cold and heat hardiness in *Drosophila melanogaster*. *Heredity* 123, 479–491. doi: 10.1038/s41437-019-0236-9

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., et al. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* 11, 4241–4257. doi: 10.1091/mbc.11.12.4241

Gaston, K. J., and Chown, S. L. (1999). Elevation and climatic tolerance: A test using dung beetles. *Oikos* 86, 584–590. doi: 10.2307/3546663

Gerken, A. R., Eller, O. C., Hahn, D. A., and Morgan, T. J. (2015). Constraints, independence, and evolution of thermal plasticity: probing genetic architecture of long- and short-term thermal acclimation. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4399–4404. doi: 10.1073/pnas.1503456112

González-Tokman, D., Córdoba-Aguilar, A., Dáttilo, W., Lira-Noriega, A., Sánchez-Guillén, R. A., and Villalobos, F. (2020). Insect responses to heat: physiological mechanisms, evolution and ecological implications in a warming world. *Biol. Rev.* 95, 802–821. doi: 10.1111/brv.12588

Goto, S. G. (2001). A novel gene that is up-regulated during recovery from cold shock in *Drosophila melanogaster*. *Gene* 270, 259–264. doi: 10.1016/S0378-1119(01)00465-6

Gracey, A. Y. (2007). Interpreting physiological responses to environmental change through gene expression profiling. *J. Exp. Biol.* 210, 1584–1592. doi: 10.1242/jeb.004333

Gu, J., and Hilser, V. J. (2009). Sequence-based analysis of protein energy landscapes reveals nonuniform thermal adaptation within the proteome. *Mol. Biol. Evol.* 26, 2217–2227. doi: 10.1093/molbev/msp140

Hoffmann, A. A., Anderson, A., and Hallas, R. (2002). Opposing clines for high and low temperature resistance in *Drosophila melanogaster*. *Ecol. Lett.* 5, 614–618. doi: 10.1046/j.1461-0248.2002.00367.x

Hoffmann, A. A., Sørensen, J. G., and Loeschcke, V. (2003). Adaptation of *Drosophila* to temperature extremes: bringing together quantitative and molecular approaches. *J. Therm. Biol.* 28, 175–216. doi: 10.1016/S0306-4565(02)00057-8

Hofmann, G. E., Buckley, B. A., Airaksinen, S., Keen, J. E., and Somero, G. N. (2000). Heat-shock protein expression is absent in the antarctic fish Trematomus bernacchii (family Nototheniidae). *J. Exp. Biol.* 203, 2331–2339.

Huey, R. B., Crill, W. D., Kingsolver, J. G., and Weber, K. E. (1992). A method for rapid measurement of heat or cold resistance of small insects. *Funct. Ecol.* 6, 489–494.

Ivanov, D. K., Escott-Price, V., Ziehm, M., Magwire, M. M., Mackay, T. F. C., Partridge, L., et al. (2015). Longevity GWAS Using the Drosophila Genetic Reference Panel. *J. Gerontol. A Biol. Sci. Med. Sci.* 70, 1470–1478. doi: 10.1093/gerona/glv047

Jørgensen, L. B., Meldrum Robertson, R., and Overgaard, J. (2019). Neural dysfunction correlates with heat coma and CTmax in Drosophila but does not set the boundaries for heat stress survival. *bioRxiv[Preprint]* doi: 10.1101/844316

Keightley, P. D., and Lynch, M. (2003). Toward a realistic model of mutations affecting fitness. *Evolution* 57, 683–685. doi: 10.1111/j.0014-3820.2003.tb01561.x

Kellermann, V., Loeschcke, V., Hoffmann, A. A., Kristensen, T. N., Fløjgaard, C., David, J. R., et al. (2012a). Phylogenetic constrains in key functional traits behind species' climate niches: patterns of desiccation and cold resistance across 95 *Drosophila* species. *Evolution* 66, 3377–3389. doi: 10.1111/j.1558-5646.2012.01685.x

Kellermann, V., Overgaard, J., Hoffmann, A. A., Fløjgaard, C., Svenning, J.-C., and Loeschcke, V. (2012b). Upper thermal limits of *Drosophila* are linked to species

distributions and strongly constrained phylogenetically. *Proc. Natl. Acad. Sci. USA* 109, 16228–16233. doi: 10.1073/pnas.1207553109

Kelsey, K. J. P., and Clark, A. G. (2017). Variation in position effect variegation within a natural population. *Genetics* 207, 1157–1166. doi: 10.1534/genetics.117.300306

Klose, M. K., and Robertson, R. M. (2004). Stress-induced thermoprotection of neuromuscular transmission. *Integr. Comp. Biol.* 44, 14–20. doi: 10.1093/icb/44.1.14

Königer, A., and Grath, S. (2018). Transcriptome analysis reveals candidate genes for cold tolerance in *Drosophila ananassae*. *Genes* 9:624. doi: 10.3390/genes9120624

Koštál, V. (2010). "Cell structural modifications in insects at low temperatures," in *Low Temperature Biology of Insects*, eds D. L. Denlinger and J. R. E. Lee (Cambridge: Cambridge University Press), 116–140. doi: 10.1017/cbo9780511675997.006

Krattinger, A., Gendre, N., Ramaekers, A., Grillenzoni, N., and Stocker, R. F. (2007). DmOAZ, the unique *Drosophila melanogaster* OAZ homologue is involved in posterior spiracle development. *Dev. Genes Evol.* 217, 197–208. doi: 10.1007/s00427-007-0134-7

Lakhotia, S. C., Mallik, M., Singh, A. K., and Ray, M. (2012). The large noncoding hsrω-n transcripts are essential for thermotolerance and remobilization of hnRNPs. HP1 and RNA polymerase II during recovery from heat shock in Drosophila. *Chromosoma* 121, 49–70. doi: 10.1007/s00412-011-0341-x

Lancaster, L. T., Dudaniec, R. Y., Chauhan, P., Wellenreuther, M., Svensson, E. I., and Hansson, B. (2016). Gene expression under thermal stress varies across a geographical range expansion front. *Mol. Ecol.* 25, 1141–1156. doi: 10.1111/mec.13548

Leemans, R., Egger, B., Loop, T., Kammermeier, L., He, H., Hartmann, B., et al. (2000). Quantitative transcript imaging in normal and heat-shocked Drosophila embryos by using high-density oligonucleotide arrays. *Proc. Natl. Acad. Sci. U.S.A.* 97, 12138–12143. doi: 10.1073/pnas.210066997

Lerman, D. N., Michalak, P., Helin, A. B., Bettencourt, B. R., and Feder, M. E. (2003). Modification of heat-shock gene expression in *Drosophila melanogaster* populations via transposable elements. *Mol. Biol. Evol.* 20, 135–144. doi: 10.1093/molbev/msg015

Li, K., Tian, Y., Yuan, Y., Fan, X., Yang, M., He, Z., et al. (2019). Insights into the Functions of LncRNAs in Drosophila. *Int. J. Mol. Sci.* 20:4646. doi: 10.3390/ijms20184646

Liao, Y., Smyth, G. K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* 47, e47–e47. doi: 10.1093/nar/gkz114

Liu, J. Z., McRae, A. F., Nyholt, D. R., Medland, S. E., Wray, N. R., Brown, K. M., et al. (2010). A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* 87, 139–145. doi: 10.1016/j.ajhg.2010.06.009

Lockwood, B. L., Sanders, J. G., and Somero, G. N. (2010). Transcriptomic responses to heat stress in invasive and native blue mussels (genus Mytilus): molecular correlates of invasive success. *J. Exp. Biol.* 213, 3548–3558. doi: 10.1242/jeb.046094

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-558

Mackay, T. F. C., and Huang, W. (2018). Charting the genotype–phenotype map: lessons from the *Drosophila melanogaster* Genetic Reference Panel. *WIREs Dev. Biol.* 7:e289. doi: 10.1002/wdev.289

Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., et al. (2012). The *Drosophila melanogaster* genetic reference panel. *Nature* 482, 173–178. doi: 10.1038/nature10811

MacMillan, H. A., Andersen, J. L., Loeschcke, V., and Overgaard, J. (2015a). Sodium distribution predicts the chill tolerance of *Drosophila melanogaster* raised in different thermal conditions. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 308, R823–R831. doi: 10.1152/ajpregu.00465.2014

MacMillan, H. A., Baatrup, E., and Overgaard, J. (2015b). Concurrent effects of cold and hyperkalaemia cause insect chilling injury. Proceedings. *Biol. Sci.* 282:20151483. doi: 10.1098/rspb.2015.1483

MacMillan, H. A., Findsen, A., Pedersen, T. H., and Overgaard, J. (2014). Cold-induced depolarization of insect muscle: differing roles of extracellular K+ during acute and chronic chilling. *J. Exp. Biol.* 217, 2930–2938. doi: 10.1242/jeb.107516

MacMillan, H. A., Knee, J. M., Dennis, A. B., Udaka, H., Marshall, K. E., Merritt, T. J. S., et al. (2016). Cold acclimation wholly reorganizes the *Drosophila melanogaster* transcriptome and metabolome. *Sci. Rep.* 6:28999. doi: 10.1038/srep28999

Manjunatha, H. B., Rajesh, R. K., and Aparna, H. S. (2010). Silkworm thermal biology: a review of heat shock response, heat shock proteins and heat acclimation in the domesticated silkworm. *Bombyx mori. J. Insect Sci.* 10:204. doi: 10.1673/031.010.20401

McMillan, D. M., Fearnley, S. L., Rank, N. E., and Dahlhoff, E. P. (2005). Natural temperature variation affects larval survival, development and Hsp70 expression in a leaf beetle. *Funct. Ecol.* 19, 844–852. doi: 10.1111/j.1365-2435.2005.01031.x

Morgan, T. J., and Mackay, T. F. C. (2006). Quantitative trait loci for thermotolerance phenotypes in *Drosophila melanogaster*. *Heredity* 96, 232–242. doi: 10.1038/sj.hdy.6800786

Nyamukondiwa, C., Terblanche, J. S., Marshall, K. E., and Sinclair, B. J. (2011). Basal cold but not heat tolerance constrains plasticity among Drosophila species (Diptera: Drosophilidae). *J. Evol. Biol.* 24, 1927–1938. doi: 10.1111/j.1420-9101.2011.02324.x

Ørsted, M., Rohde, P. D., Hoffmann, A. A., Sørensen, P., and Kristensen, T. N. (2018). Environmental variation partitioned into separate heritable components. *Evolution* 72, 136–152. doi: 10.1111/evo.13391

Overgaard, J., and MacMillan, H. A. (2017). The Integrative Physiology of Insect Chill Tolerance. *Annu. Rev. Physiol.* 79, 187–208. doi: 10.1146/annurev-physiol-022516-034142

Parankusam, S., Adimulam, S. S., Bhatnagar-Mathur, P., and Sharma, K. K. (2017). Nitric oxide (NO) in plant heat stress tolerance: current knowledge and perspectives. *Front. Plant Sci.* 8:1582. doi: 10.3389/fpls.2017.01582

Patterson, N., Price, A. L., and Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet.* 2:e190. doi: 10.1371/journal.pgen.0020190

Podrabsky, J. E., and Somero, G. N. (2004). Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. *J. Exp. Biol.* 207, 2237–2254. doi: 10.1242/jeb.01016

Pörtner, H. O. (2002). Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. *Comp. Biochem. Physiol. Part A* 132, 739–761. doi: 10.1016/S1095-6433(02)00045-4

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Qin, W., Neal, S. J., Robertson, R. M., Westwood, J. T., and Walker, V. K. (2005). Cold hardening and transcriptional change in *Drosophila melanogaster*. *Insect Mol. Biol.* 14, 607–613. doi: 10.1111/j.1365-2583.2005.00589.x

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rako, L., Blacket, M. J., McKechnie, S. W., and Hoffmann, A. A. (2007). Candidate genes and thermal phenotypes: identifying ecologically important genetic variation for thermotolerance in the Australian *Drosophila melanogaster* cline. *Mol. Ecol.* 16, 2948–2957. doi: 10.1111/j.1365-294X.2007.03332.x

Robertson, R. M., and Money, T. G. A. (2012). Temperature and neuronal circuit function: compensation, tuning and tolerance. *Curr. Opin. Neurobiol.* 22, 724–734. doi: 10.1016/j.conb.2012.01.008

Rohde, P. D., Krag, K., Loeschcke, V., Overgaard, J., Sørensen, P., and Kristensen, T. N. (2016). A quantitative genomic approach for analysis of fitness and stress related traits in a *Drosophila melanogaster* model population. *Int. J. Genom.* 2016, 1–11. doi: 10.1155/2016/2157494

Rolandi, C., Lighton, J. R. B., de la Vega, G. J., Schilman, P. E., and Mensch, J. (2018). Genetic variation for tolerance to high temperatures in a population of *Drosophila melanogaster*. *Ecol. Evol.* 8, 10374–10383. doi: 10.1002/ece3.4409

Roy, B. A., and Kirchner, J. W. (2000). Evolutionary dynamics of pathogen resistance and tolerance. *Evolution* 54, 51–63. doi: 10.1111/j.0014-3820.2000.tb00007.x

Sato, K., and Yamamoto, D. (2020). The mode of action of Fruitless: is it an easy matter to switch the sex? *Genes Brain Behav.* 19:e12606. doi: 10.1111/gbb.12606

Schou, M. F., Mouridsen, M. B., Sørensen, J. G., and Loeschcke, V. (2017). Linear reaction norms of thermal limits in Drosophila: predictable plasticity in cold

but not in heat tolerance. *Funct. Ecol.* 31, 934–945. doi: 10.1111/1365-2435.12782

Sgrò, C. M., Overgaard, J., Kristensen, T. N., Mitchell, K. A., Cockerell, F. E., and Hoffmann, A. A. (2010). A comprehensive assessment of geographic variation in heat tolerance and hardening capacity in populations of *Drosophila melanogaster* from eastern Australia. *J. Evol. Biol.* 23, 2484–2493. doi: 10.1111/j.1420-9101.2010.02110.x

Sinclair, B. J., Coello Alvarado, L. E., and Ferguson, L. V. (2015). An invitation to measure insect cold tolerance: methods, approaches, and workflow. *J. Ther. Biol.* 53, 180–197. doi: 10.1016/j.jtherbio.2015.11.003

Sinclair, B. J., Gibbs, A. G., and Roberts, S. P. (2007). Gene transcription during exposure to, and recovery from, cold and desiccation stress in *Drosophila melanogaster*. *Insect Mol. Biol.* 16, 435–443. doi: 10.1111/j.1365-2583.2007.00739.x

Sinensky, M. (1974). Homeoviscous adaptation–a homeostatic process that regulates the viscosity of membrane lipids in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 71, 522–525. doi: 10.1073/pnas.71.2.522

Sørensen, J. G., Dahlgaard, J., and Loeschcke, V. (2001). Genetic variation in thermal tolerance among natural populations of *Drosophila buzzatii*: down regulation of Hsp70 expression and variation in heat stress resistance traits. *Funct. Ecol.* 15, 289–296. doi: 10.1046/j.1365-2435.2001.00525.x

Sørensen, J. G., and Loeschcke, V. (2001). Larval crowding in *Drosophila melanogaster* induces Hsp70 expression, and leads to increased adult longevity and adult thermal stress resistance. *J. Insect Physiol.* 47, 1301–1307. doi: 10.1016/S0022-1910(01)00119-6

Sørensen, J. G., Nielsen, M. M., Kruhøffer, M., Justesen, J., and Loeschcke, V. (2005). Full genome gene expression analysis of the heat stress response in *Drosophila melanogaster*. *Cell Stress Chaperones* 10, 312–328. doi: 10.1379/csc-128r1.1

Sørensen, J. G., Schou, M. F., Kristensen, T. N., and Loeschcke, V. (2016). Thermal fluctuations affect the transcriptome through mechanisms independent of average temperature. *Sci. Rep.* 6:30975. doi: 10.1038/srep30975

Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9440–9445. doi: 10.1073/pnas.1530509100

Storey, K. B., and Storey, J. M. (2012). Insect cold hardiness: metabolic, gene, and protein adaptation. *Can. J. Zool.* 90, 456–475. doi: 10.1139/z2012-011

Štros, M. (2010). HMGB proteins: Interactions with DNA and chromatin. *Biochim. Biophys. Acta* 1799, 101–113. doi: 10.1016/j.bbagrm.2009.09.008

Stucki, D., Freitak, D., and Sundström, L. (2017). Survival and gene expression under different temperature and humidity regimes in ants. *PLoS One* 12:e0181137. doi: 10.1371/journal.pone.0181137

Sunday, J., Bennett, J. M., Calosi, P., Clusella-Trullas, S., Gravel, S., Hargreaves, A. L., et al. (2019). Thermal tolerance patterns across latitude and elevation. *Philos. Trans. R. Soc. B* 374:20190036. doi: 10.1098/rstb.2019.0036

Sunday, J. M., Bates, A. E., and Dulvy, N. K. (2011). Global analysis of thermal tolerance and latitude in ectotherms. *Proc. R. Soc. B* 278, 1823–1830. doi: 10.1098/rspb.2010.1295

Svetec, N., Werzner, A., Wilches, R., Pavlidis, P., Álvarez-Castro, J. M., Broman, K. W., et al. (2011). Identification of X-linked quantitative trait loci affecting cold tolerance in *Drosophila melanogaster* and fine mapping by selective sweep analysis. *Mol. Ecol.* 20, 530–544. doi: 10.1111/j.1365-294X.2010.04951.x

Teets, N. M., and Hahn, D. A. (2018). Genetic variation in the shape of cold-survival curves in a single fly population suggests potential for selection from climate variability. *J. Evol. Biol.* 31, 543–555. doi: 10.1111/jeb.13244

Teets, N. M., Peyton, J. T., Ragland, G. J., Colinet, H., Renault, D., Hahn, D. A., et al. (2012). Combined transcriptomic and metabolomic approach uncovers molecular mechanisms of cold tolerance in a temperate flesh fly. *Physiol. Genomics* 44, 764–777. doi: 10.1152/physiolgenomics.00042.2012

Telonis-Scott, M., Hallas, R., McKechnie, S. W., Wee, C. W., and Hoffmann, A. A. (2009). Selection for cold resistance alters gene transcript levels in *Drosophila melanogaster*. *J. Insect Physiol.* 55, 549–555. doi: 10.1016/j.jinsphys.2009.01.010

Telonis-Scott, M., van Heerwaarden, B., Johnson, T. K., Hoffmann, A. A., and Sgrò, C. M. (2013). New levels of transcriptome complexity at upper thermal limits in wild Drosophila revealed by exon expression analysis. *Genetics* 195, 809–830. doi: 10.1534/genetics.113.156224

Thran, J., Poeck, B., and Strauss, R. (2013). Serum response factor-mediated gene regulation in a Drosophila visual working memory. *Curr. Biol.* 23, 1756–1763. doi: 10.1016/j.cub.2013.07.034

Thurmond, J., Goodman, J. L., Strelets, V. B., Attrill, H., Gramates, L. S., Marygold, S. J., et al. (2019). FlyBase 2.0: *the next generation*. *Nucleic Acids Res.* 47, D759–D765. doi: 10.1093/nar/gky1003

Usha, N., and Shashidhara, L. S. (2010). Interaction between Ataxin-2 Binding Protein 1 and Cubitus-interruptus during wing development in Drosophila. *Dev. Biol.* 341, 389–399. doi: 10.1016/j.ydbio.2010.02.039

von Heckel, K., Stephan, W., and Hutter, S. (2016). Canalization of gene expression is a major signature of regulatory cold adaptation in temperate *Drosophila melanogaster*. *BMC Genomics* 17:574. doi: 10.1186/s12864-016-2866-0

Wang, J., Vasaikar, S., Shi, Z., Greer, M., and Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* 45, W130–W137. doi: 10.1093/nar/gkx356

Wang, L., Zhang, W., Li, Q., and Zhu, W. (2015). *AssocTests: Genetic Association Studies. R package version 0.0-3*. Available online at: https://CRAN.R-project.org/package=AssocTests (accessed April, 2020).

Wang, T., Sinha, A. S., Akita, T., Yanagawa, Y., and Fukuda, A. (2018). Alterations of GABAergic neuron-associated extracellular matrix and synaptic responses in Gad1-heterozygous mice subjected to prenatal stress. *Front. Cell. Neurosci.* 12:284. doi: 10.3389/fncel.2018.00284

Weber, A. L., Khan, G. F., Magwire, M. M., Tabor, C. L., Mackay, T. F. C., and Anholt, R. R. H. (2012). Genome-wide association analysis of oxidative stress resistance in *Drosophila melanogaster*. *PLoS One* 7:e34745. doi: 10.1371/journal.pone.0034745

Welte, M. A., Tetrault, J. M., Dellavalle, R. P., and Lindquist, S. L. (1993). A new method for manipulating transgenes: engineering heat tolerance in a complex, multicellular organism. *Curr. Biol.* 3, 842–853. doi: 10.1016/0960-9822(93)90218-D

Werren, J. H. (1997). Biology of Wolbachia. *Annu. Rev. Entomol.* 42, 587–609. doi: 10.1146/annurev.ento.42.1.587

Williams, C. M., Henry, H. A. L., and Sinclair, B. J. (2015). Cold truths: how winter drives responses of terrestrial organisms to climate change. *Biol. Rev.* 90, 214–235. doi: 10.1111/brv.12105

Willot, Q., Gueydan, C., and Aron, S. (2017). Proteome stability, heat hardening and heat-shock protein expression profiles in Cataglyphis desert ants. *J. Exp. Biol.* 220, 1721–1728. doi: 10.1242/jeb.154161

Wos, G., and Willi, Y. (2015). Temperature-stress resistance and tolerance along a latitudinal cline in North American *Arabidopsis lyrata*. *PLoS One* 10:e0131808. doi: 10.1371/journal.pone.0131808

Yancey, P. H. (2005). Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J. Exp. Biol.* 208, 2819–2830. doi: 10.1242/jeb.01730

Zatsepina, O. G., Velikodvorskaia, V. V., Molodtsov, V. B., Garbuz, D., Lerman, D. N., Bettencourt, B. R., et al. (2001). A *Drosophila melanogaster* strain from sub-equatorial Africa has exceptional thermotolerance but decreased Hsp70 expression. *J. Exp. Biol.* 204, 1869–1881.

Zhou, S., Campbell, T. G., Stone, E. A., Mackay, T. F. C., and Anholt, R. R. H. (2012). Phenotypic plasticity of the Drosophila transcriptome. *PLoS Genet.* 8:e1002593. doi: 10.1371/journal.pgen.1002593

Zhu, G., Xue, M., Luo, Y., Ji, G., Liu, F., Zhao, H., et al. (2017). Effects of short-term heat shock and physiological responses to heat stress in two *Bradysia adults*, *Bradysia odoriphaga* and *Bradysia difformis*. *Sci. Rep.* 7:13381. doi: 10.1038/s41598-017-13560-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

PERSPECTIVE
published: 21 July 2020
doi: 10.3389/fevo.2020.00242

Climate Change Genomics Calls for Standardized Data Reporting

Ann-Marie Waldvogel[1,2], Dennis Schreiber[1,3], Markus Pfenninger[1,3,4] and Barbara Feldmeyer[1]*

[1] Molecular Ecology Group, Senckenberg Biodiversity and Climate Research Centre, Frankfurt, Germany, [2] Institute of Zoology, University of Cologne, Cologne, Germany, [3] Institute for Organismic and Molecular Evolution, Johannes Gutenberg University, Mainz, Germany, [4] LOEWE Centre for Translational Biodiversity Genomics, Senckenberg Biodiversity and Climate Research Centre, Frankfurt, Germany

The advent of new and affordable high-throughput sequencing techniques allows for the investigation of the genetic basis of environmental adaptation throughout the plant and animal kingdom. The framework of genotype-environment associations (GEA) provides a powerful link by correlating the geographic distribution of genotype patterns of individuals or populations with environmental factors on a spatial scale. We coarsely review the short history of GEA studies, summarizing available studies, organisms, data type, and data availability for these studies. GEA is a powerful tool in climate change research and we therefore focus on climate variables as environmental factors. While our initial aim was to compare results of existing studies to identify common patterns or differences in climate adaptation, we quickly realized that such a meta-analysis approach is currently unfeasible. Based on our literature review we discuss the current shortcomings and lack of data accessibility which impede meta-analyses. Such meta-analyses would allow to draw conclusions on traits and functions crucial to adapt to different environmental, e.g., climate conditions, across species. We thus make a strong call for standardized data and reposition structure for GEA studies. Moreover, the coordinated documentation of candidate genes associated to environmental factors could allow the establishment of a new and additional gene ontology domain "environmental association." This would systematically link fitness relevant genes to the corresponding environmental factor.

Keywords: meta-anaylses, literature survey, environmental association analysis, gene ontology category, candidate genes

OPEN ACCESS

Edited by:
Inês Fragata,
University of Lisbon, Portugal

Reviewed by:
Helena Kristiina Wirta,
University of Helsinki, Finland
Ricardo Alia,
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Spain
Amanda De La Torre,
Northern Arizona University, United States

****Correspondence:***
Markus Pfenninger
Markus.Pfenninger@senckenberg.de

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Ecology and Evolution

Received: *17 March 2020*
Accepted: *06 July 2020*
Published: *21 July 2020*

Citation:
Waldvogel A-M, Schreiber D, Pfenninger M and Feldmeyer B (2020) Climate Change Genomics Calls for Standardized Data Reporting. Front. Ecol. Evol. 8:242. doi: 10.3389/fevo.2020.00242

THE POWER OF GENOTYPE-ENVIRONMENT ASSOCIATION ANALYSES

Studying the genomic and molecular underpinnings of adaptation is a central aim in evolutionary biology. As abiotic and biotic conditions vary over space and time, organisms adapt to various local environmental conditions (Bradshaw and Holzapfel, 2001; Kawecki and Ebert, 2004). Alternatively, organisms may also be phenotypically plastic and able to thrive in variable environments. The degree of adaptability and/or plasticity is a crucial parameter in the face of global climate change, which is and will be affecting almost every organism and community across the globe. To make inferences on adaptability versus plasticity, scientists conduct life history experiments, imitating the environmental conditions of interest and determining the organisms' responses in

order to reveal the breadth of their genetic and phenotypic response spectrum. Over decades, laborious quantitative trait locus (QTL) approaches were the only means of obtaining information on the underlying genetic basis targeted by various scenarios of selection (Lynch and Walsh, 1998). However, such quantitative genetic studies are restricted to a limited number of model organisms with sufficient genetic resources (e.g., Mackay, 2014). With the advent of new sequencing technologies, it is now possible to investigate the genomic basis of adaptation to environmental drivers in model organisms, but even more importantly, also in a broad range of non-model organisms (Waldvogel et al., 2020). We can now determine which genes, which regulatory regions, which epigenetic mechanism, etc. play a role in adaptation to different environmental selection pressures. With the accumulation of such studies, we expect that general common patterns will emerge, yielding a deeper understanding of the genomic mechanisms of environmental adaptation and at the same time increasing the predictive power toward new selective challenges. This, however, requires the comparison of studies. In this paper, we will first rehearse the principles and approaches of genotype-environment association (GEA) studies, then review the existing literature, highlight the shortcomings and incompatibilities of current data report practises and finish with suggestions and recommendations that would allow meta-analyses in future.

The Concept of GEA

Genotype-environment association studies, also called environmental association analyses (EAA), provide an approach to detect genetic signatures of selection that result from environmental factors by correlating geographic and genome-wide distributions of allele frequencies with environmental variation (Rellstab et al., 2015; Hoban et al., 2016; Forester et al., 2018). In contrast to classical F_{ST}-outlier approaches (see Whitlock and Lotterhos, 2015 for details), GEA allows for the detection of weakly selected loci which only show moderate or even subtle allele frequency shifts (Hancock et al., 2010; De La Torre et al., 2019), a pattern that is characteristic for polygenic adaptation (Pritchard and Di Rienzo, 2010; Berg and Coop, 2014). As in any association study, one needs to keep in mind that identified single nucleotide polymorphisms (SNPs) do not necessarily represent the functional site under selection (Wang et al., 2010). Candidate SNPs can simply be in close linkage to the actual functional site under selection, and the functional site may not be represented in the candidate list, e.g., due to coverage issues at the according site. It is thus important to combine several genomic levels of investigation, e.g., the candidate SNPs as such, candidate SNPs located within genes or in regions up- or downstream of genes with presumed regulatory function, as well as local covariation of candidate SNPs within the range of a defined genomic windows (e.g., resulting from variation in local recombination sites).

The power of GEA increases with (a) the completeness of genotypic information, and (b) the number of populations that cover and possibly replicate the environmental gradient on a broad geographical scale. Genotypic information is derived from genome-wide sequence data of multiple individuals allowing the estimation of allele frequencies per populations. Three major sequencing strategies can be distinguished, resulting in different types of genome-wide sequence information (hereafter referred to as genomic data): (1) reduced-representation sequencing (RRS) data of single individuals, (2) pooled sequencing (Pool-Seq) data of multiple individuals per population, (3) whole genome sequencing (WGS) data of single individuals (more details on pros and cons of the three data types in Waldvogel et al., 2020). Most GEA studies, and especially the pioneer studies of this approach, are based on RRS using SNP-arrays or restriction site associated DNA markers sequencing (RAD-Seq) data (**Figure 1**). These are highly cost effective and provide a snapshot in targeted marker regions or around restriction enzyme sites (Catchen et al., 2017), at the cost of mainly identifying more or less closely linked variation instead of the functional region itself. Another means to reduce sequencing effort but at the same time cover the complete genome are Pool-Seq approaches, in which multiple individuals ($N \sim 100$) per population are pooled for sequencing (Schlötterer et al., 2014). This approach is a cost-effective way of obtaining genome wide information with the only downside of not being able to infer linkage (but see Feder et al., 2012). The third option, WGS, allows to obtain the most detailed information. However, it is characterized by large amounts of sequence data and also the most cost intense option. Here, whole genomes of single individuals are usually sequenced at low coverage. A crucial prerequisite for the latter two approaches is the existence of a reference genome of at least intermediate quality in terms of contiguity and annotation completeness.

GEA Data Structure

The constantly increasing number of GEA studies (**Figure 1**) share the common goal of identifying the genomic basis of environmental adaptation. Currently, however, each study focuses on its specific research questions, working on a single or a limited number of species and thus a small fraction of existing biodiversity. To obtain the overall picture on the evolutionary responses of biodiversity to environmental change, we need to compare studies and conciliate results within animal, plant and fungi species or even across the tree of life. A promising way to identify such general patterns are comprehensive meta-analyses (Nakagawa and Poulin, 2012). Indeed, our aim was to conduct such a meta-analysis to identify common patterns of climate adaptation, but we quickly realized that such an approach is currently unfeasible. Even though several recommendation papers covering GEA study designs already exist (especially Rellstab et al., 2015), compatibility among studies was not given. A multitude of data formats and incomplete data sets seem to be a common problem in ecology and evolution research (Whitlock, 2011; Parker et al., 2016; Culina et al., 2018; Poisot et al., 2019). We therefore compared the structure and accessibility of data from a representative fraction of GEA studies published between 2014 and mid-2019 (see assessment criteria below), which revealed the incompatibility or mainly inaccessibility of data, a prerequisite for any meta-analysis.

We focused on studies using mixed effect models as statistical framework implemented in the four currently most widely used tools: Bayenv (Coop et al., 2010), Bayenv2 (Günther and Coop, 2013), Baypass (Gautier, 2015), and LFMM (Latent factor

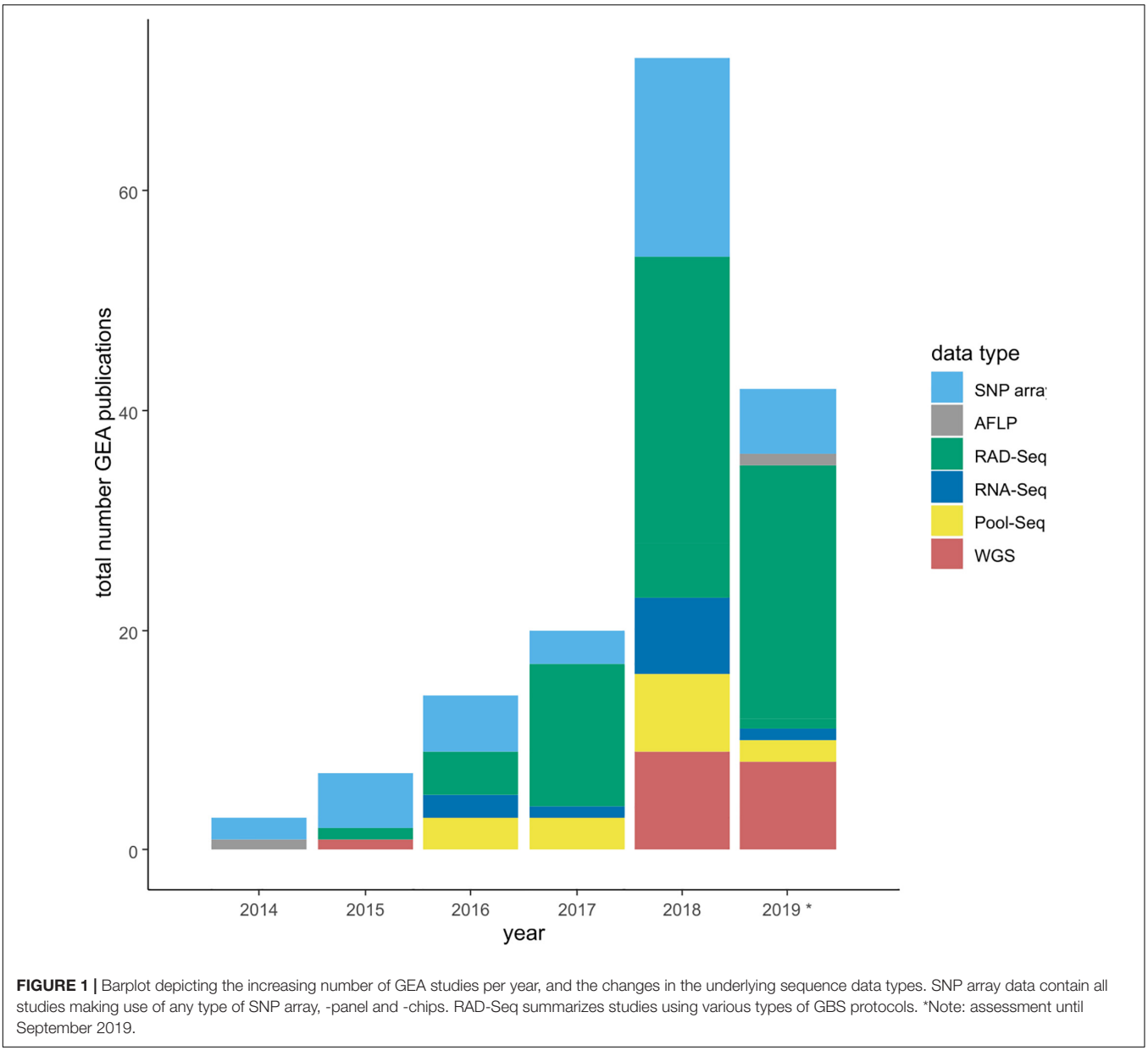


FIGURE 1 | Barplot depicting the increasing number of GEA studies per year, and the changes in the underlying sequence data types. SNP array data contain all studies making use of any type of SNP array, -panel and -chips. RAD-Seq summarizes studies using various types of GBS protocols. *Note: assessment until September 2019.

mixed models; Frichot et al., 2013; details in **Box 1**). In short, Bayenv, a Bayesian approach, tests whether the null model including the environmental factor better fits the data when compared to a model that is purely based on neutral genetic structure. Bayenv2 uses the same approach but is optimized for Pool-Seq data (Günther and Coop, 2013). Baypass (Gautier, 2015) is another Bayesian framework to identify differentiated markers, but correcting for demographic effects. LFMM (Latent factor mixed models; Frichot et al., 2013) introduces neutral genetic structure as a random factor, with the advantage of simultaneously estimating the effects of environmental factors and neutral genetic structure.

To accumulate a list of studies (**Supplementary Table 1**) for our meta-statistics, we used the original articles in which the four tools have been published (see above for references). From these we followed the "cited by" option on Google Scholar as link to all citing articles (as of September 2019). These articles were manually curated to only retain GEA studies. We decided to include these four tools only, since they appeared to sufficiently reflect the broad patterns. Other GEA-tools that follow the approach of a redundancy analysis (RDA) are widely used in a broad ecological context, resulting in hundreds of citations of which only a small fraction was relevant for our purpose. Our assessment resulted in 159 empirical GEA studies (**Figure 1**; **Supplementary Table 1**), covering multiple data types, and various organisms (**Figure 2**). Data type, i.e., the type of genomic data used for the GEA, obviously reflected the progress of DNA sequencing technology (**Figure 1**): starting off with amplified fragment length polymorphism (AFLP) and SNP array data as RRS strategies, progressing toward WGS data of single

BOX 1 | Statistical approaches to genotype-environment associations.

Statistical approaches for the inference of genotype-environment associations in genomic data sets are manifold. Aiming at the detection of multilocus selection patterns in response to environmental predictors, multivariate approaches that analyze many loci simultaneously can be considered most promising. We here outline a non-exhaustive selection of some commonly applied methods (and corresponding tools):

Differentiation-based methods

Allele frequencies of multiple populations are correlated with environmental variables. Statistical methods are based on mixed effect model, fitted to Bayesian or latent frameworks, to test correlations among multilocus allele frequencies of individuals or populations (response variable) and environmental factors (fixed factors), while accounting for population structure and relatedness between populations (random factor).

Bayesian models (implemented in BAYENV, BAYENV2 and BAYPASS; respectively Coop et al., 2010; Günther and Coop, 2013; Gautier, 2015) test for a correlation between allele frequencies and an environmental variable, while accounting for differences in sample size and population structure. The univariate approach to calculated Bayes factors per locus (Bayenv) was further extended by a differentiation-based approach in Bayenv2 (X^TX statistics) and robustness of models was further refined in Baypass.

Latent factor mixed models (implemented in LFMM, Frichot et al., 2013) detect correlations between genetic and environmental variation while simultaneously inferring background levels of population structure using unobserved variables (latent factors).

Methods based on constrained ordinations

Multivariate statistical approaches like principal component analyses (PCA) have a long tradition in genetic data analysis (Cavalli-Sforza, 1966). While classical PCA do not use predictors (indirect ordination), methods based on constrained ordinations can find covariation of multiple loci with multivariate environmental patterns. *Redundancy analysis* (RDA) makes use of constrained ordinations by multivariate linear regression of genetic and environmental data (Forester et al., 2016).

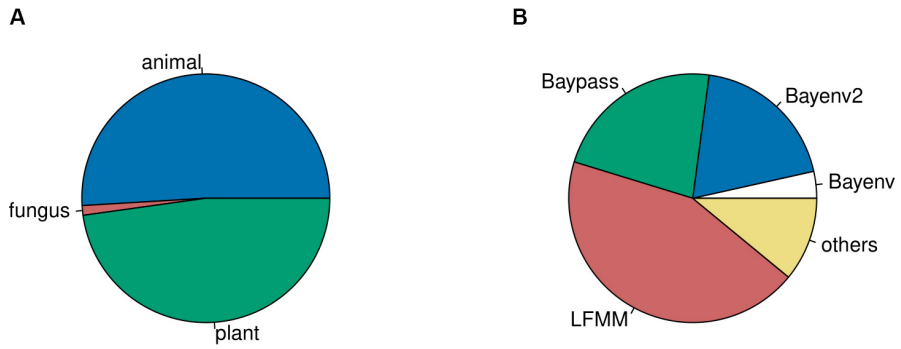


FIGURE 2 | The fraction of studies based on animals and plants is almost equal **(A)**. Frequency of tools used in GEA studies **(B)**. Tools which were used in less than four studies are summarized under "others": RDA, AutoLM, GEMMA, GLMM, Lositan, MLM, partial Mantel test, Moran spectral randomization, Samβada, Selestim, and Tassel.

individuals or pooled populations. Within our assessment, the majority of studies applied RRS strategies to generate the genomic input data and especially RAD-Seq was most commonly used in recent studies. Whilst animal and plant species are almost equally addressed in our assessment, fungal species are heavily underrepresented while other domains of the tree of life are missing (**Figure 2A**). The application of Bayesian modeling and latent factor mixed modeling is more or less balanced among the assessed GEA studies, nevertheless, LFMM is the most commonly applied tool (**Figure 2B**). Most studies based their GEA on a combination of multiple tools; according to our search criteria at least one belonged to the group of mixed effect models, the other(s) may have included additional statistical methods (**Supplementary Table 1**).

Molecular ecology studies mainly applied GEAs to investigate the genomic basis underlying local adaptation, leading to a steady increase of GEA data (**Figure 1**). Numerous statistical frameworks and tools to perform GEAs are available including categorical tests, logistic regressions, matrix correlations, general linear models, and mixed effects models, nowadays accounting for confounding factors such as population structure (reviewed e.g., in Jones et al., 2013; Rellstab et al., 2015; Forester et al., 2018, and see **Box 1**). Keeping track of state-of-the-art methods, empirical studies therefore generally differ in their choice of the applied approaches. As a consequence, and due to variation in required input data types (especially the genomic data, see below), compatibility of results among GEA studies is currently not given. Due to lacking standards for data deposition of GEA results, both in content and format, a meaningful intersection of results in a meta-analysis framework is not possible. Among the 159 publications that we collected for this study, 19 studies qualified our pre-selection criteria: WGS data (Pool-Seq or WGS of single individuals) and investigation of genome-wide SNPs (in contrast to pre-specified candidate loci or candidate regions). Of these 19 studies only six (32%) contained the minimum information necessary to perform a meta-analysis, namely the gene IDs of candidate genes that were found to be significantly associated to an environmental factor, and the precise environmental factor itself. All other studies only report summary statistics of putative candidates, but do not give candidate specific information. Of these six studies, three studies investigated at least one temperature-related

parameter [Fahrenkrog et al., 2017 (*Populus deltoides*); Henriques et al., 2018 (*Apis mellifera*); and Tabas-Madrid et al., 2018 (*Arabidopsis thaliana*)]. For proof of principle, we conducted a mini-meta analysis on these three studies (**Supplementary Table 1**). The number of temperature-associated candidate genes ranged between 108 and 466. These candidate genes did not overlap among the three species (based on gene annotation), however, we did find an overlap in functions between the candidate genes. There were 26 functions shared between all three species, as for example which DNA binding and oxidoreductase activity (**Supplementary Table 1** for complete information on shared functions in three- and two-species comparisons). Interestingly, "cell wall modification" was amongst the functions shared between the two plant species as common temperature adaptation element. Our mini-analysis indicates that different genes are selected in response to different temperature regimes, i.e., no congruence on gene level between species, these genes however share similar functions. Please note that this mini-analysis by all means does not give any conclusive information, but it does show that even between plants and insects common temperature adaptation patterns on a functional level might be expected.

CALL FOR STANDARDIZED DATA CONVENTIONS

Building on existing guidelines (especially Rellstab et al., 2015) and considering the stepwise pipeline of GEA studies including different options for downstream analyses, we suggest the following standards for the deposition of GEA results to allow for a better compatibility across studies (see **Table 1**). Hereafter, we summarize the different data types needed for GEA studies, how they are currently obtained and give recommendations on how to deposit the data in a standardized fashion. Standardized data reposition is key to extrapolate results of single GEA studies to more general patterns and conclusions.

Environmental Data for GEA

Data Acquisition

Whilst *in situ* measurements of abiotic data are not available in most cases, public databases provide access to topo-climatic factors globally interpolated over large areas (e.g., WorldClim2, Fick and Hijmans, 2017; CHELSA, Karger et al., 2017), to global hydro-environmental data for watersheds and rivers at high spatial resolution (HydroATLAS, Linke et al., 2019), or to high resolution data on regional scale. Due to high levels of covariance among abiotic, and particularly climate factors, it is common to use the linearly uncorrelated principal components of the complete set of environmental variables. The handling and preparation of environmental input data, also for multivariate approaches, for GEA are detailed in Rellstab et al. (2015).

Data Deposition Recommendation

Irrespective of the source and choice of environmental factors used in a GEA, it is of major importance to deposit the matrix of sample ID, sample location and environmental variables (also including eigenvalues of principal components if applicable; for more details on ecological metadata handling see also Fegraus et al., 2005; Madin et al., 2007; Whitlock, 2011; Michener, 2015). We recommend including a comprehensive variable table in a processable format (not pdf) in the supplement, or as upload to public data archives, such as Dryad[1] or gfbio[2] (**Table 1**).

[1]https://datadryad.org/stash
[2]https://www.gfbio.org/

TABLE 1 | Suggested standards for deposition of GEA input and results data.

Step in GEA pipeline	Data type	Data format	Deposition platform
Tool implementation	Matrix of environmental input (sample ID, sample location, environmental variables, eigenvalues if applicable)	Processable text-table format (not pdf)	Dryad; gfbio; supplement; *intended integration in NCBI BioProject or ENA Project (EMBL-EBI)*
Tool implementation	Genomic raw or trimmed reads	Fastq format	Integration in NCBI BioProject or ENA Project (EMBL-EBI)
Tool implementation	Genomic reads mapped to reference genome	Bam format	Integration in NCBI BioProject or ENA Project (EMBL-EBI)
Tool implementation	Genomic variant table	Vcf format	NCBI SNP database or EMBL-EBI EGA; *intended integration in NCBI BioProject or ENA Project (EMBL-EBI)*
Tool implementation	Final genomic input table to specific EAA tool	e.g., lfmm format	Dryad; gfbio; supplement; *intended integration in NCBI BioProject or ENA Project (EMBL-EBI)*
Structural annotation	Gene ID lists of annotated loci resulting from EAA	Processable text-table format (not pdf)	Dryad; gfbio; supplement; *intended integration in NCBI BioProject or ENA Project (EMBL-EBI)*
Functional annotation	Full set of protein sequences corresponding to the structural annotation of the reference genome	Fasta format	Integration in NCBI BioProject or ENA Project (EMBL-EBI)
Validation	Experimentally or phylogenetically validated gene with association to an environmental factor		*Intended integration in GO database referring to novel GO domain "environmental association"*

New deposition platforms suggested here in italics.

Genomic Data for GEA

Data Acquisition

The choice of the genomic data type largely depends on a cost-benefit ratio among sample and genome size. To reveal genomic signatures of selection in association to environmental variability, it is highly important that sampled populations sufficiently cover the geographic area of interest. If, for example, the aim is to investigate genetic variability along a continental climatic gradient, not only the two extremes, but multiple populations along the gradient should be sampled, optimally even in replicates. Moreover, the number of individuals sampled per population needs to be sufficiently high to obtain reliable allele frequency estimates. Sequencing budgets have to be distributed in a way to satisfy both requirements: the number of populations across space and the number of individuals per population. Especially for organisms with large genome sizes, the sampling design can be a challenge and different sequencing strategies can be considered.

Since GEAs are of exploratory nature, we generally do not have pre-knowledge on the targets of selection in respect to the environmental variables of interest. We are thus interested in covering as much of an organism's genome as possible. Whole genome individual resequencing is the recommended data type of choice, since it comprises individual information along the whole genome. Sequencing a pool of individuals is a cost-effective alternative for all organisms that have intermediate to large population sizes and small to intermediate genome sizes (Schlötterer et al., 2014) while still covering the whole genome (see details above). If, however, RRS is inevitable due to a limited number of individuals per population or very large genome sizes, we recommend targeted exome capture sequencing (e.g., Yeaman et al., 2016), or RNA-Seq (e.g., Roschanski et al., 2016; but see Knight, 2004) over RAD-Seq, SNP-arrays or the sequencing of previously known candidate genes. The rationale is that in GEA studies the targeted entities are candidate genes, which are the basis to infer the biological relevance and function of the selection targets. If gene sets are highly incomplete in the fragmented genomic data, as is the case with RAD-Seq, etc., there is a high chance that the actual target site of selection is not represented in the data. Incomplete results based on insufficient genomic resolution may thus produce misleading patterns (Lowry et al., 2017). However, even in WGS data, many candidate SNPs are just linked sites of variation without functional significance, and the "true" target of selection may also be missed (e.g., due to variance in the coverage distribution). Being able to investigate up- and downstream regions along the genome, and/or using a several kb spanning window-approach increases the reliability of identified putative candidates. Populations samples should cover the distribution range, sufficient number of individuals per population, and use WGS data whenever possible.

Data Deposition Recommendation

Rules and guidelines for data deposition of genomic sequences (including transcriptome sequences) are generally well coordinated for all relevant journals and publishers: genomic sequences as raw or trimmed read data and, if applicable, mapped data to genomic reference sequences have to be uploaded to either of the two major public platforms, NCBI (National Center for Biotechnology Information, United States), or the European Nucleotide Archive (ENA run by EMBL-EBI, United Kingdom). However, GEAs and downstream analyses depend on genomic variants between populations and criteria for variant calling are more or less arbitrary depending on custom settings. For a reliable replicability of results, we therefore recommend the deposition of primary variant tables (e.g., vcf format for WGS data, sync format for Pool-Seq data) as well as the final genomic input table (e.g., lfmm format) for the GEA implementation (**Table 1**). As for now, variant tables (vcf format) can be submitted to the NCBI SNP database or to EGA (European Genome-phenome Archive) by EMBL-EBI, and all other data types can be uploaded to Dryad or gfbio. Ideally, all data, from sequences to genomic and environmental metadata of a single study, would be deposited in a single database. For example, NCBI BioProjects, and/or ENA Projects could be developed to be more flexible, i.e., accepting various types of metadata.

Functional Inferences of GEA

Data Acquisition

Genotype-environment associations implementations will ultimately deliver information on loci significantly associated to variation in environmental factors across space (this holds for the actual target site of selection as well as closely linked sites). Structural annotation of individual loci delivers information about whether these loci are part of or contribute to the coding part of the genome, and this information is embedded in the annotation-file (gff) file. To obtain information on the function of genes, to investigate a higher level of organization, and to allow for a deeper biological interpretation of the GEA results, functional annotation is the next step. Gene ontology (GO) databases provide controlled vocabularies for the classification of gene products, and entries are manually curated (Gene Ontology Consortium, 2004). The database is structured in a "loosely hierarchical" manner, with three top ("parent") domains: molecular function (MF), biological process (BP), and cellular component (CC). The obtained functional information can be used to perform a gene set enrichment analysis to obtain information on significantly overrepresented functions in the candidate gene list versus all genes in the genome. Similarly, information on covered pathways, the position of pathways, etc. can be obtained from the reactome database (GO), as well as reactome information[3] can be obtained via a search of the protein sequences versus the interproscan database[4].

Data Deposition Recommendation

For a reliable replicability of results and meta-analyses of GEA studies across organisms, the deposition of the gene ID list of selected candidate loci in table format and the full set of protein sequences in fasta format (both corresponding and referencing to the respective genome annotation version used),

[3]https://reactome.org/

[4]http://www.ebi.ac.uk/interpro/

is crucial (**Table 1**). An additional column with according GO-IDs for each candidate is desirable, but can also be acquired based on the gene ID.

Experimental Validation of GEA

Re-sequencing

Data acquisition

A first step to validate significant polymorphisms is the verification of allele frequencies by, e.g., Sanger resequencing of candidate SNPs in individuals of the target populations, or experimental populations. Such resequencing approaches can help to decrease the false positive rates, even if conducted for a subset of candidates only. Results of association studies are inherently correlative and consequently, validation of candidate genes requires experimentation (Pardo-Diaz et al., 2015).

Data deposition recommendation

Allele frequency information for target populations should be added as supplementary table to the study including individual specific genotypes.

Molecular Profiling

Data Acquisition

Molecular validation approaches involve gene expression profiling and direct assays to test the molecular and/or ecological function (described in more detail in Pardo-Diaz et al., 2015). The detection of differential gene expression is especially important when significant loci are located up- or downstream of protein-coding regions, indicative of regulatory regions. Assays of MF mostly rely on transgenics, knockouts, knockdowns (e.g., with RNA interference, RNAi) and gene replacements. All of these assays are developed and optimized for model organisms and application to non-model species is more difficult. Nevertheless, constant development of functional tools can open doors for functional characterization also in non-model species, as e.g., with RNAi or CRISPR/Cas systems (Russell et al., 2017).

Data Deposition Recommendation

Optimally, this data would be of sufficient quality to be deposited on a public curated database such as Uniprot (uniprot.org), to make this information publicly available.

Fitness Estimation

Data Acquisition

Finally, assays of ecological function aim at testing the fitness consequences of allelic substitutions at causal genes (Barrett and Hoekstra, 2011). Such selection experiments (e.g., experimental evolution or evolve and resequencing studies), however, suffer from being highly artificial due to laboratory conditions and being mostly restricted to few established model organisms (Pfenninger and Foucault, 2019). Performing analogous experiments in natural systems (e.g., transplant experiments, more details described in Pardo-Diaz et al., 2015) and adapting them to a broad range of taxa constitutes a major challenge for molecular ecology research. However, given the resources required in terms of work force, money and experimental facilities to causally link genotype with fitness at a single locus (e.g., Rosenblum et al., 2010), it appears unrealistic that more than a tiny fraction of the thousands of already and increasingly identified candidate loci will ever be validated as described above. Thus, meta-analyses of large numbers of taxonomically diverse organisms with similar selection regimes could be an effective means to cross-validate candidate loci playing a role in ecological adaptation.

Data Deposition Recommendation

This list of candidate genes should follow the above-mentioned criteria (also see **Table 1**), and include gene ID and associated environmental factor.

CALL FOR A NOVEL GO DOMAIN "ENVIRONMENTAL ASSOCIATION"

Experimental validation of candidate genes significantly associated to variation of environmental factors, or candidates obtained from meta-analyses will finally deliver the link from genotype to the environment. The continuous increase in genomic and transcriptomic resources will fuel the accumulation of GEA studies for more and more organisms. The development of novel methodologies for experimental validation of candidate genes will advance the accumulation of knowledge on genes contributing to adaptive responses to environmental variation. With this perspective as a guiding principle, we propose the initiation of a novel GO domain to be called "environmental association (EA)" for the standardized categorization of genes causally associated to environmental variation (**Table 1**). This GO domain could become the fourth parent domain alongside MF, BP, and CC, adding ecologically relevant information to gene products. Future GO enrichment analyses on the basis of this novel domain will generate a more structured insight in the molecular basis of environmental adaptation, potentially revealing so far hidden relations.

GEA IN LIGHT OF CLIMATE CHANGE

Global climate change and the associated environmental changes will heavily impact ecosystems in their current state. While knowing about the inescapability of these changes, we are largely lacking an understanding of the mechanisms of environmental adaptation and the adaptive potential of organisms (Fitzpatrick and Edelsparre, 2018). Changing climatic conditions impose new selective forces to many ecosystems, yet, with the exception of some few model species, the affected key traits of the majority of organisms remain unknown (Alberto et al., 2013; Gienapp et al., 2014). In order to understand how biodiversity will respond to climate change we will thus need methodological approaches that keep up with the pace of climate change (Waldvogel et al., 2020). To this end, GEA clearly bring the advantage that via the correlation of genomic variation and environmental variation, pre-knowledge of specific traits is not required, but rather target traits can be inferred from the resulting candidate loci (see above). The long list of GEA studies included in this review highlights the timeliness and broad applicability of GEA, especially in non-model species (**Supplementary Table 1**).

So far, we learned from these studies that species indeed adapt to different climatic conditions and that species have multiple options to adapt to the same environmental variables (e.g., temperature). However, we are still lacking meta-analyses which would enable to extend the results of single-species studies to a more comprehensive picture allowing for global conclusions. It would also be highly desirable to include genetic variability and dispersal potential to the analytical framework to actually refine predictions by including species' potential of rapid adaptation.

GEA are based on the idea of space-for-time and thus approximate the extent of climate change given the range of climate variation across the investigated geographic space (Rellstab et al., 2015). This indirect approximation can only inform about the extent of current climate variation observed in the investigated space and will be blind for new dimensions as are expected for many areas. The only solution to overcome this limitation is measuring genomic change over the actual course of climate change, i.e., tracking populations through time. Such time-for-time approaches are currently being implemented and already allowed the tracking of adaptive trajectories and quantification of the selection regime in natural populations (Pfenninger and Foucault, 2020). The GEA space for time approach can also be embedded in the time-for-time framework by building up time-series of repeated GEA. The repetition of GEA for a given system across a climate change-relevant time horizon can allow to relate changes across space (within single GEA) with changes across time (across repeated GEA) and thus identify the effects of changing climate conditions. Such GEA time-series that are ideally based on WGS of multiple populations across wide distribution ranges will deliver invaluable molecular ecological resources to build accurate prediction models of how species can respond to climate change (Waldvogel et al., 2020). Granted the here proposed standardization of data, GEA in combination with time-series data is a powerful and most promising tool to take on the challenge of understanding the effects of climate change bevor its consequences have brought too much damage.

CONCLUSION

Systematic deposition of GEA data in a standardized and structured format will set the ground for meta analyses to assess the associations of genotypes and the environment across species, phyla or even the tree of life. Our mini analysis already shows that interesting patterns are to be expected from this data. We here suggest standards for deposition of GEA results and call for a novel GO domain to be included in the gene ontology database. By the implementation of these standards, individual GEA studies will contribute to the growth of a powerful data resource which generates insight in the adaptability of species to environmental variables, especially climate variables. Building these data in a standardized way can furthermore help us to widen the perspective from single to multiple species or even phyla. This can be away forward in investigating biodiversity responses to changing climate conditions and can be the key to improved prediction models.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

A-MW, BF, and MP conceived the article. A-MW drafted the manuscript. All authors contributed to data survey and writing of the manuscript.

ACKNOWLEDGMENTS

We thank Nikolaus Wohlenberg, Quentin Foucault, and Andreas Wieser for their help in the literature summary. Three reviewers provided constructive recommendations for improving our manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.00242/full#supplementary-material

REFERENCES

Alberto, F. J., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., et al. (2013). Potential for evolutionary responses to climate change - evidence from tree populations. *Glob. Chang. Biol.* 19, 1645–1661. doi: 10.1111/gcb.12181

Barrett, R. D. H., and Hoekstra, H. E. (2011). Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* 12, 767–780. doi: 10.1038/nrg3015

Berg, J. J., and Coop, G. (2014). A population genetic signal of polygenic adaptation. *PLoS Genet.* 10:e1004412. doi: 10.1371/journal.pgen.1004412

Bradshaw, W. E., and Holzapfel, C. M. (2001). Genetic shift in photoperiodic response correlated with global warming. *Proc. Natl. Acad. Sci. U.S.A.* 98, 14509–14511. doi: 10.1073/pnas.241391498

Catchen, J. M., Hohenlohe, P. A., Bernatchez, L., Funk, W. C., Andrews, K. R., and Allendorf, F. (2017). Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Mol. Ecol. Resour.* 17, 362–365. doi: 10.1111/1755-0998.12669

Cavalli-Sforza, L. L. (1966). Population structure and human variation. *Proc. R. Soc. B Biol. Sci.* 164, 362–379. doi: 10.2307/2801262

Coop, G., Witonsky, D., Di Rienzo, A., and Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics* 185, 1411–1423. doi: 10.1534/genetics.110.114819

Culina, A., Baglioni, M., Crowther, T. W., Visser, M. E., Woutersen-Windhouwer, S., and Manghi, P. (2018). Navigating the unfolding open data landscape in ecology and evolution. *Nat. Ecol. Evol.* 2, 420–426. doi: 10.1038/s41559-017-0458-2

De La Torre, A. R., Wilhite, B., and Neale, D. B. (2019). Environmental genome-wide association reveals climate adaptation is shaped by subtle to moderate allele frequency shifts in loblolly pine. *Genome Biol. Evol.* 11, 2976–2989. doi: 10.1093/gbe/evz220

Fahrenkrog, A. M., Neves, L. G., Resende, M. F. R., Dervinis, C., Davenport, R., Barbazuk, W. B., et al. (2017). Population genomics of the eastern cottonwood (Populus deltoides). *Ecol. Evol.* 7, 9426–9440. doi: 10.1002/ece3.3466

Feder, A. F., Petrov, D. A., and Bergland, A. O. (2012). LDx: estimation of linkage disequilibrium from high-throughput pooled resequencing data. *PLoS One* 7:e0048588. doi: 10.1371/journal.pone.0048588

Fegraus, E. H., Andelman, S., Jones, M. B., and Schildhauer, M. (2005). Emerging Technologies. *Bull. Ecol. Soc. Am.* 86, 158–168.

Fick, S. E., and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. doi: 10.1002/joc.5086

Fitzpatrick, M. J., and Edelsparre, A. H. (2018). The genomics of climate change. *Science* 359, 29–30. doi: 10.1126/science.aar3920

Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., and Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Mol. Ecol.* 25, 104–120. doi: 10.1111/mec.13476

Forester, B. R., Lasky, J. R., Wagner, H. H., and Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Mol. Ecol.* 27, 2215–2233. doi: 10.1111/mec.14584

Frichot, E., Schoville, S. D., Bouchard, G., and François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol. Biol. Evol.* 30, 1687–1699. doi: 10.1093/molbev/mst063

Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics* 201, 1555–1579. doi: 10.1534/genetics.115.181453/-/DC1

Gene Ontology Consortium (2004). The gene ontology (GO) database and informatics resource. *Nucleic Acids Res.* 32, 258–261. doi: 10.1093/nar/gkh036

Gienapp, P., Reed, T. E., and Visser, M. E. (2014). Why climate change will invariably alter selection pressures on phenology. *Proc. R. Soc. B Biol. Sci.* 281:20141611. doi: 10.1098/rspb.2014.1611

Günther, T., and Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics* 195, 205–220. doi: 10.1534/genetics.113.152462

Hancock, A. M., Alkorta-Aranburu, G., Witonsky, D. B., and Di Rienzo, A. (2010). Adaptations to new environments in humans: the role of subtle allele frequency shifts. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 2459–2468. doi: 10.1098/rstb.2010.0032

Henriques, D., Wallberg, A., Chávez-Galarza, J., Johnston, J. S., Webster, M. T., and Pinto, M. A. (2018). Whole genome SNP-associated signatures of local adaptation in honeybees of the Iberian Peninsula. *Sci. Rep.* 8, 1–14. doi: 10.1038/s41598-018-29469-5

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., et al. (2016). Finding the genomic basis of local adaptation: pitfalls. practical solutions, and future directions. *Am. Nat.* 188, 379–397. doi: 10.1086/688018

Jones, M. R., Forester, B. R., Teufel, A. I., Adams, R. V., Anstett, D. N., Goodrich, B. A., et al. (2013). Integrating landscape genomics and spatially explicit approaches to detect loci under selection in clinal populations. *Evolution* 67, 3455–3468. doi: 10.1111/evo.12237

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., et al. (2017). Climatologies at high resolution for the earth's land surface areas. *Sci. Data* 4, 1–20. doi: 10.1038/sdata.2017.122

Kawecki, T. J., and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecol. Lett.* 7, 1225–1241. doi: 10.1111/j.1461-0248.2004.00684.x

Knight, J. C. (2004). Allele-specific gene expression uncovered. *Trends Genet.* 20, 113–116. doi: 10.1016/j.tig.2004.01.001

Linke, S., Lehner, B., Dallaire, C. O., Ariwi, J., Grill, G., Anand, M., et al. (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Sci Data.* 6:283. doi: 10.1038/s41597-019-0300-6

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., et al. (2017). Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Mol. Ecol. Resour.* 17, 142–152. doi: 10.1111/1755-0998.12635

Lynch, M., and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer, 980.

Mackay, T. F. C. (2014). Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat. Rev. Genet.* 15, 22–33. doi: 10.1038/nrg3627

Madin, J., Bowers, S., Schildhauer, M., Krivov, S., Pennington, D., and Villa, F. (2007). An ontology for describing and synthesizing ecological observation data. *Ecol. Inform.* 2, 279–296. doi: 10.1016/j.ecoinf.2007.05.004

Michener, W. K. (2015). Ecological data sharing. *Ecol. Inform.* 29, 33–44. doi: 10.1016/j.ecoinf.2015.06.010

Nakagawa, S., and Poulin, R. (2012). Meta-analytic insights into evolutionary ecology: an introduction and synthesis. *Evol. Ecol.* 26, 1085–1099. doi: 10.1007/s10682-012-9593-z

Pardo-Diaz, C., Salazar, C., and Jiggins, C. D. (2015). Towards the identification of the loci of adaptive evolution. *Methods Ecol. Evol.* 6, 445–464. doi: 10.1111/2041-210X.12324

Parker, T. H., Forstmeier, W., Koricheva, J., Fidler, F., Hadfield, J. D., Chee, Y. E., et al. (2016). Transparency in ecology and evolution: real problems. *Real Solutions. Trends Ecol. Evol.* 31, 711–719. doi: 10.1016/j.tree.2016.07.002

Pfenninger, M., and Foucault, Q. (2019). Genomic processes underlying rapid adaptation of a natural *Chironomus riparius* population to unintendedly applied experimental selection pressures. *Mol. Ecol.* 29, 536–548. doi: 10.1111/mec.15347

Pfenninger, M., and Foucault, Q. (2020). Quantifying the selection regime in a natural *Chironomus riparius* population. *bioRxiv* [Preprint]. doi: 10.1101/2020.06.16.154054

Poisot, T., Bruneau, A., Gonzalez, A., Gravel, D., and Peres-Neto, P. (2019). Ecological data should not be so hard to find and reuse. *Trends Ecol. Evol.* 34, 494–496. doi: 10.1016/j.tree.2019.04.005

Pritchard, J. K., and Di Rienzo, A. (2010). Adaptation - not by sweeps alone. *Nat. Rev. Genet.* 11, 665–667. doi: 10.1038/nrg2880

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., and Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Mol. Ecol.* 24, 4348–4370. doi: 10.1111/mec.13322

Roschanski, A. M., Csilléry, K., Liepelt, S., Oddou-Muratorio, S., Ziegenhagen, B., Huard, F., et al. (2016). Evidence of divergent selection for drought and cold tolerance at landscape and local scales in Abies alba Mill. in the French Mediterranean Alps. *Mol. Ecol.* 25, 776–794. doi: 10.1111/mec.13516

Rosenblum, E. B., Römpler, H., Schöneberg, T., and Hoekstra, H. E. (2010). Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2113–2117. doi: 10.1073/pnas.0911042107

Russell, J. J., Theriot, J. A., Sood, P., Marshall, W. F., Landweber, L. F., Fritz-Laylin, L., et al. (2017). Non-model model organisms. *BMC Biol.* 15:55. doi: 10.1186/s12915-017-0391-5

Schlötterer, C., Tobler, R., Kofler, R., and Nolte, V. (2014). Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. *Nat. Rev. Genet.* 15, 749–763. doi: 10.1038/nrg3803

Tabas-Madrid, D., Méndez-Vigo, B., Arteaga, N., Marcer, A., Pascual-Montano, A., Weigel, D., et al. (2018). Genome-wide signatures of flowering adaptation to climate temperature: regional analyses in a highly diverse native range of *Arabidopsis thaliana*. *Plant Cell Environ.* 41, 1806–1820. doi: 10.1111/pce.13189

Waldvogel, A.-M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R., et al. (2020). Evolutionary genomics can improve prediction of species' responses to climate change. *Evol. Lett.* 4, 4–18. doi: 10.1002/evl3.154

Wang, K., Dickson, S. P., Stolle, C. A., Krantz, I. D., Goldstein, D. B., and Hakonarson, H. (2010). Interpretation of association signals and identification of causal variants from genome-wide association studies. *Am. J. Hum. Genet.* 86, 730–742. doi: 10.1016/j.ajhg.2010.04.003

Whitlock, M. C. (2011). Data archiving in ecology and evolution: best practices. *Trends Ecol. Evol.* 26, 61–65. doi: 10.1016/j.tree.2010.11.006

Whitlock, M. C., and Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F_{ST}. *Am. Nat.* 186, S24–S36. doi: 10.1086/682949

Yeaman, S., Hodgins, K. A., Lotterhos, K. E., Suren, H., Nadeau, S., Degner, J. C., et al. (2016). Convergent local adaptation to climate in distantly related conifers. *Science* 353, 23–26.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

MINI REVIEW
published: 18 August 2020
doi: 10.3389/fevo.2020.00271

Thermal Plasticity in Insects' Response to Climate Change and to Multifactorial Environments

Yara Katia Rodrigues[1]* *and Patrícia Beldade*[2,3,4]*

[1] Porgrama de Pós-Graduação Ciência para o Desenvolvimento (PGCD), Instituto Gulbenkian de Ciência, Oeiras, Portugal, [2] Instituto Gulbenkian de Ciência, Oeiras, Portugal, [3] CNRS-UMR 5174, EDB, Université Paul Sabatier, Toulouse, France, [4] Faculty of Sciences, CE3C: Centre for Ecology, Evolution, and Environmental Changes, University of Lisbon, Lisbon, Portugal

Phenotypic plasticity, the property by which living organisms express different phenotypes depending on environmental conditions, can impact their response to environmental perturbation, including that resulting from climate change. When exposed to altered environmental conditions, phenotypic plasticity might help or might hinder both immediate survival and future adaptation. Because climate change will cause more than a global rise in mean temperatures, it is valuable to consider the combined effects of temperature and other environmental variables on trait expression (thermal plasticity), as well as trait evolution (thermal adaptation). In this review, we focus primarily on thermal developmental plasticity in insects. We discuss the genomics of thermal plasticity and its relationship to thermal adaptation and thermal tolerance, and to climate change and multifactorial environments.

Keywords: developmental plasticity, thermal adaptation, multifactorial environments, environment-by-environment interactions, climate change

OPEN ACCESS

Edited by:
Mauro Santos, Autonomous University of Barcelona, Spain

Reviewed by:
Vincenzo Trotta, Università degli Studi della Basilicata, Italy
Luis E. Castañeda, University of Chile, Chile

***Correspondence:**
Yara Katia Rodrigues
yrodrigues@igc.gulbenkian.pt
Patrícia Beldade
pbeldade@fc.ul.pt

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Ecology and Evolution

Received: *01 June 2020*
Accepted: *30 July 2020*
Published: *18 August 2020*

Citation:
Rodrigues YK and Beldade P (2020) Thermal Plasticity in Insects' Response to Climate Change and to Multifactorial Environments. Front. Ecol. Evol. 8:271. doi: 10.3389/fevo.2020.00271

The ability of natural populations to react to environmental change will depend on the level and type of perturbation organisms experience, and also on their intrinsic capability to respond to it (Parmesan, 2006; Johnston et al., 2019). Phenotypic plasticity, the property by which living organisms express different phenotypes depending on environmental conditions, can impact their response to environmental perturbation, including that resulting from global climate change (Reed et al., 2011; Chevin et al., 2013; Merilä and Hendry, 2014; Sgrò et al., 2016; Bonamour et al., 2019). Considering thermal plasticity, in addition to thermal tolerance and thermal adaptation, will be crucial to assessing how organisms might cope with climate change. And because climate change is not only about increasing mean ambient temperature, it is also clear that it is important to consider effects of multifactorial environments, combining temperature with other environmental variables, both on trait expression and on trait evolution (Kaunisto et al., 2016; Westneat et al., 2019).

Here, we focus on effects of temperature and its combination with other environmental factors on phenotypic plasticity in terrestrial/flying insects, a taxon of ectothermal animals that includes many compelling examples of thermal plasticity. This is a large and ecologically central group of organisms whose geographic ranges, behaviors, and life histories are very much affected by ambient temperature (Colinet et al., 2015). It is also a group with recent worrying trends: steep global population declines (Sánchez-Bayo and Wyckhuys, 2019; Didham et al., 2020; Wagner, 2020), as well as expansions of agricultural pests and disease vectors (Song et al., 2017; Ryan et al., 2019). We direct our attention primarily toward recent examples, and to studies focused on the genomics of thermal plasticity of potential relevance to responses to climate change.

PHENOTYPIC PLASTICITY AND ENVIRONMENTAL PERTURBATION

Phenotypic diversity, within and across species, is shaped by interactions between organisms and their environment, which occur at different levels and different time scales. Environmental conditions determine cross-generation changes in phenotype frequencies in populations (notably through natural selection), and affect intra-generation phenotype expression of individuals (via phenotypic plasticity). In this section, we will focus on examples of phenotypic plasticity and how it can evolve and impact adaptive evolution, including in the context of climate change.

Ecological Significance and Evolution of Plasticity

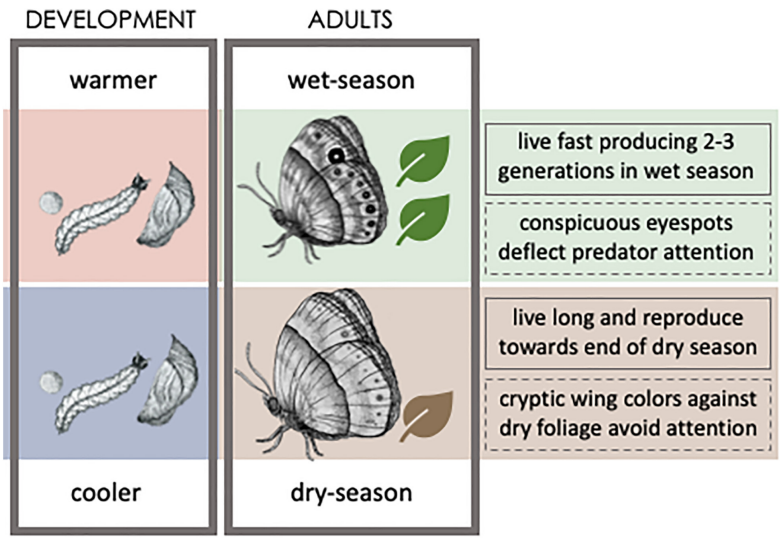
The effect of external environmental conditions on phenotype expression can happen at distinct time-scales: (1) change in progeny phenotype that depends on parental environment (trans-generational plasticity; Woestmann and Saastamoinen, 2016; Donelson et al., 2018), (2) change in adult phenotype in response to adult environment, often reversible changes in labile traits, such as behavior, and including what is called acclimation (Stillman, 2003; Sgrò et al., 2016), and (3) change in phenotype that depends on the conditions experienced during development, often leading to irreversible adult phenotypes (the main focus of this review). Indeed, the environmental conditions experienced during development can alter developmental rates and/or trajectories and result in the production of different adult phenotypes from the same genotype, in a phenomenon called developmental plasticity (reviewed in Beldade et al., 2011; Nettle and Bateson, 2015). The study of developmental plasticity, which integrates ecology, evolutionary biology, and developmental biology (eco-evo-devo, Gilbert et al., 2015), is key to understanding how organisms interact with their changing environments.

Plasticity can match organismal phenotypes to their ecological conditions and, as such, be favored by natural selection (Nettle and Bateson, 2015). Plasticity is thought to benefit populations that face distinct challenges imposed by environmental heterogeneity (e.g., Chevin et al., 2010), such as that resulting from alternating seasons (Buckley et al., 2017). Seasonal polyphenism, where alternative seasonal conditions lead to the production of distinct seasonal phenotypes, is common in insects (Nijhout, 2003; Moczek, 2010; Simpson et al., 2011; Yang and Pospisilik, 2019). Their relatively short life cycles allow for multiple generations within the year and, consequently, exposure to conditions that can differ substantially between generations. Seasonally variable environmental factors, often temperature, can induce changes in sets of integrated traits associated to distinct strategies for survival and/or reproduction, suited to the respective seasonal conditions. For example, in the butterfly *Bicyclus anynana*, the temperature during development anticipates upcoming seasonal conditions in vegetation cover, and induces changes in various adult traits associated with distinct seasonal strategies for predator avoidance and pace-of-life (**Box 1**).

Phenotypic Plasticity and Climate Change

Plasticity can be targeted by selection and evolve, and can, in turn, impact adaptive evolution (reviewed in Lafuente and Beldade, 2019). It has been argued that developmental plasticity can help (or hinder; e.g., Cenzer, 2017; Oostra et al., 2018) not only the immediate survival, but also future adaptation of populations facing environmental perturbation (Reed et al., 2011; Bonamour et al., 2019) and colonizing novel environments (Ghalambor et al., 2007; Wang and Althoff, 2019; Bilandžija et al., 2020). In addition, it has been proposed that plasticity can promote phenotypic and taxonomic diversification (Moczek, 2010; Pfennig et al., 2010; Schneider and Meyer, 2017). Whether plasticity can have an impact specifically in responses to climate change has also raised significant attention (Sgrò et al., 2016; Bonamour et al., 2019). Upon change in local environmental conditions, particularly of temperature, organisms that are thermally plastic might display phenotypic change that allows them to rapidly adjust to the new conditions, without genetic change. This type of phenotypic adjustment has been reported for some insect populations, along with other types of population responses to climate change (**Figure 1**): (1) phenotypic change resulting from genetic change, as populations adapt to new local conditions, (2) shifts in distribution range, as populations track favorable conditions, and (3) population declines that might lead to extinctions. These responses are not mutually exclusive scenarios; they can be combined in different manners (Valladares et al., 2014) and can also be hard to disentangle, as illustrated in the examples below. Shifts in species distributions can result from populations actually migrating to new locations, but can also result from population extinctions on one or multiple distribution edges. Occupation of new locations is generally followed by adaptation to the local conditions (e.g., butterflies that move up along an altitudinal gradient adapted to a host plant in the new habitat; Parmesan et al., 2015). Adaptation to climate change can involve changes in plasticity (e.g., Kingsolver and Buckley, 2018), and plastic responses can facilitate adaptation involving genetic change (e.g., Kelly, 2019) or anticipate extinction (e.g., Manfredini et al., 2019).

Phenotypic plasticity can impact species distribution and vulnerability (Foden et al., 2019), and might also impact (positively or negatively) population persistence and ability to adapt to challenges arising from climate change (Leonard and Lancaster, 2020). If plasticity leads to changes in phenotype expression in a direction that maintains/improves fitness in the new conditions, it can, indeed, allow organisms to keep pace with environmental change, preventing immediate population extinction (Merilä and Hendry, 2014) and effectively "buy time" for adaptation to occur (Chevin et al., 2010; Snell-Rood et al., 2018). While the positive impact of plasticity in a response to climate change might go beyond buying time (Levis and Pfennig, 2016; Fox et al., 2019), it is also apparent that plasticity



BOX 1 | Seasonal polyphenism in *B. anynana*.

Development under cooler temperatures leads to the production of dry-season form adults, while development under warmer temperatures leads to wet-season form adults. The seasons differ in vegetation cover and the alternative seasonal forms differ in their strategies for reproduction (solid line box on right side) and for escaping predators (dashed line box). *B. anynana* drawings by Joana Carvalho (joana_gcc).

A case study of thermal plasticity: *Bicyclus anynana* butterflies.

The afro-tropical butterfly *Bicyclus anynana* has become a valuable model of seasonal polyphenism, where an understanding of the ecological significance of alternative seasonal phenotypes can be integrated with knowledge about the developmental basis and evolution of thermal plasticity (Brakefield et al., 2009; Beldade and Peralta, 2017). *B. anynana* seasonal plasticity is believed to be an adaptation to the strongly contrasting wet versus dry seasons of its African savannah habitat. The temperature experienced during development, which anticipates the upcoming season and conditions adults will have to live in, results in adult phenotypes adjusted to each of the seasons' conditions (Figure).

Developmental temperature affects a suite of traits, including wing pigmentation (e.g., Brakefield, 1996; Mateus et al., 2014; Wasik et al., 2014), life histories (e.g., Pijpe et al., 2006; Fischer et al., 2007; Oostra et al., 2011, 2014), secondary sexual traits (e.g., Muller et al., 2019), and various behavioral traits (e.g., Prudic et al., 2011; Bear and Monteiro, 2013; Westerman and Monteiro, 2016; van Bergen and Beldade, 2019). Alternative phenotypes correspond to distinct seasonal strategies for predator avoidance and for reproduction, associated to the distinct seasonal status of the habitat's vegetation cover on which adults perch and larvae feed. Dry-season form adults have dull brown wings, which are cryptic against the background of dry foliage, and have increased body reserves, which sustain longer lifespan and the delay reproduction until the end of the season (Brakefield and Reitsma, 1991; Halali et al., 2020). A raise in ambient temperature anticipates the rainy season, when abundant vegetation provides food for rapid larval growth and adult reproduction. Wet-season form butterflies have a fast pace of life (Brakefield et al., 2009; Halali et al., 2020) and display wings with conspicuous marginal eyespots believed to deflect the attack of predators away from the more fragile body (Lyytinen et al., 2004; Olofsson et al., 2010; Prudic et al., 2015).

Laboratory studies have characterized thermal reaction norms for various traits (e.g., Oostra et al., 2011). Revealingly, lab thermal phenotypes do not include phenotypes as extreme as those seen in nature, where other variables combined with temperature might affect development outcomes (Bauerfeind and Fischer, 2005; Rodrigues et al., 2018; Singh et al., 2019), but do include intermediate phenotypes between the typical dry- and wet-season forms, which are rarely found in nature (Brakefield and Reitsma, 1991; Windig et al., 1994; Muller et al., 2019). Lab studies have also allowed characterization of the physiological and genetic basis of thermal plasticity. Measurement and manipulation of ecdysone levels in pupae developing in different temperatures implicated temperature-induced changes in the dynamics of this hormone in the regulation of *B. anynana* plasticity (Brakefield et al., 1998; Mateus et al., 2014; Oostra et al., 2014; Monteiro et al., 2015; Bear et al., 2017; Bhardwaj et al., 2020). Expression of ecdysone receptor in eyespot organizers has, furthermore, been proposed to account for differences in levels of plasticity between eyespots (Brakefield, 1996; Mateus et al., 2014; Monteiro et al., 2015), and for the evolutionary origin of thermal plasticity in eyespot development (Bhardwaj et al., 2020). Transcriptomic studies targeting individuals from different developmental temperatures have started identifying temperature-regulated genes, potentially responsible different seasonally plastic traits (Macias-Munoz et al., 2016; Oostra et al., 2018). Finally, lab studies have also shed light onto the genetic architecture and constraints on the evolution of thermal reaction norms (Holloway and Brakefield, 1995; Brakefield, 1996; Wijngaarden and Brakefield, 2001; Wijngaarden et al., 2002). On the other hand, field collections allowed for characterization of differences in reaction norms between geographical populations of *B. anynana* (de Jong et al., 2010) and between *Bicyclus* species (van Bergen et al., 2017; Balmer et al., 2018).

A number of studies have explored ideas about *B. anynana's* vulnerability to climate change. These include assessing effects of increased temperature on organismal performance (Klockmann et al., 2017) and the evolutionary potential for populations to cope with warming (Fischer et al., 2010). They also include discussions about the species' thermal developmental plasticity possibly becoming a disadvantage if climate change breaks the correlation between the inducing (temperature) and selective (vegetation) environmental variables, and leads to a mismatch between phenotype and environment (de Jong et al., 2010; Oostra et al., 2018), especially if genetic variation for plasticity is depleted (Oostra et al., 2018). Studies for this and other thermal plasticity models, which allow integration of temperature effects across levels (from gene expression, to physiology and development, to multiple adult traits, to individual fitness, to population performance), will be valuable to achieve a better understanding of the impact of phenotypic plasticity in the response to climate change.

can have a negative impact, both by compromising immediate survival (Ghalambor et al., 2007; Manfredini et al., 2019) or by slowing-down future adaptation (discussed in Beldade et al., 2011). Adaptation will be slower if developmental plasticity somehow shields genetic variation from the action of natural selection, but this can be hard to assess (Fox et al., 2019).

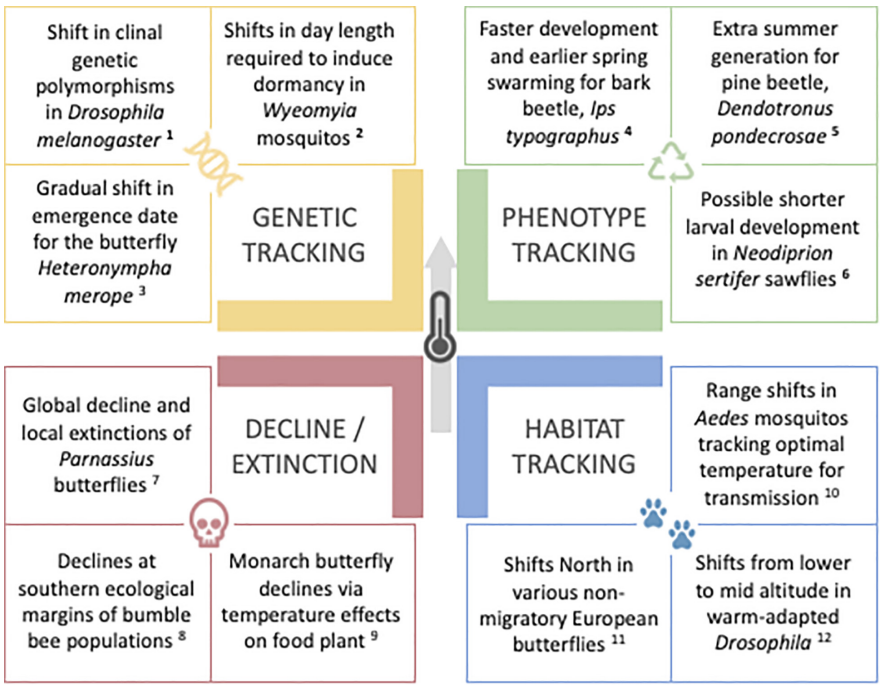


FIGURE 1 | Insect examples of responses to climate change. Populations facing warming temperatures that result from climate change can respond in different manners, including moving to new locations where temperature is closer to their optimum (blue), changing genetic composition as they adapt to new local conditions (yellow), expressing suitable temperature-induced phenotype without genetic change (green), or failing to adjust leading to decline and potential extinction (red). Examples from: [1]Umina et al., 2005; [2]Bradshaw and Holzapfel, 2001; 2001; [3]Kearney et al., 2010; [4]Jönsson et al., 2007; [5]Bentz and Powell, 2014; [6]Kollberg et al., 2013; [7]Condamine and Sperling, 2018; [8]Soroye et al., 2020; [9]Faldyn et al., 2018; [10]Ryan et al., 2019; [11]Parmesan et al., 1999; [12]Parkash et al., 2013.

Survival will be compromised when plasticity leads to expression of phenotypes that, while possibly adaptive in the historical context, are maladaptive in the new conditions (Manfredini et al., 2019). For example, under unusually warm conditions, thermal plasticity in developmental rate in the bark beetle *Ips typographus* may result in a second generation of beetles consisting of immature stages that are poorly adapted to winter conditions (Dworschak et al., 2014). Maladaptive plasticity seems more common in new habitats, presumably because there has been no evolutionary adjustment of the link between the environmental cues and physiological responses (Ghalambor et al., 2007; Chevin and Hoffmann, 2017). If the environmental cues that leads to change in phenotype expression no longer accurately predict future selective environment, plasticity actually can result in a mis-match between phenotype and environmental conditions. Climate change-related failure in the accuracy of cue predictions can lead to an aggravation of maladaptive phenotype-environment mismatches (e.g., Ghalambor et al., 2007; Bonamour et al., 2019).

THERMAL DEVELOPMENTAL PLASTICITY

Temperature is a key factor determining the geographical distribution, abundance and physiology of insects (Colinet et al., 2015). As small ectotherms whose body temperature closely matches ambient temperature, insects are particularly susceptible to thermal perturbation. Climate change-related temperature variation has been implicated in altered phenology, distribution range, and population abundance of many insect species around the world (Parmesan, 2006; Buckley et al., 2017; Cohen et al., 2018; Macgregor et al., 2019). In this section, we focus on insects' capability to tolerate, adjust, and adapt to temperature change, which their response to climate change will greatly depend on. We address the relationship between the processes and between their genomic bases.

Thermal Plasticity, Thermal Adaptation, Thermal Tolerance

Temperature acts both as an agent of natural selection (resulting in thermal adaptation), and as a factor affecting phenotype expression (in cases of thermal plasticity). There are any examples of thermal plasticity in insects, including developmental effects and adult acclimation. Temperature-dependence has been described for many processes and traits, including sex determination (Blackmon et al., 2017), induction of diapause (Saunders, 2014), body pigmentation (Sibilia et al., 2018), behavior (Abram et al., 2017). Likewise, thermal adaptation and thermal tolerance have also been extensively studied in various insect species (Tobler et al., 2015; Mallard et al., 2018; Kellermann and van Heerwaarden, 2019). Thermal tolerance, corresponding to a favorable range of temperatures for performance, can be

assessed by measuring survival and/or recovery from acute or chronic exposure to temperature extremes (e.g., Kingsolver et al., 2016). As it reflects the capability to cope with adverse temperature conditions, thermal tolerance is very obviously and very directly relevant to how organisms respond to climate change. Moreover, thermal tolerance has also been shown to be associated to other fitness related traits (e.g., tolerance of high temperatures affects dispersal in the Glanville fritillary; Saastamoinen and Hanski, 2008; Mattila, 2015), and to vary between populations and between species (e.g., Hamblin et al., 2017; Oyen and Dillon, 2018).

Thermal plasticity, thermal tolerance, and thermal adaptation are very closely intertwined. Thermal tolerance can be thermally plastic (Schou et al., 2017), but it is unclear how much plasticity in thermal tolerance will impact insects' response to climate change (Mitchell et al., 2011; Gunderson and Stillman, 2015). Thermal plasticity and thermal tolerance can facilitate thermal adaptation (e.g., Mitchell et al., 2011; Noh et al., 2017). Conversely, thermal adaptation can entail changes in thermal plasticity (discussed above), as well as in thermal tolerance. The evolution of thermal tolerance as a result of adaptation to different thermal regimes is compellingly illustrated by differences between populations along climatic clines, including the negative correlation between heat tolerance and both altitude (e.g., in *Heliconius* butterflies; Montejo-Kovacevich et al., 2020) and latitude (e.g., in *Drosophila* flies; van Heerwaarden et al., 2014).

Genomics of Thermal Plasticity

Deciphering the genetic basis of thermal plasticity involves asking about the genes involved in regulating the expression of thermally-dependent phenotypes, as well as about the genes contributing to inter-genotype variation in plasticity that can fuel its evolution (Lafuente and Beldade, 2019). Genomic-level studies of different types have made crucial contributions to both ends. First, investigating the genetic basis of the regulation of thermal plasticity requires identifying genes whose expression and/or function depends on temperature, and, among those, the genes that actually account for changes in thermally-sensitive phenotype expression. Transcriptome-profiling studies in a variety of species have documented thermal plasticity in gene expression levels, including assessment of how many and which genes are differentially expressed between temperatures. The important effect of temperature on transcription has been particularly well studied in the genetic model *Drosophila melanogaster* (e.g., Chen et al., 2015; Sørensen et al., 2016), but also in other insect examples of thermal plasticity (e.g., Oostra et al., 2018). Importantly, while transcriptome-wide scans allow us to identify many genes whose expression depends on temperature, targeted candidate gene analysis facilitates making the connection between differential gene expression and plastic trait development (e.g., thermal plasticity for body pigmentation in *D. melanogaster*; Gibert et al., 2016). Second, investigating the genetic basis of the variation in thermal plasticity involves identifying genes that harbor allelic variation contributing for differences in plasticity, and, among those, which actually fuel the evolution of plasticity. Differences between genotypes in levels of thermal developmental plastic, which correspond to significant genotype-by-environment interactions, document the existence of genetic variation for plasticity and offer the opportunity to characterize its nature. Here too, candidate gene studies are quickly being replaced by less-biased whole genome analysis, including genome-wide association studies that identify QTLs for inter-genotype differences in thermal plasticity for specific plastic traits (e.g., QTLs for thermal plasticity for body size in *D. melanogaster*; Lafuente et al., 2018). These loci can provide the raw material for the evolution of plasticity, including level, direction and inducing cues (discussed in (Lafuente and Beldade, 2019).

The rapid rate of current global climate change, with strong effects on many species, provides both a unique opportunity and a pressing need to study the genetic bases of adaptation, tolerance, and plasticity in natural populations (Franks and Hoffmann, 2012). The extent to which the same genes are involved in thermal adaptation, thermal plasticity, and thermal tolerance has also been addressed both by focusing on candidate genes and by using genomic-level approaches. Perhaps unsurprisingly, several candidate gene studies have focused on genes encoding heat-shock proteins, which have been shown to be thermally plastic, impact thermal tolerance, and differ between populations from different thermal conditions (e.g., Sørensen et al., 2001, 2019; Mattila, 2015; Liu et al., 2017). Heat-shock genes also come up as significant hits in some (e.g., Wang et al., 2019) but not all (e.g., Mallard et al., 2018) genomic-level searches. Accumulating genomic studies in *Drosophila melanogaster*, using different natural and experimental populations and different approaches (Klepsatel et al., 2013; Tobler et al., 2014; Gerken et al., 2015; Machado et al., 2016; Porcelli et al., 2016; Fabian et al., 2017; Lafuente et al., 2018; Rolandi et al., 2018; Kapun et al., 2020), are building an unprecedently powerful body of data to assess the genomic basis of thermal adaptation, and its repeatability and relationship to thermal plasticity and thermal tolerance. In the future, integration of studies covering different species, different geographical and temporal scales, and different approaches will undoubtedly help shed much needed light onto the genomics of thermal plasticity, as well as its overlap with the genomics of thermal tolerance and thermal adaptation.

MULTIFACTORIAL COMPLEX ENVIRONMENTS

Climate change entails changes in mean global temperature, but also in temperature dynamics and in other environmental variables. As such, to assess the potential impact of climate change on natural populations, it is relevant to consider the combined effects of change in temperature with change in other variables. This section considers effects of temperature and other environmental variables on both plasticity and adaptation.

Phenotypic Plasticity in Complex Environments

As illustrated above, effects of the environment on developmental outcome have been amply documented for various phenotypes and species. Indeed, phenotypic plasticity (phenotypic differences

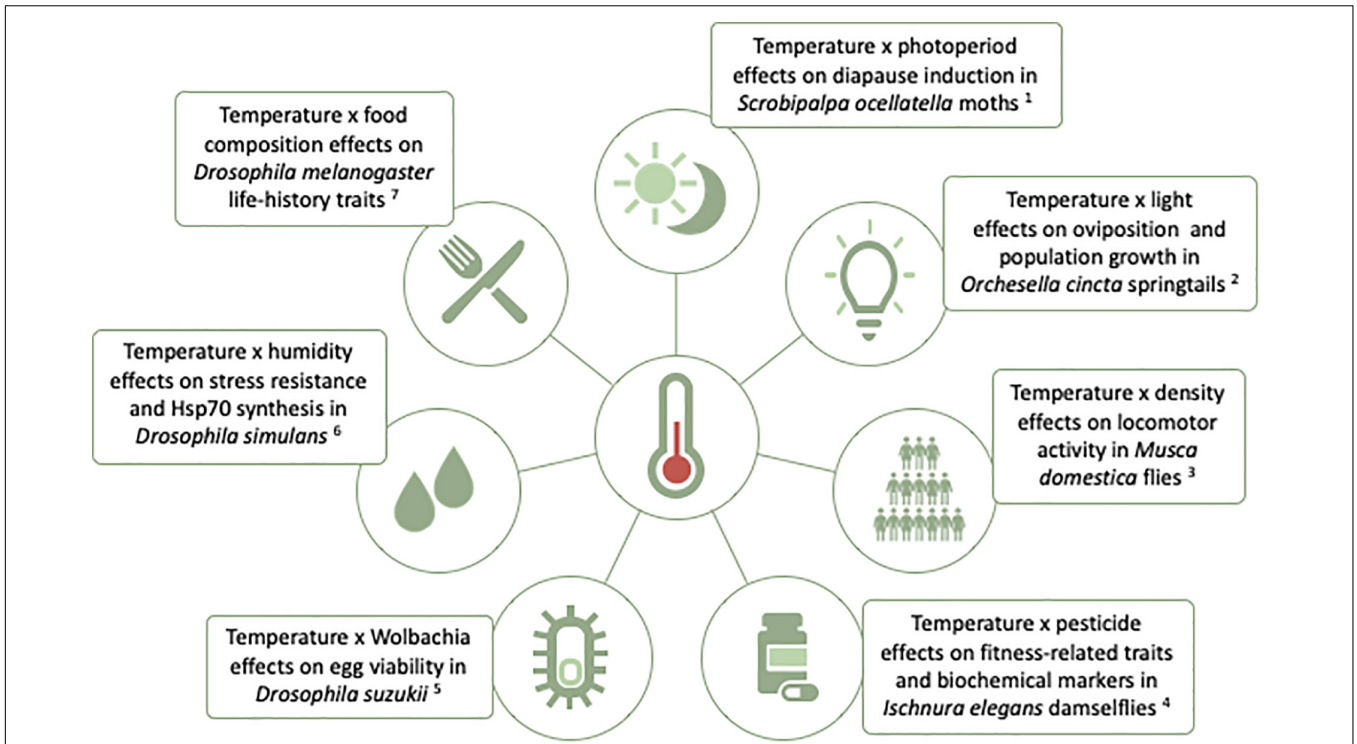


FIGURE 2 | Insect examples of studies of effects of temperature combined with other environmental factors on phenotypic plasticity. Examples from: [1]Ahmadi et al., 2018; [2]Liefting et al., 2017; [3]Schou et al., 2013; [4]Arambourou and Stoks, 2015; [5]Saeed et al., 2018; [6]Bubliy et al., 2013; [7]Kutz et al., 2019.

attributable to environmental variation) and genotype-by-environment interactions (i.e., genetic differences in how organisms respond to environmental conditions) are very common. Unlike what happens for genetic variation, though, where evolutionary biology explicitly considers interaction effects (dominance and epistasis), potential environment-by-environment interactions received considerable less attention. Traditionally, experimental studies of plasticity focused on the analysis of single, isolated environmental factors, held constant during the time it takes organisms to complete their life-cycle. This is in stark contrast with natural situations, where complex environments include variation in multiple and highly dynamic environmental cues. These different variables may act additively on phenotype expression, but may also act redundantly, synergistically, or antagonistically (Piggott et al., 2015; Westneat et al., 2019). Climate change has brought substantial attention to the analysis of multi-stressor effects in populations (Kaunisto et al., 2016), albeit with the majority of studies focused on plants or aquatic systems (e.g., Byrne and Przeslawski, 2013; Gunderson et al., 2016).

Focusing exclusively on environmental factors considered to be individually (i.e., on their own, independently of other environmental factors) and universally (i.e., always, for all genotypes) stressful fails to acknowledge that what is and is not "stressful" might depend on environmental and genetic context. For example, what is a stressful temperature under some photoperiod (or for some genotype) might not be stressful under another photoperiod (or for another genotype). Studies of thermal plasticity in multifactorial environments are increasing, including for different insect species. These studies search to investigate phenotypic effects when variation in temperature is combined with variation in other environmental variables (**Figure 2**), including biotic and abiotic factors (Bubliy et al., 2013; Schou et al., 2013; Arambourou and Stoks, 2015; Saeed et al., 2018; Kutz et al., 2019). Some studies extend the analysis of plasticity in multifactorial environments to include: (1) multiple traits and/or to multiple genotypes (e.g.,, Saastamoinen et al., 2013; Verspagen et al., 2020), (2) three-way environmental interactions (e.g., temperature × humidity × food; Bomble and Nath, 2019), and (3) quantifying underlying changes in gene expression (e.g., candidate genes, Rivas et al., 2018, and whole transcriptome, Koch and Guillaume, 2020). The results to date paint a complex picture, with distinct types of additive (e.g., Koch and Guillaume, 2020) and non-additive (e.g., Yoshii et al., 2009; Arambourou and Stoks, 2015; Piggott et al., 2015) effects of multifactorial environments, and differences between traits and between genotypes. This is an area that will, undoubtedly, know much progress in the near future.

Adaptation to Complex Environments

Aside effects on phenotype expression, multifactorial environments will obviously also affect adaptive evolution in ways that might be unpredictable based on variation for single environmental factors. Adaptation to novel combinations of

environmental variables might be harder or impossible – for example, if phenotypic change favored by one cue is at odds with that favored in relation to the other cue. Such trade-offs are illustrated by studies where adaptation to specific environments entailed costs in performance in other environments (e.g. Callahan et al., 2008; Nunney, 2016; Fox et al., 2019). In natural populations, different environmental factors act in concert as agents of selection, and can co-vary more or less independently and unpredictably. The fact that associations between environmental variables, as well as their dynamics, are likely to change as a result of climate change further endorses the interest in studying the impact of complex environments on the tempo and mode of adaptive evolution.

Our understanding of the phenotypic and genotypic change that accompanies adaptation of insects to complex environments relies on different types of studies. Studies of natural populations include both "snap-shot" and longitudinal comparisons between populations living in different environments (Reinhardt et al., 2014; Manenti et al., 2017; Lerat et al., 2019; Kapun et al., 2020). While studies of natural populations make it possible to detect genetic and phenotypic differentiation and, sometimes, associate the two, it is generally very difficult to know exactly which environmental variables explain divergence and how. Conversely, in studies of experimental populations forced to evolve in different complex environment (e.g., Tomkins et al., 2011; Tobler et al., 2015; Mallard et al., 2018), we typically know exactly which environmental variables differ between selection lines and can identify genetic differences between those lines, but it is not always easy to know which organismal phenotypes were altered and how. It will be valuable to be able to integrate studies from different types of approaches, and for different species and species groups, to have a better understanding of the mechanisms and limitations of adaptation to complex environments.

OVERVIEW AND PERSPECTIVES

Throughout the review, we highlighted what we believe are some areas of particular interest for our understanding of the relevance of thermal plasticity to climate change biology. In light of the topic of this special issue, we discussed recent studies on the genomics of thermal plasticity, distinguishing between those identifying the genes whose expression depends on temperature (and might underlie temperature-induced change in developmental outcome), and in terms of the genes that harbor allelic variants contributing to inter-genotype variation in plasticity (and can feed the evolution of thermal plasticity) (see Lafuente and Beldade, 2019). As data accumulates for different systems, we can hope to deepen our knowledge about what those genes are and about the overlap between them, as well as the overlap between them and the those underlying thermal tolerance and thermal adaptation. We also emphasized the relevance of focusing on temperature in the context of complex multifactorial environments, and the importance of considering that variation in response to temperature can depend on genetic and environmental context.

We focused on thermal plasticity in insects, its potential role in response to climate change, its genomic basis, and the interactions between temperature and other environmental factors. Each of these issues, along with related topics that we did not cover at all, is attracting substantial research attention and, we expect, will know much progress in the near future. Below we highlight the topics complementary to those we covered that are also relevant for the discussion about the relevance of thermal plasticity to climate change biology.

First, we focused primarily on developmental plasticity, which, especially in holometabolous insects, often leads to fixed adult phenotypes. We paid less attention to effects of temperature directly on adult traits, which often lead to reversible phenotypes. These include phenomena that are key to climate change biology, such as acclimation, through physiological and/or behavioral plasticity (Huey et al., 2003; Stillman, 2003). These can mitigate the immediate effect of variation in thermal environments, but can also constrain adaptation to permanent/directional temperature perturbation.

Second, we focused on effects of climate change and of multifactorial environments on molecular-level processes (e.g., thermal plasticity in gene expression), organismal-level processes (e.g., thermal plasticity in developmental outcomes), and population-level processes (e.g., thermal adaptation in experimental and natural populations). We did not discuss supra-population effects of climate change or of multifactorial environments (Fordyce, 2006), such as effects on the species composition of communities (e.g., Chown et al., 2015; de Vries et al., 2019) and on inter-specific interactions (e.g., Williams et al., 2008; Cornelissen, 2011; Wernegreen, 2012; Cahill et al., 2013), both of which can have substantial ramification effects (Grimm et al., 2013).

Finally, we focused exclusively on insect examples, but effects of climate change and multifactorial environments on phenotype expression and adaptation are also being studied in other groups (e.g., Byrne and Przeslawski, 2013; Gunderson et al., 2016; Lange and Marshall, 2017). It will be crucial to integrate different examples both to uncover unique responses, as well as to derive general principles about biological responses to climate change.

AUTHOR CONTRIBUTIONS

YR researched the literature. YR and PB conceived and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

Financial support for this work was provided by the Portuguese science funding agency, Fundação para a Ciência e Tecnologia, FCT: Ph.D. fellowship to YR (FCT SFRH/BD/114404/2016), and research support to PB ((PTDC/BIA-EVF/0017/2014 and PTDC/BEX-BID/5340/2014).

ACKNOWLEDGMENTS

We wish to thank the issue and journal editors for the invitation and for accommodating delays due to the pandemics, Joana Carvalho (@joana_gcc) for the *Bicyclus anynana* drawings in **Box 1**, and various colleagues for discussions who helped shape our ideas on this topic, including Christian Braendle, Roberto Arbore, Elvira Lafuente, and Erik van Bergen.

REFERENCES

Abram, P. K., Boivin, G., Moiroux, J., and Brodeur, J. (2017). Behavioural effects of temperature on ectothermic animals: unifying thermal physiology and behavioural plasticity. *Biol. Rev.* 92, 1859–1876. doi: 10.1111/brv.12312

Ahmadi, F., Moharramipour, S., and Mikani, A. (2018). The effect of temperature and photoperiod on diapause induction in pupae of *Scrobipalpa ocellatella* (Lepidoptera: Gelechiidae). *Environ. Entomol.* 47, 1314–1322. doi: 10.1093/ee/nvy082

Arambourou, H., and Stoks, R. (2015). Combined effects of larval exposure to a heat wave and chlorpyrifos in northern and southern populations of the damselfly *Ischnura elegans*. *Chemosphere* 128, 148–154. doi: 10.1016/j.chemosphere.2015.01.044

Balmer, A. J., Brakefield, P. M., Brattström, O., and van Bergen, E. (2018). Developmental plasticity for male secondary sexual traits in a group of polyphenic tropical butterflies. *Oikos* 127, 1812–1821. doi: 10.1111/oik.05291

Bauerfeind, S. S., and Fischer, K. (2005). Effects of food stress and density in different life stages on reproduction in a butterfly. *Oikos* 111, 514–524. doi: 10.1111/j.0030-1299.2005.13888.x

Bear, A., and Monteiro, A. (2013). Male Courtship rate plasticity in the butterfly *Bicyclus anynana* is controlled by temperature experienced during the pupal and adult Stages. *PLoS One* 8:e064061. doi: 10.1371/journal.pone.0064061

Bear, A., Prudic, K. L., and Monteiro, A. (2017). Steroid hormone signaling during development has a latent effect on adult male sexual behavior in the butterfly *Bicyclus anynana*. *PLoS One* 12:e0174403. doi: 10.1371/journal.pone.0174403

Beldade, P., Mateus, A. R. A., and Keller, R. A. (2011). Evolution and molecular mechanisms of adaptive developmental plasticity. *Mol. Ecol.* 20, 1347–1363. doi: 10.1111/j.1365-294X.2011.05016.x

Beldade, P., and Peralta, C. M. (2017). Developmental and evolutionary mechanisms shaping butterfly eyespots. *Curr. Opin. Insect Sci.* 19, 22–29. doi: 10.1016/j.cois.2016.10.006

Bentz, B. J., and Powell, J. A. (2014). Mountain pine beetle seasonal timing and constraints to bivoltinism: a comment on Mitton and Ferrenberg, Mountain pine beetle develops an unprecedented summer generation in response to climate warming. *Am. Nat.* 184, 787–796. doi: 10.1086/678405

Bhardwaj, S., Jolander, L. S. H., Wenk, M. R., Oliver, J. C., Nijhout, H. F., and Monteiro, A. (2020). Origin of the mechanism of phenotypic plasticity in satyrid butterfly eyespots. *eLife* 9, 1–13. doi: 10.7554/eLife.49544

Bilandžija, H., Hollifield, B., Steck, M., Meng, G., Ng, M., Koch, A. D., et al. (2020). Phenotypic plasticity as a mechanism of cave colonization and adaptation. *eLife* 9, 1–27. doi: 10.7554/eLife.51830

Blackmon, H., Ross, L., and Bachtrog, D. (2017). Sex determination, sex chromosomes, and karyotype evolution in insects. *J. Hered.* 108, 78–93. doi: 10.1093/jhered/esw047

Bomble, P., and Nath, B. B. (2019). Comparative susceptibility of chironomus and *Drosophila* to exposure to each and combinations of the following stressors: desiccation, heat stress and starvation. *Eur. J. Environ. Sci.* 9, 41–46. doi: 10.14712/23361964.2019.5

Bonamour, S., Chevin, L. M., Charmantier, A., and Teplitsky, C. (2019). Phenotypic plasticity in response to climate change: the importance of cue variation. *Philos. Trans. R. Soc. B Biol. Sci.* 374:20180178. doi: 10.1098/rstb.2018.0178

Bradshaw, W. E., and Holzapfel, C. M. (2001). Genetic shift in photoperiodic response correlated with global warming. *Proc. Natl. Acad. Sci. U.S.A.* 98, 14509–14511. doi: 10.1073/PNAS.241391498

Brakefield, P. M. (1996). Seasonal polyphenism in butterflies and natural selection. *Trends Ecol. Evol.* 11, 275–277. doi: 10.1016/0169-5347(96)30025-30026

Brakefield, P. M., Beldade, P., and Zwaan, B. J. (2009). The African butterfly *Bicyclus anynana*: a model for evolutionary genetics and evolutionary developmental biology. *Cold Spring Harb. Protoc.* 4:pdb.emo122. doi: 10.1101/pdb.emo122

Brakefield, P. M., Kesbeke, F., and Koch, P. B. (1998). The regulation of phenotypic plasticity of eyespots in the butterfly *Bicyclus anynana*. *Am. Nat.* 152, 853–860. doi: 10.1086/286213

Brakefield, P. M., and Reitsma, N. (1991). Phenotypic plasticity, seasonal climate and the population biology of Bicyclus butterflies (Satyridae) in Malawi. *Ecol. Entomol.* 16, 291–303. doi: 10.1111/j.1365-2311.1991.tb00220.x

Bubliy, O. A., Kristensen, T. N., and Loeschcke, V. (2013). Stress-induced plastic responses in *Drosophila simulans* following exposure to combinations of temperature and humidity levels. *J. Exp. Biol.* 216, 4601–4607. doi: 10.1242/jeb.092502

Buckley, L. B., Arakaki, A. J., Cannistra, A. F., Kharouba, H. M., and Kingsolver, J. G. (2017). Insect development, thermal plasticity and fitness implications in changing, seasonal environments. *Integr. Comp. Biol.* 57, 988–998. doi: 10.1093/icb/icx032

Byrne, M., and Przeslawski, R. (2013). Multistressor impacts of warming and acidification of the ocean on marine invertebrates' life histories. *Integr. Comp. Biol.* 53, 582–596. doi: 10.1093/icb/ict049

Cahill, A. E., Aiello-Lammens, M. E., Caitlin Fisher-Reid, M., Hua, X., Karanewsky, C. J., Ryu, H. Y., et al. (2013). How does climate change cause extinction? *Proc. R. Soc. B Biol. Sci.* 280:890. doi: 10.1098/rspb.2012.1890

Callahan, H. S., Maughan, H., and Steiner, U. K. (2008). Phenotypic plasticity, costs of phenotypes, and costs of plasticity: toward an integrative view. *Ann. N. Y. Acad. Sci.* 1133, 44–66. doi: 10.1196/annals.1438.008

Cenzer, M. L. (2017). Maladaptive plasticity masks the effects of natural selection in the red-shouldered soapberry bug. *Am. Nat.* 190, 521–533. doi: 10.1086/693456

Chen, J., Nolte, V., and Schlotterer, C. (2015). Temperature-related reaction norms of gene expression: regulatory architecture and functional implications. *Mol. Biol. Evol.* 32, 2393–2402. doi: 10.1093/molbev/msv120

Chevin, L. M., Collins, S., and Lefèvre, F. (2013). Phenotypic plasticity and evolutionary demographic responses to climate change: taking theory out to the field. *Funct. Ecol.* 27, 967–979. doi: 10.1111/j.1365-2435.2012.02043.x

Chevin, L. M., and Hoffmann, A. A. (2017). Evolution of phenotypic plasticity in extreme environments. *Philos. Trans. R. Soc. B Biol. Sci.* 372:20160138. doi: 10.1098/rstb.2016.0138

Chevin, L. M., Lande, R., and Mace, G. M. (2010). Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biol.* 8:1000357. doi: 10.1371/journal.pbio.1000357

Chown, S. L., Hodgins, K. A., Griffin, P. C., Oakeshott, J. G., Byrne, M., and Hoffmann, A. A. (2015). Biological invasions, climate change and genomics. *Evol. Appl.* 8, 23–46. doi: 10.1111/eva.12234

Cohen, J. M., Lajeunesse, M. J., and Rohr, J. R. (2018). A global synthesis of animal phenological responses to climate change/631/158/2165/2457/631/158/2039/129/141/139 letter. *Nat. Clim. Chang.* 8, 224–228. doi: 10.1038/s41558-018-0067-63

Colinet, H., Sinclair, B. J., Vernon, P., and Renault, D. (2015). Insects in fluctuating thermal environments. *Annu. Rev. Entomol.* 60, 123–140. doi: 10.1146/annurev-ento-010814-021017

Condamine, F. L., and Sperling, F. A. H. (2018). Anthropogenic threats to high-altitude parnassian diversity. *News Lepid. Soc.* 60, 94–99.

Cornelissen, T. (2011). Climate change and its effects on terrestrial insects and herbivory patterns. *Neotrop. Entomol.* 40, 155–163. doi: 10.1590/S1519-566X2011000200001

de Jong, M. A., Kesbeke, F. M. N. H., Brakefield, P. M., and Zwaan, B. J. (2010). Geographic variation in thermal plasticity of life history and wing pattern in *Bicyclus anynana*. *Clim. Res.* 43, 91–102. doi: 10.3354/cr00881

de Vries, J., Kraak, M. H. S., Verdonschot, R. C. M., and Verdonschot, P. F. M. (2019). Quantifying cumulative stress acting on macroinvertebrate assemblages in lowland streams. *Sci. Total Environ.* 694:133630. doi: 10.1016/j.scitotenv.2019.133630

Didham, R. K., Basset, Y., Collins, C. M., Leather, S. R., Littlewood, N. A., Menz, M. H. M., et al. (2020). Interpreting insect declines: seven challenges and a way forward. *Insect. Conserv. Divers.* 13, 103–114. doi: 10.1111/icad.12408

Donelson, J. M., Salinas, S., Munday, P. L., and Shama, L. N. S. (2018). Transgenerational plasticity and climate change experiments: where do we go from here? *Glob. Chang. Biol.* 24, 13–34. doi: 10.1111/gcb.13903

Dworschak, K., Gruppe, A., and Schopf, R. (2014). Survivability and post-diapause fitness in a scolytid beetle as a function of overwintering developmental stage and the implications for population dynamics. *Ecol. Entomol.* 39, 519–526. doi: 10.1111/een.12127

Fabian, D. K., Lack, J. B., Mathur, V., Schlötterer, C., Schmidt, P. S., Pool, J. E., et al. (2017). Spatially varying selection shapes life history clines among populations of *Drosophila melanogaster* from sub-Saharan Africa. *J. Evol. Biol.* 28, 826–840. doi: 10.1016/j.physbeh.2017.03.040

Faldyn, M. J., Hunter, M. D., and Elderd, B. D. (2018). Climate change and an invasive, tropical milkweed: an ecological trap for monarch butterflies. *Ecology* 99, 1031–1038. doi: 10.1002/ecy.2198

Fischer, K., Dierks, A., Franke, K., Geister, T. L., Liszka, M., Winter, S., et al. (2010). Environmental effects on temperature stress resistance in the tropical butterfly *Bicyclus anynana*. *PLoS One* 5:e15284. doi: 10.1371/journal.pone.0015284

Fischer, K., Zwaan, B. J., and Brakefield, P. M. (2007). Realized correlated responses to artificial selection on pre-adult life-history traits in a butterfly. *Heredity* 98, 157–164. doi: 10.1038/sj.hdy.6800919

Foden, W. B., Young, B. E., Akçakaya, H. R., Garcia, R. A., Hoffmann, A. A., Stein, B. A., et al. (2019). Climate change vulnerability assessment of species. *Wiley Interdiscip. Rev. Clim. Chang.* 10, 1–36. doi: 10.1002/wcc.551

Fordyce, J. A. (2006). The evolutionary consequences of ecological interactions mediated through phenotypic plasticity. *J. Exp. Biol.* 209, 2377–2383. doi: 10.1242/jeb.02271

Fox, R. J., Donelson, J. M., Schunter, C., Ravasi, T., and Gaitán-Espitia, J. D. (2019). Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. *Philos. Trans. R. Soc. B Biol. Sci.* 374:174. doi: 10.1098/rstb.2018.0174

Franks, S. J., and Hoffmann, A. A. (2012). Genetics of climate change adaptation. *Annu. Rev. Genet.* 46, 185–208. doi: 10.1146/annurev-genet-110711-155511

Gerken, A. R., Eller, O. C., Hahn, D. A., and Morgan, T. J. (2015). Constraints, independence, and evolution of thermal plasticity: probing genetic architecture of long-and short-term thermal acclimation. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4399–4404. doi: 10.1073/pnas.1503456112

Ghalambor, C. K., McKay, J. K., Carroll, S. P., and Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct. Ecol.* 21, 394–407. doi: 10.1111/j.1365-2435.2007.01283.x

Gibert, J. M., Mouchel-Vielh, E., De Castro, S., and Peronnet, F. (2016). Phenotypic plasticity through transcriptional regulation of the evolutionary hotspot gene tan in *Drosophila melanogaster*. *PLoS Genet.* 12:1006218. doi: 10.1371/journal.pgen.1006218

Gilbert, S. F., Bosch, T. C. G., and Ledón-Rettig, C. (2015). Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nat. Rev. Genet.* 16, 611–622. doi: 10.1038/nrg3982

Grimm, N. B., Chapin, F. S., Bierwagen, B., Gonzalez, P., Groffman, P. M., Luo, Y., et al. (2013). The impacts of climate change on ecosystem structure and function. *Front. Ecol. Environ.* 11, 474–482. doi: 10.1890/120282

Gunderson, A. R., Armstrong, E. J., and Stillman, J. H. (2016). Multiple stressors in a changing world: The need for an improved perspective on physiological responses to the dynamic marine environment. *Ann. Rev. Mar. Sci.* 8, 357–378. doi: 10.1146/annurev-marine-122414-133953

Gunderson, A. R., and Stillman, J. H. (2015). Plasticity in thermal tolerance has limited potential to buffer ectotherms from global warming. *Proc. R. Soc. B Biol. Sci.* 282:401. doi: 10.1098/rspb.2015.0401

Halali, S., Brakefield, P. M., Collins, S. C., and Brattström, O. (2020). To mate, or not to mate: the evolution of reproductive diapause facilitates insect radiation into African savannahs in the Late Miocene. *J. Anim. Ecol.* 89, 1230–1241. doi: 10.1111/1365-2656.13178

Hamblin, A. L., Youngsteadt, E., López-Uribe, M. M., and Frank, S. D. (2017). Physiological thermal limits predict differential responses of bees to urban heat-island effects. *Biol. Lett.* 13:125. doi: 10.1098/rsbl.2017.0125

Holloway, G. J., and Brakefield, P. M. (1995). Artificial selection of reaction norms of wing pattern elements in *Bicyclus anynana*. *Heredity* 74, 91–99. doi: 10.1038/hdy.1995.11

Huey, R. B., Hertz, P. E., and Sinervo, B. (2003). Behavioral drive versus behavioral inertia in evolution: a null model approach. *Amer. Nat.* 161, 357–366.

Johnston, A. S. A., Boyd, R. J., Watson, J. W., Paul, A., Evans, L. C., Gardner, E. L., et al. (2019). Predicting population responses to environmental change from individual-level mechanisms: Towards a standardized mechanistic approach. *Proc. R. Soc. B Biol. Sci.* 286:1916. doi: 10.1098/rspb.2019.1916

Jönsson, A. M., Harding, S., Bärring, L., and Ravn, H. P. (2007). Impact of climate change on the population dynamics of Ips typographus in southern Sweden. *Agric. For. Meteorol.* 146, 70–81. doi: 10.1016/j.agrformet.2007.05.006

Kapun, M., Barrón, M. G., Staubach, F., Obbard, D. J., Wiberg, R. A. W., Vieira, J., et al. (2020). Genomic analysis of European *Drosophila melanogaster* populations reveals longitudinal structure, continent-wide selection, and previously unknown DNA viruses. *Mol. Biol. Evol.* 2020:msaa120. doi: 10.1093/molbev/msaa120

Kaunisto, S., Ferguson, L. V., and Sinclair, B. J. (2016). Can we predict the effects of multiple stressors on insects in a changing climate? *Curr. Opin. Insect. Sci.* 17, 55–61. doi: 10.1016/j.cois.2016.07.001

Kearney, M. R., Briscoe, N. J., Karoly, D. J., Porter, W. P., Norgate, M., and Sunnucks, P. (2010). Early emergence in a butterfly causally linked to anthropogenic warming. *Biol. Lett.* 6, 674–677. doi: 10.1098/rsbl.2010.0053

Kellermann, V., and van Heerwaarden, B. (2019). Terrestrial insects and climate change: adaptive responses in key traits. *Physiol. Entomol.* 44, 99–115. doi: 10.1111/phen.12282

Kelly, M. (2019). Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. *Philos. Trans. R. Soc. B Biol. Sci.* 374:176. doi: 10.1098/rstb.2018.0176

Kingsolver, J. G., and Buckley, L. B. (2018). How do phenology, plasticity, and evolution determine the fitness consequences of climate change for montane butterflies? *Evol. Appl.* 11, 1231–1244. doi: 10.1111/eva.12618

Kingsolver, J. G., MacLean, H. J., Goddin, S. B., and Augustine, K. E. (2016). Plasticity of upper thermal limits to acute and chronic temperature variation in *Manduca sexta* larvae Joel. *J. Exp. Biol.* 219, 1290–1294. doi: 10.1242/jeb.138321

Klepsatel, P., Gáliková, M., De Maio, N., Huber, C. D., Schlötterer, C., and Flatt, T. (2013). Variation in thermal performance and reaction norms among populations of *Drosophila melanogaster*. *Evolution* 67, 3573–3587. doi: 10.1111/evo.12221

Klockmann, M., Kleinschmidt, F., and Fischer, K. (2017). Carried over: heat stress in the egg stage reduces subsequent performance in a butterfly. *PLoS One* 12:180968. doi: 10.1371/journal.pone.0180968

Koch, E. L., and Guillaume, F. (2020). Additive and mostly adaptive plastic responses of gene expression to multiple stress in *Tribolium castaneum*. *PLoS Genet.* 16:e1008768. doi: 10.1371/journal.pgen.1008768

Kollberg, I., Bylund, H., Schmidt, A., Gershenzon, J., and Björkman, C. (2013). Multiple effects of temperature, photoperiod and food quality on the performance of a pine sawfly. *Ecol. Entomol.* 38, 201–208. doi: 10.1111/een.12005

Kutz, T. C., Sgrò, C. M., and Mirth, C. K. (2019). Interacting with change: diet mediates how larvae respond to their thermal environment. *Funct. Ecol.* 33, 1940–1951. doi: 10.1111/1365-2435.13414

Lafuente, E., and Beldade, P. (2019). Genomics of developmental plasticity in animals. *Front. Genet.* 10:720. doi: 10.3389/fgene.2019.00720

Lafuente, E., Duneau, D., and Beldade, P. (2018). Genetic basis of thermal plasticity variation in *Drosophila melanogaster* body size. *PLoS Genet.* 14:e1007686. doi: 10.1371/journal.pgen.1007686

Lange, R., and Marshall, D. (2017). Ecologically relevant levels of multiple, common marine stressors suggest antagonistic effects. *Sci. Rep.* 7, 1–9. doi: 10.1038/s41598-017-06373-y

Leonard, A. M., and Lancaster, L. T. (2020). Maladaptive plasticity facilitates evolution of thermal tolerance during an experimental range shift. *BMC Evol. Biol.* 20:1587. doi: 10.1186/s12862-020-1589-1587

Lerat, E., Goubert, C., Guirao-Rico, S., Merenciano, M., Dufour, A. B., Vieira, C., et al. (2019). Population-specific dynamics and selection patterns of transposable element insertions in European natural populations. *Mol. Ecol.* 28, 1506–1522. doi: 10.1111/mec.14963

Levis, N. A., and Pfennig, D. W. (2016). Evaluating "Plasticity-First" evolution in Nature: key criteria and empirical approaches. *Trends Ecol. Evol.* 31, 563–574. doi: 10.1016/j.tree.2016.03.012

Liefting, M., Cosijn, J., and Ellers, J. (2017). Synergistic effect of daily temperature fluctuations and matching light-dark cycle enhances population growth and synchronizes oviposition behavior in a soil arthropod. *J. Insect Physiol.* 96, 108–114. doi: 10.1016/j.jinsphys.2016.10.002

Liu, Y., Su, H., Li, R., Li, X., Xu, Y., Dai, X., et al. (2017). Comparative transcriptome analysis of *Glyphodes pyloalis* walker (Lepidoptera: Pyralidae) reveals novel insights into heat stress tolerance in insects. *BMC Genomics* 18:974. doi: 10.1186/s12864-017-4355-4355

Lyytinen, A., Brakefield, P. M., Lindstrom, L., and Mappes, J. (2004). Does predation maintain eyespot plasticity in *Bicyclus anynana*? *Proc. R. Soc. B Biol. Sci.* 271, 279–283. doi: 10.1098/rspb.2003.2571

Macgregor, C. J., Thomas, C. D., Roy, D. B., Beaumont, M. A., Bell, J. R., Brereton, T., et al. (2019). Climate-induced phenology shifts linked to range expansions in species with multiple reproductive cycles per year. *Nat. Commun.* 10:4455. doi: 10.1038/s41467-019-12479-w

Machado, H. E., Bergland, A. O., O'Brien, K. R., Behrman, E. L., Schmidt, P. S., and Petrov, D. A. (2016). Comparative population genomics of latitudinal variation in *Drosophila simulans* and *Drosophila melanogaster*. *Mol. Ecol.* 25, 723–740. doi: 10.1111/mec.13446

Macias-Munoz, A., Smith, G., Monteiro, A., and Briscoe, A. D. (2016). Transcriptome-wide differential gene expression in *Bicyclus anynana* butterflies: female vision-related genes are more plastic. *Mol. Biol. Evol.* 33, 79–92. doi: 10.1093/molbev/msv197

Mallard, F., Nolte, V., Tobler, R., Kapun, M., and Schlötterer, C. (2018). A simple genetic basis of adaptation to a novel thermal environment results in complex metabolic rewiring in *Drosophila*. *Genome Biol.* 19, 1–15. doi: 10.1186/s13059-018-1503-1504

Manenti, T., Sørensen, J. G., and Loeschcke, V. (2017). Environmental heterogeneity does not affect levels of phenotypic plasticity in natural populations of three *Drosophila* species. *Ecol. Evol.* 7, 2716–2724. doi: 10.1002/ece3.2904

Manfredini, F., Arbetman, M., and Toth, A. L. (2019). A potential role for phenotypic plasticity in invasions and declines of social insects. *Front. Ecol. Evol.* 7:375. doi: 10.3389/fevo.2019.00375

Mateus, A. R. A., Marques-Pita, M., Oostra, V., Lafuente, E., Brakefield, P. M., Zwaan, B. J., et al. (2014). Adaptive developmental plasticity: compartmentalized responses to environmental cues and to corresponding internal signals provide phenotypic flexibility. *BMC Biol.* 12:97. doi: 10.1186/s12915-014-0097-x

Mattila, A. L. K. (2015). Thermal biology of flight in a butterfly: genotype, flight metabolism, and environmental conditions. *Ecol. Evol.* 5, 5539–5551. doi: 10.1002/ece3.1758

Merilä, J., and Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol. Appl.* 7, 1–14. doi: 10.1111/eva.12137

Mitchell, K. A., Sgrò, C. M., and Hoffmann, A. A. (2011). Phenotypic plasticity in upper thermal limits is weakly related to *Drosophila* species distributions. *Funct. Ecol.* 25, 661–670. doi: 10.1111/j.1365-2435.2010.01821.x

Moczek, A. P. (2010). Phenotypic plasticity and diversity in insects. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 593–603. doi: 10.1098/rstb.2009.0263

Monteiro, A., Tong, X., Bear, A., Liew, S. F., Bhardwaj, S., Wasik, B. R., et al. (2015). Differential expression of ecdysone receptor leads to variation in phenotypic plasticity across serial homologs. *PLoS Genet.* 11:1005529. doi: 10.1371/journal.pgen.1005529

Montejo-Kovacevich, G., Martin, S. H., Meier, J. I., Bacquet, C. N., Monllor, M., Jiggins, C. D., et al. (2020). Microclimate buffering and thermal tolerance across elevations in a tropical butterfly. *J. Exp. Biol.* 223:jeb220426. doi: 10.1242/jeb.220426

Muller, D., Elias, B., Collard, L., Pels, C., Holveck, M. J., and Nieberding, C. M. (2019). Polyphenism of visual and chemical secondary sexually-selected wing traits in the butterfly *Bicyclus anynana*: how different is the intermediate phenotype? *PLoS One* 14:225003. doi: 10.1371/journal.pone.0225003

Nettle, D., and Bateson, M. (2015). Adaptive developmental plasticity: what is it, how can we recognize it and when can it evolve? *Proc. R. Soc. B Biol. Sci.* 282:1005. doi: 10.1098/rspb.2015.1005

Nijhout, H. F. (2003). Development and evolution of adaptive polyphenisms. *Evol. Dev.* 5, 9–18. doi: 10.1046/j.1525-142X.2003.03003.x

Noh, S., Everman, E. R., Berger, C. M., and Morgan, T. J. (2017). Seasonal variation in basal and plastic cold tolerance: adaptation is influenced by both long- and short-term phenotypic plasticity. *Ecol. Evol.* 7, 5248–5257. doi: 10.1002/ece3.3112

Nunney, L. (2016). Adapting to a changing environment: modeling the interaction of directional selection and plasticity. *J. Hered.* 107, 15–24. doi: 10.1093/jhered/esv084

Olofsson, M., Vallin, A., Jakobsson, S., and Wiklund, C. (2010). Marginal eyespots on butterfly wings deflect bird attacks under low light intensities with UV wavelengths. *PLoS One* 5:10798. doi: 10.1371/journal.pone.0010798

Oostra, V., de Jong, M. A., Invergo, B. M., Kesbeke, F., Wende, F., Brakefield, P. M., et al. (2011). Translating environmental gradients into discontinuous reaction norms via hormone signalling in a polyphenic butterfly. *Proc. R. Soc. B Biol. Sci.* 278, 789–797. doi: 10.1098/rspb.2010.1560

Oostra, V., Mateus, A. R. A., van der Burg, K. R. L., Piessens, T., Eijk, M., Van Brakefield, P. M., et al. (2014). Ecdysteroid hormones link the juvenile environment to alternative adult life histories in a seasonal insect. *Am. Nat.* 184, E79–E92. doi: 10.1086/677260

Oostra, V., Saastamoinen, M., Zwaan, B. J., and Wheat, C. W. (2018). Strong phenotypic plasticity limits potential for evolutionary responses to climate change. *Nat. Commun.* 9:1005. doi: 10.1038/s41467-018-03384-3389

Oyen, K. J., and Dillon, M. E. (2018). Critical thermal limits of bumblebees (*Bombus impatiens*) are marked by stereotypical behaviors and are unchanged by acclimation, age or feeding status. *J. Exp. Biol.* 221:jeb165589. doi: 10.1242/jeb.165589

Parkash, R., Ramniwas, S., and Kajla, B. (2013). Climate warming mediates range shift of two differentially adapted stenothermal *Drosophila* species in the Western Himalayas. *J. Asia. Pac. Entomol.* 16, 147–153. doi: 10.1016/j.aspen.2012.12.004

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* 37, 637–669. doi: 10.2307/annurev.ecolsys.37.091305.30000024

Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J. K., Thomas, C. D., Descimon, H., et al. (1999). Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* 399, 579–583. doi: 10.1038/21181

Parmesan, C., Williams-Anderson, A., Moskwik, M., Mikheyev, A. S., and Singer, M. C. (2015). Endangered Quino checkerspot butterfly and climate change: short-term success but long-term vulnerability? *J. Insect Conserv.* 19, 185–204. doi: 10.1007/s10841-014-9743-9744

Pfennig, D. W., Wund, M. A., Snell-Rood, E. C., Cruickshank, T., Schlichting, C. D., and Moczek, A. P. (2010). Phenotypic plasticity's impacts on diversification and speciation. *Trends Ecol. Evol.* 25, 459–467. doi: 10.1016/j.tree.2010.05.006

Piggott, J. J., Townsend, C. R., and Matthaei, C. D. (2015). Reconceptualizing synergism and antagonism among multiple stressors. *Ecol. Evol.* 5, 1538–1547. doi: 10.1002/ece3.1465

Pijpe, J., Fischer, K., Brakefield, P. M., and Zwaan, B. J. (2006). Consequences of artificial selection on pre-adult development for adult lifespan under benign conditions in the butterfly *Bicyclus anynana*. *Mech. Ageing Dev.* 127, 802–807. doi: 10.1016/j.mad.2006.07.006

Porcelli, D., Westram, A. M., Pascual, M., Gaston, K. J., Butlin, R. K., and Snook, R. R. (2016). Gene expression clines reveal local adaptation and associated trade-offs at a continental scale. *Sci. Rep.* 6, 1–12. doi: 10.1038/srep32975

Prudic, K. L., Jeon, C., Cao, H., and Monteiro, A. (2011). Developmental plasticity in sexual roles of butterfly species drives mutual sexual ornamentation. *Science* 331, 73–75. doi: 10.1126/science.1197114

Prudic, K. L., Stoehr, A. M., Wasik, B. R., and Monteiro, A. (2015). Eyespots deflect predator attack increasing fitness and promoting the evolution of phenotypic plasticity. *Proc. R. Soc. B Biol. Sci.* 282:1531. doi: 10.1098/rspb.2014.1531

Reed, T. E., Schindler, D. E., and Waples, R. S. (2011). Interacting effects of phenotypic plasticity and evolution on population persistence in a changing climate. *Conserv. Biol.* 25, 56–63. doi: 10.1111/j.1523-1739.2010.01552.x

Reinhardt, J. A., Kolaczkowski, B., Jones, C. D., Begun, D. J., and Kern, A. D. (2014). Parallel geographic variation in *Drosophila melanogaster*. *Genetics* 197, 361–373. doi: 10.1534/genetics.114.161463

Rivas, M. J., Saura, M., Pérez-Figueroa, A., Panova, M., Johansson, T., André, C., et al. (2018). Population genomics of parallel evolution in gene expression and gene sequence during ecological adaptation. *Sci. Rep.* 8, 1–12. doi: 10.1038/s41598-018-33897-33898

Rodrigues, Y. K., van Bergen, E., Alves, F., Duneau, D., and Beldade, P. (2018). Complex effects of day and night temperature fluctuations on thermally plastic traits in an experimental model of adaptive seasonal plasticity. *bioRxiv* [Preprint], doi: 10.1101/207258

Rolandi, C., Lighton, J. R. B., de la Vega, G. J., Schilman, P. E., and Mensch, J. (2018). Genetic variation for tolerance to high temperatures in a population of *Drosophila melanogaster*. *Ecol. Evol.* 8, 10374–10383. doi: 10.1002/ece3.4409

Ryan, S. J., Carlson, C. J., Mordecai, E. A., and Johnson, L. R. (2019). Global expansion and redistribution of Aedes-borne virus transmission risk with climate change. *PLoS Negl. Trop. Dis.* 13:7213. doi: 10.1371/journal.pntd.0007213

Saastamoinen, M., Brommer, J. E., Brakefield, P. M., and Zwaan, B. J. (2013). Quantitative genetic analysis of responses to larval food limitation in a polyphenic butterfly indicates environmentand trait-specific effects. *Ecol. Evol.* 3, 3565–3575. doi: 10.1002/ece3.684

Saastamoinen, M., and Hanski, I. (2008). Genotypic and environmental effects on flight activity and oviposition in the glanville fritillary butterfly. *Am. Nat.* 171, 701–712. doi: 10.1086/587531

Saeed, N., Battisti, A., Martinez-Sañudo, I., and Mori, N. (2018). Combined effect of temperature and Wolbachia infection on the fitness of *Drosophila suzukii*. *Bull. Insectol.* 71, 161–169.

Sánchez-Bayo, F., and Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: a review of its drivers. *Biol. Conserv.* 232, 8–27. doi: 10.1016/j.biocon.2019.01.020

Saunders, D. S. (2014). Insect photoperiodism: effects of temperature on the induction of insect diapause and diverse roles for the circadian system in the photoperiodic response. *Entomol. Sci.* 17, 25–40. doi: 10.1111/ens.12059

Schneider, R. F., and Meyer, A. (2017). How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. *Mol. Ecol.* 26, 330–350. doi: 10.1111/mec.13880

Schou, M. F., Mouridsen, M. B., Sørensen, J. G., and Loeschcke, V. (2017). Linear reaction norms of thermal limits in *Drosophila*: predictable plasticity in cold but not in heat tolerance. *Funct. Ecol.* 31, 934–945. doi: 10.1111/1365-2435.12782

Schou, T. M., Faurby, S., Kjaersgaard, A., Pertoldi, C., Loeschcke, V., Hald, B., et al. (2013). Temperature and population density effects on locomotor activity of *Musca domestica* (Diptera: Muscidae). *Environ. Entomol.* 42, 1322–1328. doi: 10.1603/en13039

Sgrò, C. M., Terblanche, J. S., and Hoffmann, A. A. (2016). What can plasticity contribute to insect responses to climate change? *Annu. Rev. Entomol.* 61, 433–451. doi: 10.1146/annurev-ento-010715-023859

Sibilia, C. D., Brosko, K. A., Hickling, C. J., Thompson, L. M., Grayson, K. L., and Olson, J. R. (2018). Thermal physiology and developmental plasticity of pigmentation in the harlequin bug (Hemiptera: Pentatomidae). *J. Insect Sci.* 18:iey066. doi: 10.1093/jisesa/iey066

Simpson, S. J., Sword, G. A., and Lo, N. (2011). Polyphenism in insects. *Curr. Biol.* 21, 738–749. doi: 10.1016/j.cub.2011.06.006

Singh, P., Bergen, E., Van Brattström, O., Osbaldeston, D., and Paul, M. (2019). Complex multi-trait responses to multivariate environmental cues in a seasonal butterfly. *bioRxiv* [Preprint], doi: 10.1101/772749

Snell-Rood, E. C., Kobiela, M. E., Sikkink, K. L., and Shephard, A. M. (2018). Mechanisms of plastic rescue in novel environments. *Annu. Rev. Ecol. Evol. Syst.* 49, 331–354. doi: 10.1146/annurev-ecolsys-110617-062622

Song, H., Foquet, B., Mariño-Pérez, R., and Woller, D. A. (2017). Phylogeny of locusts and grasshoppers reveals complex evolution of density-dependent phenotypic plasticity. *Sci. Rep.* 7, 1–13. doi: 10.1038/s41598-017-07105-y

Sørensen, J. G., Dahlgaard, J., and Loeschcke, V. (2001). Genetic variation in thermal tolerance among natural populations of *Drosophila buzzatii*: down regulation of Hsp70 expression and variation in heat stress resistance traits. *Funct. Ecol.* 15, 289–296. doi: 10.1046/j.1365-2435.2001.00525.x

Sørensen, J. G., Giribets, M. P., Tarrío, R., Rodríguez-Trelles, F., Schou, M. F., and Loeschcke, V. (2019). Expression of thermal tolerance genes in two *Drosophila* species with different acclimation capacities. *J. Therm. Biol.* 84, 200–207. doi: 10.1016/j.jtherbio.2019.07.005

Sørensen, J. G., Schou, M. F., Kristensen, T. N., and Loeschcke, V. (2016). Thermal fluctuations affect the transcriptome through mechanisms independent of average temperature. *Sci. Rep.* 6, 1–11. doi: 10.1038/srep30975

Soroye, P., Newbold, T., and Kerr, J. (2020). Climate change contributes to widespread declines among bumble bees across continents. *Science* 367, 685–688. doi: 10.1126/science.aax8591

Stillman, J. H. (2003). Acclimation capacity underlies susceptibility to climate change. *Science* 301, 65. doi: 10.1126/science.1083073

Tobler, R., Franssen, S. U., Kofler, R., Orozco-Terwengel, P., Nolte, V., Hermisson, J., et al. (2014). Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. *Mol. Biol. Evol.* 31, 364–375. doi: 10.1093/molbev/mst205

Tobler, R., Hermisson, J., and Schlötterer, C. (2015). Parallel trait adaptation across valaopposing thermal environments in experimental *Drosophila melanogaster* populations. *Evolution* 69, 1745–1759. doi: 10.1111/evo.12705

Tomkins, J. L., Hazel, W. N., Penrose, M. A., Radwan, J. W., and Lebas, N. R. (2011). Habitat complexity drives experimental evolution of a conditionally expressed secondary sexual trait. *Curr. Biol.* 21, 569–573. doi: 10.1016/j.cub.2011.02.032

Umina, P. A., Weeks, A. R., Kearney, M. R., McKechnie, S. W., and Hoffmann, A. A. (2005). Evolution: a rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. *Science* 308, 691–693. doi: 10.1126/science.1109523

Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M. B., Balaguer, L., Benito-Garzón, M., et al. (2014). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecol. Lett.* 17, 1351–1364. doi: 10.1111/ele.12348

van Bergen, E., and Beldade, P. (2019). Seasonal plasticity in anti-predatory strategies: matching of color and color preference for effective crypsis. *Evol. Lett.* 3, 313–320. doi: 10.1002/evl3.113

van Bergen, E., Osbaldeston, D., Kodandaramaiah, U., Brattström, O., Aduse-Poku, K., and Brakefield, P. M. (2017). Conserved patterns of integrated developmental plasticity in a group of polyphenic tropical butterflies. *BMC Evol. Biol.* 17:59. doi: 10.1186/s12862-017-0907-901

van Heerwaarden, B., Lee, R. F. H., Overgaard, J., and Sgrò, C. M. (2014). No patterns in thermal plasticity along a latitudinal gradient in *Drosophila simulans* from eastern Australia. *J. Evol. Biol.* 27, 2541–2553. doi: 10.1111/jeb.12510

Verspagen, N., Suvi, I., Marjo, S., and Erik van, B. (2020). Multidimensional plasticity in the glanville fritillary butterfly: larval performance curves are temperature, host and family specific. *bioRxiv* [Preprint], doi: 10.1101/2020.05.05.065698

Wagner, D. L. (2020). Insect declines in the anthropocene. *Annu. Rev. Entomol.* 65, 457–480. doi: 10.1146/annurev-ento-011019-025151

Wang, S. P., and Althoff, D. M. (2019). Phenotypic plasticity facilitates initial colonization of a novel environment. *Evolution* 73, 303–316. doi: 10.1111/evo.13676

Wang, X. R., Wang, C., Ban, F. X., Zhu, D. T., Liu, S. S., and Wang, X. W. (2019). Genome-wide identification and characterization of HSP gene superfamily in whitefly (*Bemisia tabaci*) and expression profiling analysis under temperature stress. *Insect Sci.* 26, 44–57. doi: 10.1111/1744-7917.12505

Wasik, B. R., Liew, S. F., Lilien, D. A., Dinwiddie, A. J., Noh, H., Cao, H., et al. (2014). Artificial selection for structural color on butterfly wings and comparison with natural evolution. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12109–12114. doi: 10.1073/pnas.1402770111

Wernegreen, J. J. (2012). Mutualism meltdown in insects: bacteria constrain thermal adaptation. *Curr. Opin. Microbiol.* 15, 255–262. doi: 10.1016/j.mib.2012.02.001

Westerman, E., and Monteiro, A. (2016). Rearing temperature influences adult response to changes in mating status. *PLoS One* 11:e0146546. doi: 10.1371/journal.pone.0146546

Westneat, D. F., Potts, L. J., Sasser, K. L., and Shaffer, J. D. (2019). Causes and consequences of phenotypic plasticity in complex environments. *Trends Ecol Evol* 34, 555–568. doi: 10.1016/j.tree.2019.02.010

Wijngaarden, P. J., and Brakefield, P. M. (2001). Lack of response to artificial selection on the slope of reaction norms for seasonal polyphenism in the butterfly *Bicyclus anynana*. *Heredity* 87, 410–420. doi: 10.1046/j.1365-2540.2001.00933.x

Wijngaarden, P. J., Koch, P. B., and Brakefield, P. M. (2002). Artificial selection on the shape of reaction norms for eyespot size in the butterfly *Bicyclus anynana*: direct and correlated responses. *J. Evol. Biol.* 15, 290–300. doi: 10.1046/j.1420-9101.2002.00380.x

Williams, S. E., Shoo, L. P., Isaac, J. L., Hoffmann, A. A., and Langham, G. (2008). Towards an integrated framework for assessing the vulnerability of species to climate change. *PLoS Biol.* 6:e0060325. doi: 10.1371/journal.pbio.0060325

Windig, J. J., Brakefield, P. M., Reitsma, N., and Wilson, J. G. M. (1994). Seasonal polyphenism in the wild: survey of wing patterns in five species of Bicyclus butterflies in Malawi. *Ecol. Entomol.* 19, 285–298. doi: 10.1111/j.1365-2311.1994.tb00420.x

Woestmann, L., and Saastamoinen, M. (2016). The importance of trans-generational effects in lepidoptera. *Curr. Zool.* 62, 489–499. doi: 10.1093/cz/zow029

Yang, C. H., and Pospisilik, J. A. (2019). Polyphenism - A window into gene-environment interactions and phenotypic plasticity. *Front. Genet.* 10:132. doi: 10.3389/fgene.2019.00132

Yoshii, T., Vanin, S., Costa, R., and Helfrich-Förster, C. (2009). Synergic entrainment of *Drosophila's* circadian clock by light and temperature. *J. Biol. Rhythms* 24, 452–464. doi: 10.1177/0748730409348551

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 18 September 2020
doi: 10.3389/fgene.2020.538226

Genetic Constraints, Transcriptome Plasticity, and the Evolutionary Response to Climate Change

Michael L. Logan[1,2]* *and Christian L. Cox*[3]

[1] Department of Biology, University of Nevada, Reno, Reno, NV, United States, [2] Smithsonian Tropical Research Institute, Panama City, Panama, [3] Department of Biological Sciences and Institute of Environment, Florida International University, Miami, FL, United States

In situ adaptation to climate change will be critical for the persistence of many ectotherm species due to their relative lack of dispersal capacity. Climate change is causing increases in both the mean and the variance of environmental temperature, each of which may act as agents of selection on different traits. Importantly, these traits may not be heritable or have the capacity to evolve independently from one another. When genetic constraints prevent the "baseline" values of thermal performance traits from evolving rapidly, phenotypic plasticity driven by gene expression might become critical. We review the literature for evidence that thermal performance traits in ectotherms are heritable and have genetic architectures that permit their unconstrained evolution. Next, we examine the relationship between gene expression and both the magnitude and duration of thermal stress. Finally, we identify genes that are likely to be important for adaptation to a changing climate and determine whether they show patterns consistent with thermal adaptation. Although few studies have measured narrow-sense heritabilities of thermal performance traits, current evidence suggests that the end points of thermal reaction norms (tolerance limits) are moderately heritable and have the potential to evolve rapidly. By contrast, performance at intermediate temperatures has substantially lower evolutionary potential. Moreover, evolution in many species appears to be constrained by genetic correlations such that populations can adapt to either increases in mean temperature or temperature variability, but not both. Finally, many species have the capacity for plastic expression of the transcriptome in response to temperature shifts, with the number of differentially expressed genes increasing with the magnitude, but not the duration, of thermal stress. We use these observations to develop a conceptual model that describes the likely trajectory of genome evolution in response to changes in environmental temperature. Our results indicate that extreme weather events, rather than gradual increases in mean temperature, are more likely to drive genetic and phenotypic change in wild ectotherms.

Keywords: climate change, contemporary evolution, gene expression, heritability, molecular evolution, phenotypic plasticity, thermal adaptation, transcriptome

OPEN ACCESS

Edited by:
Torsten Nygaard Kristensen, Aalborg University, Denmark

Reviewed by:
Andrés Pérez-Figueroa, University of Porto, Portugal
Carla Sgro, Monash University, Australia

***Correspondence:**
Michael L. Logan
mike.logan1983@gmail.com

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *26 February 2020*
Accepted: *20 August 2020*
Published: *18 September 2020*

Citation:
Logan ML and Cox CL (2020) Genetic Constraints, Transcriptome Plasticity, and the Evolutionary Response to Climate Change. Front. Genet. 11:538226. doi: 10.3389/fgene.2020.538226

CLIMATE CHANGE AS AN AGENT OF SELECTION

The majority of species are dispersal-limited and must adapt to climate change *in situ* if they are to avoid extinction (Hoffmann and Sgro, 2011). The first response of many ectothermic animals will be to adjust their behavior to reduce exposure to stressful temperatures (Kearney et al., 2009; Logan et al., 2013, 2015; Cox et al., 2018; Fey et al., 2019). Nevertheless, behavioral adjustments on their own may be insufficient to maintain fitness, requiring populations to track shifting fitness optima through genetic adaptation and phenotypic plasticity (Berger et al., 2013; Logan et al., 2014, 2019; Buckley et al., 2015; Geerts et al., 2015). A major question that remains is whether populations have heritable variation in climate-related traits such that they may adapt to environmental change over short time scales (Leal and Gunderson, 2012; Walters et al., 2012).

Historical data and climate forecasts suggest that shifts in environmental temperature associated with climate change has occurred (and will continue to occur) along two distinct axes (Alley, 2000; IPCC, 2013). First, mean temperature is increasing, primarily as a result of days and seasons that are gradually warming (**Figure 1A**). Second, the variance of environmental temperature is increasing, primarily because of a rise in the frequency of extreme weather events such as heat waves and cold snaps (**Figure 1B**). These two axes of thermal change are likely to generate selection on different components of thermal reaction norms (Gabriel and Lynch, 1992; Gilchrist, 1995; Angilletta, 2009). For example, gradual increases in mean temperature will favor genotypes that confer higher thermal optima for ecologically important activities (e.g., genotypes associated with the ability to digest food more effectively at warmer temperatures; Fontaine et al., 2018). By contrast, increases in temperature variability will favor genotypes that boost phenotypic plasticity or whose fitness values are insensitive to temperature (Lynch and Gabriel, 1987; Gabriel and Lynch, 1992).

THE EVOLUTIONARY POTENTIAL OF THE THERMAL NICHE

While theory indicates that the mean or variance of environmental temperature should select for changes in different thermal performance traits, these traits will not evolve unless they are heritable and unconstrained by genetic correlations (Lande and Arnold, 1983; Lynch and Walsh, 1998). In practice, the thermal niche of a given population is usually approximated with a "thermal performance curve" (TPC; **Figure 2A**, inset). TPCs relate a fitness-proxy (usually an ecologically relevant trait such as locomotor performance) to body temperature (Huey and Hertz, 1984), and often follow an archetypical shape whereby performance increases with body temperature to some optimum (T_{opt}) and then sharply declines above that optimum (a pattern driven by the thermodynamics of enzyme function; Hochachka and Somero, 2002). The thermal optimum is expected to be under selection primarily as a result of gradually increasing mean temperatures (Logan et al., 2014). The ends of the TPC (where performance drops to zero) are referred to as the critical thermal limits (critical thermal minimum = CT_{min}; critical thermal maximum = CT_{max}), and these are closely related to the breadth of the TPC (T_{br}). The performance breadth and critical thermal limits are thought to be under selection primarily as a result of changes in the variance of environmental temperature, although performance breadth is probably also affected by selection for changes in performance at intermediate temperatures (Lynch and Gabriel, 1987; Gabriel and Lynch, 1992; Logan et al., 2014). Finally, the height of the TPC describes the maximal performance capacity (P_{max}) of the population. These five components of thermal performance curves can be thought of as "thermal performance traits" that combine to define the shape of the thermal niche and may or may not have the capacity to evolve independently of one another (Gomulkiewicz and Kirkpatrick, 1992; Stinchcombe and Kirkpatrick, 2012; Martins et al., 2018; Logan et al., 2020).

Indeed, studies of thermal performance curves across environmental gradients suggest that their shapes may be constrained (Knies et al., 2009; Angilletta et al., 2010; Logan et al., 2013; Phillips et al., 2014). For example, when measured at the level of the phenotype, the area under the curve tends to remain constant even as the shape of the curve changes (Gilchrist, 1996; Kingsolver and Gomulkiewicz, 2003; Izem and Kingsolver, 2005; Phillips et al., 2014). This represents a "specialist-generalist tradeoff" whereby a species can either perform well over a narrow range of temperatures or poorly over a broad range of temperatures (**Figure 1C**). Specialist-generalist tradeoffs arise from the inability of organisms to optimize biochemical performance across a broad range of temperatures at the subcellular level and often manifest as a negative correlation between whole-organism performance breadth and maximal performance (or as a positive correlation between the critical thermal limits; Hochachka and Somero, 2002). Another pattern commonly observed at the phenotypic level is the "thermodynamic effect" (also referred to as the "hotter-is-better" hypothesis; Angilletta et al., 2010). This effect occurs because biochemical reactions are typically more efficient at warmer temperatures (Hochachka and Somero, 2002), and leads to a positive correlation between the thermal optimum and maximal performance at the whole-organism level (**Figure 1D**).

If both the specialist-generalist tradeoff and the thermodynamic effect are driven by underlying genetic correlations and occur in the same populations, they represent true evolutionary constraints that can give rise to non-intuitive evolutionary dynamics depending on whether average thermal conditions or extreme weather events are more important sources of selection. For example, if the mean environmental temperature changes faster than the variance, selection should first favor an increase in the thermal optimum, which should then indirectly cause an increase in maximal performance via the thermodynamic effect. This increase in maximal performance should then drive a decrease in performance breadth as a result of a specialist-generalist tradeoff. Thus, adaptation to higher mean temperature can lead to maladaptation with respect to temperature variability (**Figure 1E**). Alternatively, if the

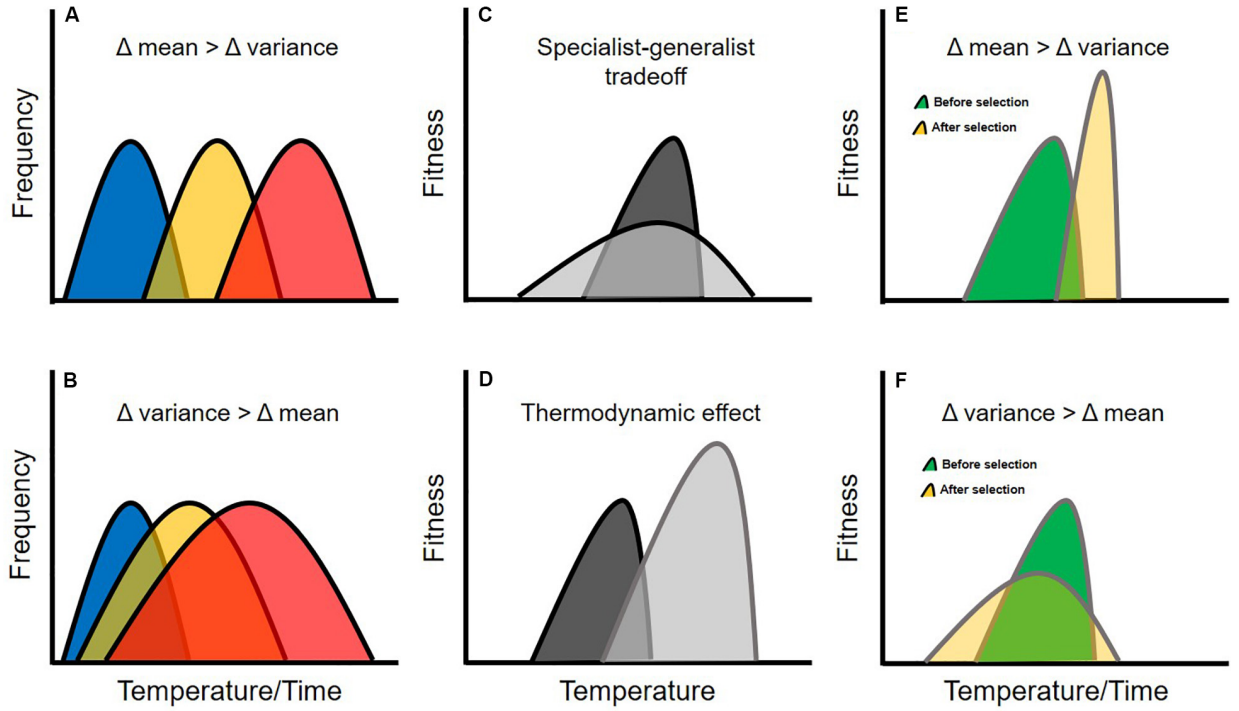


FIGURE 1 | Genetic correlations can constrain the evolution of thermal performance curves, and evolutionary trajectories will likely depend on the specific nature of environmental change. For example, different moments of the environmental temperature distribution can change at different rates, with mean temperature increasing faster than the variance **(A)**, or vise-versa **(B)**. Traits may be constrained in their evolution via a specialist-generalist tradeoff **(C)** which occurs when maximal performance is negatively genetically correlated with performance breadth, or a thermodynamic effect **(D)**, which occurs when maximal performance is positively genetically correlated with the thermal optimum. If these evolutionary constraints occur in the same population, complex evolutionary dynamics can result from selection on thermal performance traits. For example, if mean environmental temperature increases faster than the variance **(E)**, selection should favor an increase in the thermal optimum, with maximal performance also increasing as an indirect result of the thermodynamic effect. As maximal performance increases, performance breadth should then decline as an indirect result of a specialist-generalist tradeoff. Thus, the population becomes well-adapted to mean temperature and maladapted to temperature variability. If the variance of environmental temperature increases faster than the mean **(F)**, selection should favor an increase in performance breadth, with maximal performance decreasing as an indirect result of a specialist-generalist tradeoff. As maximal performance decreases, the thermal optimum should then decline as an indirect result of the thermodynamic effect. Thus, the population becomes well-adapted to temperature variability and maladapted to mean temperature. The colors of the curves in this figure are arbitrary and meant to help increase readability.

variance in environmental temperature increases faster than the mean, selection should first favor an increase in performance breadth which should indirectly cause a decrease in maximal performance as a result of a specialist-generalist tradeoff. This decrease in maximal performance would then result in a decline in the thermal optimum due to the thermodynamic effect. In this case, adaptation to temperature variability will lead to maladaptation with respect to mean temperature (**Figure 1F**). Clearly, understanding the extent to which thermal niche evolution is constrained by genetic correlations is critical for generating accurate climate-impact forecasts.

To understand genetic constraints underlying the evolution of the thermal niche, we canvassed the literature for primary, peer-reviewed studies reporting heritabilities (broad and narrow-sense) and genetic correlations underlying the thermal performance traits that make up the thermal niches of animals. We searched the terms "quantitative genetics AND thermal physiology," "genetic correlations AND thermal physiology," "heritability AND thermal trait," "genetics AND specialist-generalist AND temperature," "genetics AND hotter-is-better," "genetics AND thermodynamic effect," "heritability of CTmax," "heritability of CTmin," "heritability of thermal optimum," "heritability AND cold tolerance," and "heritability AND chill-coma" in Google Scholar in October 2019. Due to the rapid decline of relevant studies after the first few pages of search results, we focused on the first 50 results for each set of search terms (ordered by relevance). To ensure that our sampling was robust, we subsequently (July 2020) included an additional 50 search results on Google Scholar (total = 100 results per search) and conducted a separate set of searches with the same search terms in Thompson Web of Science, again ordered by relevance. In total, we examined more than 1400 results from these databases for possible heritability and genetic correlation estimates. Finally, we included additional studies that we were aware of but that did not come up in our literature searches. These various search avenues likely uncovered the majority of quantitative genetic parameter estimates for our target traits that were available in the literature. Our full database contained 98 independent heritability and genetic correlation estimates from 55 studies. Note that the temperature ramping rates used in

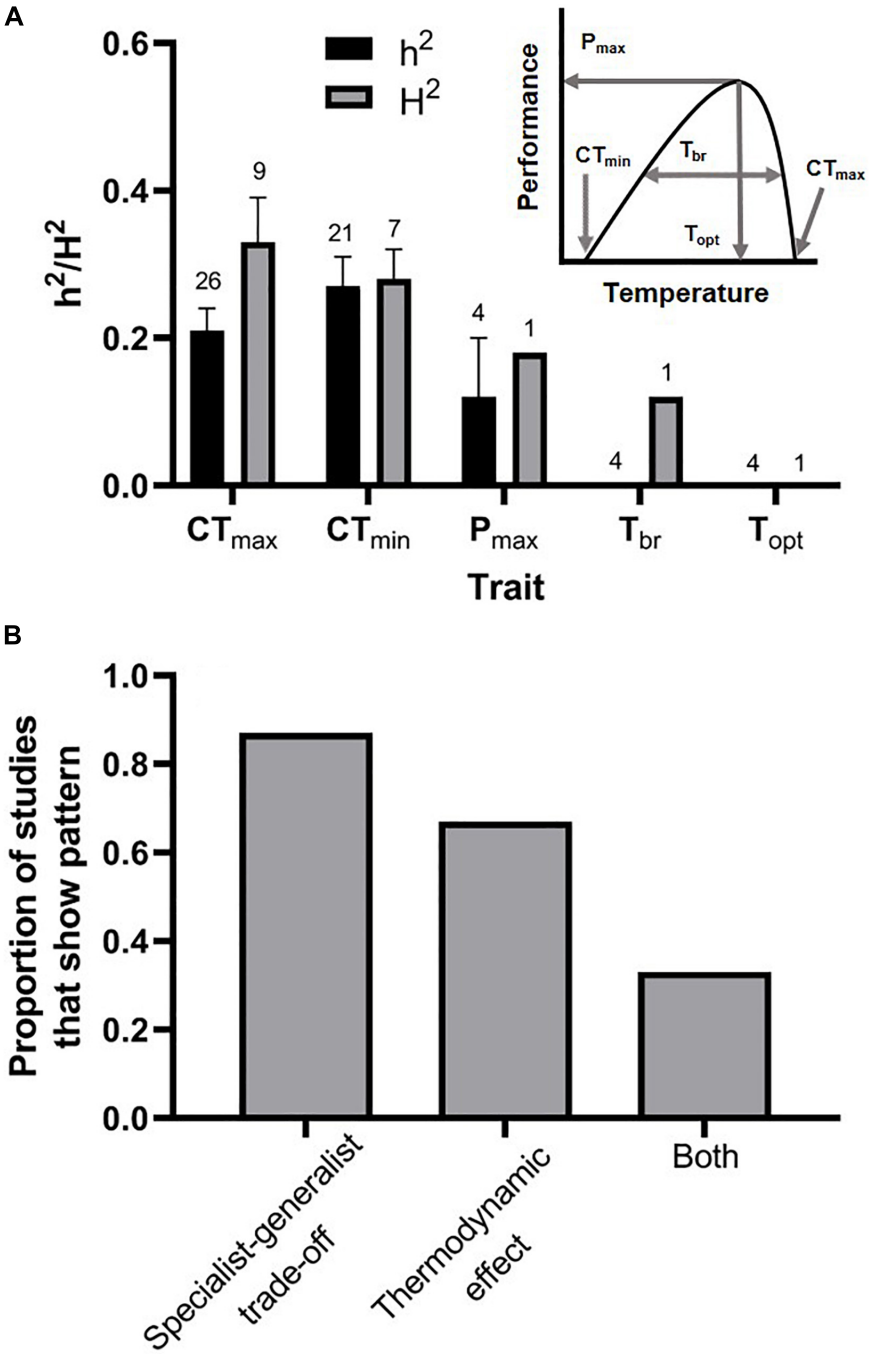


FIGURE 2 | Patterns of genetic constraint on the thermal performance traits that combine to define the shape of thermal performance curves (**A**, inset). **(A)** The critical thermal limits (CT_{min} and CT_{max}) are moderately heritable (h^2 = narrow-sense heritability; H^2 = broad-sense heritability), whereas the limited evidence that is available suggests that performance breadth (T_{br}) and the thermal optimum (T_{opt}) lack the capacity to respond rapidly to selection (error bars represent standard errors, and samples sizes are above each bar). **(B)** The majority of studies that tested for either a specialist-generalist tradeoff or a thermodynamic effect underlying the evolution of thermal performance curves found evidence for either one or the other pattern. Two of six studies that tested for both types of constraints in the same population found evidence suggesting that both were operating. **Supplementary Table S1** contains the list of studies from which we extracted the values included in this figure.

these studies varied by several orders of magnitude, and ramping rate is known to affect heritability estimates (Terblanche et al., 2007; Chown et al., 2009). Namely, slow-ramping protocols tend to produce lower heritability estimates, and there is evidence from simulation studies that this may be due to error introduced during longer ramping procedures (Rezende et al., 2011; Santos

et al., 2011, 2012). Thus, while most of the heritability estimates included in our analyses were taken from ramping protocols of some kind, when the heritability from both "slow" and "fast" ramping protocols were reported for the same population or species (total of four studies), we only included the latter in our analyses to eliminate pseudoreplication and reduce error as much as possible. For a detailed explanation of how we collated and assessed data from these papers, please see "Extended Methods" in the Online **Supplementary Information**. We have uploaded the full list of studies included in our analyses of trait heritability and genetic correlations in an online supplementary data file (**Supplementary Table S1**).

Of the five thermal performance traits that define the shape of the thermal performance curve (**Figure 2A**, inset), only the critical thermal limits (CT_{min} and CT_{max}) were consistently and substantially heritable (**Figure 2A**). The average broad and narrow-sense heritabilities of CT_{min} were 0.27 and 0.28, respectively. The average broad and narrow-sense heritabilities of CT_{max} were 0.33 and 0.21, respectively. It is interesting to note that phylogenetic studies on some taxa have led to the conclusion that upper thermal limits, but not lower thermal limits, are evolutionarily conserved (Araújo et al., 2013; Grigg and Buckley, 2013; Diamond and Chick, 2017), and this appears to conflict with the relatively high heritability of upper thermal limits observed in controlled breeding studies. The resolution of this conflict may arise from the fact that many species behaviorally thermoregulate during the hottest times of the day or during heat waves, leading to a reduction in the strength of selection on upper thermal tolerance (Muñoz et al., 2014). Thus, even though upper thermal tolerance may be infrequently exposed to selection, this trait may retain its ability to respond to selection in many populations. Indeed, laboratory evolution experiments that expose organisms to selection in warmer environments frequently demonstrate rapid evolutionary change in upper thermal limits (Bettencourt et al., 1999; Gilchrist and Huey, 1999; Sambucetti et al., 2010; Hangartner and Hoffmann, 2016; but see Schou et al., 2014).

To our knowledge, there are only five estimates (from four studies) of the quantitative genetic parameters underlying the other major thermal performance traits: maximal performance, performance breadth, and the thermal optimum. Maximal performance was moderately heritable at an average narrow-sense heritability of 0.12. Every study that examined the performance breadth and the thermal optimum found zero additive genetic variation underlying these traits. Due to the low sample sizes for most of these traits, we did not conduct formal statistical comparisons. Of the studies ($N = 15$) that tested for genetic correlations corresponding to either a specialist-generalist tradeoff or a thermodynamic effect, the majority found evidence of one or the other. 87% of studies found evidence of a specialist-generalist tradeoff, while 67% of studies found evidence of a thermodynamic effect (**Figure 2B**). Additionally, of the six studies that tested for both a specialist-generalist tradeoff and thermodynamic effect in the same population, two of those studies detected both patterns (**Figure 2B**). All else remaining equal, these results suggest that the endpoints of the thermal niche (the critical thermal limits) can respond relatively rapidly to selection, although they are likely constrained to some extent by genetic correlations. By contrast, the traits which describe performance at intermediate temperatures (e.g., T_{opt}) appear to have minimal capacity for rapid evolution.

GENE EXPRESSION PLASTICITY

For most organisms, thermal performance traits are not fixed across environmental conditions, but instead can exhibit adaptive or non-adaptive phenotypic plasticity (Scheiner, 1993; Via et al., 1995; Ghalambor et al., 2007, 2015). For example, previous exposure to cool temperatures reduced the recovery time after induction of chill- coma in fruit flies (*Drosophila melanogaster*) compared to flies reared at intermediate temperatures (Ayrinhac et al., 2004). Similarly, acclimation to warmer temperatures increased time to immobilization (a measure of heat tolerance) in the freshwater crustacean *Daphnia magna* (Yampolsky et al., 2014a).

The mechanism driving most phenotypic plasticity is changes in gene expression (Scheiner, 1993; Schlichting and Pigliucci, 1993; Schlichting and Smith, 2002; Chen et al., 2017). Shifts in gene expression can involve only a few genes (Hamdoun et al., 2003), or can occur across the entire transcriptome (Bay and Palumbi, 2015). For example, shifts in the expression of genes in the heat-shock protein (*hsp*) 70 family seem to underlie phenotypic plasticity in thermal tolerance limits in the oyster *Crassostrea gigas* (Hamdoun et al., 2003), whereas exposure to warm temperatures was associated with alterations of whole-transcriptome expression and increased heat tolerance in the coral *Acropora nana* (Bay and Palumbi, 2015). Broadly, this suggests that phenotypic plasticity, mediated by gene expression, is important for the adaptive response to global climate change.

To understand how gene expression might be involved in the response to climate change, we canvassed the literature for studies that measured transcriptomic responses to thermal stress in ectothermic animals. We searched the terms "transcriptome heat stress," "transcriptome expression temperature vertebrate," "gene expression heat vertebrate," "transcriptome expression thermal," "transcriptome thermal," and "gene expression thermal" in Google Scholar during October 2019. We conducted a subsequent, deeper search (100 results for each set of search terms) in both Google Scholar and Thompson Web of Science during July 2020. These queries returned hundreds of journal articles, each of which we evaluated for relevance. Ultimately, this process yielded 36 articles containing 42 independent estimates of the effects of temperature on the transcriptomic response in ectotherms. These studies spanned early microarray work to recent experiments that leveraged high-throughput RNA sequencing, and they focused on acute, reversible gene expression responses rather than fixed changes that may occur over development (**Table 1**).

All species in these experiments, which range from arthropods to vertebrates and occur in diverse habitats across the globe, shift expression of their transcriptome in response to thermal changes (**Table 1**). However, the temperature changes experienced by organisms in these studies varied greatly in

TABLE 1 | Studies of transcriptomic responses to temperature change.

Study	Organism	Species	Data type	Change in transcriptome expression?	Altered *hsp* expression?
Akashi et al. (2016)	Lizard	*Anolis allogus*	RNAseq	Y	Y
Akashi et al. (2016)	Lizard	*Anolis homolechis*	RNAseq	Y	Y
Akashi et al. (2016)	Lizard	*Anolis sagrei*	RNAseq	Y	Y
Coughlin et al. (2019)	Fish	*Osmerus mordax*	RNASEq	Y	Y
Cui et al. (2019)	Insect	*Megacopta cribaria*	RNAseq	Y	Y
Etges et al. (2017)	Insect	*Drosophila mojavensis*	RNAseq	Y	Y
Gleason and Burton (2015)	Mollusc	*Chlorostoma funebralis*	RNAseq	Y	Y
Gracey et al. (2004)	Fish	*Cyprinus carpio*	Microarray	Y	Y
Hu et al. (2016)	Fish	*Danio rerio*	RNAseq	Y	N/A
Hu et al. (2016)	Fish	*Oreochromis niloticus*	RNAseq	Y	N/A
Jayasundara et al. (2013)	Fish	*Thunnus orientalis*	Microarray	Y	Y
Jesus et al. (2016)	Fish	*Squalius carolitertii*	RNAseq	Y	Y
Jesus et al. (2016)	Fish	*Squalius torgalensis*	RNAseq	Y	Y
Kassahn et al. (2007)	Fish	*Pomacentrus moluccensis*	Microarray	Y	Y
Kim et al. (2017)	Mollusc	*Crassostrea gigas*	RNAseq	Y	Y
Lewis et al. (2010)	Fish	*Onchorhyncus mykiss*	Microarray	Y	Y
Li et al. (2017)	Fish	*Onchorhyncus mykiss*	RNAseq	Y	Y
Li et al. (2019)	Fish	*Megalobroma amblycephala*	RNAseq	Y	Y
Lim et al. (2016)	Mollusc	*Crassostrea gigas*	RNAseq	Y	Y
Liu et al. (2013)	Fish	*Ictalurus* hybrids	RNAseq	Y	N/A
Lockwood et al. (2010)	Mollusc	*Mytilus trossulus*	Microarray	Y	Y
Lockwood et al. (2010)	Mollusc	*Mytilus galloprovincialis*	Microarray	Y	Y
Logan and Somero (2011)	Fish	*Gillichthys mirabilis*	Microarray	Y	Y
Moskalev et al. (2015)	Insect	*Drosophila melanogaster*	RNAseq	Y	N/A
Moya et al. (2012)	Cnidarian	*Anemonia viridis*	Microarray	Y	Y
Narum and Campbell (2015)	Fish	*Oncorhynchus mykiss*	RNAseq	Y	Y
Qian and Xue (2016)	Fish	*Larimichthys crocea*	RNAseq	Y	Y
Quinn et al. (2011)	Fish	*Salvelinus alpinus*	Microarray	Y	Y
Semmouri et al. (2019)	Crustacean	*Temora longicornis*	RNAseq	Y	Y
Shi et al. (2019)	Fish	*Salmo salar*	RNAseq	Y	Y
Smith et al. (2013)	Fish	*Melanotaenia duboulayi*	RNAseq	Y	Y
Smolina et al. (2015)	Crustacean	*Calanus finmarchius*	RNAseq	Y	Y
Smolina et al. (2015)	Crustacean	*Calanus glacialis*	RNAseq	Y	N/A
Sørensen et al. (2016)	Insect	*Drosophila melanogaster*	RNAseq	Y	Y
Stillman and Tagmount (2009)	Crustacean	*Petrolisthes cinctipes*	Microarray	Y	Y
Vornanen et al. (2005)	Fish	*Onchorhyncus mykiss*	Microarray	Y	Y
Wang et al. (2014)	Mollusc	*Echinolittoria malacaria*	RNAseq	Y	Y
Wellenreuther et al. (2019)	Fish	*Chrysophus auratus*	RNAseq	Y	Y
Xiao et al. (2016)	Spider	*Pardosa pseudoannulata*	RNAseq	Y	Y
Yampolsky et al. (2014b)	Crustacean	*Daphnia* sp.	Microarray	Y	N/A
Yang et al. (2016)	Fish	*Ctenopharyngodon idellus*	RNAseq	Y	Y
Zheng et al. (2019)	Crustacean	*Marsupaenus japonicus*	RNASeq	Y	Y

Irrespective of the methodology or focal taxon, all experiments detected changes in gene expression when the organism was exposed to a change in temperature. Additionally, all studies that reported analyses of heat-shock protein (hsp) genes detected shifts in the expression of these genes.

their magnitude and duration. When restricting the analyses to RNAseq studies and excluding whole-organism studies (**Supplementary Table S2**), we found that the magnitude of temperature change [$F_{(3,\ 33)}$ = 13.0448, P = 0.0010; **Figure 3A**], but not the duration of exposure [$F_{(3,\ 33)}$ = 2.1269, P = 0.1542: **Figure 3B**] predicted the number of log-transformed differentially expressed genes when controlling for log-transformed transcriptome size [$F_{(3,33)}$ = 3.3718, P = 0.0753] using linear regression models. These results indicate that brief, severe weather events could impact gene expression and phenotypic plasticity more profoundly than longer-term changes in thermal conditions.

A previous study by Gunderson and Stillman (2015) reported limited potential for plastic responses to warming across a broad

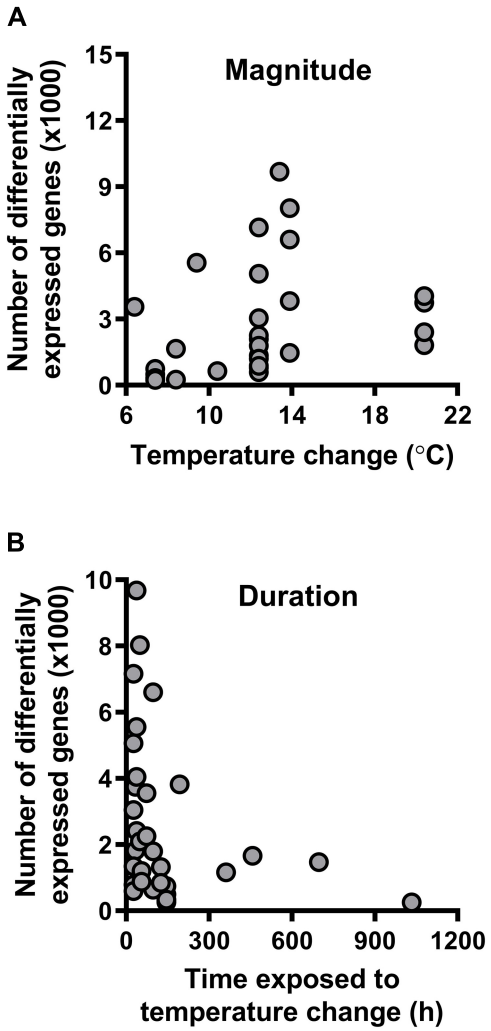


FIGURE 3 | Transcriptomic responses to temperature change. **(A)** The number of differentially expressed genes increases with the magnitude of temperature change to which an organism is exposed. **(B)** The number of differentially expressed genes does not vary with duration of exposure to a given magnitude of temperature change. See **Table 1** for the list of studies from which we extracted the values included in this figure.

range of organisms. This suggests that there may not be a one-to-one correspondence between the magnitude of gene expression shifts and physiological plasticity in the typical whole-organism traits that investigators measure (e.g., CT_{max} and CT_{min}). Indeed, not all mRNAs that are transcribed will be translated into proteins (Liu et al., 2016), possibly leading to a discordance between the magnitude of gene expression plasticity and phenotypic plasticity. Additionally, many of the studies cited in Gunderson and Stillman (2015) involved ramping or constant-exposure thermal stress experiments, which may be less likely to result in large-scale changes in gene expression. In general, further work is needed to understand the link between gene expression plasticity under large magnitude shifts in temperature and phenotypic plasticity in thermal tolerance limits.

Among the genes that were differentially expressed in response to temperature, gene ontology (and similar) analyses have found that biological processes associated with protein synthesis, folding and degradation, oxygen transport, and biological and cellular responses to heat and other stress-stimuli are often significantly enriched (**Supplementary Table S3**). Heat shock proteins, which are a conserved set of molecular chaperone proteins with important roles for responding to stress in general, and heat stress in particular (Feder and Hofmann, 1999), were especially important. Genes for heat-shock proteins were frequently (94% of species, **Table 1**) affected by changes in temperature, with shifts in expression often occurring in well-characterized canonical genes such as hsp40 (or DNAJ), hsp70, and hsp90 (**Supplementary Table S4**). Because expression of heat shock proteins is usually altered in response to changing temperature and has been linked to phenotypic plasticity (Hamdoun et al., 2003), these proteins are likely to be important targets of selection as global climate change progresses.

HOW WILL GENOMES RESPOND TO SELECTION WHEN THERMAL ENVIRONMENTS SHIFT?

Rapid environmental change can induce selection on the genome in two major ways. First, selection can target sequence variation in crucial protein-coding genes (Hoekstra et al., 2004; Rosenblum et al., 2010). This is most likely to occur when the capacity for gene expression plasticity is minimal or under weak selection and may manifest as changes in loci that affect the "baseline" values of thermal performance traits like the thermal optimum or the critical thermal limits. Second, if variation among individuals in gene expression plasticity is high or under strong selection, the primary adaptive response to a changing climate may be shifts in loci that are associated with variation in gene expression (Behera and Nanjundiah, 1995; Ghalambor et al., 2015; Campbell-Staton et al., 2020). Selection on gene expression could target trans-regulatory pathways or the upstream and downstream cis-regulatory regions that affect expression of individual genes (Schlichting and Pigliucci, 1993; Via, 1993; Campbell-Staton et al., 2020), and is likely to increase the frequency of genotypes with broad phenotypic reaction norms. Alternatively, selection could target genes that regulate epigenetic mechanisms such as histone modification or methylation (Johannes et al., 2009; Furrow and Feldman, 2014).

Our review of the literature suggests several pathways by which shifts in environmental temperature distributions should impact genomic variation (**Figure 4**). To date, studies suggest that the endpoints of the thermal niche (the critical thermal limits) are heritable, whereas performance at intermediate temperatures (e.g., T_{opt}) are not (**Figure 2**). The critical thermal limits are most important under extreme weather conditions such as heat waves and cold snaps (Campbell-Staton et al., 2017), indicating that baseline genetic variation for thermal performance may be more capable of responding to these extreme events than to gradual changes in mean temperature (although adaptation to extreme weather events may still be constrained by genetic

correlations; **Figures 1**, **2**). Similarly, most species appear to alter gene expression when they are exposed to short-term shifts in temperature (**Table 1**), and the number of genes that are differentially expressed increases with the magnitude of the temperature shift (akin to a short-term extreme weather event; **Figure 3A**). In contrast, the number of differentially expressed genes did not vary with duration of exposure to these temperature shifts (**Figure 3B**). This pattern may reflect a reduced importance of gene expression plasticity when environmental change is dominated by longer-term increases in mean temperature. Taken together, these data suggest that genomic responses will be more rapid and pronounced in response to changes in the frequency of extreme weather events than in response to gradual warming.

Emerging patterns from genomic and transcriptomic studies also suggest that the specific nature of environmental change will be important for determining trajectories of molecular evolution (**Figure 4**). As environmental temperature distributions change, different moments of the distribution can shift at different rates. Because the mean and variance of environmental temperatures drive selection on separate traits that have varying levels of additive genetic variation underlying them, we would expect "baseline" genetic adaptation and changes in gene expression plasticity to make up different components of the adaptive response depending on the details of environmental change (**Figure 4A**). If both the change in mean temperature and the change in variance are low, then selection will be weak or non-existent on all traits (compensatory responses might be entirely behavioral, for example), leading to zero molecular and phenotypic evolution. However, if the change in mean temperature is higher than the change in variance, we would expect shifts in alleles underlying variation in gene expression instead of shifts in alleles underlying "baseline" values of thermal traits. This is because, even though selection favoring better performance at intermediate temperatures should be high (**Figure 4B**), we would predict minimal evolution (**Figure 4C**) since the relevant traits (e.g., T_{opt}) appear to lack additive genetic variation (**Figure 2A**).

Patterns of molecular and phenotypic evolution should be different if the variance of environmental temperature changes faster than the mean. There appears to be substantial genetic variation in both the critical thermal limits and the gene expression response to thermal stress within populations (**Figures 2**, **3** and **Table 1**). Thus, selection for performance at extreme temperatures should favor loci that correspond to high and low baseline values of CT_{max} and CT_{min}, respectively (**Figures 4A,D**). Selection for increased gene expression plasticity should also increase in strength as extreme weather events become more common (**Figure 4D**), but the rate of change in plasticity should decline as baseline trait values become locally adapted (**Figure 4E**).

FUTURE RESEARCH DIRECTIONS

The literature on the genomic and transcriptomic basis of thermal adaptation hints at multiple potential evolutionary outcomes depending on the nature of environmental change. Nevertheless, these observations should be considered preliminary, as comparatively few studies have investigated the quantitative genetic basis of full thermal performance curves. Thus, estimates of heritabilities and genetic correlations underlying performance at intermediate temperatures (temperatures at or close to the thermal optimum) are exceedingly rare. To our knowledge, only three studies have estimated narrow-sense heritabilities of the thermal optimum and performance breadth. Two of these were on lizards (Logan et al., 2018; Martins et al., 2018) and the third was on an invasive population of harlequin beetles (Logan et al., 2020). A fourth study reported broad-sense heritabilities of the performance breadth and the thermal optimum in parasitoid wasps (Gilchrist, 1996). Although a general pattern of low genetic variation in these traits is starting to emerge from this research, we need many more studies of the quantitative genetics of full thermal performance curves to understand whether performance at intermediate temperatures truly lacks rapid evolutionary potential, or whether the patterns we report here are an artifact of insufficient sampling.

Most studies have examined genetic variation in either the baseline values of thermal traits or their plasticity, but rarely both. Future work should focus on the genetic basis of baseline values of thermal traits and their plasticity in the same populations to tease apart the independent contribution of both to local adaptation under environmental change. A rare example of such a study is Gerken et al. (2015), who assessed the heritability and genomic basis of both basal cold tolerance and its plasticity in laboratory lines of fruit flies. They found that baseline thermal tolerance was genetically correlated with its plasticity, implying that adaptation is constrained when both the mean and variance of temperature are increasing.

Our review suggests that genes in the heat shock protein family are a likely target for selection when environments first shift, and the evolutionary potential of these genes may be a major determinant of populations' resilience in the face of climate change. Past evolution of heat shock proteins is dominated by repeated duplications and insertion events, which might have been followed by neofunctionalization (Waters, 1995; Franck et al., 2004; Yamashita et al., 2004; Huang et al., 2008). At least in some contexts, there is evidence of directional selection on heat-shock proteins (Bettencourt et al., 2002; Fares et al., 2002). However, we do not know whether selection acts primarily on the coding sequences of these genes or on their upstream and downstream regulatory regions. Future work should determine the level of functional sequence variation underlying this family of genes in wild populations, and the relationship between heat-shock protein evolution and population mean fitness.

Our results suggest that the evolution of gene expression plasticity may be particularly important in maintaining fitness under climate change, not only because a number of thermal traits appear to lack genetic variation in their baseline values, but also because extreme weather events are rising in frequency. Moreover, past research has revealed that the capacity for gene expression plasticity can be heritable and evolve rapidly (Gerken et al., 2015; Leder et al., 2015). Additionally, variation in plasticity that is not genetic may persist across generations due

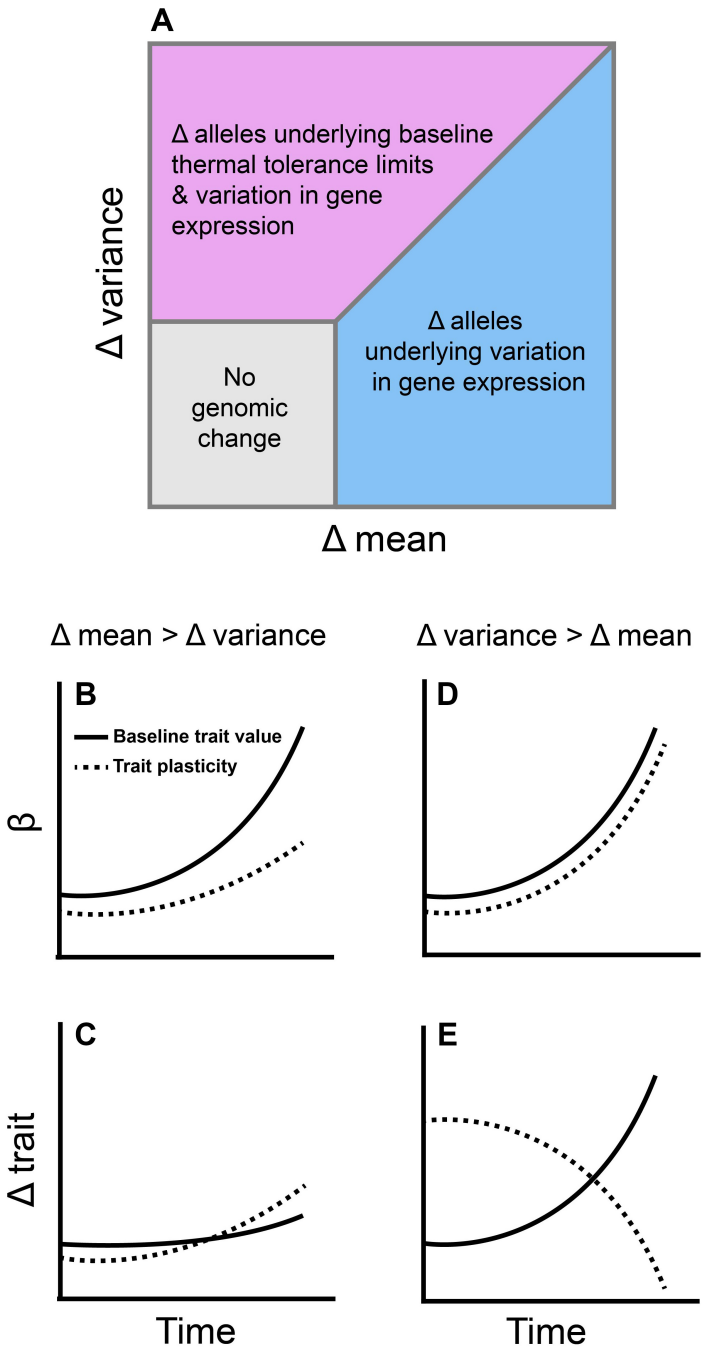


FIGURE 4 | The rates at which different moments of the environmental temperature distribution change are likely to impact observed patterns of genomic and phenotypic evolution. **(A)** Based on patterns of genetic variation reported in the literature, alleles underlying variation in gene expression (blue and purple regions) are more likely to change in frequency during environmental change than alleles underlying baseline thermal tolerance. Only in cases where the change in the variance of temperature is equal to or higher than the change in mean temperature should alleles associated with baseline trait values shift substantially (purple region). **(B)** If mean environmental temperature changes faster than the variance, selection (β) should favor an increase in baseline values of traits like the thermal optimum, while selection for increased plasticity should gradually rise as baseline values fail to evolve due to genetic constraints. **(C)** When mean environmental temperature increases faster than the variance, traits that correspond to performance at intermediate temperatures (such as the thermal optimum) should evolve slowly while plasticity increases to compensate for the lack of heritability in the baseline values of these traits. **(D)** When the variance of environmental temperature increases faster than the mean, selection should favor an increase in both the baseline values of traits which correspond to performance at extreme temperatures (e.g., the critical thermal limits) and the plasticity of such traits. **(E)** Because the critical thermal limits are heritable in most species, they should evolve in response to selection. This should lead to a reduction in the rate of change in plasticity as baseline trait values become locally adapted. Note that this assumes more genetic variation is initially present in baseline thermal tolerance than in its plasticity. The dashed and solid lines in **(E)** would be flipped if there was more genetic variation underlying the plasticity of thermal tolerance than in their baseline values.

to epigenetic mechanisms and can be important for population persistence in the initial stages of environmental change (Geng et al., 2013; Schlichting and Wund, 2014). Despite growing evidence that the evolution of phenotypic plasticity may be critical for organismal responses to climate change, it is still unclear how selection on plasticity is manifested at the level of the genome. Related questions that should be addressed by future research include 1) If extreme weather events select for higher gene expression plasticity, should we expect fast changes in regulatory regions of the genome, non-coding regions, or both? 2) Does selection for increased phenotypic plasticity constrain the evolution of baseline thermal tolerance (or vice versa)? Additional studies of within-population variation in baseline thermal tolerance and plasticity, and the genetic loci associated with each, are sorely needed.

CONCLUSION

Our review suggests that several general rules may be emerging from studies of the genetic and transcriptomic basis of thermal performance:

1. In many species, there is more genetic variation in performance at extremely high or low temperatures than in performance at intermediate temperatures.
2. Gene expression plasticity is rampant when organisms are exposed to acute thermal stress.
3. Patterns (1) and (2) indicate that populations are more likely to evolve rapidly in response to extreme weather events than in response to gradual changes in mean temperature, and the rate at which different moments of the temperature distribution change will determine the dominant trajectory of phenotypic and genetic evolution.
4. Gene regulatory networks linked to heat shock proteins are likely to be a major target of selection as environmental temperatures become warmer and more variable.

Finally, our work highlights the need for further studies on the quantitative genetic basis of thermal performance curves and the interactions between baseline thermal tolerance and gene expression plasticity. Continued advances in this field should lead to substantial improvements in our ability to predict the viability of animal populations as our planet continues to change.

AUTHOR CONTRIBUTIONS

Both authors conceived of the study, reviewed the literature, and wrote the manuscript.

ACKNOWLEDGMENTS

We thank the EvolDoers reading group at the University of Nevada, Reno for comments which helped us improve the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.538226/full#supplementary-material

REFERENCES

Akashi, H. D., Cádiz Díaz, A., Shigenobu, S., Makino, T., and Kawata, M. (2016). Differentially expressed genes associated with adaptation to different thermal environments in three sympatric Cuban Anolis lizards. *Mol. Ecol.* 25, 2273–2285. doi: 10.1111/mec.13625

Alley, R. B. (2000). Ice-core evidence of abrupt climate changes. *Proc. Natl. Acad. Sci. U.S.A.* 97:1331. doi: 10.1073/pnas.97.4.1331

Angilletta, M. J. (2009). *Thermal Adaptation*. Oxford: Oxford University Press.

Angilletta, M. J., Huey, R. B., and Frazier, M. R. (2010). Thermodynamic effects on organismal performance: is hotter better? *Physiol. Biochem. Zool.* 83, 197–206. doi: 10.1086/648567

Araújo, M. B., Ferri-Yá ñez, F., Bozinovic, F., Marquet, P. A., Valladares, F., and Chown, S. L. (2013). Heat freezes niche evolution. *Ecol. Lett.* 16, 1206–1219. doi: 10.1111/ele.12155

Ayrinhac, A., Debat, V., Gibert, P., Kister, A. G., Legout, H., Moreteau, B., et al. (2004). Cold adaptation in geographical populations of *Drosophila melanogaster*: phenotypic plasticity is more important than genetic variability. *Funct. Ecol.* 18, 700–706. doi: 10.1111/j.0269-8463.2004.00904.x

Bay, R. A., and Palumbi, S. R. (2015). Rapid acclimation ability mediated by transcriptome changes in reef-building corals. *Genome Biol. Evol.* 7, 1602–1612. doi: 10.1093/gbe/evv085

Behera, N., and Nanjundiah, V. (1995). An investigation into the role of phenotypic plasticity in evolution. *J. Theor. Biol.* 172, 225–234. doi: 10.1006/jtbi.1995.0019

Berger, D., Postma, E., Blanckenhorn, W. U., and Walters, R. J. (2013). Quantitative genetic divergence and standing genetic (co)variance in thermal reaction norms along latitude. *Evolution* 67, 2385–2399. doi: 10.1111/evo.12138

Bettencourt, B. R., Feder, M. E., and Cavicchi, S. (1999). Experimental evolution of hsp70 expression and thermotolerance in *Drosophila melanogaster*. *Evolution* 53, 484–492. doi: 10.2307/2640784

Bettencourt, B. R., Kim, I., Hoffmann, A. A., and Feder, M. E. (2002). Response to natural and laboratory selection at the *Drosophila* hsp70 genes. *Evolution* 56, 1796–1801. doi: 10.1554/0014-3820(2002)056%5B1796:rtnals%5D2.0.co;2

Buckley, L. B., Ehrenberger, J. C., and Angilletta, M. J. (2015). Thermoregulatory behaviour limits local adaptation of thermal niches and confers sensitivity to climate change. *Funct. Ecol.* 29, 1038–1047. doi: 10.1111/1365-2435.12406

Campbell-Staton, S. C., Cheviron, Z. A., Rochette, N., Catchen, J., Losos, J. B., and Edwards, S. V. (2017). Winter storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard. *Science* 357:495. doi: 10.1126/science.aam5512

Campbell-Staton, S. C., Winchell, K. M., Rochette, N. C., Fredette, J., Maayan, I., Schweizer, R. M., et al. (2020). Parallel selection on thermal physiology facilitates repeated adaptation of city lizards to urban heat islands. *Nat. Ecol. Evol.* 4, 652–658. doi: 10.1038/s41559-020-1131-8

Chen, E.-H., Hou, Q.-L., Wei, D.-D., Jiang, H.-B., and Wang, J.-J. (2017). Phenotypic plasticity, trade-offs and gene expression changes accompanying dietary restriction and switches in *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae). *Sci. Rep.* 7, 1988–1988.

Chown, S. L., Jumbam, K. R., Sørensen, J. G., and Terblanche, J. S. (2009). Phenotypic variance, plasticity and heritability estimates of critical thermal limits depend on methodological context. *Funct. Ecol.* 23, 133–140. doi: 10.1111/j.1365-2435.2008.01481.x

Coughlin, D., Nicastro, L., Brookes, P., Bradley, M., Shuman, J., Steirer, E., et al. (2019). Thermal acclimation and gene expression in rainbow smelt: changes in

the myotomal transcriptome in the cold. *Comp. Biochem. Physiol. D Genom. Proteomics* 31:100610. doi: 10.1016/j.cbd.2019.100610

Cox, C. L., Logan, M. L., Bryan, O., Kaur, D., Leung, E., Mccormack, J., et al. (2018). Do ring-necked snakes choose retreat sites based upon thermal preferences? *J. Therm. Biol.* 71, 232–236. doi: 10.1016/j.jtherbio.2017.11.020

Cui, J., Zhu, S.-Y., Gao, Y., Bi, R., Xu, Z., and Shi, S.-S. (2019). Comparative transcriptome analysis of *Megacopta cribraria* (Hemiptera: Plataspidae) in response to high-temperature stress. *J. Econ. Entomol.* 112, 407–415. doi: 10.1093/jee/toy330

Diamond, S. E., and Chick, L. D. (2017). The Janus of macrophysiology: stronger effects of evolutionary history, but weaker effects of climate on upper thermal limits are reversed for lower thermal limits in ants. *Curr. Zool.* 64, 223–230. doi: 10.1093/cz/zox072

Etges, W. J., De Oliveira, C. C., Rajpurohit, S., and Gibbs, A. G. (2017). Effects of temperature on transcriptome and cuticular hydrocarbon expression in ecologically differentiated populations of desert *Drosophila*. *Ecol. Evol.* 7, 619–637. doi: 10.1002/ece3.2653

Fares, M. A., Barrio, E., Sabater-Munoz, B., and Moya, A. (2002). The evolution of the heat-shock protein GroEL from Buchnera, the primary endosymbiont of aphids, is governed by positive selection. *Mol. Biol. Evol.* 19, 1162–1170. doi: 10.1093/oxfordjournals.molbev.a004174

Feder, M. E., and Hofmann, G. E. (1999). Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu. Rev. Physiol.* 61, 243–282. doi: 10.1146/annurev.physiol.61.1.243

Fey, S. B., Vasseur, D. A., Alujević, K., Kroeker, K. J., Logan, M. L., O'connor, M. I., et al. (2019). Opportunities for behavioral rescue under rapid environmental change. *Glob. Change Biol.* 25, 3110–3120. doi: 10.1111/gcb.14712

Fontaine, S. S., Novarro, A. J., and Kohl, K. D. (2018). Environmental temperature alters the digestive performance and gut microbiota of a terrestrial amphibian. *J. Exp. Biol.* 221:jeb187559. doi: 10.1242/jeb.187559

Franck, E., Madsen, O., Van Rheede, T., Ricard, G., Huynen, M. A., and De Jong, W. W. (2004). Evolutionary diversity of vertebrate small heat shock proteins. *J. Mol. Evol.* 59, 792–805. doi: 10.1007/s00239-004-0013-z

Furrow, R. E., and Feldman, M. W. (2014). Genetic variation and the evolution of epigenetic regulation. *Evolution* 68, 673–683. doi: 10.1111/evo.12225

Gabriel, W., and Lynch, M. (1992). The selective advantage of reaction norms for environmental tolerance. *J. Evol. Biol.* 5, 41–59. doi: 10.1046/j.1420-9101.1992.5010041.x

Geerts, A. N., Vanoverbeke, J., Vanschoenwinkel, B., Van Doorslaer, W., Feuchtmayr, H., Atkinson, D., et al. (2015). Rapid evolution of thermal tolerance in the water flea *Daphnia*. *Nat. Clim. Change* 5:665. doi: 10.1038/nclimate2628

Geng, Y., Gao, L., and Yang, J. (2013). "Epigenetic flexibility underlying phenotypic plasticity," in *Progress in Botany*, eds U. Lüttge, W. Beyschlag, and J. Cushman (Cham: Springer), 153–163. doi: 10.1007/978-3-642-30967-0_5

Gerken, A. R., Eller, O. C., Hahn, D. A., and Morgan, T. J. (2015). Constraints, independence, and evolution of thermal plasticity: probing genetic architecture of long-and short-term thermal acclimation. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4399–4404. doi: 10.1073/pnas.1503456112

Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N., and Hughes, K. A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* 525, 372–375. doi: 10.1038/nature15256

Ghalambor, C. K., Mckay, J. K., Carroll, S. P., and Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct. Ecol.* 21, 394–407. doi: 10.1111/j.1365-2435.2007.01283.x

Gilchrist, G. W. (1995). Specialists and generalists in changing environments 1. Fitness landscapes of thermal sensitivity. *Am. Nat.* 146, 252–270. doi: 10.1086/285797

Gilchrist, G. W. (1996). A quantitative genetic analysis of thermal sensitivity in the locomotor performance curve of *Aphidius ervi*. *Evolution* 50, 1560–1572. doi: 10.1111/j.1558-5646.1996.tb03928.x

Gilchrist, G. W., and Huey, R. B. (1999). The direct response of *Drosophila melanogaster* to selection on knockdown temperature. *Heredity* 83, 15–29. doi: 10.1038/sj.hdy.6885330

Gleason, L. U., and Burton, R. S. (2015). RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail C hlorostoma funebralis. *Mol. Ecol.* 24, 610–627. doi: 10.1111/mec.13047

Gomulkiewicz, R., and Kirkpatrick, M. (1992). Quantitative genetics and the evolution of reaction norms. *Evolution* 46, 390–411. doi: 10.2307/2409860

Gracey, A. Y., Fraser, E. J., Li, W., Fang, Y., Taylor, R. R., Rogers, J., et al. (2004). Coping with cold: an integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. *Proc. Natl. Acad. Sci. U.S.A.* 101, 16970–16975. doi: 10.1073/pnas.0403627101

Grigg, J. W., and Buckley, L. B. (2013). Conservatism of lizard thermal tolerances and body temperatures across evolutionary history and geography. *Biol. Lett.* 9:20121056. doi: 10.1098/rsbl.2012.1056

Gunderson, A. R., and Stillman, J. H. (2015). Plasticity in thermal tolerance has limited potential to buffer ectotherms from global warming. *Proc. Biol. Sci.* 282:20150401. doi: 10.1098/rspb.2015.0401

Hamdoun, A. M., Cheney, D. P., and Cherr, G. N. (2003). Phenotypic plasticity of HSP70 and HSP70 gene expression in the *Pacific oyster* (*Crassostrea gigas*): implications for thermal limits and induction of thermal tolerance. *Biol. Bull.* 205, 160–169. doi: 10.2307/1543236

Hangartner, S., and Hoffmann, A. A. (2016). Evolutionary potential of multiple measures of upper thermal tolerance in *Drosophila melanogaster*. *Funct. Ecol.* 30, 442–452. doi: 10.1111/1365-2435.12499

Hochachka, P. W., and Somero, G. N. (2002). *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. New York: Oxford University Press.

Hoekstra, H. E., Drumm, K. E., and Nachman, M. W. (2004). Ecological genetics of adaptive color polymorphism in pocket mice: geographic variation in selected and neutral genes. *Evolution* 58, 1329–1341. doi: 10.1554/03-418

Hoffmann, A. A., and Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature* 470, 479–485.

Hu, P., Liu, M., Liu, Y., Wang, J., Zhang, D., Niu, H., et al. (2016). Transcriptome comparison reveals a genetic network regulating the lower temperature limit in fish. *Sci. Rep.* 6:28952.

Huang, L.-H., Wang, H.-S., and Kang, L. (2008). Different evolutionary lineages of large and small heat shock proteins in eukaryotes. *Cell Res.* 18, 1074–1076. doi: 10.1038/cr.2008.283

Huey, R. B., and Hertz, P. E. (1984). Is a jack-of-all-temperatures a master of none? *Evolution* 38, 441–444. doi: 10.1111/j.1558-5646.1984.tb00302.x

IPCC (2013). *IPCC, 2014: Climate Change 2014: Synthesis Report*. Geneva: IPCC.

Izem, R., and Kingsolver, J. G. (2005). Variation in continuous reaction norms: quantifying directions of biological interest. *Am. Nat.* 166, 277–289. doi: 10.1086/431314

Jayasundara, N., Gardner, L. D., and Block, B. A. (2013). Effects of temperature acclimation on *Pacific bluefin* tuna (*Thunnus orientalis*) cardiac transcriptome. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 305, R1010–R1020.

Jesus, T. F., Grosso, A. R., Almeida-Val, V. M. F., and Coelho, M. M. (2016). Transcriptome profiling of two Iberian freshwater fish exposed to thermal stress. *J. Therm. Biol.* 55, 54–61. doi: 10.1016/j.jtherbio.2015.11.009

Johannes, F., Porcher, E., Teixeira, F. K., Saliba-Colombani, V., Simon, M., Agier, N., et al. (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet.* 5:e1000530. doi: 10.1371/journal.pgen.1000530

Kassahn, K. S., Caley, M. J., Ward, A. C., Connolly, A. R., Stone, G., and Crozier, R. H. (2007). Heterologous microarray experiments used to identify the early gene response to heat stress in a coral reef fish. *Mol. Ecol.* 16, 1749–1763. doi: 10.1111/j.1365-294x.2006.03178.x

Kearney, M., Shine, R., and Porter, W. P. (2009). The potential for behavioral thermoregulation to buffer "cold-blooded" animals against climate warming. *Proc. Natl. Acad. Sci. U.S.A.* 106, 3835–3840. doi: 10.1073/pnas.0808913106

Kim, B.-M., Kim, K., Choi, I.-Y., and Rhee, J.-S. (2017). Transcriptome response of the *Pacific oyster*, *Crassostrea gigas* susceptible to thermal stress: a comparison with the response of tolerant oyster. *Mol. Cell. Toxicol.* 13, 105–113. doi: 10.1007/s13273-017-0011-z

Kingsolver, J. G., and Gomulkiewicz, R. (2003). Environmental variation and selection on performance curves. *Integr. Comp. Biol.* 43, 470–477. doi: 10.1093/icb/43.3.470

Knies, J. L., Kingsolver, J. G., and Burch, C. L. (2009). Hotter is better and broader: thermal sensitivity of fitness in a population of bacteriophages. *Am. Nat.* 173, 419–430. doi: 10.1086/597224

Lande, R., and Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution* 37, 1210–1226. doi: 10.2307/2408842

Leal, M., and Gunderson, A. R. (2012). Rapid change in the thermal tolerance of a tropical lizard. *Am. Nat.* 180, 815–822. doi: 10.1086/668077

Leder, E. H., Mccairns, R. S., Leinonen, T., Cano, J. M., Viitaniemi, H. M., Nikinmaa, M., et al. (2015). The evolution and adaptive potential of transcriptional variation in sticklebacks—signatures of selection and widespread heritability. *Mol. Biol. Evol.* 32, 674–689. doi: 10.1093/molbev/msu328

Lewis, J. M., Hori, T. S., Rise, M. L., Walsh, P. J., and Currie, S. (2010). Transcriptome responses to heat stress in the nucleated red blood cells of the rainbow trout (*Oncorhynchus mykiss*). *Physiol. Genomics* 42, 361–373. doi: 10.1152/physiolgenomics.00067.2010

Li, B., Sun, S., Zhu, J., Yanli, S., Wuxiao, Z., and Ge, X. (2019). Transcriptome profiling and histology changes in juvenile blunt snout bream (*Megalobrama amblycephala*) liver tissue in response to acute thermal stress. *Genomics* 111, 242–250. doi: 10.1016/j.ygeno.2018.11.011

Li, Y., Huang, J., Liu, Z., Zhou, Y., Xia, B., Wang, Y., et al. (2017). Transcriptome analysis provides insights into hepatic responses to moderate heat stress in the rainbow trout (*Oncorhynchus mykiss*). *Gene* 619, 1–9. doi: 10.1016/j.gene.2017.03.041

Lim, H.-J., Kim, B.-M., Hwang, I. J., Lee, J.-S., Choi, I.-Y., Kim, Y.-J., et al. (2016). Thermal stress induces a distinct transcriptome profile in the *Pacific oyster Crassostrea gigas*. *Comp. Biochem. Physiol. D Genomics Proteomics* 19, 62–70. doi: 10.1016/j.cbd.2016.06.006

Liu, S., Wang, X., Sun, F., Zhang, J., Feng, J., Liu, H., et al. (2013). RNA-Seq reveals expression signatures of genes involved in oxygen transport, protein synthesis, folding, and degradation in response to heat stress in catfish. *Physiol. Genomics* 45, 462–476. doi: 10.1152/physiolgenomics.00026.2013

Liu, Y., Beyer, A., and Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. *Cell* 165, 535–550. doi: 10.1016/j.cell.2016.03.014

Lockwood, B. L., Sanders, J. G., and Somero, G. N. (2010). Transcriptomic responses to heat stress in invasive and native blue mussels (genus *Mytilus*): molecular correlates of invasive success. *J. Exp. Biol.* 213, 3548–3558. doi: 10.1242/jeb.046094

Logan, C. A., and Somero, G. N. (2011). Effects of thermal acclimation on transcriptional responses to acute heat stress in the eurythermal fish *Gillichthys mirabilis* (Cooper). *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 300, R1373–R1383.

Logan, M. L., Cox, R. M., and Calsbeek, R. (2014). Natural selection on thermal performance in a novel thermal environment. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14165–14169. doi: 10.1073/pnas.1404885111

Logan, M. L., Curlis, J. D., Gilbert, A. L., Miles, D. B., Chung, A. K., Mcglothlin, J. W., et al. (2018). Thermal physiology and thermoregulatory behaviour exhibit low heritability despite genetic divergence between lizard populations. *Proc. R. Soc. B Biol. Sci.* 285:20180697. doi: 10.1098/rspb.2018.0697

Logan, M. L., Fernandez, S. G., and Calsbeek, R. (2015). Abiotic constraints on the activity of tropical lizards. *Funct. Ecol.* 29, 694–700. doi: 10.1111/1365-2435.12379

Logan, M. L., Huynh, R. K., Precious, R. A., and Calsbeek, R. G. (2013). The impact of climate change measured at relevant spatial scales: new hope for tropical lizards. *Glob. Change Biol.* 19, 3093–3102. doi: 10.1111/gcb.12253

Logan, M. L., Minnaar, I. A., Keegan, K. M., and Clusella-Trullas, S. (2020). The evolutionary potential of an insect invader under climate change. *Evolution* 74, 132–144. doi: 10.1111/evo.13862

Logan, M. L., Van Berkel, J., and Clusella-Trullas, S. (2019). The bogert effect and environmental heterogeneity. *Oecologia* 191, 817–827. doi: 10.1007/s00442-019-04541-7

Lynch, M., and Gabriel, W. (1987). Environmental tolerance. *Am. Nat.* 129, 283–303.

Lynch, M., and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Oxford: Oxford University Press.

Martins, F., Kruuk, L., Llewelyn, J., Moritz, C., and Phillips, B. (2018). Heritability of climate-relevant traits in a rainforest skink. *Heredity* 122, 41–52. doi: 10.1038/s41437-018-0085-y

Moskalev, A., Zhikrivetskaya, S., Krasnov, G., Shaposhnikov, M., Proshkina, E., Borisoglebsky, D., et al. (2015). A comparison of the transcriptome of *Drosophila melanogaster* in response to entomopathogenic fungus, ionizing radiation, starvation and cold shock. *BMC Genomics* 16:S8. doi: 10.1186/1471-2164-16-S13-S8

Moya, A., Ganot, P., Furla, P., and Sabourault, C. (2012). The transcriptomic response to thermal stress is immediate, transient and potentiated by ultraviolet radiation in the sea anemone *Anemonia viridis*. *Mol. Ecol.* 21, 1158–1174. doi: 10.1111/j.1365-294x.2012.05458.x

Muñoz, M. M., Stimola, M. A., Algar, A. C., Conover, A., Rodriguez, A. J., Landestoy, M. A., et al. (2014). Evolutionary stasis and lability in thermal physiology in a group of tropical lizards. *Proc. R. Soc. B Biol. Sci.* 281:20132433. doi: 10.1098/rspb.2013.2433

Narum, S. R., and Campbell, N. R. (2015). Transcriptomic response to heat stress among ecologically divergent populations of redband trout. *BMC Genomics* 16:103. doi: 10.1186/s12864-015-1246-5

Phillips, B. L., Llewelyn, J., Hatcher, A., Macdonald, S., and Moritz, C. (2014). Do evolutionary constraints on thermal performance manifest at different organizational scales? *J. Evol. Biol.* 27, 2687–2694. doi: 10.1111/jeb.12526

Qian, B., and Xue, L. (2016). Liver transcriptome sequencing and de novo annotation of the large yellow croaker (*Larimichthy crocea*) under heat and cold stress. *Mar. Genomics* 25, 95–102. doi: 10.1016/j.margen.2015.12.001

Quinn, N. L., Mcgowan, C. R., Cooper, G. A., Koop, B. F., and Davidson, W. S. (2011). Identification of genes associated with heat tolerance in Arctic charr exposed to acute thermal stress. *Physiol. Genomics* 43, 685–696. doi: 10.1152/physiolgenomics.00008.2011

Rezende, E. L., Tejedo, M., and Santos, M. (2011). Estimating the adaptive potential of critical thermal limits: methodological problems and evolutionary implications. *Funct. Ecol.* 25, 111–121. doi: 10.1111/j.1365-2435.2010.01778.x

Rosenblum, E. B., Römpler, H., Schöneberg, T., and Hoekstra, H. E. (2010). Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2113–2117. doi: 10.1073/pnas.0911042107

Sambucetti, P., Scannapieco, A. C., and Norry, F. M. (2010). Direct and correlated responses to artificial selection for high and low knockdown resistance to high temperature in *Drosophila buzzatii*. *J. Therm. Biol.* 35, 232–238. doi: 10.1016/j.jtherbio.2010.05.006

Santos, M., Castaneda, L. E., and Rezende, E. L. (2011). Making sense of heat tolerance estimates in ectotherms: lessons from *Drosophila*. *Funct. Ecol.* 25, 1169–1180. doi: 10.1111/j.1365-2435.2011.01908.x

Santos, M., Castaneda, L. E., and Rezende, E. L. (2012). Keeping pace with climate change: what is wrong with the evolutionary potential of upper thermal limits? *Ecol. Evol.* 2, 2866–2880. doi: 10.1002/ece3.385

Scheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annu. Rev. Ecol. Syst.* 24, 35–68.

Schlichting, C. D., and Pigliucci, M. (1993). Control of phenotypic plasticity via regulatory genes. *Am. Nat.* 142, 366–370. doi: 10.1086/285543

Schlichting, C. D., and Smith, H. (2002). Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evol. Ecol.* 16, 189–211. doi: 10.1023/a:1019624425971

Schlichting, C. D., and Wund, M. A. (2014). Phenotypic plasticity and epigenetic marking: an assessment of evidence for genetic accommodation. *Evolution* 68, 656–672. doi: 10.1111/evo.12348

Schou, M. F., Kristensen, T. N., Kellermann, V., Schlötterer, C., and Loeschcke, V. (2014). A *Drosophila* laboratory evolution experiment points to low evolutionary potential under increased temperatures likely to be experienced in the future. *J. Evol. Biol.* 27, 1859–1868. doi: 10.1111/jeb.12436

Semmouri, I., Asselman, J., Van Nieuwerburgh, F., Deforce, D., Janssen, C. R., and De Schamphelaere, K. A. (2019). The transcriptome of the marine calanoid copepod Temora longicornis under heat stress and recovery. *Mar. Environ. Res.* 143, 10–23. doi: 10.1016/j.marenvres.2018.10.017

Shi, K.-P., Dong, S.-L., Zhou, Y.-G., Li, Y., Gao, Q.-F., and Sun, D.-J. (2019). RNA-seq reveals temporal differences in the transcriptome response to acute heat stress in the Atlantic salmon (*Salmo salar*). *Comp. Biochem. Physiol. D Genomics Proteomics* 30, 169–178. doi: 10.1016/j.cbd.2018.12.011

Smith, S., Bernatchez, L., and Beheregaray, L. B. (2013). RNA-seq analysis reveals extensive transcriptional plasticity to temperature stress in a freshwater fish species. *BMC Genomics* 14:375. doi: 10.1186/1471-2164-14-375

Smolina, I., Kollias, S., Møller, E. F., Lindeque, P., Sundaram, A. Y., Fernandes, J. M., et al. (2015). Contrasting transcriptome response to thermal stress in two key zooplankton species, *Calanus finmarchicus* and *C. glacialis*. *Mar. Ecol. Prog. Ser.* 534, 79–93. doi: 10.3354/meps11398

Sørensen, J. G., Schou, M. F., Kristensen, T. N., and Loeschcke, V. (2016). Thermal fluctuations affect the transcriptome through mechanisms independent of average temperature. *Sci. Rep.* 6:30975.

Stillman, J. H., and Tagmount, A. (2009). Seasonal and latitudinal acclimatization of cardiac transcriptome responses to thermal stress in porcelain crabs, *Petrolisthes cinctipes*. *Mol. Ecol.* 18, 4206–4226. doi: 10.1111/j.1365-294x.2009.04354.x

Stinchcombe, J. R., and Kirkpatrick, M. (2012). Genetics and evolution of function-valued traits: understanding environmentally responsive phenotypes. *Trends Ecol. Evol.* 27, 637–647. doi: 10.1016/j.tree.2012.07.002

Terblanche, J. S., Deere, J. A., Clusella-Trullas, S., Janion, C., and Chown, S. L. (2007). Critical thermal limits depend on methodological context. *Proc. R. Soc. B Biol. Sci.* 274, 2935–2943. doi: 10.1098/rspb.2007.0985

Via, S. (1993). Adaptive phenotypic plasticity: target or by-product of selection in a variable environment? *Am. Nat.* 142, 352–365. doi: 10.1086/285542

Via, S., Gomulkiewicz, R., De Jong, G., Scheiner, S. M., Schlichting, C. D., and Van Tienderen, P. H. (1995). Adaptive phenotypic plasticity: consensus and controversy. *Trends Ecol. Evol.* 10, 212–217. doi: 10.1016/s0169-5347(00)89061-8

Vornanen, M., Hassinen, M., Koskinen, H., and Krasnov, A. (2005). Steady-state effects of temperature acclimation on the transcriptome of the rainbow trout heart. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 289, R1177–R1184.

Walters, R. J., Blanckenhorn, W. U., and Berger, D. (2012). Forecasting extinction risk of ectotherms under climate warming: an evolutionary perspective. *Funct. Ecol.* 26, 1324–1338.

Wang, W., Hui, J. H., Chan, T. F., and Chu, K. H. (2014). De novo transcriptome sequencing of the snail *Echinolittorina malaccana*: Identification of genes responsive to thermal stress and development of genetic markers for population studies. *Mar. Biotechnol.* 16, 547–559.

Waters, E. R. (1995). The molecular evolution of the small heat-shock proteins in plants. *Genetics* 141, 785–795.

Wellenreuther, M., Le Luyer, J., Cook, D., Ritchie, P. A., and Bernatchez, L. (2019). Domestication and temperature modulate gene expression signatures and growth in the Australasian snapper *Chrysophrys auratus*. *G3 Genes Genomes Genetics* 9, 105–116.

Xiao, R., Wang, L., Cao, Y., and Zhang, G. (2016). Transcriptome response to temperature stress in the wolf spider *Pardosa pseudoannulata* (Araneae: Lycosidae). *Ecol. Evol.* 6, 3540–3554.

Yamashita, M., Hirayoshi, K., and Nagata, K. (2004). Characterization of multiple members of the HSP70 family in platyfish culture cells: molecular evolution of stress protein HSP70 in vertebrates. *Gene* 336, 207–218.

Yampolsky, L. Y., Schaer, T. M., and Ebert, D. (2014a). Adaptive phenotypic plasticity and local adaptation for temperature tolerance in freshwater zooplankton. *Proc. R. Soc. B Biol. Sci.* 281:20132744.

Yampolsky, L. Y., Zeng, E., Lopez, J., Williams, P. J., Dick, K. B., Colbourne, J. K., et al. (2014b). Functional genomics of acclimation and adaptation in response to thermal stress in *Daphnia*. *BMC Genomics* 15:859. doi: 10.1186/1471-2164-15-859

Yang, Y., Yu, H., Li, H., Wang, A., and Yu, H.-Y. (2016). Effect of high temperature on immune response of grass carp (*Ctenopharyngodon idellus*) by transcriptome analysis. *Fish Shellfish Immunol.* 58, 89–95.

Zheng, J., Cao, J., Mao, Y., Su, Y., and Wang, J. (2019). Comparative transcriptome analysis provides comprehensive insights into the heat stress response of *Marsupenaeus japonicus*. *Aquaculture* 502, 338–346.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

GLOSSARY OF TERMS

Thermal performance curve:	The mathematical relationship between an ecologically relevant metric of performance (e.g., locomotion, energy assimilation, immune function, etc.) and organismal body temperature. These curves are often used to approximate a populations' thermal niche and can be sub-divided into "thermal performance traits" that describe different aspects of its shape.
Thermal performance trait:	A phenotypic trait that describes performance (e.g., locomotion, energy assimilation, immune function, etc.) at one or a range of temperatures. These traits combine to describe the shape of a population's thermal performance curve.
Narrow-sense heritability (h^2):	The component of phenotypic variation in a trait that is comprised of additive genetic variation. Narrow-sense heritability describes the capacity for a trait to respond efficiently to selection.
Broad-sense heritability (H^2):	The component of phenotypic variation in a trait that is comprised of both additive and non-additive genetic variation, including the effects of dominance and epistasis. Broad-sense heritability includes forms of genetic variation that do not respond efficiently to selection (e.g., recessive alleles that can remain hidden from selection in the heterozygous state).
Genetic correlation:	Positive or negative statistical correlation between genes underlying different phenotypic traits. Genetic correlations often arise from linkage disequilibrium or pleiotropy and can cause correlated evolution of a trait that is not itself under direct selection, but rather is genetically correlated with a different trait that is under direct selection.
Gene expression:	Transcription of mRNA from the genome, which can later be translated into a protein. All mRNA transcripts expressed in a cell, tissue, or organism are referred to as the transcriptome.
Gene expression plasticity:	The ability to alter gene expression in response to an environmental cue. This could be measured at the level of the organism (i.e., the total number of genes that shift their expression) or at the level of an individual gene (i.e., the number and persistence of gene transcripts).
Phenotypic plasticity:	The capacity of the same genotype to produce different phenotypes in different environments. The functional basis of phenotypic plasticity is usually gene expression plasticity.

frontiers in Genetics

MINI REVIEW
published: 25 September 2020
doi: 10.3389/fgene.2020.564515

Predicting Thermal Adaptation by Looking Into Populations' Genomic Past

Andrés J. Cortés[1,2], Felipe López-Hernández[1] and Daniela Osorio-Rodriguez[3]*

[1] Corporación Colombiana de Investigación Agropecuaria AGROSAVIA, C.I. La Selva, Rionegro, Colombia, [2] Departamento de Ciencias Forestales, Facultad de Ciencias Agrarias, Universidad Nacional de Colombia – Sede Medellín, Medellín, Colombia, [3] Division of Geological and Planetary Sciences, California Institute of Technology (Caltech), Pasadena, CA, United States

Molecular evolution offers an insightful theory to interpret the genomic consequences of thermal adaptation to previous events of climate change beyond range shifts. However, disentangling often mixed footprints of selective and demographic processes from those due to lineage sorting, recombination rate variation, and genomic constrains is not trivial. Therefore, here we condense current and historical population genomic tools to study thermal adaptation and outline key developments (genomic prediction, machine learning) that might assist their utilization for improving forecasts of populations' responses to thermal variation. We start by summarizing how recent thermal-driven selective and demographic responses can be inferred by coalescent methods and in turn how quantitative genetic theory offers suitable multi-trait predictions over a few generations via the breeder's equation. We later assume that enough generations have passed as to display genomic signatures of divergent selection to thermal variation and describe how these footprints can be reconstructed using genome-wide association and selection scans or, alternatively, may be used for forward prediction over multiple generations under an infinitesimal genomic prediction model. Finally, we move deeper in time to comprehend the genomic consequences of thermal shifts at an evolutionary time scale by relying on phylogeographic approaches that allow for reticulate evolution and ecological parapatric speciation, and end by envisioning the potential of modern machine learning techniques to better inform long-term predictions. We conclude that foreseeing future thermal adaptive responses requires bridging the multiple spatial scales of historical and predictive environmental change research under modern cohesive approaches such as genomic prediction and machine learning frameworks.

Keywords: coalescent theory, genome-wide association studies, genome-wide selection scans, genome–environment associations, phylogeography, breeder's equation, genomic prediction, machine learning

OPEN ACCESS

Edited by:
Margarida Matos, University of Lisbon, Portugal

Reviewed by:
Anti Vasemägi, Swedish University of Agricultural Sciences, Sweden
Takeshi Kawakami, Independent Researcher, Boston, United States

****Correspondence:***
Andrés J. Cortés
acortes@agrosavia.co

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *21 May 2020*
Accepted: *24 August 2020*
Published: *25 September 2020*

Citation:
Cortés AJ, López-Hernández F and Osorio-Rodriguez D (2020) Predicting Thermal Adaptation by Looking Into Populations' Genomic Past. Front. Genet. 11:564515. doi: 10.3389/fgene.2020.564515

ON THE CHALLENGES OF STUDYING GENOMIC THERMAL ADAPTATION

Warming is imposing an unprecedented climate emergency on nature, food, energy supply, and economy around the world (Ripple et al., 2020). While evolutionary genomics may improve prediction of populations' responses to thermal change (Waldvogel et al., 2020a), geologic records of temperature and carbon dioxide (CO_2) variations (**Supplementary Figure S1**) are also insightful into the coupling of biodiversity, climate, and the carbon cycle and hence may help predicting the

consequences of future carbon emissions (Zachos et al., 2008). For instance, several reports of fire activity (Whitlock and Bartlein, 2003; Bush et al., 2008) and hydroclimate changes (Wang et al., 2017) as records of thermal changes during the Holocene have taught us that extinction is a slow process and that many species may already be functionally extinct (Cronk, 2016). A key modern advance has precisely been to couple the extinction risk with the migratory potential under an ecological niche conservatism scenario (Steinbauer et al., 2018), and predictions of population-level genomic and phenotypic responses to thermal change (Hoffmann and Sgro, 2011). Although atmospheric CO_2 has been found to be better correlated with richness of (plant) species (**Supplementary Figure S1C**) than temperature itself throughout the Cenozoic up until 20 Mya (Jaramillo et al., 2006; Royer and Chernoff, 2013), we need to improve our understanding on how thermal change vulnerability impacts current and historical adaptive genetic variation in order to enhance populations response projections (Razgour et al., 2019).

Genomes are diverse in signatures of the populations' evolutionary past across timescales (Wolf and Ellegren, 2017) and therefore are informative on historical adaptive responses to ancient and more recent events of climate change (**Figure 1** and **Table 1**). By revealing the nature of these signatures and learning from previous reactions to environmental change, genomics can truly assist modern predictions aimed at incorporating responses beyond migration. Yet, disentangling often confused selective and demographic signatures from those due to genetic drift and genomic constrains is challenging (Ellegren and Galtier, 2016), consequently delaying the factual utilization of genomics for forecasting. Therefore, in this mini-review we envision summarizing modern tools from the genomic era that are enriching our comprehension of the genetic consequences of past and recent climate change, while offering a perspective on how to improve predictive models that incorporate thermal adaptation. Specifically, we aim prospecting how genomic prediction (GP) and machine learning (ML) approaches may offer cohesive frameworks to (1) integrate more traditional, but heterogeneous, genomic, and ecological datasets across temporal scales, by (2) maximizing prediction accuracies, while (3) understating the relative contribution of the underlying genomic processes. This is still a future avenue of research, and so we close by offering perspectives. Different drivers of the genomic landscape to thermal adaption (Gompert et al., 2014; Ravinet et al., 2017; Cortés and Blair, 2018; López-Hernández and Cortés, 2019), such as disruptive and background selection, gene flow (Miller et al., 2020), shared ancestral polymorphism, and mutation/recombination rate variation (Feder et al., 2012; Ellegren and Wolf, 2017; Cortés et al., 2018b), have been identified. In order to discern among them, a first necessary step toward the evaluation of the adaptive potential involves typifying the genomic landscape by using summary statistics like nucleotide diversity, π (Nei, 1987), and relative, F_{ST} (Weir and Cockerham, 1984), and absolute, D_{XY} (Nei, 1987), divergence. F_{ST} vs. D_{XY} contrasts inform population divergence in the presence of gene flow (co-occurrence of peaks in both profiles), recurrent selection across subpopulations (F_{ST} peaks match shallow D_{XY} valleys), and selective sweeps predating the subpopulations' split (F_{ST} peaks match deep D_{XY} valleys) (Nachman and Payseur, 2012; Cruickshank and Hahn, 2014; Irwin et al., 2016). Inferences are more robust if carried out across replicated samplings of contrasting populations (e.g., in terms of thermal variation) within a hierarchically nested framework of divergence (Cortés et al., 2018b). A second step refers to the detection of selection signatures, if any – i.e., hard vs. soft selection sweeps (Pritchard et al., 2010; Zahn and Purnell, 2016), which must be followed by a third validation step across replicated demographics (Roesti et al., 2014; Lotterhos and Whitlock, 2015) and temporal levels (Nosil and Feder, 2011; Matos et al., 2015; Fragata et al., 2018).

Exclusively phenotypic empirical methods (**Figure 1A**), such as *in situ* monitoring, growth chamber experiments, and "common garden" (provenance) tests (Miller et al., 2020), constitute baseline evidence of thermal adaptation and should therefore inform more advanced genomic approaches. Naturally available environmental gradients (e.g., elevation or latitudinal clines) can also be used as proxies for climate change (Wheeler et al., 2016; Cortés and Wheeler, 2018), which is known as space-for-time (SFT) substitution. Replicated "common garden" tests (a.k.a. reciprocal transplants) carried out in an SFT framework are in turn useful to test whether populations can cope with changes through local adaptation (standing variation) or via phenotypic plasticity, especially in long-living species (Bridle and Vines, 2007; Sedlacek et al., 2015). Within an SFT framework, restricted gene flow can lead to small-scale genetic structures (Stanton et al., 1997) or distorted source/sink-like patterns (e.g., Cortés et al., 2014) driven by environmental factors (Nathan and Muller-Landau, 2000). Asymmetric migratory potential in a local scale may provide suitable habitats within only a few meters of the current locations (Yamagishi et al., 2005; Scherrer and Körner, 2011) but may also lead to narrowly adapted populations, even in the face of gene flow (Fitzpatrick et al., 2015), that may respond poorly to future conditions (North et al., 2011; Miller et al., 2020).

FROM RECENT GENETIC RESPONSES TO SHORT-TERM PREDICTIONS

Coalescence Informs on Contemporary Thermal-Driven Selective and Demographic Changes

In order to trace back thermal-driven selective and demographic changes at recent temporal scales (**Figure 1B**), coalescent theory (Wakeley, 2008) helps in discriminating among authentic signatures of selection and those related to demography (e.g., bottlenecks and among populations reduced gene flow), from spurious covariates (Yeaman and Otto, 2011) such as lineage sorting (Wolf and Ellegren, 2017; Becher et al., 2020) and inversions (Dolgova et al., 2010; Fragata et al., 2014). Recursive simulation-based tools to incorporate the mutation/selection balance (Bustamante et al., 2001) across various scenarios of divergence and gene flow are approximate Bayesian computation – ABC (Csilléry et al., 2010; Cornuet et al., 2014), and pairwise sequentially Markovian

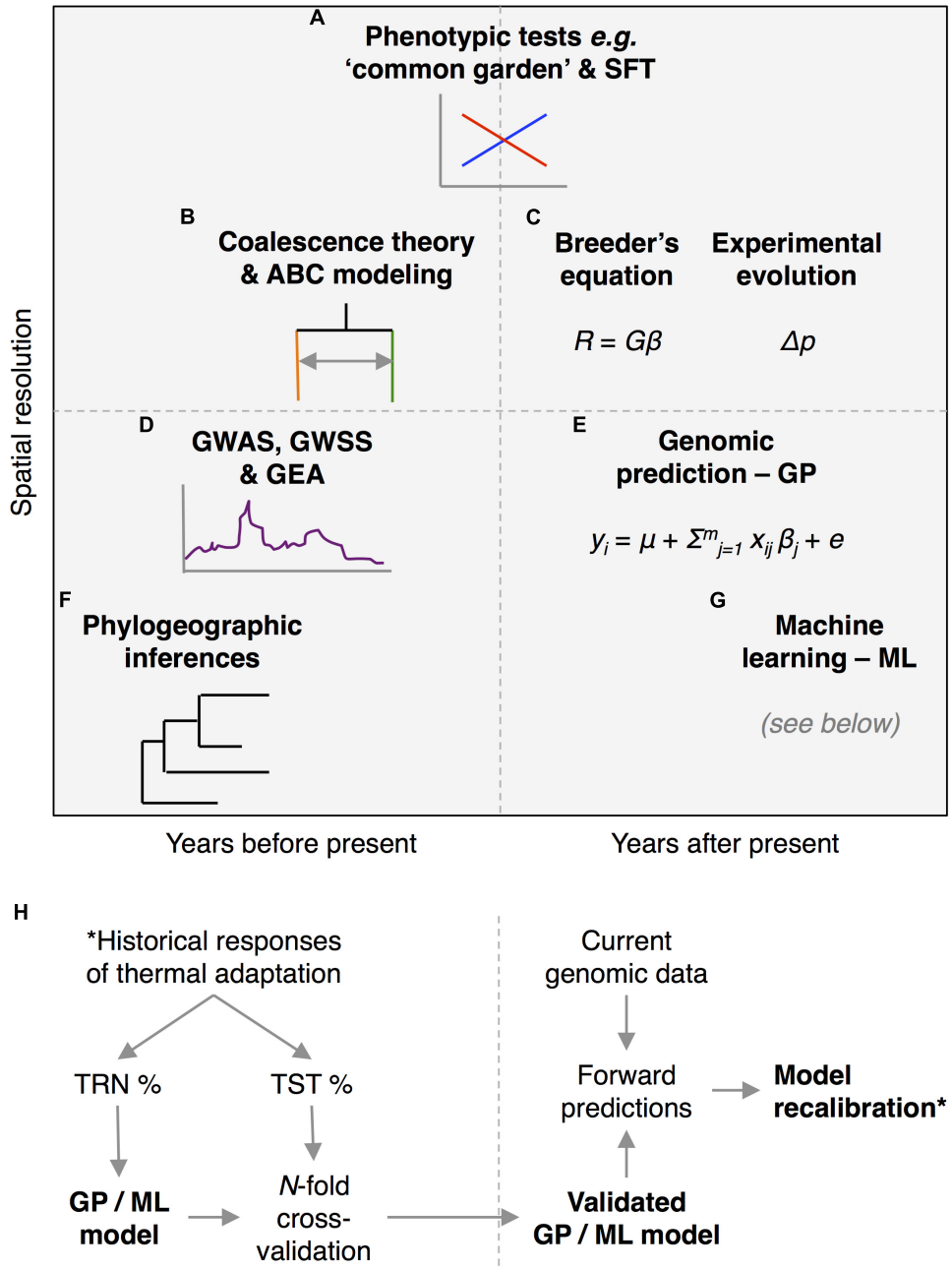


FIGURE 1 | Potential approaches to assess populations' thermal adaptation by looking into their genomic past. Genomic analyses allow reconstructing populations' adaptive responses to previous events of climate change across various temporal scales **(A,B,D,F)**, as a tool to improve forecasting **(C,E,G,H)**. **(A)** Empirical approaches such as replicated "common garden" (provenance) tests and space-for-time (SFT) substitution allow studying *in situ* ongoing genomic thermal adaptation. The inset plot exemplifies a significant genotype-by-environment (*G*x*E*) interaction, as can be quantified using reciprocal transplant experiments between habitat types that differ in their thermal stress. **(B)** Coalescent and approximate Bayesian computation (ABC) analyses help infer recent thermal-driven selective and demographic responses. The inset diagram shows a typical coalescent genealogy depicting divergence with gene flow. **(C)** The breeder's equation predicts responses of genetically correlated traits over one generation (vector *R*) given standardized selection gradients to thermal stress (vector β) by means of the variance–covariance matrix (*G*) of additive genetic parameter estimates. Alternatively, experimental evolution traces real-time changes in allele frequencies (Δ*p*) across generations. **(D)** When genomic signatures of thermal selection are under divergent selection after several generations, genome-wide association (GWAS), and selection (GWSS) scans, as well as genome–environment associations (GEA), allow characterizing the genomic architecture of thermal adaptation. The inset Manhattan plot schematizes a hypothetical genomic scan between populations that contrast in their thermal adaptation. **(E)** Modern high-throughput genotyping may facilitate predictions of the thermal adaptive potential over multiple generations using infinitesimal models under a genomic prediction (GP) framework. **(F)** Phylogeographic approaches offer an understanding of the genomic consequences of deep-time thermal shifts at an evolutionary time scale. The inset tree represents an imaginary phylogeny. Finally, **(G)** machine learning (ML) approaches **(H)** trained using heterogeneous past responses to thermal variation may enhance long-term predictions of the thermal adaptive potential. ML's *modus operandi*, as GPs, requires partitioning the calibrating historical dataset between training (TRN) and testing (TST) subsets that are iteratively imputed into a *N*-fold cross-validation scheme.

TABLE 1 | Case studies that have addressed thermal adaptation at different temporal scales using diverse genetic analyses.

Analytical approach	Diagram	Data sources	Main finding	References
Coalescence theory and ancestry distribution models	**Figure 1B**	20 alpine plant species across the European Alps genotyped with AFLP markers and analyzed with ancestry distribution models	Ancestry distribution models open new perspectives to forecast population genetic changes within species	Jay et al., 2012
Coalescence theory in a SFT framework	**Figure 1B**	273 *Salix* genets in 12 SFT populations genotyped with 7 SSRs	There is asymmetric gene flow across a thermal gradient that may be affected under future climate conditions	Cortés et al., 2014
Coalescence theory	**Figure 1B**	Exome re-sequencing of 48 *Populus trichocarpa* individuals	Effective population size has varied in concert with atmospheric temperature deviation from the past c. 120,000 years	Zhou et al., 2014
Quantitative genetics	**Figure 1C**	Review of models on whether evolutionary changes within species can contribute to species adapting to global thermal change	Evolutionary processes and trait trade-offs (*Q* matrix) need to be incorporated into schemes that try to manage thermal impacts	Hoffmann and Sgro, 2011
Quantitative genetics	**Figure 1C**	Review discussing thermal adaptation to climate change from an evolutionary physiological perspective	Species' physiological, genetic and plastic (Nicotra et al., 2010) capacities can aid in forecasting their response to thermal change	Chown et al., 2010
Quantitative genetics	**Figure 1C**	Physiological model that simulates thermal tolerance assays for multilocus quantitative traits in *D. melanogaster*	Realized heritabilities of knockdown temperature may underestimate the true heritability of the upper thermal limit	Rezende et al., 2010; Santos et al., 2012
Breeder's equation in 2-habitats SFT design	**Figure 1C**	1,061 *Salix herbacea* genotypes, from 2 habitats in a SFT design, screened for 6 thermally influenced traits and 7 SSRs	Significant heritable variation in morphology and phenology might help *S. herbacea* adapt to thermal stress	Sedlacek et al., 2016
Quantitative genetics and breeder's equation	**Figure 1C**	166 lines of *D. melanogaster* assessed for cold tolerance at 5 temperatures	Low thermal tolerance is environment specific and evolvability decreases with increasing developmental temperatures	Ørsted et al., 2019
Quantitative genetics and breeder's equation	**Figure 1C**	4,267 25- to 35-year-old European larch trees growing in 21 reforestation installations across 4 distinct climatic regions in Austria	Genetic evaluation across broad thermal gradients permits delineation of suitable reforestation areas under future climates	Lstiburek et al., 2020
GWAS	**Figure 1D**	Review on molecular-level regulation of the annual growth cycle in temperate and boreal regions	Merging genomic analyses with more quantitative approaches will aid studies on how species cope with thermal changes	Singh et al., 2017
eGWAS	**Figure 1D**	Whole-genome transcriptional responses in *D. subobscura* subjected to threefold replicated laboratory thermal shocks	Many genes appear to be involved in thermal adaptation, as expected for the adaptive evolution of a complex trait	Laayouni et al., 2007
GWAS across a SFT latitudinal gradient	**Figure 1D**	446 *Populus trichocarpa* trees from a latitudinal gradient screened for bud-break in 2 provenance trials and with 2.2-M SNPs	Variation in bud-break reflects differential selection for thermal functions likely to be affected by climate warming	McKown et al., 2018
GWSS across a SFT latitudinal gradient	**Figure 1D**	Two populations of *D. subobscura* from different latitudes introduced to a new common laboratory environment and WGS	Populations followed different genetic routes to reach predictable and similar adaptive phenotypic outcomes	Seabra et al., 2017
GWSS given a modern heat wave	**Figure 1D**	Long-term time series of seasonal genetic data in *D. subobscura*	Genetic constitution of the populations transiently shifted to summer-like frequencies during the 2011 heat wave	Rodriguez-Trelles et al., 2013
GWSS in 2 postglacial lineages	**Figure 1D**	48 *Populus alba* ramets from 2 postglacial recolonization lineages genotyped with GWS for 1.7-M SNP markers	Selection from standing variation implies the potential for rapid evolution of *P. alba* populations in the face of thermal change	Stölting et al., 2015
GEA at a continental scale	**Figure 1D**	78 Andean and Mesoamerican wild bean accessions with 23,373 GBS-derived SNPs and 3 bioclimatic heat stress indices	24 associated loci with contrasting habitat types flank 22 heat shock protein genes (Simões et al., 2003; Sørensen et al., 2003)	López-Hernández and Cortés, 2019
GEA at a latitudinal gradient	**Figure 1D**	Four populations of *D. subobscura* from different latitudes screened for 4 candidate loci for thermal adaptation in inversions	Inversion frequency clines are being maintained by local thermal adaptation in face of gene flow	Simões and Pascual, 2018
GEA at a regional scale	**Figure 1D**	79 natural *Fagus sylvatica* populations, 144 SNPs out of 52 thermal candidate genes, and 87 environmental predictors	*F. sylvatica* exhibits local genetic adaptation to thermal heterogeneity at the regional scale (Swiss Alps)	Pluess et al., 2016
GEA at a regional scale	**Figure 1D**	140 wild tomato accessions, 6,830 SNPs, and redundancy analysis (RDA), structural equation modeling (SEM), and generalized dissimilarity modeling (GDM)	Regional differences in the abiotic environment contribute to genomic divergence within a wild tomato species	Gibson and Moyle, 2020

(Continued)

TABLE 1 | Continued

Analytical approach	Diagram	Data sources	Main finding	References
Genomic prediction (GP)	**Figure 1E**	48 cows genotypes with a BovineLD BeadChip and studied in climate-controlled chambers that simulate a heat wave event	GP for heat tolerance may increase resilience and welfare in animal breeding to increased incidence and duration of heat events	Garner et al., 2016
Backward genomic prediction (GP)	**Figure 1E**	Re-sequencing of 15 1900-year-old maize cobs from Turkey Pen Shelter, and GBS data of 1,316 modern landraces for training	Thermal adaptation drove modern maize divergence and was selected *in situ* from ancient standing variation 2000 years ago	Swarts et al., 2017
Genomic prediction (GP)	**Figure 1E**	287 elite spring wheat lines assessed in a 90K Illumina array for traits as thermal time to flowering in 18 heat/drought environments	GP is capable to predict complex traits and find the best environments to adapt new crop lines to heat and drought stress events	Sukumaran et al., 2017
Genomic prediction (GP)	**Figure 1E**	3,485 wheat lines genotyped with 9,285 GBS-derived SNPs and phenotyped for grain yield in heat and drought environments	GP can be used to increase the size of plant nurseries by considering un-phenotyped lines for heat and drought stress-resilience	Juliana et al., 2019
Fossil record	**Figure 1F**	Palynological neotropical plant diversity of 1,411 morpho-species and 287,736 occurrences (65–20 million years ago)	Low Paleocene flora diversity, more diverse early Eocene flora exceeding Holocene levels, and a decline at early Oligocene	Jaramillo et al., 2006
Phylogenetics	**Figure 1F**	Thoreau's dataset of the Concord (MA) flora that provides data on changes in species abundance and flowering time (150 years)	Thermal change has shaped the phylo-genetically biased pattern of species loss in species that do not respond to temperature	Willis et al., 2008
Fossil record	**Figure 1F**	Pollen and macroscopic charcoal from the Erazo profile (Ecuador)	Global Pleistocene temperature change can radically alter vegetation communities on the Andean flank in western Amazonia	Cardenas et al., 2011
Phylogeographic inferences – fossils	**Figure 1F**	Long-term ecological records and their relevance to climate change predictions for a warmer world	Range shifts, community turnover, genetic adaptation, and an increase in diversity are observed during warmer intervals	Willis and MacDonald, 2011
Phylogeographic inferences	**Figure 1F**	17 time-calibrated phylogenies of major tetrapod clades and climatic data from distributions of > 500 extant species	Rates of projected climate change dramatically exceed past rates of thermal niche evolution among vertebrate species	Quintero and Wiens, 2013
Phylogeographic inferences	**Figure 1F**	Niche shifts among populations within 56 plant and animal species using time-calibrated phylogenetic trees	Rates of change in thermal niches in plant and animal populations have been much slower than projected climate change	Jezkova and Wiens, 2016
Phylogenetic-assisted modeling	**Figure 1F**	9,737 records for 1,312 plant species and phylogenetic correlation matrix as an additional random effect	Tropical plants do not have narrower heat tolerances, but are more at risk due to their upper thermal limits (Feeley et al., 2020)	Sentinella et al., 2020
Dynamic eco-evolutionary modeling	**Figure 1G**	Four endemic Alpine plant species analyzed with niche modeling, and individual-based demographic and genetic simulations	Monitoring species' local abundance instead of their range better informs on species' extinction risks under thermal change	Cotto et al., 2017
Machine learning (ML)	**Figure 1G**	Species geographic distributions modeling using maximum entropy (MaxEnt)	ML modeling can be used for discrimination of suitable vs. unsuitable areas for the species with presence-only datasets	Phillips et al., 2017
Machine learning (ML)	**Figure 1G**	Temporal uncertainty framework to assess when and where cultivation of key crops in sub-Saharan Africa will become unviable	Incremental, preparatory and transformational adaptation phases enable projected crop transformational changes	Rippke et al., 2016
Machine learning (ML)	**Figure 1G**	Random forest in Himalaya's *Betula* for last inter-glaciation, present (1970–2000) and future (2061–2080) conditions	Biodiversity in high elevation ecosystems is sensitive to global environmental changes, especially temperature warming	Mohapatra et al., 2019
Machine learning (ML)	**Figure 1G**	Modeling of the spatiotemporal distribution in the present and the future of pine in heat scenarios (RCP 4.5 y RCP 8.5) by MaxEnt	There were good predictions for both climate change scenarios, and two contrasted tendencies of progressive evolution	Garah and Bentouati, 2019
Machine learning (ML)	**Figure 1G**	Association between gene expression and critical temperature in divergent trout populations was measured by random forest	The "gradient boosting" approach showed that evolution for higher upper thermal tolerance is possible	Chen et al., 2018
Machine learning (ML) + phylogenetic diversity	**Figure 1G**	Predictive models of taxonomic and phylogenetic diversity using vascular plant database for the United States	Native phylogenetic diversity is likely to decrease over the next half century despite increases in species richness	Park et al., 2020

(Continued)

TABLE 1 | Continued

Analytical approach	Diagram	Data sources	Main finding	References
The potential of big data	**Figure 1G**	Special issue inspired by the symposium "Fitness landscapes, big data, and the predictability of evolution"	Understanding evolutionary adaptive responses in the face of epistasis is a major need that could benefit from big data	Visser et al., 2018
Genomic prediction (GP) + machine learning (ML)	**Figures 1E,G**	*ca.* 11,000 wheat landrace accessions assessed for 40,000 GBS-derived SNPs and traits possibly related with heat stress	Deep learning should be integrated with GBLUP for the study of complex traits and the *GxE* interaction	Montesinos-Lopez et al., 2018
Genomic prediction (GP) + machine learning (ML)	**Figures 1E,G**	*ca.* 3,500 wheat landrace accessions examined for 2,038 GBS-derived SNPs in 4 environments of drought and 2 of heat stress	MLP and SVM were competitive in genomic prediction of complex traits possibly related to heat stress as days to heading	Montesinos-Lopez et al., 2019

Examples enlighten how analytical approaches that try to reconstruct populations' past genetic adaptive responses to previous events of climate change could be proxies for better forecasting. This compilation is built for illustrative purposes and is not meant to be exhaustive. Examples are sorted as in ***Figure 1****. SFT, space-for-time substitution, GWAS, genome-wide association study; eGWAS, expression GWAS; GWSS, genome-wide selection scans; GEA, genome–environment associations; SSRs, simple sequence repeats; SNP, single-nucleotide polymorphism; WGS, whole-genome sequencing; GBS, genotyping-by-sequencing; SVM, support vector machine; MLP, multilayer perceptron; GP, genomic prediction; ML, machine learning.*

coalescent – PSMC (Nadachowska-Brzyska et al., 2016). These approaches can inform how isolated populations that usually occupy climates with scarce habitat complexity (Flantua et al., 2019) may favor thermal generalists, while intricate local-scale heterogeneity at larger scales could trigger (Hughes, 2006; Cortés et al., 2018a) thermal specialists with limited migration potential (Cuesta et al., 2019). They can also model population sizes (Beerli, 2006) in concert with thermal changes (Zhou et al., 2014; Lehnert et al., 2019). Yet, these approaches may be limited by computational burden as they rely on simulation-based rejection sampling, while much effort is gone into the design of multiple scenarios, dimensionality reduction, and feature selection (Schrider and Kern, 2018).

The Breeder's Equation Assists Multi-Trait Predictions Over a Few Generations

In order for thermal adaptation to happen, there must be heritable trait variation upon which selection, enforced by climate change, acts (Darwin, 1874). A simple deterministic model that condenses this evolutionary paradigm, aiding in the forecast of adaptive trait responses across few generations, comes from the quantitative genetic discipline and is known as the breeder's equation (**Figure 1C**). Its multivariate form (Walsh, 2008) allows predicting responses of genetically correlated traits (vector R) to standardized thermal selection gradients (vector β) over one generation, so that $R = G\beta$, where G is the variance–covariance matrix of additive genetic parameter estimates – a proxy for traits' heritabilities and trade-offs (Falconer and Mackay, 1996). The potential evolutionary response can therefore be computed using selection-gradient estimates derived from fitness proxies (i.e., fitness values regressed as a function of standardized trait values) and marker-based heritabilities (Lynch and Ritland, 1999). This approach by itself is not novel, but what makes it powerful is that it can be coupled with SFT (Wheeler et al., 2014), among other trials, to predict thermal responses to thermal change (Sedlacek et al., 2016). Yet, a major drawback is that selection gradients heavily depend on the nature of the fitness proxies (Sedlacek et al., 2016). Alternatively, experimental evolution studies (Exposito-Alonso et al., 2019) could test more explicitly how rapidly growing populations may respond to different thermal scenarios (Kawecki et al., 2012) that, together with evolve and re-sequence analyses (Turner and Miller, 2012), may contribute to understand the genetic basis of short-term thermal adaptation.

FROM DEEPER GENOMIC SIGNATURES OF SELECTION TO MID-TERM PREDICTIONS

Genome-Wide Scans Reveal Signatures of Divergent Selection to Past Thermal Adaptation

Assuming that enough generations have passed as to exhibit divergent selection to thermal changes, genome-wide association (GWAS) (Hirschhorn and Daly, 2005) and selection (GWSS) (Sabeti et al., 2007) scans (**Figure 1D**) are essential analytical tools to reconstruct the genomic architecture of adaptive trait divergence to thermal stress (Lecheta et al., 2020; Zwoinska et al., 2020). These methods assume that some allele variants are in linkage disequilibrium (LD) (Slatkin, 2008) with causal variants that influence the adaptive phenotype (Morris and Borevitz, 2011; Tam et al., 2019), a.k.a. genetic "hitchhiking" (Maynard Smith and Haigh, 1974; Feder and Nosil, 2010). An interface between GWAS and GWSS studies where *loci* are directly correlated with niche's thermal variables is named genome–environment association (GEA) (Forester et al., 2016) and is insightful to infer past thermal adaptation, too (Hancock et al., 2011; Pluess et al., 2016; López-Hernández and Cortés, 2019). Yet, these approaches partly disregard non-additive and highly polygenic architectures (Stephan, 2016; Csillery et al., 2018; Barghi et al., 2020) and may be misleading (Maher, 2008; Pennisi, 2014) if standardized data (Waldvogel et al., 2020b) and statistical covariates (Lambert and Black, 2012), such as population stratification (Barton et al., 2019) and genomic constrains (Wray et al., 2013; Huber et al., 2016), are incorrectly accounted for.

Genomic Prediction May Assist Forecasting of Adaptive Traits Over Multiple Generations

A cutting-edge development that materialized after bringing genomics into quantitative genetics theory is genomic prediction (GP) (Desta and Ortiz, 2014; Crossa et al., 2017; Grattapaglia et al., 2018). GP uses historical phenotypic data to adjust marker-based infinitesimal (**Figure 1E**) models (Meuwissen et al., 2001; Gianola et al., 2006; de los Campos et al., 2013) that may overcome some of the restraints described in the previous section. GP may offer a more thoughtful picture of complex traits (e.g., thermal adaptation), presumably regulated by many low-effect *loci* (Pritchard et al., 2010). GP has so far informed predictions of single adaptive traits in populations with known pedigrees (Saint Pierre et al., 2012; Cros et al., 2019) and hybrid origins (Technow et al., 2014; Tan et al., 2017), as well as multi-trait inferences across diverse unrelated populations (Crossa et al., 2007, 2016; Resende et al., 2012; Suontama et al., 2019) under genotype by environment interactions (GxE) (Montesinos-Lopez et al., 2018; Crossa et al., 2019) facing polygenic climate adaptation (Isabel et al., 2020). GP of thermal adaptive traits across multiple generations and populations may be incipient (**Table 1**), yet it harbors a promising potential, as was demonstrated by reversely predicting unobserved thermal phenology in 1900-year-old ancient corn (Swarts et al., 2017), and as we prospect in the last section of this mini-review.

FROM DEEP-TIME GENOMIC CONSEQUENCES OF THERMAL SHIFTS TO LONG-TERM PREDICTIONS

Phylogeography Offers Insights Into Past Responses at an Evolutionary Scale

Phylogeographic inferences (**Figure 1F**) offer insights into how species (1) diversify (Quintero and Wiens, 2013) and (2) face the effects of past thermal variation (Jezkova and Wiens, 2016; Richardson et al., 2019) by boosting complex interactions such as species facilitation (Wheeler et al., 2015), adaptive introgression, and hybrid speciation (Coyne and Orr, 2004; Abbott et al., 2013; Payseur and Rieseberg, 2016; Marques et al., 2019). For instance, interspecific hybrids with intermediate niche requirements may rescue population's gene pools in the face of climate change, while they can also display signals of heterosis for thermal adaption due to dominance on recessive alleles or overdominance via novel allele combinations (Abdelmula et al., 1999; Leinonen et al., 2011). Modern phylogeographic inferences currently rely on abundant and unlinked genetic markers (Bryant et al., 2012) that are capable of bypassing traditional assumptions of single gene mutation models (Caliebe, 2008) while accounting for scenarios of reticulate evolution (Vargas et al., 2017). Marker-based inferences also offer higher resolution to validate cases where adaptive radiation (Madriñán et al., 2013), and ecological parapatric speciation resulted from local patterns of environmental variation (Cortés et al., 2018a) that may resemble those expected by thermal change. Mosaics of local-habitat heterogeneity can ultimately enforce thermal pre-adaptation (Cortés and Wheeler, 2018). Distance-based phylogenic reconstruction without proper out-groups (Baum et al., 2005; Cortés, 2013) is yet a major risk of these approaches.

Machine Learning May Bridge Historical Genomics and Long-Term Predictions

A promising way to simultaneously make sense of multiple sources of historical genomic data that can be utilized to predict populations' adaptive responses is by merging them into a machine learning (ML) framework (**Figures 1G,H**). ML bypasses the "curse of dimensionality" and benefits from high-dimensional inputs of heterogeneous dependent variables ("features") without *a priori* knowledge of their joint probability distribution (Schrider and Kern, 2018). This improves predictions' "recall" (true positive) rate among a set of possible responses, especially when the classification is iteratively trained using "labeled" data (i.e., historical thermal responses may offer novel calibration datasets, **Table 1**) via *N*-fold cross-validation. ML has been routinely used to make ecological niche modeling (Phillips et al., 2017; Valencia et al., 2020) and functional predictions across genomes (Libbrecht and Noble, 2015). Yet, ML may likely displace other tools useful to characterize the genomic consequences of thermal adaptation, already introduced in this mini-review, such as ABC modeling (Liu et al., 2019) and GWSS (Schrider and Kern, 2018).

CONCLUDING REMARKS

Thermal adaptation is a complex polygenic trait well-described in terms of its genetic architecture and selection footprints across a wide range of phylogenetically diverse taxa (Way and Oren, 2010; Valladares et al., 2014; López-Hernández and Cortés, 2019). While genomics has enabled these achievements that rely on past events of thermal variation, forward predictions remain one step behind partly because (1) disentangling selective and demographic drivers of the genomic landscape from fortuitous genomic constrains (Logan and Cox, 2020) is puzzling (Ellegren and Galtier, 2016) and (2) merging these heterogeneous signatures and data sources into a cohesive predictive framework was unfeasible, until recently. In this mini-review, we advocated for novel approaches that may enhance our understanding of the genetic consequences of past climate change, while offering new avenues to calibrate more accurate predictive models of the thermal adaptive potential. For instance, ML advances are likely to now move beyond species distribution modeling (Phillips et al., 2017) and functional genomics (Libbrecht and Noble, 2015) to permeate the backward interpretation of recent genetic demographic responses and genomic signatures to historical thermal selection by updating popular but sometimes intractable methods such as ABC modeling and GWSS (Schrider and Kern, 2018). Meanwhile, GP and ML might boost forward predictions of the adaptive potential beyond a single generation by training multifactorial models that can try incorporating genomic heterogeneous evidence of historical thermal adaption across a

wide spectrum of temporal scales. Ultimately, understanding how biotas formed in response to historical environmental change may improve our ability to predict and mitigate the threats to species posed by global warming (Ding et al., 2020).

Despite GP's and ML's being useful to comprehend and predict thermal adaptation, these new paradigms are not exempt of criticism. A reiterative misconception is that because these methodologies aim at strengthening predictions and classification boundaries, they do not offer a mechanistic understanding of the subjacent processes. However, even though GP and ML rely on algorithmically generated models, both are far from "black boxes" because they allow direct measurement of the contribution of each genetic marker (Resende et al., 2012; Spindel et al., 2016) and "feature" (Schrider and Kern, 2018), to the point that they can offer higher resolution than traditional genetic mapping (Hirschhorn and Daly, 2005) and deterministic model building (Otto and Day, 2007) techniques. A second misconception assumes computational burden. Although both GP and ML require a large number of simulations, they do not depend on rejection sampling, which means they may efficiently use all of the simulations to inform the mapping of historical thermal data to parameters (Schrider and Kern, 2018).

FUTURE DIRECTIONS

So far, GP and ML have been mostly utilized to address thermal adaptation individually (**Table 1**). For instance, GP has been used to project heat tolerance in diverse wheat lines (Sukumaran et al., 2017; Juliana et al., 2019), and bovine genotypes (Garner et al., 2016), in all cases more as a proof of concept. Similarly, ML approaches have not only deepened our understating on populations' range shifts in the light of thermal variation (Rippke et al., 2016; Garah and Bentouati, 2019; Mohapatra et al., 2019) but also assisted eGWAS of critical temperature thresholds (Chen et al., 2018) and phylogenetic forecasting in plants (Park et al., 2020). However, since GP and ML are both cutting-edge tools, there is still room and need for new developments. For instance, merging more cohesively past adaptive responses to previous events of environmental change into cutting-edge analytical frameworks like GP and ML will ultimately allow predicting whether populations' adaptive potential may keep up with the pace of current thermal increase (Franks and Hoffmann, 2012; Franks et al., 2014). Swarts et al. (2017) illustrates that across-temporal predictions may be useful not only to improve forecasting (Sweet et al., 2019) but also to better understand previous responses to thermal variation, since they used backward GP to demonstrate that thermal adaptation in maize was selected *in situ* from ancient standing variation 2000 years ago. By enlightening on the nature of these historical genetic signatures to past climate change, genomics can also enhance predictions that aim at incorporating adaptive responses beyond extirpation and range shifts (Chen et al., 2011).

Data sources incorporated into GP and ML can transcend those with a direct genomic connotation and involve others that can modulate or be informative of the thermal responses. For instance, from an abiotic point of view, nutrient availability (Little et al., 2016), absorption (Wu et al., 2020), and soil interactions (Sedlacek et al., 2014) could act as enhancers or limiting factors of the adaptive responses. From a biotic perspective, among-ecotype differentiation (Cortés et al., 2012a,b, 2013; Blair et al., 2016), intrapopulation divergence (Cortés et al., 2011; Blair et al., 2012, 2018; Kelleher et al., 2012), and within-family variation (Galeano et al., 2012; Blair et al., 2013) could encourage or coerce adaptation. Population's functioning, abundance, distribution, and diversity, as predicted from controlled experiments (Way and Oren, 2010; Elmendorf et al., 2012; Wolkovich et al., 2012; Andresen et al., 2016; Becklin et al., 2017; Singh et al., 2017), experimental evolution (Tenaillon et al., 2012; Mallard et al., 2018; Pfenninger and Foucault, 2020), biological monitoring (Walther et al., 2002; Franks et al., 2013; Wipf et al., 2013; Reichstein et al., 2014; Hällfors et al., 2020), and shifts observed in the fossil record (Alsos et al., 2009; Willis and MacDonald, 2011; Lyons et al., 2016; Bruelheide et al., 2018), can feed back on climate change (Pearson et al., 2013) and so be considered as drivers themselves. Regardless of the exact nature and extent of the data type, both GP and ML may offer suitable scenarios to merge diverse, and even conflicting, data sources in order to pinpoint emergent properties (Street et al., 2011) out of a complex system, as is thermal genomic adaptation. Therefore, a key guideline for new developments concerns a better coupling of GP and ML approaches. Until now, only a few works have relied on both methodologies, in the context of thermal adaptation in wheat landraces (Montesinos-Lopez et al., 2018, 2019), but have not gone beyond technical comparisons/recommendations, nor have designed integrated pipelines. Also, reconciling modern genomics with last-generation predictive inferences of the thermal adaptive potential and stochastic demographic modeling (Jenouvrier et al., 2009) is necessary. Open-access resources and data sharing platforms are as crucial in this effort as new integrated analytical pipelines. We are looking forward to seeing more cohesive (Beyer et al., 2020) and systematic studies and predictions across the rich and informative temporal spectrum (Kristensen et al., 2018) of past and future environmental variation (Franks et al., 2013). These efforts should be carried out through a wide range of spatial scales (Parmesan and Hanley, 2015; Way et al., 2015; Gonzalez et al., 2020) spanning contrasting ecosystems (Lenoir et al., 2020), microhabitats (Zellweger et al., 2020), and unrelated taxa, which together may already be keeping heritable adaptive trait differentiation valuable for long-term thermal responses and informative for conservation prioritizations (Barnosky et al., 2017; Elsen et al., 2020).

AUTHOR CONTRIBUTIONS

AC conceived this mini-review. FL-H collected the literature and prepared diagrams. DO-R compiled the historical climate data. AC wrote the first draft of the mini-review with further contributions from FL-H and DO-R. All authors contributed to the article and approved the submitted version.

FUNDING

AC was supported by grants 4.1-2016-00418 and BS2017-0036 from Vetenskapsrådet (VR) and Kungliga Vetenskapsakademien (KVA), respectively. The National Science Foundation (NSF) and the SIMONS Collaboration on the Origins of Life support DO-R. The editorial fund from the Colombian Corporation for Agricultural Research (AGROSAVIA) was thanked for subsidizing the mini-review B-type processing charge.

ACKNOWLEDGMENTS

We acknowledge thoughtful discussions with M. W. Blair regarding the genetic basis of thermal adaptation that took place with AC and DO-R during the Erasmus funded workshop "Molecular Breeding for Abiotic Constraints in Plants" held in Montpellier (France) during the summer of 2012. Some of the ideas discussed here were also framed into perspective, thanks to suggestions from A. A. Hoffmann to AC as part of the "Climate Change and Evolution" symposium during the XIV Congress of the European Society for Evolutionary Biology (ESEB) held at Lisbon (Portugal) in August 2013. The Evolutionary Biology Centre (EBC) Graduate School on Genomes and Phenotypes from Uppsala University is recognized for promoting AC participation in this meeting. AGROSAVIA's Department for Research Capacity Building is credited for granting time to AC to carry out synergistic discussions and progress meetings during 2016 and 2017 in order to pursue this mini-review, as well as for sponsoring FL-H's internship during 2018. We thank D. Royer for the Cenozoic temperature, CO_2, and species richness dataset. Special thanks are given to M. J. Torres-Urrego for support while drafting and revising this mini-review. The topic editor and the two reviewers are recognized for their thoughtful suggestions to improve the scope of the mini-review, as well as for making possible the insightful special issue on "Coping with Climate Change: A Genomic Perspective on Thermal Adaptation."

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.564515/full#supplementary-material

FIGURE S1 | Past and future of thermal and CO_2 variation, and their correlates with past biodiversity. **(A)** Temperature and richness of plant species (from pollen) for the Cenozoic Era (65 Mya – present). Temperature estimates (**Supplementary Table S1**) were computed by Hansen et al. (2013) using the original $\delta^{18}O$ record from Zachos et al. (2008). Richness of plant species from pollen data (**Supplementary Table S1**) is based on 15 Neotropical stratigraphic sections inspected by Jaramillo et al. (2006). This profile goes from 65 to 20 Mya due to a lack of more recent suitable sampling records. **(B)** Projections of the near-surface temperature anomalies to 2,050 (**Supplementary Table S2**), which follow the CIMP5 RCP 8.5 scenario from the KNMI (http://climexp.knmi.nl/) repository averaged from an original 5-min resolution. Light gray shaded areas depict minimum and maximum estimates. **(C)** Atmospheric CO_2 and richness of plant species (as in A) for the Cenozoic Era (65 Mya – present). CO_2 records are an updated version (**Supplementary Table S1**) derived from Royer and Chernoff (2013), originally compiled by Beerling and Royer (2011). **(D)** Projected CO_2 concentration (ppm) to 2,050 also follow the CIMP5 RCP 8.5 scenario, as in B (**Supplementary Table S3**).

TABLE S1 | Dataset of temperature, atmospheric CO_2, and richness of plant species for the Cenozoic Era (65 Mya – present for temperature and CO_2, and 65–20 Mya for richness of plant species). Temperature estimates were computed by Hansen et al. (2013) from five-point running means of the original temporal resolution of the $\delta^{18}O$ record from Zachos et al. (2008), a profile of surface low-magnesium calcitic fossils (including planktonic foraminifera, belemintes, brachiopods, and bivalves) that was lower during periods with warmer seawater. Atmospheric CO_2 corresponds to an updated version from Royer and Chernoff (2013), originally compiled by Beerling and Royer (2011). Richness of plant species is based on pollen data from Jaramillo et al. (2006), who analyzed 1,530 samples from 15 stratigraphic sections in Colombia and Venezuela (Neotropics).

TABLE S2 | Projections of thermal variation to 2,050. Simulation of Near-Surface Air Temperature Anomalies (°C) from 1,860 to 2,050 follow the CIMP5 RCP 8.5 scenario from the KNMI (http://climexp.knmi.nl/) database averaged from an original 5 min resolution. Minimum and maximum temperature estimates were generated by the coupled ACCESS *v.*1.0 model specifically designed for the CIMP5 project (Kowalczyk et al., 2013).

TABLE S3 | Projections of CO_2 concentration (ppm) to 2,050. Simulations follow the CIMP5 RCP 8.5 scenario from 1,860 to 2,050 available at KNMI (http://climexp.knmi.nl/) database averaged from an original 5 min resolution.

REFERENCES

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., et al. (2013). Hybridization and speciation. *J. Evol. Biol.* 26, 229–246.

Abdelmula, A. A., Link, W., Von Kittlitz, E., and Stelling, D. (1999). Heterosis and inheritance of drought tolerance in Faba Bean, *Vicia Faba* L. *Plant Breed.* 118, 485–490. doi: 10.1046/j.1439-0523.1999.00411.x

Alsos, I. G., Alm, T., Normand, S., and Brochmann, C. (2009). Past and future range shifts and loss of diversity in Dwarf Willow (*Salix Herbacea* L.). Inferred from genetics, fossils and modelling. *Glob. Ecol. Biogeogr.* 18, 223–239. doi: 10.1111/j.1466-8238.2008.00439.x

Andresen, L. C., Müller, C., De Dato, G., Dukes, J. S., Emmett, B. A., Estiarte, M., et al. (2016). Shifting impacts of climate change. *Adv. Ecol. Res.* 55, 437–473.

Barghi, N., Hermisson, J., and Schlötterer, C. (2020). Polygenic adaptation: a unifying framework to understand positive selection. *Nat. Rev. Genet.* doi: 10.1038/s41576-020-0250-z

Barnosky, A. D., Hadly, E. A., Gonzalez, P., Head, J., Polly, P. D., Lawing, A. M., et al. (2017). Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems. *Science* 355: eaah4787.

Barton, N., Hermisson, J., and Nordborg, M. (2019). Why structure matters. *eLife* 8:e45380.

Baum, D. A., Smith, S. D., and Donovan, S. S. (2005). The tree-thinking challenge. *Science* 310, 979–970. doi: 10.1126/science.1117727

Becher, H., Jackson, B. C., and Charlesworth, B. (2020). Patterns of genetic variability in genomic regions with low rates of recombination. *Curr. Biol.* 30, 94.e3–100.e3.

Becklin, K. M., Walker, S. M. II, Way, D. A., and Ward, J. K. (2017). Co2 studies remain key to understanding a future world. *New Phytol.* 214, 34–40. doi: 10.1111/nph.14336

Beerli, P. (2006). Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22, 341–345. doi: 10.1093/bioinformatics/bti803

Beerling, D. J., and Royer, D. L. (2011). Convergent cenozoic Co2 history. *Nat. Geosci.* 4, 418–420. doi: 10.1038/ngeo1186

Beyer, R. M., Krapp, M., and Manica, A. (2020). High-resolution terrestrial climate, bioclimate and vegetation for the last 120,000 years. *Sci. Data* 7:236. doi: 10.1038/s41597-020-0552-1

Blair, M. W., Cortes, A. J., Farmer, A. D., Huang, W., Ambachew, D., Penmetsa, R. V., et al. (2018). Uneven recombination rate and linkage disequilibrium

across a reference Snp map for common Bean (*Phaseolus Vulgaris* L.). *PLoS One* 13:e0189597. doi: 10.1371/journal.pone.0189597

Blair, M. W., Cortés, A. J., Penmetsa, R. V., Farmer, A., Carrasquilla-Garcia, N., and Cook, D. R. (2013). A high-throughput snp marker system for parental polymorphism screening, and diversity analysis in common Bean (*Phaseolus Vulgaris* L.). *Theoret. Appl. Genet.* 126, 535–548. doi: 10.1007/s00122-012-1999-z

Blair, M. W., Cortés, A. J., and This, D. (2016). Identification of an *Erecta* gene and its drought adaptation associations with wild and cultivated common Bean. *Plant Sci.* 242, 250–259. doi: 10.1016/j.plantsci.2015.08.004

Blair, M. W., Soler, A., and Cortés, A. J. (2012). Diversification and population structure in common Beans (*Phaseolus Vulgaris* L.). *PLoS One* 7:e49488. doi: 10.1371/journal.pone.0049488

Bridle, J. R., and Vines, T. H. (2007). Limits to evolution at range margins: when and why does adaptation fail? *Trends Ecol. Evol.* 22, 140–147. doi: 10.1016/j.tree.2006.11.002

Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S. M., et al. (2018). Global trait–environment relationships of plant communities. *Nat. Ecol. Evol.* 2, 1906–1917.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., and Roychoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol. Biol. Evol.* 29, 1917–1932. doi: 10.1093/molbev/mss086

Bush, M. B., Silman, M. R., Mcmichael, C., and Saatchi, S. (2008). Fire, climate change and biodiversity in amazonia: a late-holocene perspective. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1795–1702. doi: 10.1098/rstb.2007.0014

Bustamante, C. D., Wakeley, J., Sawyer, S., and Hartl, D. L. (2001). Directional selection and the site-frequency spectrum. *Genetics* 159, 1779–1788.

Caliebe, A. (2008). Mathematical models in population genetics. *Medizinische Genetik* 20, 282–287.

Cardenas, M. L., Gosling, W. D., Sherlock, S. C., Poole, I., Pennington, R. T., and Mothes, P. (2011). The response of vegetation on the Andean Flank in Western Amazonia to pleistocene climate change. *Science* 331, 1055–1058. doi: 10.1126/science.1197947

Chen, I. C., Hill, J. K., Ohlemuller, R., Roy, D. B., and Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science* 333, 1024–1026. doi: 10.1126/science.1206432

Chen, Z., Farrell, A. P., Matala, A., and Narum, S. R. (2018). Mechanisms of thermal adaptation and evolutionary potential of conspecific populations to changing environments. *Mol. Ecol.* 27, 659–674. doi: 10.1111/mec.14475

Chown, S. L., Hoffmann, A. A., Kristensen, T. N., Angilletta, M. J., Stenseth, N. C., and Pertoldi, C. (2010). Adapting to climate change: a perspective from evolutionary physiology. *Clim. Res.* 43, 3–15. doi: 10.3354/cr00879

Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., et al. (2014). Diyabc V2.0: a software to make approximate bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30, 1187–1189. doi: 10.1093/bioinformatics/btt763

Cortés, A. J. (2013). On the origin of the common Bean (*Phaseolus Vulgaris* L.). *Am. J. Plant Sci.* 4, 1998–2000.

Cortés, A. J., and Blair, M. W. (2018). Genotyping by sequencing and genome – environment associations in wild common bean predict widespread divergent adaptation to drought. *Front. Plant Sci.* 9:128. doi: 10.3389/fpls.2018.00128

Cortés, A. J., Chavarro, M. C., and Blair, M. W. (2011). Snp marker diversity in common Bean (*Phaseolus Vulgaris* L.). *Theoret. Appl. Genet.* 123, 827–845. doi: 10.1007/s00122-011-1630-8

Cortés, A. J., Chavarro, M. C., Madriñán, S., This, D., and Blair, M. W. (2012a). Molecular ecology and selection in the drought-related *Asr* gene polymorphisms in wild and cultivated common Bean (*Phaseolus Vulgaris* L.). *BMC Genet.* 13:58. doi: 10.1186/1471-2156-13-58

Cortés, A. J., This, D., Chavarro, C., Madriñán, S., and Blair, M. W. (2012b). Nucleotide diversity patterns at the drought-related *Dreb2* encoding genes in wild and cultivated common Bean (*Phaseolus Vulgaris* L.). *Theoret. Appl. Genet.* 125, 1069–1085. doi: 10.1007/s00122-012-1896-5

Cortés, A. J., Garzón, L. N., Valencia, J. B., and Madriñán, S. (2018a). On the causes of rapid diversification in the páramos: isolation by ecology and genomic divergence in *Espeletia*. *Front. Plant Sci.* 9:1700. doi: 10.3389/fpls.2018.01700

Cortés, A. J., Skeen, P., Blair, M. W., and Chacón-Sánchez, M. I. (2018b). Does the genomic landscape of species divergence in *Phaseolus* Beans coerce parallel signatures of adaptation and domestication? *Front. Plant Sci.* 9:1816. doi: 10.3389/fpls.2018.01816

Cortés, A. J., Monserrate, F., Ramírez-Villegas, J., Madriñán, S., and Blair, M. W. (2013). Drought tolerance in wild plant populations: the case of common Beans (*Phaseolus Vulgaris* L.). *PLoS One* 8:e62898. doi: 10.1371/journal.pone.0062898

Cortés, A. J., Waeber, S., Lexer, C., Sedlacek, J., Wheeler, J. A., Van Kleunen, M., et al. (2014). Small-scale patterns in snowmelt timing affect gene flow and the distribution of genetic diversity in the Alpine Dwarf Shrub *Salix Herbacea*. *Heredity* 113, 233–239. doi: 10.1038/hdy.2014.19

Cortés, A. J., and Wheeler, J. A. (2018). "The environmental heterogeneity of mountains at a fine scale in a changing world," in *Mountains, Climate, and Biodiversity*, eds C. Hoorn, A. Perrigo, and A. Antonelli (New York, NY: Wiley).

Cotto, O., Wessely, J., Georges, D., Klonner, G., Schmid, M., Dullinger, S., et al. (2017). A dynamic eco-evolutionary model predicts slow response of alpine plants to climate warming. *Nat. Commun.* 8:15399.

Coyne, J. A., and Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer.

Cronk, Q. (2016). Plant extinctions take time. *Science* 353, 446–447. doi: 10.1126/science.aag1794

Cros, D., Mbo-Nkoulou, L., Bell, J. M., Oum, J., Masson, A., Soumahoro, M., et al. (2019). Within-family genomic selection in rubber tree (*Hevea Brasiliensis*) increases genetic gain for rubber production. *Ind. Crops Prod.* 138:111464. doi: 10.1016/j.indcrop.2019.111464

Crossa, J., Burgueno, J., Dreisigacker, S., Vargas, M., Herrera-Foessel, S. A., Lillemo, M., et al. (2007). Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* 177, 1889–1913. doi: 10.1534/genetics.107.078659

Crossa, J., Jarquin, D., Franco, J., Perez-Rodriguez, P., Burgueno, J., Saint-Pierre, C., et al. (2016). Genomic prediction of gene bank wheat landraces. *G3* 6, 1819–1834. doi: 10.1534/g3.116.029637

Crossa, J., Martini, J. W. R., Gianola, D., Perez-Rodriguez, P., Jarquin, D., Juliana, P., et al. (2019). Deep Kernel and deep learning for genome-based prediction of single traits in multienvironment breeding trials. *Front. Genet.* 10:1168. doi: 10.3389/fgene.2019.01168

Crossa, J., Perez-Rodriguez, P., Cuevas, J., Montesinos-Lopez, O., Jarquin, D., De Los Campos, G., et al. (2017). Genomic selection in plant breeding: methods, models, and perspectives. *Trends Plant Sci.* 22, 961–975.

Cruickshank, T. E., and Hahn, M. W. (2014). Reanalysis suggests that genomic Islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* 23, 3133–3157. doi: 10.1111/mec.12796

Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., and François, O. (2010). Approximate Bayesian computation (Abc) in practice. *Trends Ecol. Evol.* 25, 410–418. doi: 10.1016/j.tree.2010.04.001

Csillery, K., Rodriguez-Verdugo, A., Rellstab, C., and Guillaume, F. (2018). Detecting the genomic signal of polygenic adaptation and the role of epistasis in evolution. *Mol. Ecol.* 27, 606–612. doi: 10.1111/mec.14499

Cuesta, F., Tovar, C., Llambí, L. D., Gosling, W. D., Halloy, S., Carilla, J., et al. (2019). Thermal Niche traits of high alpine plant species and communities across the tropical andes and their vulnerability to global warming. *J. Biogeogr.* 47, 408–420. doi: 10.1111/jbi.13759

Darwin, C. R. (1874). *The Descent of Man and Selection in Relation to Sex*. New York, NY: Hurst and Company.

de los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., and Calus, M. P. (2013). Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* 193, 327–345. doi: 10.1534/genetics.112.143313

Desta, Z. A., and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19, 592–601. doi: 10.1016/j.tplants.2014.05.006

Ding, W. N., Ree, R. H., Spicer, R. A., and Xing, Y. W. (2020). Ancient orogenic and monsoon-driven assembly of the world's richest temperate alpine flora. *Science* 369, 578–581. doi: 10.1126/science.abb4484

Dolgova, O., Rego, C., Calabria, G., Balanya, J., Pascual, M., Rezende, E. L., et al. (2010). Genetic constraints for thermal coadaptation in *Drosophila Subobscura*. *BMC Evol. Biol.* 10:363. doi: 10.1186/1471-2148-10-363

Ellegren, H., and Galtier, N. (2016). Determinants of genetic diversity. *Nat. Rev. Genet.* 17, 422–433.

Ellegren, H., and Wolf, J. B. W. (2017). Parallelism in genomic landscapes of differentiation, conserved genomic features and the role of linked selection. *J. Evol. Biol.* 30, 1516–1518. doi: 10.1111/jeb.13113

Elmendorf, S. C., Henry, G. H. R., Hollister, R. D., Björk, R. G., Bjorkman, A. D., Callaghan, T. V., et al. (2012). Global assessment of experimental climate warming on tundra vegetation: heterogeneity over space and time. *Ecol. Lett.* 15, 164–175. doi: 10.1111/j.1461-0248.2011.01716.x

Elsen, P. R., Monahan, W. B., Dougherty, E. R., and Merenlender, A. M. (2020). Keeping pace with climate change in global terrestrial protected areas. *Sci. Adv.* 6:eaay0814. doi: 10.1126/sciadv.aay0814

Exposito-Alonso, M., Burbano, H. A., Bossdorf, O., Nielsen, R., and Weigel, D. (2019). Natural selection on the *Arabidopsis thaliana* genome in present and future climates. *Nature* 573, 126–129. doi: 10.1038/s41586-019-1520-9 doi: 10.1038/s41586-019-1520-9

Falconer, D. S., and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Essex: Longman.

Feder, J. L., Gejji, R., Yeaman, S., and Nosil, P. (2012). Establishment of new mutations under divergence and genome hitchhiking. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 461–474. doi: 10.1098/rstb.2011.0256

Feder, J. L., and Nosil, P. (2010). The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution* 64, 1729–1747. doi: 10.1111/j.1558-5646.2009.00943.x

Feeley, K., Martinez-Villa, J., Perez, T., Silva Duque, A., Triviño Gonzalez, D., and Duque, A. (2020). The thermal tolerances, distributions, and performances of tropical montane tree species. *Front. For. Glob. Change* 3:25. doi: 10.3389/ffgc.2020.00025

Fitzpatrick, S. W., Gerberich, J. C., Kronenberger, J. A., Angeloni, L. M., and Funk, W. C. (2015). Locally adapted traits maintained in the face of high gene flow. *Ecol. Lett.* 18, 37–47. doi: 10.1111/ele.12388

Flantua, S. G. A., O'dea, A., Onstein, R. E., Giraldo, C., and Hooghiemstra, H. (2019). The flickering connectivity system of the North Andean Páramos. *J. Biogeogr.* 46, 1808–1825. doi: 10.1111/jbi.13607

Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., and Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Mol. Ecol.* 25, 104–120. doi: 10.1111/mec.13476

Fragata, I., Lopes-Cunha, M., Barbaro, M., Kellen, B., Lima, M., Santos, M. A., et al. (2014). How much can history constrain adaptive evolution? A real-time evolutionary approach of inversion polymorphisms in *Drosophila Subobscura*. *J. Evol. Biol.* 27, 2727–2738. doi: 10.1111/jeb.12533

Fragata, I., Simões, P., Matos, M., Szathmáry, E., and Santos, M. (2018). playing evolution in the laboratory: from the first major evolutionary transition to global warming. *Europhys. Lett.* 122:38001. doi: 10.1209/0295-5075/122/38001

Franks, P. J., Adams, M. A., Amthor, J. S., Barbour, M. M., Berry, J. A., Ellsworth, D. S., et al. (2013). Sensitivity of plants to changing atmospheric Co2 concentration: from the geological past to the next century. *New Phytol.* 197, 1077–1094. doi: 10.1111/nph.12104

Franks, S. J., and Hoffmann, A. A. (2012). Genetics of climate change adaptation. *Annu. Rev. Genet.* 46, 185–208.

Franks, S. J., Weber, J. J., and Aitken, S. N. (2014). Evolutionary and plastic responses to climate change in terrestrial plant populations. *Evol. Appl.* 7, 123–139. doi: 10.1111/eva.12112

Galeano, C. H., Cortés, A. J., Fernandez, A. C., Soler, A., Franco-Herrera, N., Makunde, G., et al. (2012). Gene-based single nucleotide polymorphism markers for genetic and association mapping in common Bean. *BMC Genet.* 13:48. doi: 10.1186/1471-2156-13-48

Garah, K., and Bentouati, A. (2019). Using the maxent model for assessing the impact of climate change on the Aurasian Aleppo pine distribution in algeria. *Afr. J. Ecol.* 57, 500–511. doi: 10.1111/aje.12630

Garner, J. B., Douglas, M. L., Williams, S. R., Wales, W. J., Marett, L. C., Nguyen, T. T., et al. (2016). Genomic selection improves heat tolerance in dairy cattle. *Sci. Rep.* 6:34114.

Gianola, D., Fernando, R. L., and Stella, A. (2006). Genomic-assisted prediction of genetic value with semiparametric procedures. *Genetics* 173, 1761–1776. doi: 10.1534/genetics.105.049510

Gibson, M. J. S., and Moyle, L. C. (2020). Regional differences in the abiotic environment contribute to genomic divergence within a wild tomato species. *Mol. Ecol.* 29, 2204–2217. doi: 10.1111/mec.15477

Gompert, Z., Comeault, A. A., Farkas, T. E., Feder, J. L., Parchman, T. L., Buerkle, C. A., et al. (2014). Experimental evidence for ecological selection on genome variation in the wild. *Ecol. Lett.* 17, 369–379. doi: 10.1111/ele.12238

Gonzalez, A., Germain, R. M., Srivastava, D. S., Filotas, E., Dee, L. E., Gravel, D., et al. (2020). Scaling-up biodiversity-ecosystem functioning research. *Ecol. Lett.* 23, 757–776.

Grattapaglia, D., Silva-Junior, O. B., Resende, R. T., Cappa, E. P., Muller, B. S. F., Tan, B., et al. (2018). Quantitative genetics and genomics converge to accelerate forest tree breeding. *Front. Plant Sci.* 9:1693. doi: 10.3389/fpls.2018.01693

Hällfors, M. H., Antao, L. H., Itter, M., Lehikoinen, A., Lindholm, T., Roslin, T., et al. (2020). Shifts in timing and duration of breeding for 73 boreal bird species over four decades. *PNAS* 117, 18557–18565. doi: 10.1073/pnas.1913579117

Hancock, A. M., Brachi, B., Faure, N., Horton, M. W., Jarymowycz, L. B., Sperone, F. G., et al. (2011). Adaptation to climate across the *Arabidopsis Thaliana* genome. *Science* 334, 83–86. doi: 10.1126/science.1209244

Hansen, J., Sato, M., Russell, G., and Kharecha, P. (2013). Climate sensitivity, sea level and atmospheric carbon dioxide. *Philos. Trans. A Math. Phys. Eng. Sci.* 371:20120294. doi: 10.1098/rsta.2012.0294

Hirschhorn, J. N., and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* 6, 95–108. doi: 10.1038/nrg1521

Hoffmann, A. A., and Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature* 470, 479–485.

Huber, C. D., Degiorgio, M., Hellmann, I., and Nielsen, R. (2016). Detecting recent selective sweeps while controlling for mutation rate and background selection. *Mol. Ecol.* 25, 142–156. doi: 10.1111/mec.13351

Hughes, C. (2006). From the cover: Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the andes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10334–10339. doi: 10.1073/pnas.0601928103

Irwin, D. E., Alcaide, M., Delmore, K. E., Irwin, J. H., and Owens, G. L. (2016). Recurrent selection explains parallel evolution of genomic regions of high relative but low absolute differentiation in a ring species. *Mol. Ecol.* 25, 4488–4507. doi: 10.1111/mec.13792

Isabel, N., Holliday, J. A., and Aitken, S. N. (2020). Forest genomics: advancing climate adaptation, forest health, productivity, and conservation. *Evol. Appl.* 13, 3–10. doi: 10.1111/eva.12902

Jaramillo, C., Rueda, M. J., and Mora, G. (2006). Cenozoic plant diversity in the neotropics. *Science* 311, 1893–1896. doi: 10.1126/science.1121380

Jay, F., Manel, S., Alvarez, N., Durand, E. Y., Thuiller, W., Holderegger, R., et al. (2012). Forecasting changes in population genetic structure of alpine plants in response to global warming. *Mol. Ecol.* 21, 2354–2368. doi: 10.1111/j.1365-294x.2012.05541.x

Jenouvrier, S., Caswell, H., Barbraud, C., Holland, M., Stroeve, J., and Weimerskirch, H. (2009). Demographic models and Ipcc climate projections predict the decline of an emperor penguin population. *PNAS* 106, 1844–1847. doi: 10.1073/pnas.0806638106

Jezkova, T., and Wiens, J. J. (2016). Rates of change in climatic niches in plant and animal populations are much slower than projected climate change. *Proc. Biol. Sci.* 283:20162104. doi: 10.1098/rspb.2016.2104

Juliana, P., Montesinos-López, O. A., Crossa, J., Mondal, S., González Pérez, L., Poland, J., et al. (2019). Integrating genomic-enabled prediction and high-throughput phenotyping in breeding for climate-resilient bread wheat. *Theoret. Appl. Genet.* 132, 177–194. doi: 10.1007/s00122-018-3206-3

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M. C. (2012). Experimental evolution. *Trends Ecol. Evol.* 27, 547–560.

Kelleher, C. T., Wilkin, J., Zhuang, J., Cortés, A. J., Quintero, ÁL. P., Gallagher, T. F., et al. (2012). Snp discovery, gene diversity, and linkage disequilibrium in wild populations of *Populus Tremuloides*. *Tree Genet. Genomes* 8, 821–829. doi: 10.1007/s11295-012-0467-x

Kowalczyk, E. A., Stevens, L., Law, R. M., Dix, M., Wang, Y. P., Harman, I. N., et al. (2013). The land surface model component of access: description and impact on the simulated surface climatology. *Aust. Meteorol. Oceanogr.* 63, 65–82. doi: 10.22499/2.6301.005

Kristensen, T. N., Ketola, T., and Kronholm, I. (2018). Adaptation to environmental stress at different timescales. *Ann. N. Y. Acad. Sci.* [Epub ahead of print] doi: 10.1111/nyas.13974

Laayouni, H., Garcia-Franco, F., Chavez-Sandoval, B. E., Trotta, V., Beltran, S., Corominas, M., et al. (2007). Thermal evolution of gene expression profiles in *Drosophila Subobscura*. *BMC Evol. Biol.* 7:42. doi: 10.1186/1471-2148-7-42

Lambert, C. G., and Black, L. J. (2012). Learning from our gwas mistakes: from experimental design to scientific method. *Biostatistics* 13, 195–203. doi: 10.1093/biostatistics/kxr055

Lecheta, M. C., Awde, D. N., O'leary, T. S., Unfried, L. N., Jacobs, N. A., Whitlock, M. H., et al. (2020). Integrating Gwas and transcriptomics to identify the molecular underpinnings of thermal stress responses in *Drosophila Melanogaster*. *Front. Genet.* 11:658. doi: 10.3389/fgene.2020.00658

Lehnert, S. J., Kess, T., Bentzen, P., Kent, M. P., Lien, S., Gilbey, J., et al. (2019). Genomic signatures and correlates of widespread population declines in salmon. *Nat. Commun.* 10:2996.

Leinonen, P. H., Remington, D. L., and Savolainen, O. (2011). Local adaptation, phenotypic differentiation, and hybrid fitness in diverged natural populations of *Arabidopsis Lyrata*. *Evolution* 65, 90–107. doi: 10.1111/j.1558-5646.2010.01119.x

Lenoir, J., Bertrand, R., Comte, L., Bourgeaud, L., Hattab, T., Murienne, J., et al. (2020). Species better track climate warming in the oceans than on land. *Nat. Ecol. Evol.* 4, 1044–1059. doi: 10.1038/s41559-020-1198-2

Libbrecht, M. W., and Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nat. Rev. Genet.* 16, 321–332. doi: 10.1038/nrg3920

Little, C. J., Wheeler, J. A., Sedlacek, J., Cortés, A. J., and Rixen, C. (2016). Small-scale drivers: the importance of nutrient availability and snowmelt timing on performance of the Alpine Shrub *Salix Herbacea*. *Oecologia* 180, 1015–1024. doi: 10.1007/s00442-015-3394-3

Liu, S., Cornille, A., Decroocq, S., Tricon, D., Chague, A., Eyquard, J. P., et al. (2019). The complex evolutionary history of apricots: species divergence, gene flow and multiple domestication events. *Mol. Ecol.* 28, 5299–5314. doi: 10.1111/mec.15296

Logan, M. L., and Cox, C. L. (2020). Genetic constraints, transcriptome plasticity, and the evolutionary response to climate change. *Front. Genet.* doi: 10.3389/fgene.2020.538226

López-Hernández, F., and Cortés, A. J. (2019). Last-generation genome–environment associations reveal the genetic basis of heat tolerance in common Bean (*Phaseolus Vulgaris* L.). *Front. Genet.* 10:22. doi: 10.3389/fgene.2019.00954

Lotterhos, K. E., and Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* 24, 1031–1046. doi: 10.1111/mec.13100

Lstiburek, M., Schueler, S., El-Kassaby, Y. A., Hodge, G. R., Stejskal, J., Korecky, J., et al. (2020). *In Situ* genetic evaluation of european larch across climatic regions using marker-based pedigree reconstruction. *Front. Genet.* 11:28. doi: 10.3389/fgene.2020.00028

Lynch, M., and Ritland, K. (1999). Estimation of pairwise relatedness with molecular markers. *Genetics* 152, 1753–1766.

Lyons, S. K., Amatangelo, K. L., Behrensmeyer, A. K., Bercovici, A., Blois, J. L., Davis, M., et al. (2016). Holocene shifts in the assembly of plant and animal communities implicate human impacts. *Nature* 529, 80–83. doi: 10.1038/nature16447

Madriñán, S., Cortés, A. J., and Richardson, J. E. (2013). Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Front. Genet.* 4:192. doi: 10.3389/fgene.2013.00192

Maher, B. (2008). Missing heritability. *Nature* 456, 18–21. doi: 10.1038/456018a

Mallard, F., Nolte, V., Tobler, R., Kapun, M., and Schlötterer, C. (2018). A simple geneticbasis of adaptation to a novel thermal environment results in complex metabolic rewiring in *Drosophila*. A simple genetic basis of adaptation to a novel thermal environment results in complex metabolic rewiring in Drosophila. *Genome Biol.* 19:119.

Marques, D. A., Meier, J. I., and Seehausen, O. (2019). A combinatorial view on speciation and adaptive radiation. *Trends Ecol. Evol.* 34, 531–544. doi: 10.1016/j.tree.2019.02.008

Matos, M., Simoes, P., Santos, M. A., Seabra, S. G., Faria, G. S., Vala, F., et al. (2015). History, chance and selection during phenotypic and genomic experimental evolution: replaying the tape of life at different levels. *Front. Genet.* 6:71. doi: 10.3389/fgene.2015.00071

Maynard Smith, J., and Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genet. Res.* 23, 23–35. doi: 10.1017/s0016672300014634

McKown, A. D., Klapste, J., Guy, R. D., El-Kassaby, Y. A., and Mansfield, S. D. (2018). Ecological genomics of variation in bud-break phenology and mechanisms of response to climate warming in *Populus Trichocarpa*. *New Phytol.* 220, 300–316. doi: 10.1111/nph.15273

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.

Miller, A. D., Coleman, M. A., Clark, J., Cook, R., Naga, Z., Doblin, M. A., et al. (2020). Local thermal adaptation and limited gene flow constrain future climate responses of a marine ecosystem engineer. *Evol. Appl.* 13, 918–934. doi: 10.1111/eva.12909

Mohapatra, J., Singh, C., Hamid, M., Verma, A., Semwal, S. C., Gajmer, B., et al. (2019). Modelling *Betula Utilis* distribution in response to climate-warming scenarios in Hindu-Kush Himalaya using random forest. *Biodiv. Conserv.* 28, 2295–2317. doi: 10.1007/s10531-019-01731-w

Montesinos-Lopez, A., Montesinos-Lopez, O. A., Gianola, D., Crossa, J., and Hernandez-Suarez, C. M. (2018). Multi-environment genomic prediction of plant traits using deep learners with dense architecture. *G3* 8, 3813–3828.

Montesinos-Lopez, O. A., Martin-Vallejo, J., Crossa, J., Gianola, D., Hernandez-Suarez, C. M., Montesinos-Lopez, A., et al. (2019). A benchmarking between deep learning, support vector machine and bayesian threshold best linear unbiased prediction for predicting ordinal traits in plant breeding. *G3* 9, 601–618.

Morris, B. B. G. P., and Borevitz, J. O. (2011). Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biol.* 12:232. doi: 10.1186/gb-2011-12-10-232

Nachman, M. W., and Payseur, B. A. (2012). Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 409–421. doi: 10.1098/rstb.2011.0249

Nadachowska-Brzyska, K., Burri, R., Smeds, L., and Ellegren, H. (2016). Psmc analysis of effective population sizes in molecular ecology and its application to black-and-white *Ficedula* Flycatchers. *Mol. Ecol.* 25, 1058–1072. doi: 10.1111/mec.13540

Nathan, R., and Muller-Landau, H. C. (2000). Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends Ecol. Evol.* 15, 278–285. doi: 10.1016/s0169-5347(00)01874-7

Nei, M. (1987). *Molecular Evolutionary Genetics*. New York, NY: Columbia University Press.

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., et al. (2010). Plant phenotypic plasticity in a changing climate. *Trends Plant Sci.* 15, 684–692. doi: 10.1016/j.tplants.2010.09.008

North, A., Pennanen, J., Ovaskainen, O., and Laine, A.-L. (2011). Local adaptation in a changing world: the roles of gene-flow, mutation, and sexual reproduction. *Evolution* 65, 79–89. doi: 10.1111/j.1558-5646.2010.01107.x

Nosil, P., and Feder, J. L. (2011). Genomic divergence during speciation: causes and consequences. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 332–342. doi: 10.1098/rstb.2011.0263

Ørsted, M., Hoffmann, A. A., Rohde, P. D., Sørensen, P., and Kristensen, T. N. (2019). Strong impact of thermal environment on the quantitative genetic basis of a key stress tolerance trait. *Heredity* 122, 315–325. doi: 10.1038/s41437-018-0117-7

Otto, S. P., and Day, T. (2007). *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton, NJ: Princeton University Press.

Park, D. S., Willis, C. G., Xi, Z., Kartesz, J. T., Davis, C. C., and Worthington, S. (2020). Machine learning predicts large scale declines in native plant phylogenetic diversity. *New Phytol.* 227, 1544–1556. doi: 10.1111/nph.16621

Parmesan, C., and Hanley, M. E. (2015). Plants and climate change: complexities and surprises. *Ann. Bot.* 116, 849–864. doi: 10.1093/aob/mcv169

Payseur, B. A., and Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Mol. Ecol.* 25, 2337–2360. doi: 10.1111/mec.13557

Pearson, R. G., Phillips, S. J., Loranty, M. M., Beck, P. S. A., Damoulas, T., Knight, S. J., et al. (2013). Shifts in arctic vegetation and associated feedbacks under climate change. *Nat. Clim. Change* 3, 673–677. doi: 10.1038/nclimate1858

Pennisi, E. (2014). Disputed Islands. *Science* 345, 611–613. doi: 10.1126/science.345.6197.611

Pfenninger, M., and Foucault, Q. (2020). Genomic processes underlying rapid adaptation of a natural *Chironomus Riparius* population to unintendedly

applied experimental selection pressures. *Mol. Ecol.* 29, 536–548. doi: 10.1111/mec.15347

Phillips, S. J., Anderson, R. P., Dudiík, M., Schapire, R. E., and Blair, M. E. (2017). Opening the black box: An open-source release of maxent. *Ecography* 40, 887–893. doi: 10.1111/ecog.03049 doi: 10.1111/ecog.03049 doi: 10.1111/ecog.03049

Pluess, A. R., Frank, A., Heiri, C., Lalague, H., Vendramin, G. G., and Oddou-Muratorio, S. (2016). Genome-environment association study suggests local adaptation to climate at the regional scale in *Fagus Sylvatica*. *New Phytol.* 210, 589–601. doi: 10.1111/nph.13809

Pritchard, J. K., Pickrell, J. K., and Coop, G. (2010). The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* 20, R208–R215.

Quintero, I., and Wiens, J. J. (2013). Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species. *Ecol. Lett.* 16, 1095–1103. doi: 10.1111/ele.12144

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlovic, M., et al. (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J. Evol. Biol.* 30, 1450–1477. doi: 10.1111/jeb.13047

Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibanez, C., et al. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *PNAS* 116, 10418–10423. doi: 10.1073/pnas.1820663116

Reichstein, M., Bahn, M., Mahecha, M. D., Kattge, J., and Baldocchi, D. D. (2014). Linking plant and ecosystem functional biogeography. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13697–13702. doi: 10.1073/pnas.1216065111

Resende, M. D. V. Jr., Resende, M. F. R., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A., Aguiar, A. M., et al. (2012). Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytol.* 194, 116–128. doi: 10.1111/j.1469-8137.2011.04038.x

Rezende, E. L., Tejedo, M., and Santos, M. (2010). Estimating the adaptive potential of critical thermal limits: methodological problems and evolutionary implications. *Funct. Ecol.* 25, 111–121. doi: 10.1111/j.1365-2435.2010.01778.x

Richardson, J. E., Madriñán, S., Gómez-Gutiérrez, M. C., Valderrama, E., Luna, J., Banda-R, K., et al. (2019). Using dated molecular phylogenies to help reconstruct geological, climatic, and biological history: Examples from Colombia. *Geol. J.* 53, 2935–2943. doi: 10.1002/gj.3133

Rippke, U., Ramirez-Villegas, J., Jarvis, A., Vermeulen, S. J., Parker, L., Mer, F., et al. (2016). Timescales of transformational climate change adaptation in Sub-Saharan African agriculture. *Nat. Clim. Change* 6, 605–609. doi: 10.1038/nclimate2947

Ripple, W. J., Wolf, C., Newsome, T. M., Barnard, P., and Moomaw, W. R. (2020). World Scientists' warning of a climate emergency. *BioScience* 70, 8–12.

Rodriguez-Trelles, F., Tarrio, R., and Santos, M. (2013). Genome-wide evolutionary response to a heat wave in *Drosophila*. *Biol. Lett.* 9:20130228. doi: 10.1098/rsbl.2013.0228

Roesti, M., Gavrilets, S., Hendry, A. P., Salzburger, W., and Berner, D. (2014). The genomic signature of parallel adaptation from shared genetic variation. *Mol. Ecol.* 23, 3944–3956. doi: 10.1111/mec.12720

Royer, D. L., and Chernoff, B. (2013). Diversity in neotropical wet forests during the cenozoic is linked more to atmospheric Co2 than temperature. *Proc. Biol. Sci.* 280:20131024. doi: 10.1098/rspb.2013.1024

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918. doi: 10.1038/nature06250

Saint Pierre, C., Crossa, J. L., Bonnett, D., Yamaguchi-Shinozaki, K., and Reynolds, M. P. (2012). Phenotyping transgenic wheat for drought resistance. *J. Exp. Bot.* 63, 1799–1808. doi: 10.1093/jxb/err385

Santos, M., Castaneda, L. E., and Rezende, E. L. (2012). Keeping pace with climate change: what is wrong with the evolutionary potential of upper thermal limits? *Ecol. Evol.* 2, 2866–2880. doi: 10.1002/ece3.385

Scherrer, D., and Körner, C. (2011). Topogaphically controlled thermal-habitat differentiation buffers Alpine plant diversity against climate warming. *J. Biogeogr.* 38, 406–416. doi: 10.1111/j.1365-2699.2010.02407.x

Schrider, D. R., and Kern, A. D. (2018). Supervised machine learning for population genetics: a new paradigm. *Trends Genet.* 34, 301–312. doi: 10.1016/j.tig.2017.12.005

Seabra, S. G., Fragata, I., Antunes, M. A., Faria, G. S., Santos, M. A., Sousa, V. C., et al. (2017). Different genomic changes underlie adaptive evolution in populations of contrasting history. *Mol. Biol. Evol.* 35, 549–563. doi: 10.1093/molbev/msx247

Sedlacek, J., Bossdorf, O., Cortés, A. J., Wheeler, J. A., and Van-Kleunen, M. (2014). What role do plant-soil interactions play in the habitat suitability and potential range expansion of the Alpine Dwarf Shrub *Salix Herbacea*? *Basic Appl. Ecol.* 15, 305–315. doi: 10.1016/j.baae.2014.05.006

Sedlacek, J., Cortés, A. J., Wheeler, J. A., Bossdorf, O., Hoch, G., Klapste, J., et al. (2016). Evolutionary potential in the Alpine: trait heritabilities and performance variation of the Dwarf Willow *Salix Herbacea* from different elevations and microhabitats. *Ecol. Evol.* 6, 3940–3952. doi: 10.1002/ece3.2171

Sedlacek, J., Wheeler, J. A., Cortés, A. J., Bossdorf, O., Hoch, G., Lexer, C., et al. (2015). The response of the Alpine Dwarf Shrub *Salix Herbacea* to altered snowmelt timing: lessons from a multi-site transplant experiment. *PLoS One* 10:e0122395. doi: 10.1371/journal.pone.0122395

Sentinella, A. T., Warton, D. I., Sherwin, W. B., Offord, C. A., Moles, A. T., and Wang, Z. (2020). Tropical plants do not have narrower temperature tolerances, but are more at risk from warming because they are close to their upper thermal limits. *Glob. Ecol. Biogeogr.* 29, 1387–1398. doi: 10.1111/geb.13117

Simões, J. L., Gouvêa, N., and Margis, M. (2003). Small heat shock proteins genes are differentially expressed in distinct varieties of common Bean. *Braz. J. Plant Physiol.* 15, 33–41. doi: 10.1590/s1677-04202003000100005

Simões, P., and Pascual, M. (2018). Patterns of geographic variation of thermal adapted candidate genes in *Drosophila Subobscura* sex chromosome arrangements. *BMC Evol. Biol.* 18:60. doi: 10.1186/s12862-018-1178-1

Singh, R. K., Svystun, T., Aldahmash, B., Jonsson, A. M., and Bhalerao, R. P. (2017). Photoperiod- and temperature-mediated control of phenology in trees - a molecular perspective. *New Phytol.* 213, 511–524. doi: 10.1111/nph.14346

Slatkin, M. (2008). Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* 9, 477–485. doi: 10.1038/nrg2361

Sørensen, J. G., Kristensen, T. N., and Loeschcke, V. (2003). The evolutionary and ecological role of heat shock proteins. *Ecol. Lett.* 6, 1025–1037. doi: 10.1046/j.1461-0248.2003.00528.x

Spindel, J. E., Begum, H., Akdemir, D., Collard, B., Redona, E., Jannink, J. L., et al. (2016). Genome-wide prediction models that incorporate de novo gwas are a powerful new tool for tropical rice improvement. *Heredity* 116, 395–408. doi: 10.1038/hdy.2015.113

Stanton, M. L., Galen, C., and Shore, J. (1997). Population structure along a steep environmental gradient: consequences of flowering time and habitat variation in the snow buttercup, *Ranunculus Adoneus*. *Evolution* 51, 79–94. doi: 10.2307/2410962

Steinbauer, M. J., Grytnes, J. A., Jurasinski, G., Kulonen, A., Lenoir, J., Pauli, H., et al. (2018). Accelerated increase in plant species richness on mountain summits is linked to warming. *Nature* 556, 231–234.

Stephan, W. (2016). Signatures of positive selection: from selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. *Mol. Ecol.* 25, 79–88. doi: 10.1111/mec.13288

Stölting, K. N., Paris, M., Meier, C., Heinze, B., Castiglione, S., Bartha, D., et al. (2015). Genome-wide patterns of differentiation and spatially varying selection between postglacial recolonization lineages of *Populus Alba* (Salicaceae), a widespread forest tree. *New Phytol.* 207, 723–734. doi: 10.1111/nph.13392

Street, N. R., Jansson, S., and Hvidsten, T. (2011). A systems biology model of the regulatory network in *Populus* leaves reveals interacting regulators and conserved regulation. *BMC Plant Biol.* 11:13. doi: 10.1186/1471-2229-11-13

Sukumaran, S., Crossa, J., Jarquin, D., Lopes, M., and Reynolds, M. P. (2017). Genomic prediction with pedigree and genotype x environment interaction in spring wheat grown in South and West Asia, North Africa, and Mexico. *G3* 7, 481–495. doi: 10.1534/g3.116.036251

Suontama, M., KláPšTě, J., Telfer, E., Graham, N., Stovold, T., Low, C., et al. (2019). Efficiency of genomic prediction across two *Eucalyptus Nitens* seed orchards with different selection histories. *Heredity* 122, 370–379. doi: 10.1038/s41437-018-0119-5

Swarts, K., Gutaker, R. M., Benz, B., Blake, M., Bukowski, R., Holland, J., et al. (2017). Genomic estimation of complex traits reveals ancient maize adaptation to temperate North America. *Science* 357, 512–515. doi: 10.1126/science.aam9425

Sweet, L. C., Green, T., Heintz, J. G. C., Frakes, N., Graver, N., Rangitsch, J. S., et al. (2019). Congruence between future distribution models and empirical data for an iconic species at Joshua Tree National Park. *Ecosphere* 10:e0276. doi: 10.1002/ecs2.2763

Tam, V., Patel, N., Turcotte, M., Bosseì, Y., Pareì, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 20, 467–484. doi: 10.1038/s41576-019-0127-1

Tan, B., Grattapaglia, D., Martins, G. S., Ferreira, K. Z., Sundberg, B. R., and Ingvarsson, P. R. K. (2017). Evaluating the accuracy of genomic prediction of growth and wood traits in two eucalyptus species and their F1 hybrids. *BMC Plant Biol.* 17:110. doi: 10.1186/s12870-017-1059-6

Technow, F., Schrag, T. A., Schipprack, W., Bauer, E., Simianer, H., and Melchinger, A. E. (2014). Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. *Genetics* 197, 1343–1355. doi: 10.1534/genetics.114.165860

Tenaillon, O., Rodriguez-Verdugo, A., Gaut, R. L., Mcdonald, P., Bennett, A. F., Long, A. D., et al. (2012). The molecular diversity of adaptive convergence. *Science* 335, 457–461. doi: 10.1126/science.1212986

Turner, T. L., and Miller, P. M. (2012). Investigating natural variation in drosophila courtship song by the evolve and resequence approach. *Genetics* 191, 633–642. doi: 10.1534/genetics.112.139337

Valencia, J. B., Mesa, J., León, J. G., Madriñán, S., and Cortés, A. J. (2020). Climate vulnerability assessment of the *Espeletia* complex in Páramo sky islands of the northern Andes. *Front. Ecol. Evol.* doi: 10.3389/fevo.2020.565708

Valladares, F., Matesanz, S., Guilhaumon, F., Araujo, M. B., Balaguer, L., Benito-Garzon, M., et al. (2014). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecol. Lett.* 17, 1351–1364. doi: 10.1111/ele.12348

Vargas, O. M., Ortiz, E. M., and Simpson, B. B. (2017). Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-andean diversification (Asteraceae: Astereae: *Diplostephium*). *New Phytol.* 214, 1736–1750. doi: 10.1111/nph.14530

Visser, J. A. G. M. D., Elena, S. F., Fragata, I. S., and Matuszewski, S. (2018). The utility of fitness landscapes and big data for predicting evolution. *Heredity* 121, 401–405. doi: 10.1038/s41437-018-0128-4

Wakeley, J. (2008). *Coalescent Theory: An Introduction*. Cambridge: Harvard University.

Waldvogel, A. M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R., et al. (2020a). Evolutionary genomics can improve prediction of species' responses to climate change. *Evol. Lett.* 4, 4–18. doi: 10.1002/evl3.154

Waldvogel, A. M., Schreiber, D., Pfenninger, M., and Feldmeyer, B. (2020b). Climate change genomics calls for standardised data reporting. *Front. Ecol. Evol.* 8:242. doi: 10.3389/fevo.2020.00242

Walsh, B. (2008). Evolutionary quantitative genetics. *Handb. Stat. Genet*. 1, 533–586.

Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., et al. (2002). Ecological responses to recent climate change. *Nature* 416, 389–395.

Wang, X., Edwards, R. L., Auler, A. S., Cheng, H., Kong, X., Wang, Y., et al. (2017). Hydroclimate changes across the amazon lowlands over the past 45,000 years. *Nature* 541, 204–207. doi: 10.1038/nature20787

Way, D. A., and Oren, R. (2010). Differential responses to changes in growth temperature between trees from different functional groups and biomes: a review and synthesis of data. *Tree Physiol.* 30, 669–688. doi: 10.1093/treephys/tpq015

Way, D. A., Oren, R., and Kroner, Y. (2015). The space-time continuum: the effects of elevated Co2 and temperature on trees and the importance of scaling. *Plant Cell Environ.* 38, 991–1007. doi: 10.1111/pce.12527

Weir, B. S., and Cockerham, C. (1984). Estimating F-Statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.2307/2408641

Wheeler, J. A., Cortés, A. J., Sedlacek, J., Karrenberg, S., Van Kleunen, M., Wipf, S., et al. (2016). The snow and the willows: accelerated spring snowmelt reduces performance in the low-lying Alpine Shrub *Salix Herbacea*. *J. Ecol.* 104, 1041–1050. doi: 10.1111/1365-2745.12579

Wheeler, J. A., Hoch, G., Cortés, A. J., Sedlacek, J., Wipf, S., and Rixen, C. (2014). Increased Spring freezing vulnerability for Alpine Shrubs under early snowmelt. *Oecologia* 175, 219–229. doi: 10.1007/s00442-013-2872-8

Wheeler, J. A., Schnider, F., Sedlacek, J., Cortés, A. J., Wipf, S., Hoch, G., et al. (2015). With a little help from my friends: community facilitation increases performance in the Dwarf Shrub *Salix Herbacea*. *Basic Appl. Ecol.* 16, 202–209. doi: 10.1016/j.baae.2015.02.004

Whitlock, C., and Bartlein, P. J. (2003). holocene fire activity as a record of past environmental change. *Dev. Quatern. Sci.* 1, 479–490. doi: 10.1016/s1571-0866(03)01022-4

Willis, C. G., Ruhfel, B., Primack, R. B., Miller-Rushing, A. J., and Davis, C. C. (2008). Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. *PNAS* 105, 17029–17033. doi: 10.1073/pnas.0806446105

Willis, K. J., and MacDonald, G. M. (2011). Long-term ecological records and their relevance to climate change predictions for a warmer world. *Annu. Rev. Ecol. Evol. Syst.* 42, 267–287. doi: 10.1146/annurev-ecolsys-102209-144704

Wipf, S., Stöckli, V., Herz, K., and Rixen, C. (2013). The oldest monitoring site of the alps revisited: accelerated increase in plant species Richness on Piz Linard summit since 1835. *Plant Ecol. Divers.* 6, 447–455. doi: 10.1080/17550874.2013.764943

Wolf, J. B., and Ellegren, H. (2017). Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* 18, 87–100. doi: 10.1038/nrg.2016.133

Wolkovich, E. M., Cook, B. I., Allen, J. M., Crimmins, T. M., Betancourt, J. L., Travers, S. E., et al. (2012). Warming experiments underpredict plant phenological responses to climate change. *Nature* 485, 494–497. doi: 10.1038/nature11014

Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., and Visscher, P. M. (2013). Pitfalls of predicting complex traits from Snps. *Nat. Rev. Genet.* 14, 507–515. doi: 10.1038/nrg3457

Wu, X., Islam, A. S. M. F., Limpot, N., Mackasmiel, L., Mierzwa, J., Cortés, A. J., et al. (2020). Genome-wide Snp identification and association mapping for seed mineral concentration in Mung Bean (*Vigna Radiata* L.). *Front. Genet.* 11:656. doi: 10.3389/fgene.2020.00656

Yamagishi, H., Allison, T. D., and Ohara, M. (2005). Effect of snowmelt timing on the genetic structure of an *Erythronium Grandiflorum* population in an Alpine environment. *Ecol. Res.* 20, 199–204. doi: 10.1007/s11284-004-0032-7

Yeaman, S., and Otto, S. P. (2011). Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution* 65, 2123–2129. doi: 10.1111/j.1558-5646.2011.01277.x

Zachos, J. C., Dickens, G. R., and Zeebe, R. E. (2008). An early cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451, 279–283. doi: 10.1038/nature06588

Zahn, L. M., and Purnell, B. A. (2016). Genes under pressure. *Science* 354:52. doi: 10.1126/science.354.6308.52

Zellweger, F., De Frenne, P., Lenoir, J., Vangansbeke, P., Verheyen, K., Bernhardt-Römermann, M., et al. (2020). Forest microclimate dynamics drive plant responses to warming. *Science* 368, 772–775.

Zhou, L., Bawa, R., and Holliday, J. A. (2014). Exome resequencing reveals signatures of demographic and adaptive processes across the genome and range of Black Cottonwood (*Populus Trichocarpa*). *Mol. Ecol.* 23, 2486–2499. doi: 10.1111/mec.12752

Zwoinska, M. K., Rodrigues, L. R., Slate, J., and Snook, R. R. (2020). Phenotypic responses to and genetic architecture of sterility following exposure to sub-lethal temperature during development. *Front. Genet.* 11:573. doi: 10.3389/fgene.2020.00573

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 09 October 2020
doi: 10.3389/fgene.2020.565836

The Cyclically Seasonal *Drosophila subobscura* Inversion O_7 Originated From Fragile Genomic Sites and Relocated Immunity and Metabolic Genes

Charikleia Karageorgiou, Rosa Tarrío* and Francisco Rodríguez-Trelles**

Grup de Genòmica, Bioinformàtica i Biologia Evolutiva (GGBE), Departament de Genètica i de Microbiologia, Universitat Autonòma de Barcelona, Barcelona, Spain

OPEN ACCESS

Edited by:
Pedro Simões,
University of Lisbon, Portugal

Reviewed by:
Stephen Wade Schaeffer,
Pennsylvania State University (PSU), United States
Priscilla Erickson,
University of Virginia, United States
Emma Berdan,
University of Gothenburg, Sweden

***Correspondence:**
Charikleia Karageorgiou
charikleia.karageorgiou@uab.cat
Rosa Tarrío
rosamaria.tarrio@uab.cat
Francisco Rodríguez-Trelles
franciscojose.rodrigueztrelles@uab.cat

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *26 May 2020*
Accepted: *09 September 2020*
Published: *09 October 2020*

Citation:
Karageorgiou C, Tarrío R and Rodríguez-Trelles F (2020) The Cyclically Seasonal Drosophila subobscura Inversion O_7 Originated From Fragile Genomic Sites and Relocated Immunity and Metabolic Genes. Front. Genet. 11:565836. doi: 10.3389/fgene.2020.565836

Chromosome inversions are important contributors to standing genetic variation in *Drosophila subobscura*. Presently, the species is experiencing a rapid replacement of high-latitude by low-latitude inversions associated with global warming. Yet not all low-latitude inversions are correlated with the ongoing warming trend. This is particularly unexpected in the case of O_7 because it shows a regular seasonal cycle that peaks in summer and rose with a heatwave. The inconsistent behavior of O_7 across components of the ambient temperature suggests that is causally more complex than simply due to temperature alone. In order to understand the dynamics of O_7, high-quality genomic data are needed to determine both the breakpoints and the genetic content. To fill this gap, here we generated a PacBio long read-based chromosome-scale genome assembly, from a highly homozygous line made isogenic for an O_{3+4+7} chromosome. Then we isolated the complete continuous sequence of O_7 by conserved synteny analysis with the available reference genome. Main findings include the following: (i) the assembled O_7 inversion stretches 9.936 Mb, containing > 1,000 annotated genes; (ii) O_7 had a complex origin, involving multiple breaks associated with non-B DNA-forming motifs, formation of a microinversion, and ectopic repair in *trans* with the two homologous chromosomes; (iii) the O_7 breakpoints carry a pre-inversion record of fragility, including a sequence insertion, and transposition with later inverted duplication of an *Attacin* immunity gene; and (iv) the O_7 inversion relocated the major insulin signaling *forkhead box subgroup O* (*foxo*) gene in tight linkage with its antagonistic regulatory partner *serine/threonine–protein kinase B* (*Akt1*) and disrupted concerted evolution of the two inverted *Attacin* duplicates, reattaching them to dFOXO metabolic enhancers. Our findings suggest that O_7 exerts antagonistic pleiotropic effects on reproduction and immunity, setting a framework to understand its relationship with climate change. Furthermore, they are relevant for fragility in genome rearrangement evolution and for current views on the contribution of breakage versus repair in shaping inversion-breakpoint junctions.

Keywords: non-B DNA, genome fragility, *foxo* (forkhead box subgroup O), *Akt1* (serine/threonine–protein kinase B), *Attacin* antibacterial genes, immunometabolism, thermal adaptation, seasonal selection

INTRODUCTION

Chromosome inversions are arguably the genetic traits with the earliest and richest record of associations with climate (Hoffmann and Rieseberg, 2008). Research into evolutionary responses to contemporary global warming (Hughes, 2000; Parmesan, 2006) is therefore faced with the challenge of understanding how inversions originate and spread in populations (Kirkpatrick, 2010), while trying to determine their roles in climatic adaptation (Gienapp et al., 2008; Messer et al., 2016).

Chromosome inversions are ubiquitous chromosomal mutations consisting in the reversal of the orientation of a chromosome segment. They originate through either of two major mechanisms, each with its associated distinctive footprints. The first mechanism is intrachromatidal non-allelic homologous recombination (NAHR) between inversely oriented repeats. This mechanism generates inversions with duplications at their ends in both the inverted and uninverted states (Cáceres et al., 1999). The second mechanism is chromosomal breakage and ectopic repair via non-homologous end joining (NHEJ). This mechanism either does not generate duplications or generates them but at the ends of the inverted state only. These two types of NHEJ footprints have been explained in terms of differences in the mode of breakage. Two modes of breakage have been advanced: "cut-and-paste" via clean double-strand breaks (DSBs) that generate blunt ends and staggered. NHEJ inversions without duplications at their ends would originate via cut-and-paste (Wesley and Eanes, 1994), whereas those with inverted duplications at their ends would originate via staggered breaks in one or the two breakpoints. Two staggering models for the origin of the inverted duplications have been proposed (Kehrer-Sawatzki et al., 2005; Matzkin et al., 2005; Ranz et al., 2007): according to the isochromatid model, the duplications would be the filled-in single-stranded overhangs that would result from paired single strand breaks (SSBs) located staggered with each other on opposite strands of the same chromatid (Kehrer-Sawatzki et al., 2005), whereas according to the chromatid model, the duplications would result from unequal exchange between paired sister chromatids, each with one of two paired staggered DSBs at each breakpoint (Matzkin et al., 2005). Note that here the terms *isochromatid* and *chromatid* have switched meanings relative to how they are used in cytogenetics (Savage, 1976). The two staggering models are chromatid models because they assume that inversions originate from either single chromatids during premeiotic mitosis (isochromatid), or paired sister chromatids from the same chromosome during meiotic prophase (chromatid) (Ranz et al., 2007). The models cannot be distinguished based on the pattern of inverted duplications. Yet the chromatid model has been favored over the isochromatid model, because of the length of DNA that would need to be unwound by enzymatic activity in the latter model (Ranz et al., 2007). The chromatid model is also not without potential caveats because NHEJ was found to be suppressed during the meiotic prophase in *Drosophila* (Joyce et al., 2012; Hughes et al., 2018). The prevalence and distribution of the NAHR and NEHJ mechanisms of inversion formation within and across lineages are currently under debate (Ranz et al., 2007; Delprat et al., 2019). The NEHJ mechanism rests upon the occurrence of two or more DSBs. But the source of the DSBs (whether environmental, such as ionizing radiation, or spontaneous, such as non-B DNA-associated sequence instability, where non-B DNA denotes any DNA conformation that is not in the canonical right-handed B form; Lobachev et al., 2007; Zhao et al., 2010; Farré et al., 2015), the relative contributions of breakage versus repair to shaping breakpoint junctions (Ranz et al., 2007; Kramara et al., 2018; Scully et al., 2019), and the relative frequency with which the joined broken ends are from the same chromatid (isochromatid model) versus two distinct sisters (chromatid model) (Ranz et al., 2007) or even, as has been more recently suggested by Orengo et al. (2019), non-sister chromatids (chromosome model) are additional open questions.

Inversions can have direct or/and indirect functional effects (Kirkpatrick, 2010). Direct effects are those ascribable to the mutation *per se*, as it altered the structure or expression of functional sequences at the breakpoints, or the functional neighborhood of genes in the cell nucleus (McBroome et al., 2020). Indirect effects emanate from their associated recombination–suppression effects when in heterozygous condition, whereby they can bind together into close linkage association particular combinations of alleles at genetically distant loci. The evolutionary significance of polymorphic inversions is often thought to chiefly stem from their indirect effects (Dobzhansky, 1947; Wasserman, 1968; Kirkpatrick and Barton, 2006). Although data have been lacking on the relative importance of the two types of effects, there has been renewed interest in using genomics to determine mechanisms for the spread, establishment, and maintenance or fixation of inversions (Corbett-Detig and Hartl, 2012; Corbett-Detig, 2016; Fuller et al., 2016, 2017, 2019; Cheng et al., 2018; Said et al., 2018; Lowry et al., 2019). Because they usually involve many genes, chromosome inversions have enhanced potential for affecting multiple traits, which should expand the opportunities for their maintenance via balancing selection. The extent to which that is the case and the types and transience of the balancing selection mechanisms involved are only beginning to be elucidated (Kapun and Flatt, 2018; Wellenreuther and Bernatchez, 2018; Faria et al., 2019). Amid these unknowns, the inversion polymorphisms of *Drosophila subobscura* emerged among the first genetic traits identified as involved in a species' adaptation to contemporary climate warming (Rodríguez-Trelles and Rodríguez, 1998, 2007; Balanyà et al., 2006; Rezende et al., 2010).

Drosophila subobscura is a native Palearctic species broadly distributed in Europe and the newly invaded areas of North and South America (reviewed in Krimbas, 1992), where it is found generally associated with woodland habitats. It belongs in the obscura group, within which it clusters with the recently derived small-island endemics *Drosophila guanche* and *Drosophila madeirensis*, forming the subobscura three-species subgroup (Bächli, 2020). *D. subobscura* has one of the smallest and least repetitive *Drosophila* reference genomes obtained thus far, which is distributed among five large telocentric chromosomes (A, J, U, E, and O) and one small dot (Karageorgiou et al., 2019). In stark contrast with its two insular relatives, the species has evolved highly rearranged chromosome sequences, which is due

to having experienced accelerated fixation rates of paracentric inversions, especially the A sex chromosome. This situation has been interpreted as indicative of the inversions' role in binding together adaptive alleles in the face of the species' intense continent-wide gene flow (Karageorgiou et al., 2019). Presently, *D. subobscura* harbors a rich inversion polymorphism, with its five major chromosomes showing parallel adaptive variation patterns across latitude (Ayala et al., 1989), seasons (Rodríguez-Trelles et al., 1996, 2013), and through a heatwave (Rodríguez-Trelles et al., 2013), while rapidly shifting in close association with the ongoing rise in global temperatures (Rodríguez-Trelles and Rodríguez, 1998, 2010; Balanyà et al., 2006). Laboratory attempts to establish the causal nature of this association have, however, largely been inconclusive (Santos et al., 2005; Fragata et al., 2014). Ultimately, a complete understanding of the role of inversions in adaptation to contemporary climate warming in *D. subobscura* will necessarily include the identities and functional properties of the genome sequences affected by them. Advances along this line include the isolation and characterization of breakpoint sequences for 11 of the more than 65 large cytologically visible inversions known for the species, including A_2 (Puerma et al., 2017), O_3 (Papaceit et al., 2012), O_4 and O_8 (Puerma et al., 2016a), E_1 and E_2 (Puerma et al., 2014), E_3 and E_9 (Orengo et al., 2015), E_{12} (Puerma et al., 2016b), and U_1 and U_2 (Karageorgiou et al., 2019). An overall conclusion is that none of these inversion breakpoints disrupted any obvious candidate gene for direct adaptation to temperature, despite the fact that all but the E_3 inversion are supposed to be involved in adaptation to climate (e.g., Menozzi and Krimbas, 1992; Rego et al., 2010; Arenas et al., 2018). Apart from the fact that thermal traits are genetically complex and that many of the genes that impinge on them are still unknown, the above conclusion supports that those inversions' role in thermal adaptation would be through either position effects, indirect linkage generation effects, or both.

As part of a wider effort to develop a high-quality reference genome for *D. subobscura* encompassing the species' rich chromosomal polymorphisms, here we focus on the O_7 inversion. The breakpoints of this inversion were located cytologically at subsections 77B/C and 85E on the Kunze–Mühl and Müller *standard* map (**Figure 1A**; Kunze-Mühl and Müller, 1958; Götz, 1965). O_7 is among the top 10% known largest *D. subobscura* inversions, stretching most of the centromere-proximal half of the O chromosome (Krimbas, 1992). In nature, it attains significant frequencies only in combination with the non-overlapping centromere-distal complex of two overlapping inversions O_{3+4}, forming the chromosome arrangement O_{3+4+7} (**Figure 1B**). The tight association between O_7 and O_{3+4} is likely maintained by an interaction between selection and the strongly reduced recombination between them (Pegueroles et al., 2010a).

O_7 could be initially classified as a warm-climate inversion. In the Palearctic, it shows a southern distribution. In northwest Spain, where it has been longitudinally monitored starting in mid-1970s (Fontdevila et al., 1983; Rodríguez-Trelles et al., 1996, 2013), it shows a pronounced regular seasonal cycle (estimated to account for more than 60% of the inversion's temporal variation; Rodríguez-Trelles et al., 1996) that peaks in summer and drops in winter (**Figure 1C**). In 2011, it rose to summer-like levels in spring during a heatwave, with the magnitude of the increase closely matching that of the thermal anomaly (**Figure 1C**; Rodríguez-Trelles et al., 2013). However, (i) the average annual frequency of O_7 in northwest Spain remains unchanged after decades of sustained climate warming experienced by the region (Rodríguez-Trelles et al., 2013; our unpublished records). (ii) Following the 2011 heatwave, the inversion reached summer-like frequencies in April, but did not continue rising through the ensuing summer (**Figure 1C**), perhaps hampered by recessive deleterious alleles (Rodríguez-Trelles et al., 2013). (iii) The Palearctic distribution of O_7 is disjointed between the peninsulas of Iberia and Turkey (Götz, 1967). These are similar latitude areas separated by ∼2,500 km within the continuous species' range. Assuming that the inversion is molecularly the same in the two areas, this spatial pattern can hardly be explained on the sole basis of a postglacial expansion scenario (Menozzi and Krimbas, 1992), considering how rapidly it spread through the recently invaded areas of the New World (Prevosti et al., 1988). And (iv) in the more studied Iberian Peninsula, the distribution of the inversion has negative or no correlations with the geographical variation in temperature. For example, the average annual frequency of the inversion declines from ∼50% to near-zero values along the > 1,000-km stretching from the northwestern-most to the northeastern-most territories, despite the latter having a warmer climate than the former (de Frutos, 1972; Solé et al., 2002; Rodríguez-Trelles et al., 2013). The same is true for the West Atlantic fringe of the peninsula along which the inversion levels remain basically the same despite the fact that it stretches seven latitudinal degrees of steep thermal gradient (Brehm and Krimbas, 1988; Solé et al., 2002; Rodríguez-Trelles et al., 2013). The inconsistent patterns of O_7 between components of the ambient temperature suggest that it is influenced by selective factors other than temperature alone.

The O chromosome offers the methodological advantage over the other *D. subobscura* chromosomes that there is an available balancer-strain called *Varicose*/*Bare* (*Va*/*Ba*) (Sperlich et al., 1977). In this study, we first used the *Va*/*Ba* strain to develop an isogenic line with two identical copies of a wild O chromosome carrying the O_{3+4+7} arrangement. Second, we used PacBio long-read technology to generate a high-quality annotated chromosome-scale genome sequence for the line. Third, we isolated the complete continuous nucleotide sequence of the inversion O_7 by conserved synteny analysis of the obtained O_{3+4+7} chromosome with the available O chromosome from the species' reference genome, which is structurally O_{3+4} (Karageorgiou et al., 2019). In addition, we also considered two other published sequences of the O chromosome, including a high-quality long-read–based sequence from *D. subobscura* (Bracewell et al., 2019), and an Illumina-based sequence from *D. guanche* (Puerma et al., 2018). We give an account of O_7 main features, together with a detailed description of its mechanism of formation. Our findings provide clues to the mixed evidence for this inversion's role in thermal adaptation.

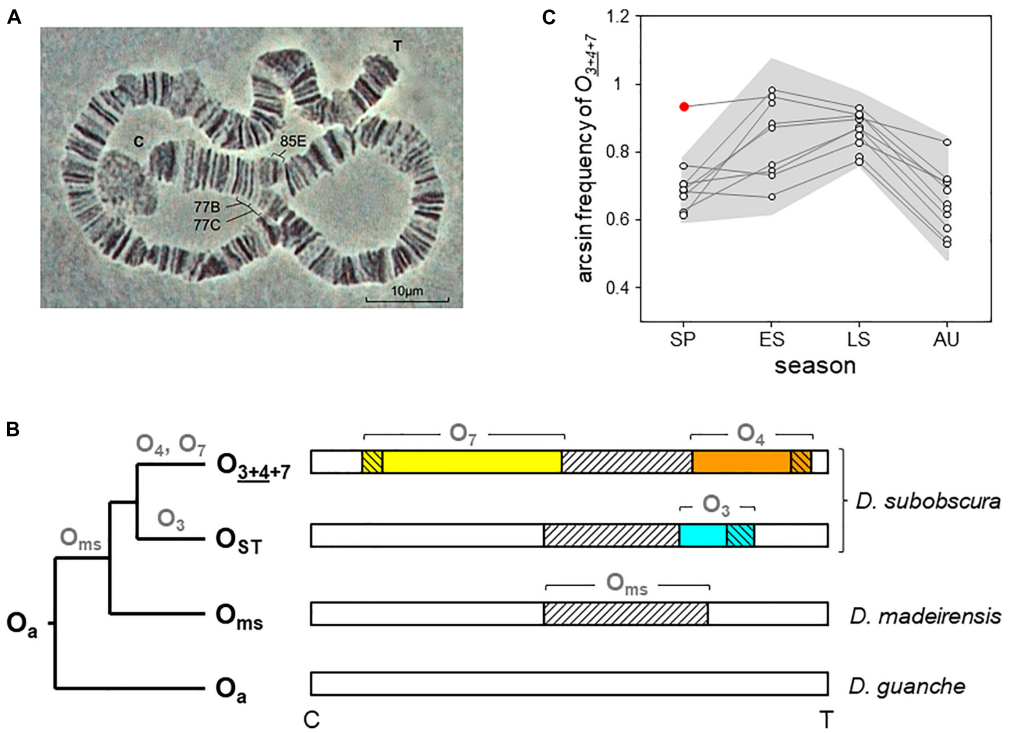


FIGURE 1 | O_7 inversion and O_{3+4+7} chromosome arrangement of *D. subobscura*. **(A)** Light micrograph (400 ×) of the O_7 diagnostic loop from two paired polytene O chromosomes of a O_{3+4+7}/O_{3+4} heterokaryotype, with indicated cytological map positions of the two inversion breakpoints (Kunze-Mühl and Müller, 1958; Götz, 1965). C and T denote centromere and telomere, respectively. **(B)** Phylogeny and chromosomal locations of the inversions forming the O_{3+4+7} arrangement in the subobscura subgroup. Names at the root and tips (bold black) and on branches (bold gray) denote chromosome arrangements and inversions, respectively. The ancestral O arrangement of the subgroup is O_a (Karageorgiou et al., 2019). The chromosome-central inversion O_{ms} (diagonally hatched) is so-called because it became fixed in the last common ancestor of *D. madeirensis* and *D. subobscura* (Karageorgiou et al., 2019). In *D. subobscura*, O_3 (blue) and O_4 (orange) are two centromere-distal inversions with overlapping cytological map positions originated independently on separate O_{ms} branches. The centromere-proximal inversion O_7 (yellow) is assumed to have originated along the branch of O_4. O_{ms} became extinct as a single inversion in *D. subobscura*. Note that O_3 is not in the path from O_a to O_{3+4+7}, being the inversion that generated the O_{ST} arrangement. **(C)** Five decades of cyclic seasonal change of O_{3+4+7} at Mount Pedroso, Spain. Consecutive seasonal data (dots) from the same year are connected by lines. The gray background plots the ± 2σ confidence band around the seasonal averages, and the red dot the summer-like value recorded during the spring 2011 heatwave. Included are published data from 1976 to 1981 (Fontdevila et al., 1983), 1988 to 1991 (Rodríguez-Trelles et al., 1996), 2011 to 2012 (Rodríguez-Trelles et al., 2013), and our 2015 unpublished arcsin-transformed records for late summer (0.845) and autumn (0.574). SP, spring; ES and LS, early and late summer; AU, autumn.

MATERIALS AND METHODS

Species Karyotype and Inversion Nomenclature

Drosophila subobscura shows the ancestral karyotype configuration of the genus *Drosophila*, consisting of five large telocentric rods (Muller elements A-E) and one dot (Muller F) (Powell, 1997). The five rods include the sex chromosome (Muller A) and four autosomes of which the O chromosome (Muller E; homologous to chromosome arm 3R from *D. melanogaster*) is the largest (∼30 Mb), comprising around 25% of the species' nuclear euchromatic genome (∼125 Mb; Karageorgiou et al., 2019).

An early landmark in the study of chromosomal inversion polymorphisms of *D. subobscura* was the development of structurally homozygous strains, as tools to identify new inversions by the location and shape of the loops formed in inversion heterozygotes (Zollinger, 1950; Maynard-Smith and Maynard-Smith, 1954; Zouros et al., 1974; Loukas et al., 1979). The "Küsnacht" strain, named after the Swiss locality of collection of the flies (Zollinger, 1950), became the first established (Koske and Maynard-Smith, 1954). The chromosomal arrangements of the strain, which happened to be those most common in Central Europe, were subscripted ST (for "standard") and from them new inversions were designated with numeral subindices following their order of discovery (Kunze-Mühl and Sperlich, 1955). This naming system was not intended to convey polarity of evolutionary change. Accordingly, O_{3+4+7} is the arrangement that can be interconverted with O_{ST} by the two centromere-distal overlapping inversions O_3 and O_4 (denoted by the underline joining the subscripts; Zouros et al., 1974) and the centromere-proximal inversion O_7. The ancestor-descendant relationships of these inversions are shown in **Figure 1B**.

Drosophila Lines

O chromosome conserved synteny analysis was based on data from four whole-genome *de novo* assemblies, including three PacBio long-read–based assemblies from *D. subobscura* and one Illumina short-read–based assembly from *D. guanche*. Of the

three *D. subobscura* assemblies, one was used as reference for inversion O_7 and was newly generated in this study. The other two were used as references for the *standard* configuration [note that the distal breakpoint of O_7 maps within inversion O_{ms} (Karageorgiou et al., 2019), whereby is expected to exhibit opposite orientation in *D. subobscura* relative to *D. guanche*; **Figure 1B**] and were already available (Karageorgiou et al., 2019; Bracewell et al., 2019). Also available was the assembly from *D. guanche* (Puerma et al., 2018), which was used as an outgroup. Henceforth, we will refer to these four assemblies as Ds_7, Ds_ch-cu, Ds_B, and Dg, respectively.

To generate the Ds_7 assembly, we developed a line that is isogenic for an O_{3+4+7} arrangement from the wild and homokaryotypic and highly homozygous for the ST arrangements of the rest of the chromosomes (i.e., A_{ST}, J_{ST}, U_{ST}, E_{ST}, and O_{3+4+7}). The O arrangement was first isolated by crossing wild males to virgin females from the *cherry-curled* (*ch-cu*) recessive marker stock; they were then submitted to nine generations of backcrossing with ch-cu females and finally isogenized using the *Va/Ba* balancer stock (Sperlich et al., 1977). The expression of the *Ba* gene is highly variable. Therefore, to prevent potential errors at sorting out phenotypically O_{3+4+7} homokaryotypes, the *Va/Ba* stock was previously selected for zero macrobristles on the scutum and scutellum. Crossing schemes and the methods for polytene chromosome staining and identification are described elsewhere (Rodríguez-Trelles et al., 1996). The assayed line was stored frozen at −80°C immediately upon obtention. The wild flies used to develop the line were derived from our survey of the natural population of Berbikiz (Spain; Lat.: 43,18949, Long.: –3,09025, Datum: WGS84, elevation: 219 m a.s.l) conducted in July 7, 2012 (Rodríguez-Trelles et al., 2013).

The remaining three assemblies were derived from strains homokaryotypic for all chromosomes. The Ds_ch-cu assembly was generated from the *ch-cu* strain of our laboratory (A_{ST}, J_{ST}, U_{ST}, E_{ST}, and O3+4; Karageorgiou et al., 2019) and the Ds_B assembly from an isofemale laboratory stock derived from a natural population from Eugene, Oregon, in 2006 (A_{ST}, J_{ST}, U_{1+2}, E_{ST}, and O_{3+4}; Bracewell et al., 2019). The Dg assembly was generated from an isofemale laboratory stock derived from a natural population from the Canary Islands, Spain, in winter 1999 (Puerma et al., 2018); it shows the chromosome configuration of the last common ancestor of the *subobscura* subgroup except for chromosome E, which carries the arrangement E_{g1} (A_a, J_a, U_{1+2}, E_{g1}, and O_a; Puerma et al., 2018; Karageorgiou et al., 2019; Bracewell et al., 2019).

High Molecular Weight Genomic DNA Isolation and PacBio Whole-Genome Sequencing

High-quality high-molecular-weight gDNA was obtained from 60 mg of −80°C frozen adult females, using a modified version of the phenol/chloroform method of Chen et al. (2010) that yields ∼25 μg of high-quality DNA per assay, as assessed by NanoDrop ND1000 (NanoDrop Technologies Inc., Wilmington, DE, United States) spectrophotometer and standard agarose gel electrophoresis. The genome of the Ds_7 isogenic line was sequenced to nominal 66-fold genome coverage using PacBio (Pacific Biosciences, Menlo Park, CA, United States) Sequel single-molecule real-time (SMRT) technology from a 20-kb SMRTbell template library, using Polymerase 3.0 chemistry and two SMRT cells. Libraries construction and PacBio sequencing were outsourced to Macrogen (Macrogen Inc., Seoul, South Korea).

Chromosome-Scale Assembly and Scaffolding

Raw PacBio reads were assembled using the Canu assembler (version 1.8; Koren et al., 2017) on recommended settings for read error correction, trimming and assembly, and genome size set at 150Mb based on previously published flow cytometry data (Karageorgiou et al., 2019). These analyses were performed on a 2.80-GHz 8-CPU Intel Xeon 64-bit 32GB-RAM computer running Ubuntu 18.04 LTS.

Chromosome-scale assembly and scaffolding followed the four steps outlined in Karageorgiou et al. (2019) as well as a fifth step, to improve genome completeness and contiguity, consisting of merging the Ds_7 assembly with a preselected set of segments from the reference Ds_ch-cu assembly using one round of quickmerge (Chakraborty et al., 2016), as follows: first, the CANU contigs that could be certainly anchored, ordered, and oriented on the nuclear chromosomes were aligned against the Ds_ch-cu reference using NUCmer (Kurtz et al., 2004). Second, the segments of Ds_ch-cu not overlapped by the CANU contigs, each extended 10 kb outward from each of its two ends, were extracted. Finally, separately for each chromosome, the extracted Ds_ch-cu segments, together with the CANU contigs set as the backbone, were fed into quickmerge. This approach was found to reduce the chances of misassembly and chimerism, while making it straightforward to trace the non-backbone sequence in the assembly. Dot plots of the merged assembly against the reference Ds_ch-cu assembly were used as a further step of misassembling correction. The obtained Ds_7 assembly was polished with 26 × mean coverage of 150–base-pair (bp) MP Illumina reads from the O_{3+4+7} isogenic line using two rounds of PILON (version 1.23; Walker et al., 2014).

Genome Annotation

Gene prediction and annotation of the assembled genome were conducted using the MAKER (version 3.01.02.-beta; Holt and Yandell, 2011; Campbell et al., 2014) annotation pipeline. Repetitive elements were identified using RepeatMasker (version 4.0.6; Smit et al., 2013/2015, at[1]) combined with three repeat libraries, including (i) the *Drosophila* genus–specific repeat library contained in the Repbase database (release 20170127; Bao et al., 2015); (ii) a library of *subobscura* subgroup specific satellites, sat290 and SGC-sat (Karageorgiou et al., 2019); and (iii) a library of *de novo* identified repeats generated using RepeatModeler (version1.0.11) on the assembly masked for the first two libraries. Novel long terminal repeats (LTRs), miniature inverted-repeat transposable elements (MITEs), tandem repeats,

[1]http://repeatmasker.org

and rDNA and tDNA genes were identified using LTRharvest (GenomeTools version1.5.10; Ellinghaus et al., 2008), MITE Tracker (version 2.7.1; Crescente et al., 2018), Tandem Repeat Finder (TRF; version4.09; Benson, 1999), RNAmmer (version 1.2; Lagesen et al., 2007), and tRNAscan-SE (version 2.0; Lowe and Chan, 2016), respectively. All tools were run on default settings, except LTRharvest, for which we set -seed 100, -similar 90.0, and -mintsd 5, following Hill and Betancourt (2018). The quality of the annotation was controlled using the Annotation Edit Distance (AED) metric (Eilbeck et al., 2005). AED values are bounded between 0 and 1. An AED value of 0 indicates perfect agreement of the annotation to aligned evidence, and conversely, a value of 1 indicates no evidence support.

Functional annotation of MAKER-predicted proteins was made by BLASTP (version 2.6.0 +) searches against the *Drosophila* UniProt-SwissProt manually curated datasets (Apweiler et al., 2004). Prediction of protein functional domains was accomplished using InterProScan (version 5.29–68.0; Jones et al., 2014) on the Pfam (Finn et al., 2016), InterPro (Finn et al., 2017), and Gene Ontology (Ashburner et al., 2000; The Gene Ontology Consortium, 2017) domain databases. Genome assembly and annotation completeness were gauged using the Benchmarking Universal Single-Copy Orthologs (BUSCO) tool [BUSCO, version 4 (Seppey et al., 2019)], with the latest update of the dipteran gene set (diptera_odb10), which contains 3,285 highly conserved, single-copy genes expected to be present in any dipteran genome.

Isolation and Characterization of the O_7 Breakpoints

Suppose that +A|+B+C|+D and +A|−C−B|+D represent two chromosome arrangements whose gene orders differ only by the orientation of the segment between A and D (with symbols denoting A and D, the segments upstream from the centromere-proximal breakpoint and downstream from the centromere-distal breakpoint, respectively; vertical bars, breakpoint junctions; and plus/minus signs, orientation of the segment relative to the uninverted sequence). We proceeded in two steps. First, we isolated the regions containing the breakpoint junctions by chromosome conserved synteny analysis between the uninverted and inverted states using the Synteny Mapping and Analysis Program (SyMAP, version 4.2.; Soderlund et al., 2011) tool on default options, and NUCmer (see Karageorgiou et al., 2019). The O_7 breakpoints were identified as the loci of interrupted synteny whose locations and distance from each other agree with the cytogenetic mapping data of the inversion (Karageorgiou et al., 2019). Second, we localized the breakpoint junctions at base-pair resolution and performed comparative analyses of their flanking sequences using the progressive guide tree-based MAFFT algorithm (version 7[2]) with the accuracy-oriented method "L-INS-i" (Katoh et al., 2019). Each of the regions +A|+B and +C|+D from the uninverted state was aligned separately, first with +A|−C and then with −B|+D from the inverted state. From each of the four resulting alignments, we used the regions showing positional homology between the uninverted and inverted states to isolate segments A, B, C, and D, correspondingly. The remaining sequences of the uninverted state were submitted to a second round of comparative analysis among them, and with segments A to D to identify the homologies missed in the first round. As representatives of the uninverted state, we used Ds_ch-cu together with the previously published assemblies Ds_B and Dg, and this last one was set as the outgroup.

[2]http://mafft.cbrc.jp/alignment/software/

Phylogenetic Inferences

MAFFT-based tree reconstruction of the *Attacin* gene family in *Drosophila* was performed via maximum likelihood. Model selection and tree inference were conducted using IQ-Tree (Kalyaanamoorthy et al., 2017; Nguyen et al., 2015). Tree searches were conducted starting from sets of 100 initial maximum parsimony trees using nearest neighbor interchange with default perturbation strength and a stopping rule settings. Branch support was assessed using the ultrafast bootstrap approximation (UFboot; 1,000 replicates) (Hoang et al., 2018), and two single-branch tests including the Shimodaira–Hasegawa-like approximate likelihood ratio test (SH-aLRT; 1,000 replicates) (Guindon et al., 2010) and the approximate Bayes parametric test (Anisimova et al., 2011).

Non-B DNA Sequence and Transcription Factor Binding Site Scans

Scans for potential non-B DNA–forming sequences considered the following features: inverted repeats (IRs) (capable of forming hairpin and/or cruciform DNA), direct/tandem repeats (slipped/hairpin structures), mirror repeats (triplexes), alternate purine-pyrimidine tracts (left-handed Z-DNA), G4 motifs (tetraplex and G−quadruplex DNA), and A−phased repeats (static bending). Searches were conducted online using for IRs Palindrome Analyzer (Brázda et al., 2016[3]; accessed January 24, 2020) with repeat length of 6-20 nt, spacer length $\leq$ 10 nt, and number of mismatches $\leq$ 1; for tandem repeats Tandem Repeat Finder (TRF version 4.09; Benson, 1999[4]; accessed Jan 24, 2020) in basic mode; and for the remaining features nBMST (Cer et al., 2012[5]; accessed January 24, 2020) with prefixed default settings. The propensity of IRs to adopt non-B conformation was assessed using the difference in free energy between the DNA sequence in the linear and cruciform structures, as implemented in Palindrome Analyser (Brázda et al., 2016).

Transcription start site (TSS) prediction was conducted using the NNPP method (Reese, 2001[6]). Searches for putative binding sites for Relish (*Rel*), the heterodimer Dif/Rel, dFOXO, Dorsal (*dl*), and Serpent (*srp*) transcription factors in the 1-kb upstream region of the *Attacin* predicted TSSs were performed using the FIMO tool (Grant et al., 2011) from the MEME suite (Bailey et al., 2015). For *Rel* and Dif/Rel, and for dFOXO,

[3]http://bioinformatics.ibp.cz:9999/#/en/palindrome

[4]https://tandem.bu.edu/trf/trf.html

[5]https://nonb-abcc.ncifcrf.gov/apps/nBMST/default

[6]https://www.fruitfly.org/seq_tools/promoter.html

we used the FootprintDB database (Sebastián and Contreras-Moreira, 2014[7]) *Drosophila melanogaster* Major Position Matrix Motifs (DMMPMM) identified, respectively, by Senger et al. (2004) and Weirauch et al. (2014). For *dl* and *srp*, we used the REDfly database (version 5.5.3; Rivera et al., 2019[8]) improved iDMMPMM motifs developed by Kulakovskiy and Makeev (2009). Searches were performed using a *p* value cutoff of 10^{-3}.

RESULTS

Chromosome-Scale Assembly and Annotation of Chromosome Arrangement O_{3+4+7}

The PacBio Sequel sequencing of the O_{3+4+7} isogenic line genome generated 2,457,493 reads, with mean and longest lengths of 11,257 bp and 117,750 bp, respectively. Canu correction and trimming retained a 42-fold genome coverage for the assembly. Of the 385 Canu-generated contigs, the 14 that could be confidently anchored, ordered, and oriented covered the complete reference genome, with an added length of 126.770 Mb and N50 of 10.587 Mb. Quickmerge of those 14 CANU contigs resulted in six chromosome-scale scaffolds, one per each of the major *D. subobscura* chromosomes (**Table 1**). Of note, chromosome O was built from two contigs only, with the centromere-proximal contig (tig00026085; 29.679 Mb) spanning almost all the chromosome length (96.9%) (**Figure 2A**). The Ds_7 assembly contained 13,459 MAKER-annotated genes, nearly all with well-supported predictions (AED_{50} = 99.3%). Only 2.6% (87) of the BUSCO genes were missing, indicating that the assembly is almost complete. The O chromosome contained 3,220 (23.9%) of the annotations of the assembly.

Identification of Inversion O_7 Using Chromosome Conserved Synteny Analysis

The structural transition between the O chromosomes of the Ds_7 and Ds_ch-cu assemblies called for one large megabase-sized inversion (**Figures 2B,C**), whose breakpoints located cytologically precisely as it would be expected if they were from O_7. Relative to the nearest of the available 140 cytologically mapped markers of the O chromosome (see Karageorgiou et al., 2019), the proximal breakpoint was located 44.5 kb downstream from *Sb* (Dmel\CG4316) and 117.4-kb upstream from microsatellite *dsub02*, and the distal breakpoint 111.8 kb downstream from *rdx* (Dmel\CG12537) and 29.3kb upstream from *Abi* (Dmel\CG9749). *Sb* and *dsub02* have been respectively mapped to subsections 77B (Dolgova, 2013) and 77C (Santos et al., 2010), and *rdx* and *Abi* to subsection 85E (Dolgova, 2013; Pegueroles et al., 2013) of the Kunze-Mühl and Müller (1958) *standard* cytological map. Other than O_7, no *D. subobscura* inversion maps to those positions.

Comparative analysis of the genes annotated in the regions immediately flanking the breakpoints in Ds_7, Ds_ch-cu and

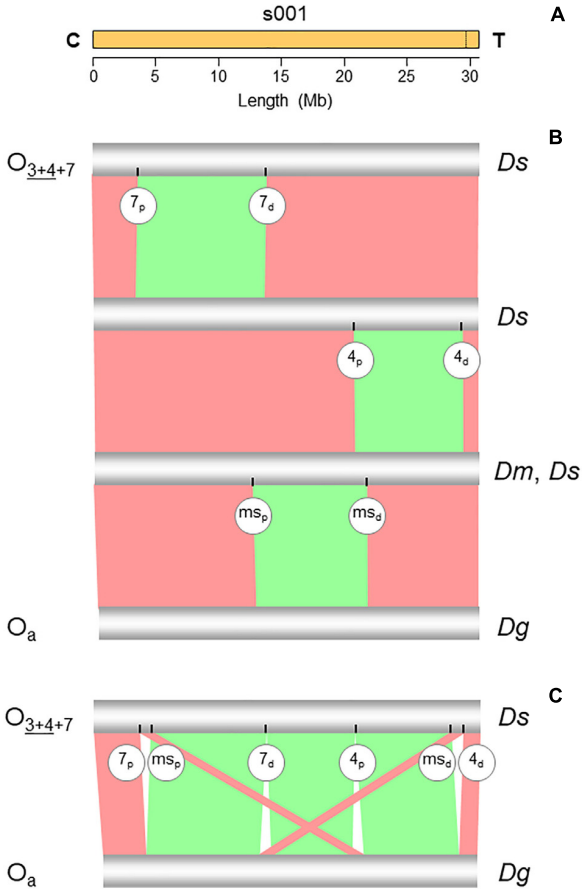


FIGURE 2 | Chromosome conserved synteny analysis of the O_7 breakpoints. **(A)** The long read-based O_{3+4+7} chromosome-scale scaffold (s001). The vertical dotted line near the telomere indicates the location of the stitch between the two Canu tigs. **(B)** SyMAP comparative synteny analysis showing the inversions found along the path from the ancestral arrangement of the subobscura subgroup (O_a) to O_{3+4+7} (see **Figure 1B**). In addition to O_7, the two chromosomes also differ by inversions O_4 and O_{ms}. **(C)** SyMAP direct comparison of the O_a and O_{3+4+7} chromosome arrangements. Bands connecting the chromosomes denote uninverted (pink) and inverted (green) synteny blocks. Labeled ticks on chromosomes indicate proximal (p) and distal (d) inversion breakpoints. The remaining symbols are as in **Figure 1**.

[7]http://floresta.eead.csic.es/footprintdb

[8]http://redfly.ccr.buffalo.edu/

TABLE 1 | Ds_7 assembly summary statistics (Muller elements are given in parenthesis, and lengths are given in megabases of sequence).

Component	Length	Scaffolds	Canu contigs	Largest Canu contig	Gene models
Nuclear genome	126.770	6	14	29.679	13,459
Dot (F)	1.412	1	1	1.412	96
A (A)	22.941	1	2	17.229	2,323
J (D)	25.018	1	3	10.587	2,652
U (B)	26.010	1	3	13.133	2,561
E (C)	20.783	1	3	9.524	2,607
O (E)	30.629	1	2	29.679	3,220

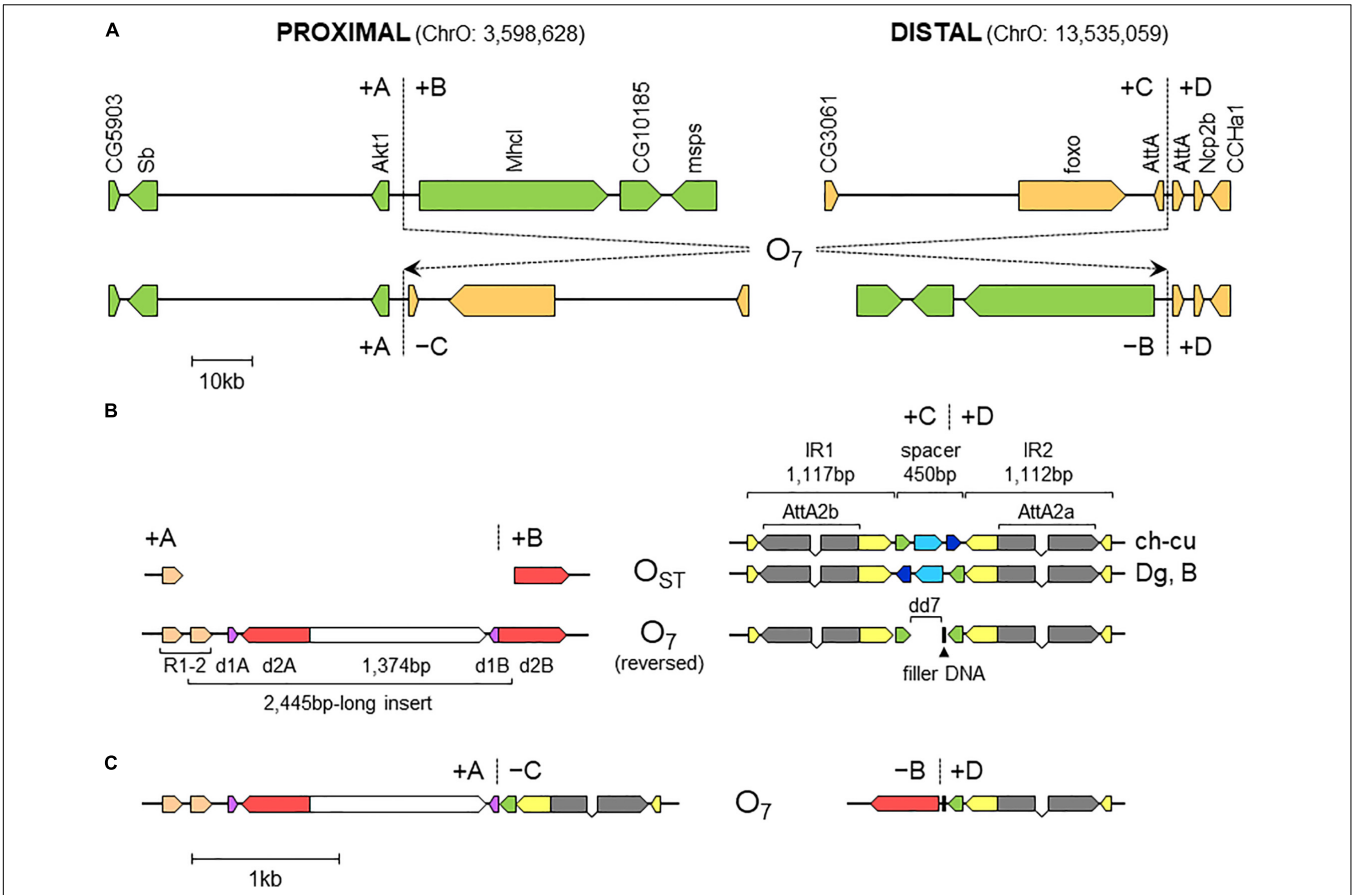


FIGURE 3 | Sequence-annotated breakpoint regions in O_7 and the standard uninverted state. **(A)** Gene scale (10-kb scale bar) depiction of the proximal and distal breakpoints in the uninverted state (+A|+B and +C|+D, respectively) versus O_7 (+A|−C and −B|+D). In green are segments A and B, and in sepia C and D. Vertical broken lines indicate break junctions, and arrow boxes the size and direction of the genes labeled vertically using the names of the corresponding *D. melanogaster* orthologs. The coordinates of O_7 in the assembled chromosome are given in parentheses. **(B)** Zoom-in (1-kb scale bar) on the regions immediately flanking the break junctions, with O_7 oriented backward (i.e., +B+C, instead of −C−B) to better track the differences with the uninverted state. Arrow boxes indicate the size and direction of the sequence elements discussed in the text. Gray boxes (exons) linked by polygonal lines (introns) represent the two AttA2 paralogs oriented in the direction of transcription. In the distal breakpoint, the two alternative haplotypes of the uninverted states, namely, that from ch-cu and that from Dg and B, are shown. Note the reversal of the spacer in ch-cu versus Dg and B, and the mirror halves flanking dd7 in O_7. **(C)** O7 represented as in **(B)**, but in its actual orientation (i.e., -C-B).

Ds_B with those in the outgroup Dg (**Figure 3A**) corroborated that Ds_ch-cu and Ds_B carried the uninverted state, whereas Ds_7 carried the inverted state. The assembled O_7 has a size of 9,936,431 bp, totaling 32.4% of the chromosome (30,629,152 bp). It has a GC content (43.8%) below that of the O chromosome (44.9%) since it is located in the chromosome centromere-proximal half, which is relatively AT-rich (Karageorgiou et al., 2019). O_7 was predicted to have 1,028 protein-coding genes, or 31.9% of the gene models of the O chromosome, in close agreement with its percent of chromosome length.

Nature and Properties of the DNA Sequences Surrounding O_7 Breakpoint Junctions

Proximal Breakpoint of O_7

The alignments used for isolation of the breakpoint junctions and their corresponding flanking regions A, B, C, and D are shown in **Supplementary Figures 1**, **2**. **Figure 3B** provides a schematic representation of the +A|+B region based on the alignment of **Supplementary Figure 3**. In the case of O_7, the region was reconstructed using the reverse complement of segment -B. The breakpoint junction is located within a 2,445-bp-long sequence stretch present only in the inverted state. The site of the insertion is flanked by multiple indels, which suggests that the insertion occurred in a region of prior sequence instability. Of the insertion length, 2,317 bp are on the + A segment and 128 bp on the + B segment. The insertion begins with a 153-bp-long direct repetition (R1-2) of the upstream flank. Proceeding downstream from this repeat, there are two inverted duplications named d1 and d2, each with copies A and B, with d1 shorter (59 bp long each of d1A and d1B) than d2 (534 and 540 bp for copies d2A and d2B, respectively). The two A copies (i.e., d1A and d2A) are separated from the two B copies (i.e., d1B and d2B) by an intervening sequence of 1,374 bp. The junction between + A and + B is precisely located between d1B and d2B. d2B extends

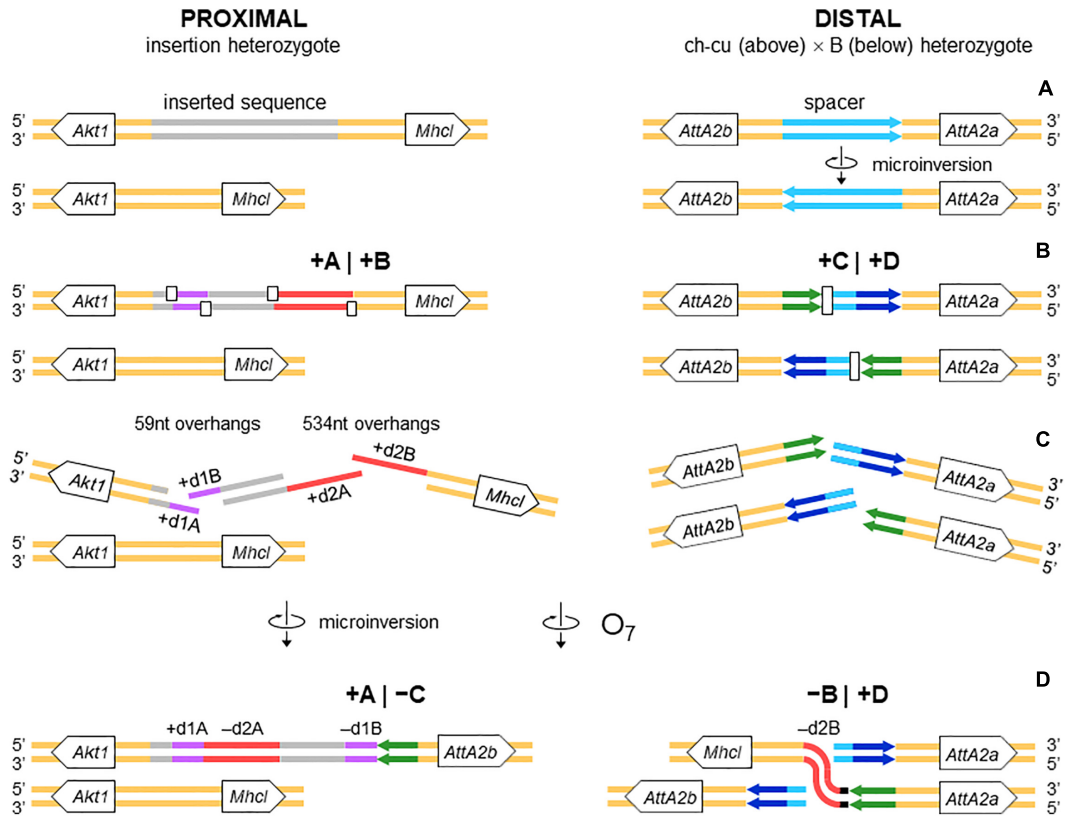


FIGURE 4 | Isochromatid-chromosome mixed staggered model of O_7 formation. **(A)** Start from an individual heterozygous (or homozygous) for the insertion at the proximal breakpoint, and heterozygous ch-cu–type (above)/B-type (below) for the orientation of the spacer at the distal breakpoint. **(B)** Occurrence of two pairs of isochromatidal staggered SSBs at the proximal breakpoint (open squares), and two staggered chromosomal blunt-ended DSBs at the distal breakpoint (open rectangles), with demarcation of breakpoints flanking segments +A|+B and +C|+D, respectively. **(C)** Emergence of broken ends with single stranded overhangs (+ d1A, + d1B and + d2A, + d2B) at the proximal breakpoint. **(D)** Microinversion formation with fill-in of the overhangs, resulting in the terminal inverted duplications + d1A and -d1B, and the reversed orientation of segment -d2A. And O_7 formation by reversal of the +B+C segment with fill-in of the -d2B overhang, and distal breakpoint repair via rejoining in *trans* with the homologous chromosome.

409 bp downward from the downstream end of the insertion into the region of resumed homology between O_7 and the uninverted state, indicating the orientation of the parent copy.

The above pattern of sequence copy number, order, and orientation most parsimoniously indicates that the proximal breakpoint of O_7 was formed on an insertion region that experienced two pairs of staggered SSBs, which resulted in two DSBs (**Figure 4**; but see section "DISCUSSION" for alternative models). The upstream-most DSB generated the proximal breakpoint of a 2,026-bp-long microinversion and the downstream-most DSB generated a junction flanked upstream by the distal end of the microinversion and downstream by the proximal end of O_7. Accordingly, duplications d1 and d2 would, respectively, represent the filled-in staggered SSB-induced terminal single-stranded overhangs of the microinversion and inversion O_7. **Figure 3C** shows O_7's segments A and B such as they are found in the inversion. That d2A and d2B show direct instead of reverse relative orientation as it would be expected if paired–staggered SSBs generate inversions with inverted repeated ends (Ranz et al., 2007) would be explained by the reversal in the orientation of d2A as a result of the microinversion.

Relative to the predicted nearest gene TSSs, the events took place in an intergenic region. Specifically, the upstream-most SSB occurred 1,364 bp downstream from *Akt1* (CG4006; *serine/threonine–protein kinase B*) gene, and the downstream-most one 2,047 nt upstream from *Mhcl* (CG31045; *myosin heavy chain-like*) gene (**Figure 3A**). From our repeat annotation pipeline, the region around the breakages is a composite of repetitive sequences [16 in total, ranging in length from 21 bp of a $(TTG)_n$ simple-repeat to 532 bp of satellite rnd-4_family-179], interspersed with traces of transposable elements [84 bp from an LTR and 72 bp from a long interspersed nuclear element (LINE)]. Overall, no evidence of open reading frames and/or specific motifs could be found pointing to the observed breakages as directly caused by the insertion/excision of other sequences.

The role of non-B DNA as source of DSBs is well-established. Generally, DSBs are expected to colocalize with their causal non-B DNA motifs (e.g., Kolb et al., 2009; Lu et al., 2015). We used this prediction to investigate whether the local DNA conformational environment of the ancestral sequence could have acted as trigger or mediator of the complex rearrangement of the proximal breakpoint region. We proceeded in two steps:

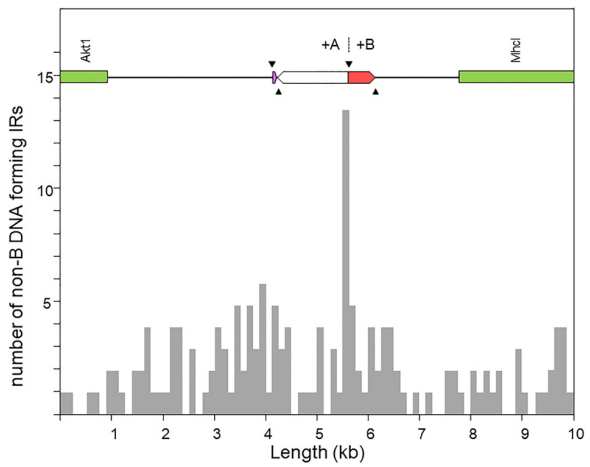


FIGURE 5 | Distribution of inverted repeats (IRs) with potential to adopt non-B DNA hairpin/cruciform structures across a 10-kb window centered at the region of the O_7 proximal breakpoint, as obtained using Palindrome Analyzer (Brázda et al., 2016). The region of interest is shown at scale above the plot. The purple and red boxes represent respectively the parental copies of d1 and d2 prior to their duplication as a result of the two pairs of staggered SSBs (black arrowheads). The highest concentration of IRs occurs around the junction of the O_7 breakpoint (+A|+B).

first, we reconstructed the region of the rearrangement before the breakages. It should be recalled that most of the rearranged sequence is embedded in an insertion that is absent in the ancestral non-rearranged state. Therefore, we reconstructed the prebreakages state by undoing the hypothetical rearrangement steps that generated the present sequence state. Specifically, we reversed the orientation of the microinversion (**Supplementary Figure 4**) and deleted one copy of each DSB-induced duplication (**Supplementary Figure 5**). The resulting sequence had the form: + d1, (+ 1,374 bp), |, + d2 (**Figure 4B**). Which copy of each of the two duplications to eliminate was inconsequential, because they are nearly identical to each other in the two cases (98.3% and 95.6%, for the identities between copies A and B of dup1 and dup2, respectively). Furthermore, the observed high level of identity (97.3%) between d2 and its homologous region in Ds_ch-cu and Ds_B suggested that the rearrangement is recent enough to allow assuming that the original conformational sequence features that could have mediated it are still observable. After establishing the prebreakage sequence, we next looked for sequences with the potential to form non-B DNA structures along a 10-kb window centered on it.

Figure 5 shows the distribution of the number of IRs capable of forming hairpin and cruciform structures along the target sequence. The highest density occurs immediately around the junction between the microinversion and inversion O_7. In particular, the breakpoint is located within a ∼150-nt-long stretch of AT-rich sequence [simple repeat (ATTT)n, from our genome annotation pipeline] containing 15 IRs, of which one located 68 nt downstream the breakpoint junction ranked in the top 5% with highest likelihood of intrastrand annealing to form a hairpin (AATTTT AAAATT; $\Delta G_S - \Delta G_L = 2.64$). In addition, embedded in the IR cluster, there is one tandem repeat of 8.7 copies of the consensus heptanucleotide AATAAAT, and one mirror repeat of two 11 nt-long repeats separated by a 30-nt spacer, indicating that the proximal breakpoint of O_7 occurred on an unstable sequence with potential for adopting multiple alternative non-B DNA conformations.

Distal Breakpoint of Inversion O_7

Figure 3B provides a schematic representation of the +C|+D region based on the alignment of **Supplementary Figure 6**. In the case of O_7, the region was reconstructed using the reverse complement of segment -C. From up to downstream, the breakpoint junction is located within a 450-aligned-sites-long gap-rich spacer region, spanning between two highly identical long IRs, IR1 and IR2, of 1,117 and 1,112 sites of alignment length, respectively. There is no evidence of duplicated sequence in Ds_7 relative to the other assemblies, indicating that the DSB either was a clean cut or did not involve significantly staggered SSBs. On the other hand, the spacer of Ds_7 was the shortest (250 nt) of all four lines (407, 317, and 343 nt for Ds_ch-cu, Ds_B, and Dg, respectively) because of a single deletion located precisely at the center of the region (hereon called dd7, for distal deletion of O_7). A closer look at the pattern of pairwise sequence similarities along the spacer revealed two findings: (i) dd7 split the Ds_7 spacer in two mirror halves. For the upstream half, Ds_7 is almost identical (96.8%) to Ds_ch-cu while bearing no detectable homology to Ds_B, whereas for the downstream half, Ds_7 is almost identical (97,6%) to Ds_B while bearing no detectable homology to Ds_ch-cu; and (ii) the spacer of Ds_ch-cu is almost identical (95.4%; excluding indels) to that of Ds_B but in reversed orientation. The reversal occurred in Ds_ch-cu, because in Ds_B the spacer is oriented as in the outgroup Dg.

Altogether, the above observations can be understood as follows (**Figure 4**). Prior to the origination of the distal breakpoint of O_7, a carrier of an uninverted chromosome of B-type experienced a reversal of the spacer region between the IRs, giving rise to the uninverted chromosome of ch-cu–type. Later on, a homokaryotype for the uninverted chromosome that was heterozygous for the microinversion of the spacer underwent at least two DSBs, one in each of two homologous non-sister chromatids, such that the DSB in the ch-cu–type chromatid occurred immediately before the first site of the dd7 and that in the B-type chromatid immediately after the last site of the dd7. Finally, the reversed + B end generated by the proximal staggered DSB in the ch-cu–type chromatid illegitimately joined with the + D end generated by the distal DSB in its homologous non-sister B-type chromatid, which resulted in a recombinant chromosome carrying the inversion O_7 with the exact observed dd7 deletion.

Like the proximal breakpoint, the distal breakpoint occurred in an intergenic region yet at comparatively much shorter distance (∼390 bp) to the nearest genes. Specifically, the breakage separated two copies of an *Attacin* gene (CG10146; *AttA*) located opposite to each other on each of the two arms of the long IR. Our repeat annotation pipeline did not identify repetitive sequences in the vicinity of the distal breakpoint in Ds_ch-cu or Ds_B.

We searched the region of the spacer for potential non-B DNA–forming sequences in the vicinity of the breakpoint

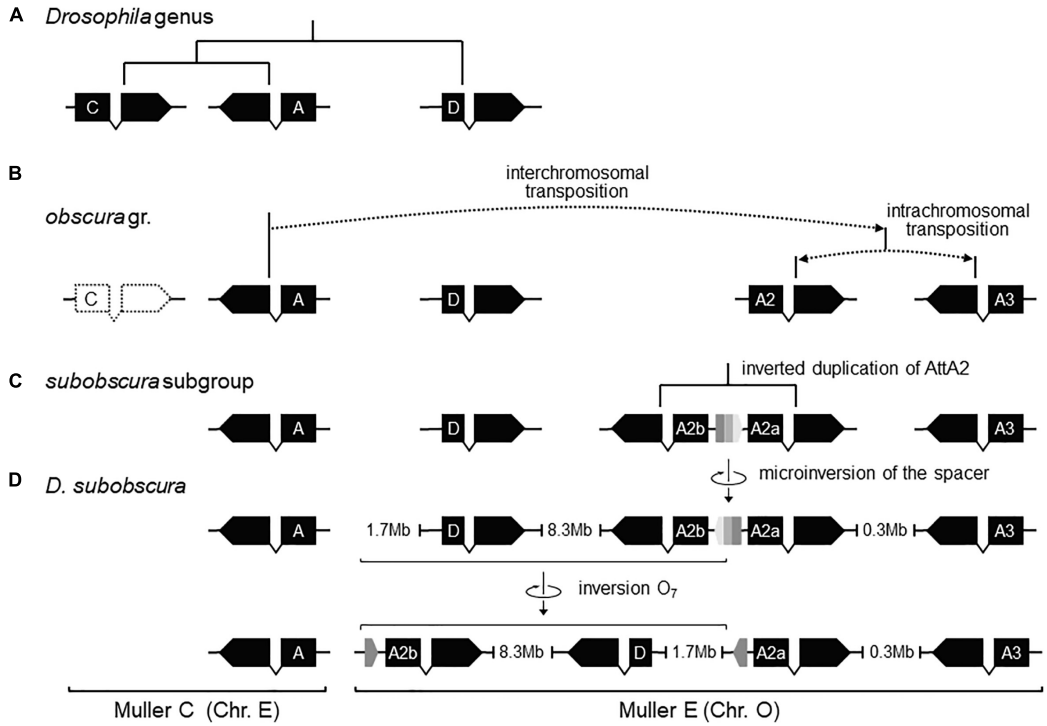


FIGURE 6 | Pre-O_7 history of instability of the distal breakpoint. **(A)** In the most recent common ancestor of the *Drosophila* genus, *AttD* was the only *Attacin* gene present in Muller element E. **(B)** Later, the ancestor of the obscura group lost *AttC*, and underwent DNA-based interchromosomal transposition of *AttA* (or a close paralog; see **Supplementary Figure S7**) from Muller C to Muller E, followed by DNA-based intrachromosomal transposition within Muller E, giving rise to *AttA2* and *AttA3* (whether simultaneously or sequentially and, if the latter, which was first is unknown). **(C)** Before the split of the subobscura subgroup, *AttA2* was duplicated, giving rise to the inverted duplicates *AttA2a* (parent copy) and *AttA2b* (daughter) separated by a short central spacer. **(D)** In *D. subobscura*, the central spacer underwent a reversal, generating a microinversion polymorphism with segregating states B-type (ancestral) and ch-cu–type (derived). Genes are represented as solid black boxes (exons) linked by polygonal lines (introns), and oriented in the direction of transcription. The central spacer is represented as a box colored in three shades of gray pointing in the direction of its orientation.

junctions in Ds_ch-cu and Ds_B. In both cases, we found that the IR with the highest propensity to form a hairpin was a perfect 14-bp-long palindromic sequence located next to the breakpoint junctions (ATGAACT AGTTCAT; $\Delta G_S - \Delta G_L = = 2.05$; located 13 and 2 bp upstream and downstream the junction in Ds_ch-cu and Ds_B, respectively). Apart from IRs, we did not detect additional potential non-B DNA sequences around the distal breakpoint.

All nucleotides in the +A|−C region of Ds_7 could be unambiguously ascribed to segment A or C. However, in the −B|+D region −B and +D are separated by 21 extra inserted nucleotides (i.e., GAGCACTCTCCACAGCAAAGT). We decided to ascribe this sequence to the distal breakpoint junction, because it contains an 8-bp substring (underlined) that resembles the beginning of the +D end (CATCAAAG), and hence it likely represents filler DNA generated by a microhomology-templated repair mechanism.

Pre-inversion Record of Rearrangement of O_7 Breakpoints

Previously, it was shown that the proximal breakage of O_7 was preceded by an insertion. Likewise, the region of the distal breakage had a pre-inversion history of rearrangement, which run closely associated with a highly dynamic evolution of the *Attacin* immunity gene family in the *obscura* species group. This conclusion is based on phylogenetic analysis of the *Attacin* family in *Drosophila* (**Supplementary Figure 7**) using synteny to distinguish orthologous from paralogous copies (**Supplementary Table 1**). The results are summarized in **Figures 6A–D**. The most recent common ancestor of the *Drosophila* genus (**Figure 6A**) carried three copies of the gene with relationships [(A,C),D], of which the more distant D was located in Muller element E, and the closer to each other A and C in Muller element C. After it split from the *melanogaster* group (**Figure 6B**), the branch leading to the *obscura* group lost copy C and underwent an interchromosomal transposition of copy A from Muller element C to E. The daughter copy then underwent another, in this case intrachromosomal, transposition, which originated two new *Attacin* copies that we called *AttA2* and *AttA3*, with *AttA2* located between *foxo* and *Npc2b*, and *AttA3* located ∼300 kb downstream from *AttA2*, between *Cul5* and *Sirt7*. The two transpositions were genome-based duplications rather than retroposition events, because the new copies conserved the intron position of their parental gene. Before the split of the *subobscura* subgroup (**Figure 6C**), copy *AttA2* underwent an inverted duplication that generated the two closely

spaced copies *AttA2b* and *AttA2a* in head-to-head orientation, and transcribed in opposite directions. In *D. subobscura* (**Figure 6D**), the spacer between the IRs experienced a reversal of orientation generating the microinversion polymorphism of the distal breakpoint. Subsequently, a heterozygote for the microinversion underwent distal DSBs that allowed the formation of the recombinant O_7 inversion via ectopic repair of non-sister chromatids.

Potentially Functional Effects of the O_7 Mutation

The distal break of O_7 disrupted concerted evolution between two subobscura subgroup-specific *AttA2* duplicates. This conclusion is based on the previous section's results, together with the phylonetwork of coding sequences shown in **Figure 7**. Accordingly, right after the duplication of *AttA2*, the two paralogs began to evolve in concert, converting each other to generate their present characteristic phylogenetic pattern of greater resemblance between paralogs from the same species (i.e., *D. guanche* and *D. subobscura*) than between orthologs from different species (e.g., Puig-Giribets et al., 2019). At one end of the resemblance, it is the ch-cu strain, whose two *AttA2* copies are identical to each other, and at the other end O_7, where the copy relocated by the inversion evolved significantly faster than the one that remained in place, owing exclusively to an acceleration of the synonymous substitution rate [$P < 0.05$; Tajima's relative rate test (Tajima, 1993) using either of the remaining six sequences as outgroup], as the two copies are identical at the amino acid level. The acceleration took place in the direction of a slight decrease in codon bias in the relocated copy (Nc = 51.2 vs. 50.7, for the comparison *AttA2b* vs. *AttA2a*, respectively; where Nc is the improved effective number of codons index; Sun et al., 2013). The increased synonymous rate can be understood, in part because the inversion released the two *Attacin* copies from evolving in concert; and in part assuming that the expression of the paralogs shifted as a result of changes in regulatory environment associated with their relocation.

Considering the short spacing between the two *AttA2* paralogs in the uninverted chromosome (∼390 bp), it appeared likely that the inversion would have detached them from part of their promoters, binding them to new potentially *cis*-acting elements. To assess this possibility, we searched 1 kb upstream of the predicted TSS of each gene for putative transcription factor binding sites (TFBSs) for five transcription factors (TFs), including the nuclear factor κB factors dorsal (dl), dorsal-immunity related factor Dif and Relish (Rel), the GATA factor Serpent (*srp*), and the forkhead factor dFOXO. The first four TFs are under control of the Toll and immune deficiency (IMD) immunity pathways and regulate *Attacin* inducible expression in response to bacterial infection (Senger et al., 2004). dFOXO TF is controlled by the insulin/insulin-like growth factor signaling (IIS) metabolic pathway and regulates constitutive *Attacin* expression in non-infected flies suffering from energy shortage or stress (Becker et al., 2010). The results are shown in **Figure 8**. The *AttA2* genes had predicted TFBSs for the immunity related factors in both uninverted and inverted chromosome states, but only the *AttA2* genes of the inverted chromosome had TFBSs for the metabolic factor dFOXO. Furthermore, the dFOXO TFBSs were all contributed by the newly attached sequence. The fact that the *AttA2* genes were conserved at the amino acid level in *D. subobscura*, together with the observed qualitative difference in predicted *cis*-acting sequence between uninverted and inverted chromosomes, suggests that the inversion O_7 brought the *AttA2* genes under the influence of the IIS metabolic pathway.

In addition to the *Attacin* immunity genes, the breakpoint regions include *Akt1* and *foxo*, two interacting core components of the IIS metabolic pathway identified by other studies as candidate for climate adaptation (Fabian et al., 2012; Paaby et al., 2014; Kapun et al., 2016; Durmaz et al., 2019). The roles of these genes and the potential impact of O_7 on them are dealt with in the *Discussion*.

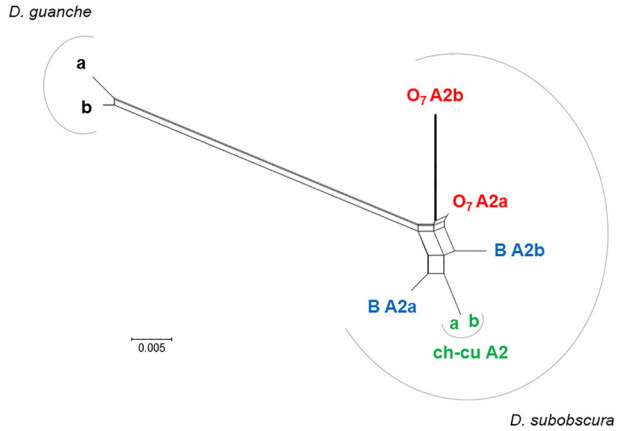


FIGURE 7 | Phylogenetic network of *AttA2* orthologous and paralogous sequences from the inverted (O_7) and uninverted (B and ch-cu) states in *D. guanche* and *D. subobscura*. The duplication is older than the species, but the paralogs cluster within species owing to concerted evolution. Depicted thicker is the branch leading to the copy relocated by O_7 (i.e., *AttA2b*), which is longer than that leading to the copy that remained in place (*AttA2a*) due to an acceleration of the synonymous substitution rate, likely as a result of having escaped concerted evolution. The split network was constructed using the NeighborNet method as implemented in SPLITSTREE version 4.14.5 (Huson and Bryant, 2006), on the JC69 + I (% of invariable sites 81.6) best-fit model distances obtained using the DIVEIN web server (https://indra.mullins.microbiol.washington.edu/DIVEIN/) (Deng et al., 2010). Sets of parallel edges represent conflicting topological signals.

DISCUSSION

Molecular Mechanism of O_7 Formation

O_7 Is a Complex Multibreak Inversion Formed via Rejoining in *trans* With the Two Homologous Chromosomes

Sequence data on inversion formation in *Drosophila* have been interpreted in terms of two major mechanisms with associated distinctive footprints. The first mechanism is intrachromatidal NAHR between inversely oriented repeats. This mechanism generates inversions with duplications at their ends in both the

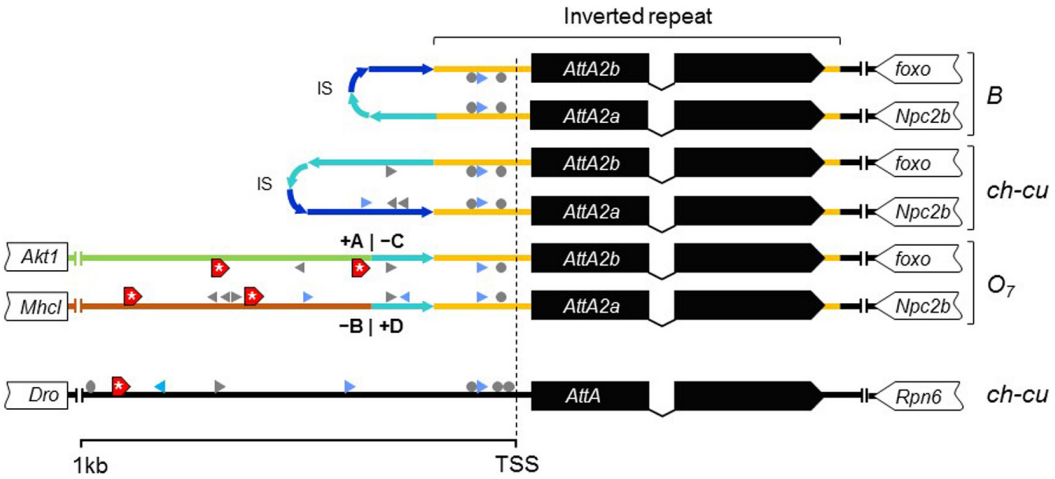


FIGURE 8 | New dFOXO binding sites. Schematic representation of up to 1-kb sequence up and downstream predicted TSSs of *AttA* and the inverted duplicates of *AttA2* (represented as in **Figure 6**) in O_7 and the uninverted (B and ch-cu) states, including also the nearest flanking genes. Colored lines connecting genes designate the following: orange, region of the inverted repeats; dark and light blue, first and second halves of the spacer of the inverted repeats, respectively, oriented as the arrowheads; green and brown, the novel sequences to which the *AttA2* copies became reattached by O_7, with corresponding breakpoints (+A|−C and −B|+D) indicated. The inverted repeats of B and ch-cu are folded over each other. Putative TFBs are symbolized: gray arrowheads and circles (palindromic sites), for Dorsal, Relish, and Diff/Relish; blue arrowheads for Serpent, and red boxes with an asterisk for dFOXO, respectively. Only *AttA* and the two *Att2A* copies of O_7 have TFBs for dFOXO.

inverted and uninverted states (Cáceres et al., 1999), which is not the case of O_7.

The second mechanism is chromosomal breakage and ectopic repair via NHEJ. This mechanism either does not generate duplications or generates them but at the ends of the inverted state only. These two types of NHEJ footprints have been explained in terms of two alternative modes of breakage: cut-and-paste via clean DSBs that generate blunt ends and staggered on the same (isochromatidal) or different (chromatidal) sister chromatids (see *Introduction*). In the case of O_7, it is not a cut-and-paste inversion, but neither is it a typical staggered breaks inversion. Thus, while the inversion proximal breakpoint could be either isochromatidal (**Figure 4**) or chromatidal (**Figure 9**), the distal breakpoint has to involve the two homologous chromosomes (**Figures 4**, **9**). This latter pattern could be deduced because of the chanceful circumstance that our two representatives of the uninverted state (i.e., Ds_ch-cu and Ds_B) segregated for the microinversion of the spacer between the IRs flanking the distal breakpoint. Alternatively, the distal breakage could have occurred in a recombinant between chromosome types ch-cu and B. This, however, appears unlikely because crossover within microinversions should be extremely rare (Greig, 2007). Our conclusion agrees with a study of the genealogical relationships between inversions of the E chromosome in *D. subobscura*, which proposed that E_9 arose in a heterokaryotype E_{ST}/E_{1+2} to accommodate a conflict between molecular and cytological data (Orengo et al., 2019). This and our results indicate that NHEJ inversions form through mechanisms that can incorporate information from the two homologous chromosomes (chromosome model), in addition to the previously proposed intrasister and intersister chromatidal exchanges.

The Breaks of the O_7 Inversion Were Likely Induced by Non-B DNA Secondary Structures

Inversion O_7 provides, to our knowledge, the first compelling evidence for a role of non-B DNA in inversion formation in *Drosophila*. Previous studies had reported the presence of AT-rich sequences around the breakpoints of some fixed (Cirera et al., 1995; Richards et al., 2005) and polymorphic (Prazeres da Costa et al., 2009) inversions. In no instance, however, were particular sequences susceptible to adopt secondary structures identified. In the case of O_7, the proximal break junction occurred just within a palindromic AT-rich repeat capable of adopting hairpin/cruciform, slipped and triplex DNA conformations. Likewise, the distal junctions are located next to perfect 14-bp-long hairpin/cruciform-forming palindromes.

The role of non-B DNA-forming sequences in causing genome instability is well-established (Wang and Vasquez, 2006; Lobachev et al., 2007; Aguilera and Gómez-González, 2008; Zhao et al., 2010). The shift from B to non-B DNA conformation occurs while DNA is in single-stranded form, e.g., behind replication forks, between Okazaki fragments, or in actively transcribed genes (Voineagu et al., 2008). Non B-DNA structures induce DSBs through, e.g., stalling replication and transcription (Mani and Chinnaiyan, 2010; Kaushal and Freudenreich, 2019). There are no specific predictions as to the type, number, and location of the DSBs generated by any given structure in any particular situation. Still, a single structure can induce multiple DSBs across hundreds of base pairs around it (Wang et al., 2006; McKinney et al., 2020), and stalled replication forks can accumulate up to 3 kb of single-stranded DNA (Sogo et al., 2002; Lopes et al.,

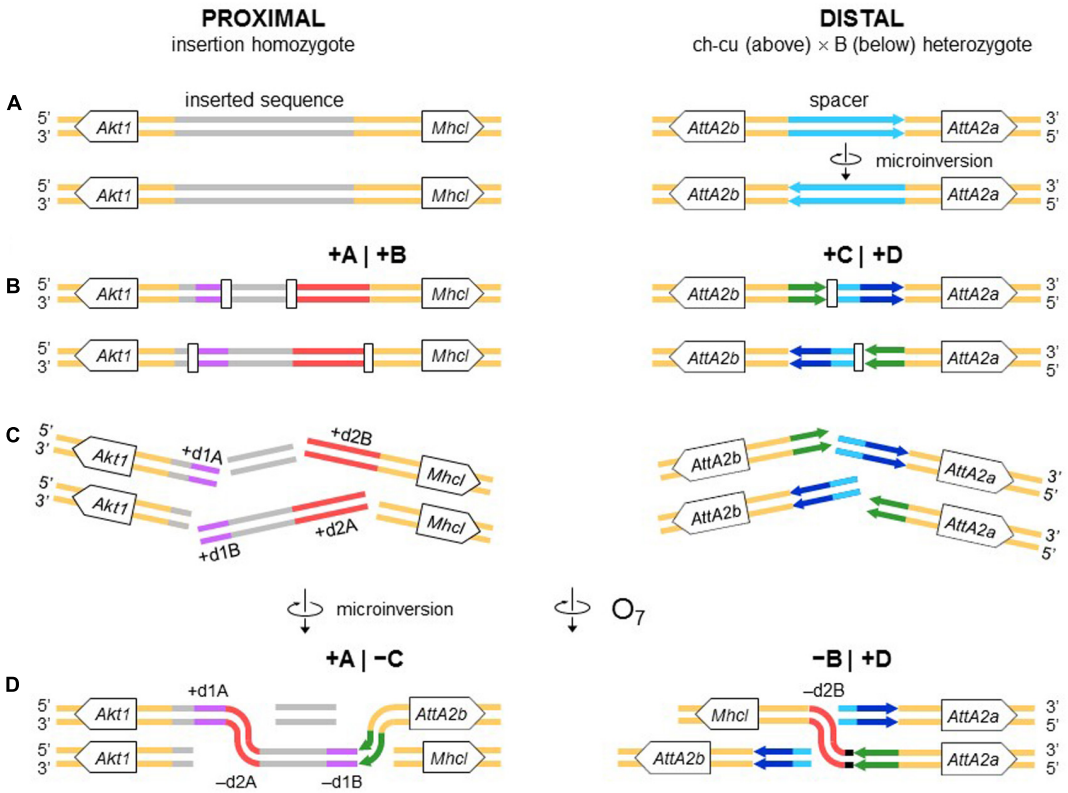


FIGURE 9 | Chromosome model of O_7 formation. **(A)** Start from an individual homozygous for the insertion at the proximal breakpoint, and as in **Figure 4** at the distal breakpoint. **(B)** Occurrence of two pairs of staggered chromosomal blunt-ended DSBs (open rectangles) at the proximal breakpoint, and as in **Figure 3** at the distal breakpoint. **(C,D)** Microinversion formation, and formation of O_7 via rejoining in *trans* with the homologous chromosome as indicated. The model results in an order of the duplications (i.e., + d1A, -d2A, -d1B, -d2B) identical to that resulting from the model of **Figure 4**.

2006). In the case of O_7, this length is well over the size of the overhangs that would be generated by an isochromatid model of the proximal breakpoint (58 and 534 nt; see **Figure 3** and **Supplementary Figure 3**).

The Inverted Duplications at the O_7 Breakpoints Could Be Footprints of Repair Instead of Staggered Breakage

All the aforementioned inverted duplication-generating NHEJ models are predicated upon the role of DNA breakage (Ranz et al., 2007). However, the inverted duplications at the ends of O_7 could also be explained as a result exclusively of repair, with no need for invoking staggering of the breaks. DNA repair has emerged as a key factor capable of generating extremely complex breakpoint sequence rearrangements (reviewed in Scully et al., 2019). The spectrum of known error-prone repair mechanisms can be grossly classified as recombination-based, such as microhomology-mediated end-joining (MMEJ), and replication-based, such as break-induced replication (BIR) and microhomology-mediated BIR (MMBIR) (Lee et al., 2007; Zhang et al., 2009; Hastings et al., 2009). Here, the term *microhomology* is used to mean a short tract (~1 – 25 bp) of chance similarity, rather than common descent. In the case of O_7, three features suggest that what appear to be footprints of breakage by the staggering models could in fact be footprints of a replication-based mode of repair (reviewed in Kramara et al., 2018; Scully et al., 2019), including (i) presence of non-B DNA-forming sequences just in, or adjacent to breakpoint junctions (see below); (ii) spatial proximity of the breakpoint regions in the nucleus, as evinced by the fact that the genes flanking the junctions are closely related functionally (Farré et al., 2015; but see Sunder and Wilson, 2019); and (iii) multiple breaks concentrated in a short sequence segment. A fourth feature, namely, presence of microhomology at the distal breakpoint junction, would be also consistent with a recombination-based mechanism such as MMEJ. Overall, these features suggest that O_7 arose as result of a non-B DNA-induced replication impairment, affecting at least its proximal breakpoint. It is known that this type of events can trigger BIR and MMBIR repair (Sakofsky et al., 2015). Of the two pathways, the second pathway has yet to be identified in *Drosophila* (Alexander et al., 2016; Bhandari et al., 2019). A possible scenario is detailed in **Figure 10**: first, non-B DNA-induced stalling of a replication fork at the proximal breakpoint of a ch-cu–type chromosome led to two DSBs generating three fragments. Second, the centromere-proximal fragment engaged in a BIR event using the homologous region of a B-type chromosome. Third, a second fork stalling triggered a switch from BIR to MMBIR with template switching to

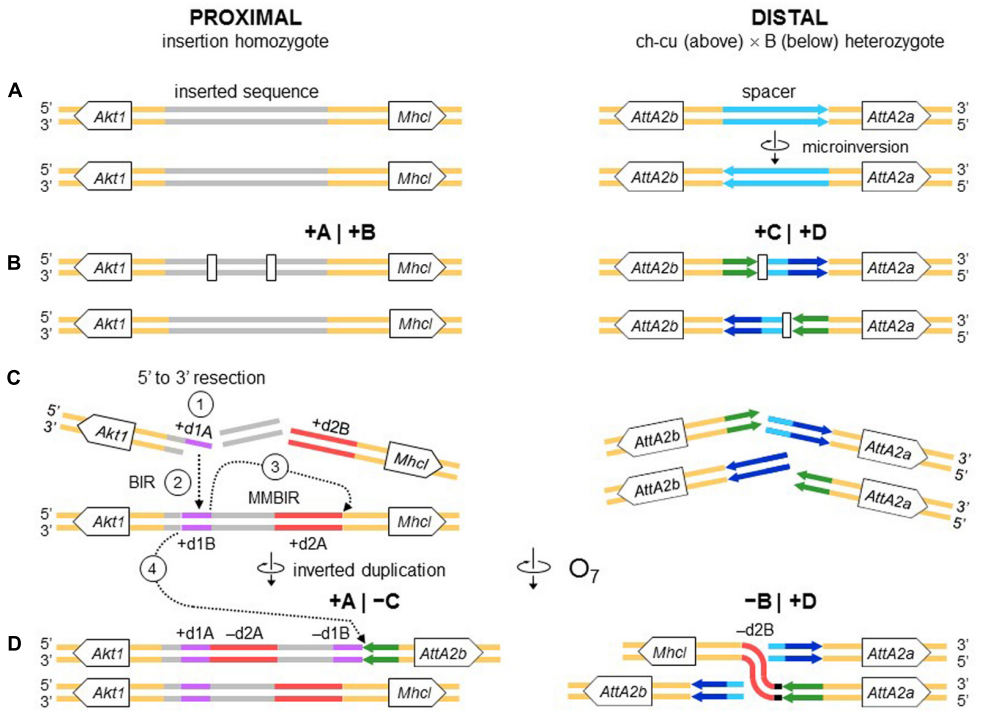


FIGURE 10 | Chromosome and BIR/MMBIR repair model of O_7 formation. **(A)** Start as in **Figure 9**. **(B)** Occurrence of one pair of blunt-ended DSBs (open rectangles) at the proximal breakpoint of the ch-cu type chromosome, and as in **Figure 4** at the distal breakpoint. **(C,D)** Step 1: 5′ to 3′ resection generating a 3′ single stranded + d1A end. Step 2: beginning of a BIR event via strand invasion into the homologous region of the B-type chromosome. Step 3: switch from BIR to MMBIR, with forward template switching to the distal end of + d2A and backward copying. Step 4: MMEJ to the distal break-end of O_7 from the original ch-cu–type chromosome. The distal breakpoint repaired as in **Figure 9**. The model results in an order of the duplications (i.e., + d1A, -d2A, -d1B, -d2B) identical to that resulting from the model of **Figures 4**, **9**.

a downstream microhomology. Copying backward from the new template resulted in the rearrangement of the proximal breakpoint, including the inverted duplication of the O_7 end (e.g., Lee et al., 2007; Smith et al., 2007; Carvalho et al., 2015; Tremblay-Belzile et al., 2015). Finally, the event was terminated by an MMEJ to the distal break-end of O_7 from the original ch-cu-type chromosome (e.g., Scully et al., 2019).

The O_7 Breakpoints Carry a Pre-inversion Record of Fragility

The breakpoint sequences of O_7 had a record of instability prior to the origin of the inversion, as evinced by the fact that they are located within sequences inserted from elsewhere in the genome. This suggests that the regions that gained those insertions were relatively exposed in the nucleus (reviewed in Farré et al., 2015). In the case of the proximal breakpoint, that could be associated with high levels of transcriptional activity at the broadly expressed *Akt1* gene (Andjelković et al., 1995; Slade and Staveley, 2016).

That the O_7 junctions arose in fragile regions, beyond the proximate effects of their associated non-B DNA (see above), may be most apparent from the pre-inversion record of recurrent rearrangement of the IR at the distal breakpoint (**Figure 6**). This record is particularly amenable to reconstruction because the IR largely consists of two copies of the *Attacin A* gene that are highly conserved. It includes at least three rearrangements that occurred in the lineage of *D. subobscura* after its separation from that of the *melanogaster* group (see section "RESULTS"; **Figure 6**), namely, (i) insertion of *AttA2* between the *foxo* and *Npc2b* genes; (ii) emergence of the IR by inverted duplication of the parental *AttA2* (**Figure 6B**), which could have occurred through an event of forward template switching and backward copying by the DNA polymerase (Smith et al., 2007; Lee et al., 2007), as discussed above; and (iii) emergence of the ch-cu–type chromosome via inversion of the spacer between the IRs in a B-type chromosome (**Figure 6D**), which could be explained as an outcome of a stem-loop formation by the IR, followed by resolution of the strand-exchange junctions between the IR arms (see Figure 4 in Leigh Brown and Ish-Horowicz, 1981; Figure 3 in Kolb et al., 2009 and Zhao et al., 2010).

The pre-O_7 insertion in the proximal breakpoint is specific to *D. subobscura* and is therefore much more recent than that of *AttA2* in the distal breakpoint. Preliminary analyses indicate that it is internally rearranged relative to other paralogous copies, supporting that it carries recombinogenic potential. The origin and evolution of this inserted sequence, as well as its possible implication in the formation of other *D. subobscura* inversions, warrant further investigation (CK, RT, and FR-T; manuscript in preparation).

O7 Breakpoints Potentially Functional Effects

O_7 Relocated *foxo* in Tight Linkage Association With Its Antagonistic Regulatory Partner of the IIS Metabolic Pathway *Akt1*

O_7 changed *foxo* from being megabases (∼10 Mb) away from *Akt1* to being tightly linked to it, with only the short *AttA2b* gene sandwiched between them. *Akt1* and *foxo* are functionally conserved genes, which, in *Drosophila*, encode the serine/threonine–protein kinase B AKT/PKB, and the forkhead-box DNA-binding domain-containing TF dFOXO, respectively. The two genes are key antagonistic regulators of the IIS pathway (Teleman, 2010; Slade and Staveley, 2016), a major trigger of shifts in anabolic versus catabolic cellular activity in response to nutritional status (de Jong and Bochdanovits, 2003) and multiple other cues (Regan et al., 2020). In abundant nutrient conditions, AKT/PKB inactivates dFOXO, thus shifting food energy allocation toward reproduction and growth (the IIS pathway). Conversely, scarce nutrient conditions prevent AKT/PKB from inactivating dFOXO, which redirects metabolism toward mobilization of energy stores for somatic maintenance (FOXO pathway). Laboratory research using large effect mutants has shown that the IIS/FOXO pathway is extensively pleiotropic, with major evolutionary conserved effects on fitness-related life-history traits, including growth, size, reproduction, lifespan, and stress resistance (reviewed in Flatt and Partridge, 2018). Research from the field found IIS loci to harbor substantial genetic variation, which frequently exhibits spatiotemporal patterns that look as if they were shaped by selection on the associated IIS traits (Fabian et al., 2012; Paaby et al., 2014; Kapun et al., 2016). In a recent laboratory assay, two *foxo* alleles showing opposite latitudinal clines in *D. melanogaster* were compared on an otherwise homogeneous genetic background. The alleles showed contrasting effects on viability, size-related traits, starvation resistance, and fat content, whose directions were overall consistent with predictions from the clinal variation of the characters (Durmaz et al., 2019).

The O_7 mutation could have altered *Akt1* and/or *foxo* function via multiple non-mutually exclusive mechanisms, such as mutual regulatory interference, considering that they are antagonistic effectors; relocation to the sides of an immunity gene (i.e., *AttA2b*) expected to be under intense purifying selection on expression (see below); and alteration of the genes' functional neighborhood at higher-order levels of chromatin organization (Farré et al., 2015; McBroome et al., 2020). It could be argued that the nuclear environment of the genes remained basically unaltered, if the reason why they became involved in the rearrangement was that they already were in close spatial proximity to each other in the nucleus. This, however, did not necessarily have to be the case, considering recent findings in yeast that rejoining of DNA break ends is not determined by the predamage spatial proximity of the DSBs (Sunder and Wilson, 2019). Be that as it may, bearing in mind that the seasonal increase of O_7 occurs from early spring to midsummer, coinciding with the growth season, it seems more likely that whatever the effect of the inversion mutation on *Akt1* and/or *foxo*, it occurred in the direction of an enhanced basal IIS versus dFOXO activity relative to the O_{ST} ancestral state. This would raise the question of why the O_7 frequencies decrease (and those of O_{ST} increase) every year from late summer to winter.

O_7 Disrupted the Concerted Evolution of Two *AttA2* Immunity Genes and Reattached Them to Putative dFOXO Metabolic Enhancers

The immune function is highly energy demanding in terms of both maintenance and, especially, rapid deployment upon infection (reviewed in Dolezal et al., 2019). Therefore, within a limited energy budget, a trade-off is expected between reproduction and immunity (Schwenke et al., 2016). The *Drosophila* innate immune response consists of a cellular and a humoral component. The humoral component involves the production of antimicrobial peptides, among which Attacins are active against gram-negative bacteria (Hanson and Lemaitre, 2020). The two main modes of Attacin production, including the induced (by a factor of even > 100) upon infection mode, and the basal in absence-of-infection mode link immunity with the *Akt1*/*foxo* IIS metabolic signaling pathway (Becker et al., 2010; Dolezal et al., 2019). The inducible mode is regulated primarily by the immunodeficiency *Imd* signaling pathway and to a lesser extent by the *Toll* signaling pathway. The two signaling pathways have the same effect of activating dFOXO, thus mobilizing resources toward the production of Attacins (Dionne et al., 2006; Dolezal et al., 2019). The basal mode is regulated directly by dFOXO activity when induced by starvation (Becker et al., 2010; Buchon et al., 2014). Immunity genes, including *Attacins*, are among the known most rapidly evolving genes and have frequently shown evidence of local adaptation in *Drosophila* (Lazzaro and Clark, 2001, 2003).

There would be a number of mechanisms by which the O_7 mutation could have reduced *Attacin* genes' expression. For example, the breakage of the invertedly transcribed *AttA2* tandem duplicates could have impaired the inducibility of one or the two paralogs, or their separation could have made them lose gene expression coregulation, as might be surmised from the observations that they halted or slowed down evolving in concert, and that *AttA2b* shows decreased codon bias. These mechanisms could have acted synergistically with each other and with those already discussed in connection with *Akt1* and *foxo*. Although this scenario could be partially offset by the increase in basal *AttA2* transcript levels that may be expected from the duplicates having been reattached to dFOXO enhancers (Becker et al., 2010), all in all, the evidence suggests that (i) at its inception, O_7 caused a rearrangement with partial disruption of a set of functionally related loci with overlapping pleiotropic effects on immunometabolic traits. If, in addition to these direct effects, there concurred indirect effects of linkage between locally, and given the functional relationship, likely epistatically interacting alleles warrant further investigation; and (ii) the resulting haplotype imparted a shifted

pattern of resource allocation toward reproduction at a cost to immunity, compared to the O_{ST} ancestor. Such an opposing antagonistic pleiotropy would result in a seasonal frequency cycle qualitatively similar to that shown by the inversions, if reproduction is favored from early spring to midsummer, when O_7 rises (and O_{ST} wanes), and immunity from late summer to winter, when it wanes (and O_{ST} rises). There is ample evidence that the qualitative and quantitative composition of temperate bacterial communities cycles seasonally (Lazzaro et al., 2015; Shigyo et al., 2019). Recently, a study using *D. melanogaster* from the eastern United States (Behrman et al., 2018) found a seasonal shift in immunocompetence, with the trait value declining every spring to autumn. The shift was interpreted as resulting from relaxed selection for immune response during the warm season, much like what we propose here for the O_7/O_{ST} inversion polymorphism. Prior data on temporal genetic variation within and between O inversions point to additional loci that would be consistent with the seasonal cycle of O_7 being mediated by immunometabolic selection (Rodríguez-Trelles, 2003). The case of the *Mpi* gene encoding the key glycolytic enzyme mannose-6-phosphate isomerase (MPI) is noteworthy. From our assembly, *Mpi* is located 2.15 Mb outward from the distal breakpoint of O_7, which is within the estimated region of the inversion-associated strong recombination–suppression effect (3.5 Mb; Pegueroles et al., 2010b). The MPI fast/slow electrophoretic polymorphism was found to be only moderately associated with the O_7/O_{ST} polymorphism. Yet (i) the magnitude of the locus-by-inversion disequilibrium cycled seasonally, and (ii) the cycling occurred because the Fast allele increased in frequency every winter only within the O_7 chromosomal class, but not within the O_{ST} class (Rodríguez-Trelles, 2003). The behavior of *Mpi* could be in part an outcome of hitch-hiking with other linked loci involved in seasonal adaptation. One such candidate could be the *Na pumpα subunit* (*Atpα*) gene, located only 0.13 Mb farther away from O_7 than *Mpi*, and recently found to be under positive selection for defense against plant secondary compounds in *D. subobscura* (Pegueroles et al., 2016). Still, immune elicitation in *Drosophila* relies upon massive upregulation of glycolysis (Dolezal et al., 2019), which should place a strong demand on MPI activity (Shtraizent et al., 2017). In addition to the evidence from *D. subobscura* just discussed, **Supplementary Table 2** provides additional loci found to exhibit seasonal variation in a genomic survey from other *Drosophila*, which may be candidates for being involved in the seasonal cycling of O_7.

CONCLUSION AND OUTLOOK

Previous work on the spatiotemporal distribution patterns of the inversion polymorphisms of *D. subobscura* indicated that O_7 is driven by selective factors other than temperature alone. Here, we addressed this issue using a genome-based approach to isolate and characterize the O_7 breakpoints. Our findings have general implications for current theories on the molecular mechanisms of formation of this common type of structural genomic change. Furthermore, they suggest that O_7 may have altered fly's immunometabolism through at least direct effects on core immunity and metabolism genes. This result could help to explain the inversion's conflicting correlations with the seasonal and decadal climate changes, taking into account recent findings from microbial ecology, which indicate that microbial community responses to short- and long-term climate changes can be largely uncorrelated (Romero-Olivares et al., 2017). Considering its large size, it seems likely that O_7's evolution is also shaped by additional direct or/and indirect effects on genes other than those near its breakpoints. Further progress along this line will include development of functional tests of the identified genes on inverted versus uninverted chromosome backgrounds and use of the obtained assembly for building a SNP panel for O chromosome-wide scans of selection. We have incorporated the chromosome-scale sequence of O_{3+4+7} obtained here into our reference genome browser[9] to facilitate the further use of this resource.

DATA AVAILABILITY STATEMENT

Datasets presented in this article are available at the European Nucleotide Archive (ENA) under the project ID: PRJEB38585.

AUTHOR CONTRIBUTIONS

CK, RT, and FR-T contributed to the design and implementation of the research, to the analysis of the results, and to the writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Spanish Ministerio de Ciencia e Innovación grant CGL2017-89160P; and Generalitat de Catalunya grant 2017SGR 1379 to the Grup de Genòmica, Bioinformàtica i Biologia Evolutiva, Universitat Autònoma de Barcelona (Spain). CK was supported by a PIF Ph.D. fellowship from the Universitat Autònoma de Barcelona (Spain). Note that the funding agencies were not involved in the design of the study or in any aspect of the collection, analysis and interpretation of the data or paper writing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.565836/full#supplementary-material

[9] http://dsubobscura.serveftp.com/

REFERENCES

Aguilera, A., and Gómez-González, B. (2008). Genome instability: a mechanistic view of its causes and consequences. *Nat. Rev. Genet.* 9, 204–217. doi: 10.1038/nrg2268

Alexander, J. L., Beagan, K., Orr-Weaver, T. L., and McVey, M. (2016). Multiple mechanisms contribute to double-strand break repair at rereplication forks in Drosophila follicle cells. *Proc. Natl. Acad. Sci. U.S.A.* 113, 13809–13814. doi: 10.1073/pnas.1617110113

Andjelković, M., Jones, P. F., Grossniklaus, U., Cron, P., Schier, A. F., Dick, M., et al. (1995). Developmental regulation of expression and activity of multiple forms of the Drosophila RAC protein kinase. *J. Biol. Chem.* 270, 4066–4075. doi: 10.1074/jbc.270.8.4066

Anisimova, M., Gil, M., Dufayard, J. F., Dessimoz, C., and Gascuel, O. (2011). Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* 60, 685–699. doi: 10.1093/sysbio/syr041

Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., et al. (2004). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 32, D115–D119. doi: 10.1093/nar/gkw1099

Arenas, C., Zivanovic, G., and Mestres, F. (2018). Chromosomal thermal index: a comprehensive way to integrate the thermal adaptation of *Drosophila subobscura* whole karyotype. *Genome* 61, 73–78.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., and Cherry, J. M. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Ayala, F. J., Serra, L., and Prevosti, A. (1989). A grand experiment in evolution: the *Drosophila subobscura* colonization of the Americas. *Genome* 31, 246–255. doi: 10.1139/g89-042

Bächli, G. (2020). *TaxoDros: The Database on Taxonomy of Drosophilidae*. Available online at: https://www.taxodros.uzh.ch (accessed February 21, 2020).

Bailey, T. L., Johnson, J., Grant, C. E., and Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Res.* 43, W39–W49. doi: 10.1093/nar/gkv416

Balanyà, J., Oller, J. M., Huey, R. B., Gilchrist, G. W., and Serra, L. (2006). Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science* 313, 1773–1775. doi: 10.1126/science.1131002

Bao, W., Kojima, K. K., and Kohany, O. (2015). Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* 6:11.

Becker, T., Loch, G., Beyer, M., Zinke, I., Aschenbrenner, A. C., Carrera, P., et al. (2010). FOXO-dependent regulation of innate immune homeostasis. *Nature* 463, 369–373. doi: 10.1038/nature08698

Behrman, E. L., Howick, V. M., Kapun, M., Staubach, F., Bergland, A. O., Petrov, D. A., et al. (2018). Rapid seasonal evolution in innate immunity of wild *Drosophila melanogaster*. *Proc. Biol. Sci.* 285:20172599. doi: 10.1098/rspb.2017.2599

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573

Bhandari, J., Karg, T., and Golic, K. G. (2019). Homolog-dependent repair following dicentric chromosome breakage in *Drosophila melanogaster*. *Genetics* 212, 615–630. doi: 10.1534/genetics.119.302247

Bracewell, R., Chatla, K., Nalley, M. J., and Bachtrog, D. (2019). Dynamic turnover of centromeres drives karyotype evolution in Drosophila. *eLife* 8:e49002. doi: 10.7554/eLife.49002

Brázda, V., Kolomazník, J., Lýsek, J., Hároníková, L., Coufal, J., and Šťastný, J. (2016). Palindrome analyser - A new web-based server for predicting and evaluating inverted repeats in nucleotide sequences. *Biochem. Biophys. Res. Commun.* 478, 1739–1745. doi: 10.1016/j.bbrc.2016.09.015

Brehm, A., and Krimbas, C. B. (1988). The inversion polymorphism of *Drosophila subobscura* natural populations from Portugal. *Genét. Ibér.* 39, 235–248.

Buchon, N., Silverman, N., and Cherry, S. (2014). Immunity in *Drosophila melanogaster* –from microbial recognition to whole-organism physiology. *Nat. Rev. Immunol.* 14, 796–810. doi: 10.1038/nri3763

Cáceres, M., Ranz, J. M., Barbadilla, A., Long, M., and Ruiz, A. (1999). Generation of a widespread Drosophila inversion by a transposable element. *Science* 285, 415–418. doi: 10.1126/science.285.5426.415

Campbell, M. S., Holt, C., Moore, B., and Yandell, M. (2014). Genome annotation and curation using MAKER and MAKER-P. *Curr. Protoc. Bioinformatics* 48, 4.11.1–4.11.39. doi: 10.1002/0471250953.bi0411s48

Carvalho, C. M., Pfundt, R., King, D. A., Lindsay, S. J., Zuccherato, L. W., and Macville, M. V. (2015). Absence of heterozygosity due to template switching during replicative rearrangements. *Am. J. Hum. Genet.* 96, 555–564. doi: 10.1016/j.ajhg.2015.01.021

Cer, R. Z., Bruce, K. H., Donohue, D. E., Temiz, N. A., Mudunuri, U. S., Yi, M., et al. (2012). Searching for non-B DNA-forming motifs using nBMST (non-B DNA motif search tool). *Curr. Protoc. Hum. Genet.* 18, 11–22. doi: 10.1002/0471142905.hg1807s73

Chakraborty, M., Baldwin-Brown, J. G., Long, A. D., and Emerson, J. J. (2016). Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. *Nucleic Acids. Res.* 44:e147. doi: 10.1093/nar/gkw654

Chen, H., Rangasamy, M., Tan, S. Y., Wang, H., and Siegfried, B. D. (2010). Evaluation of five methods for total DNA extraction from Western corn rootworm beetles. *PLoS One* 5:e11963. doi: 10.1371/journal.pone.0011963

Cheng, C., Tan, J. C., Hahn, M. W., and Besansky, N. J. (2018). Systems genetic analysis of inversion polymorphisms in the malaria mosquito Anopheles gambiae. *Proc. Natl. Acad. Sci. U.S.A.* 115, E7005–E7014. doi: 10.1073/pnas.1806760115

Cirera, S., Martín-Campos, J. M., Segarra, C., and Aguadé, M. (1995). Molecular characterization of the breakpoints of an inversion fixed between *Drosophila melanogaster* and *D. subobscura*. *Genetics* 139, 321–326.

Corbett-Detig, R. B. (2016). Selection on inversion breakpoints favors proximity to pairing sensitive sites in *Drosophila melanogaster*. *Genetics* 204, 259–265. doi: 10.1534/genetics.116.190389

Corbett-Detig, R. B., and Hartl, D. L. (2012). Population genomics of inversion polymorphisms in *Drosophila melanogaster*. *PLoS Genet.* 8:e1003056. doi: 10.1371/journal.pgen.1003056

Crescente, J. M., Zavallo, D., Helguera, M., and Vanzetti, L. S. (2018). MITE Tracker: an accurate approach to identify miniature inverted-repeat transposable elements in large genomes. *BMC Bioinformatics* 19:348. doi: 10.1186/s12859-018-2376-y

de Frutos, R. (1972). Contribution to the study of chromosomal polymorphism in the Spanish populations of *Drosophila subobscura*. *Genét. Ibér.* 24, 123–140.

de Jong, G., and Bochdanovits, Z. (2003). Latitudinal clines in *Drosophila melanogaster*: body size, allozyme frequencies, inversion frequencies, and the insulin-signalling pathway. *J. Genet.* 82, 207–223. doi: 10.1007/BF02715819

Delprat, A., Guillén, Y., and Ruiz, A. (2019). Computational sequence analysis of inversion breakpoint regions in the cactophilic Drosophila mojavensis lineage. *J. Hered.* 110, 102–117. doi: 10.1093/jhered/esy057

Deng, W., Maust, B. S., Nickle, D. C., Learn, G. H., Liu, Y., Heath, L., et al. (2010). DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. *Biotechniques* 48, 405–408. doi: 10.2144/000113370

Dionne, M. S., Pham, L. N., Shirasu-Hiza, M., and Schneider, D. S. (2006). Akt and foxo dysregulation contribute to infection-induced wasting in Drosophila. *Curr. Biol.* 16, 1977–1985. doi: 10.1016/j.cub.2006.08.052

Dobzhansky, T. (1947). Genetics of natural populations. XIV. A response of certain gene arrangements in the third chromosome of *Drosophila pseudoobscura* to natural selection. *Genetics* 32, 142–160.

Dolezal, T., Krejcova, G., Bajgar, A., Nedbalova, P., and Strasser, P. (2019). Molecular regulations of metabolism during immune response in insects. *Insect. Biochem. Mol. Biol.* 109, 31–42. doi: 10.1016/j.ibmb.2019.04.005

Dolgova, O. (2013). *Genetic and Phenotypic Differentiation in three Chromosomal Arrangements of Drosophila subobscura*. Ph.D. thesis, Universitat Autònoma de Barcelona, Barcelona.

Durmaz, E., Rajpurohit, S., Betancourt, N., Fabian, D. K., Kapun, M., Schmidt, P., et al. (2019). A clinal polymorphism in the insulin signaling transcription factor foxo contributes to life-history adaptation in Drosophila. *Evolution* 73, 1774–1792. doi: 10.1111/evo.13759

Eilbeck, K., Lewis, S., Mungall, C., Yandell, M., Stein, L., Durbin, R., et al. (2005). The sequence ontology: a tool for the unification of genome annotations. *Genome Biol.* 6:R44. doi: 10.1186/gb-2005-6-5-r44

Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* 9:18. doi: 10.1186/1471-2105-9-18

Fabian, D. K., Kapun, M., Nolte, V., Kofler, R., Schmidt, P. S., Schlötterer, C., et al. (2012). Genome-wide patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North America. *Mol. Ecol.* 21, 4748–4769. doi: 10.1111/j.1365-294X.2012.05731.x

Faria, R., Johannesson, K., Butlin, R. K., and Westram, A. M. (2019). Evolving inversions. *Trends Ecol. Evol.* 34, 239–248. doi: 10.1016/j.tree.2018.12.005

Farré, M., Robinson, T. J., and Ruiz-Herrera, A. (2015). An integrative breakage model of genome architecture, reshuffling and evolution. *Bioessays* 37, 479–488. doi: 10.1002/bies.201400174

Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., et al. (2017). InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res.* 45, D190–D199. doi: 10.1093/nar/gkw1107

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344

Flatt, T., and Partridge, L. (2018). Horizons in the evolution of aging. *BMC Biol.* 16:93. doi: 10.1186/s12915-018-0562-z

Fontdevila, A., Zapata, C., Alvarez, G., Sanchez, L., Méndez, J., and Enriquez, I. (1983). Genetic coadaptation in the chromosomal polymorphism of *Drosophila subobscura*. I. Seasonal changes of gametic disequilibrium in a natural population. *Genetics* 105, 935–955.

Fragata, I., Lopes-Cunha, M., Bárbaro, M., Kellen, B., Lima, M., Santos, M. A., et al. (2014). How much can history constrain adaptive evolution? A real-time evolutionary approach of inversion polymorphisms in *Drosophila subobscura*. *J. Evol. Biol.* 27, 2727–2738. doi: 10.1111/jeb.12533

Fuller, Z. L., Haynes, G. D., Richards, S., and Schaeffer, S. W. (2016). Genomics of natural populations: How differentially expressed genes shape the evolution of chromosomal inversions in *Drosophila pseudoobscura*. *Genetics* 204, 287–301. doi: 10.1534/genetics.116.191429

Fuller, Z. L., Haynes, G. D., Richards, S., and Schaeffer, S. W. (2017). Genomics of natural populations: evolutionary forces that establish and maintain gene arrangements in *Drosophila pseudoobscura*. *Mol. Ecol.* 26, 6539–6562. doi: 10.1111/mec.14381

Fuller, Z. L., Koury, S. A., Phadnis, N., and Schaeffer, S. W. (2019). How chromosomal rearrangements shape adaptation and speciation: case studies in *Drosophila pseudoobscura* and its sibling species *Drosophila persimilis*. *Mol. Ecol.* 28, 1283–1301. doi: 10.1111/mec.14923

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A., and Merilä, J. (2008). Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.* 17, 167–178. doi: 10.1111/j.1365-294X.2007.03413.x

Götz, W. (1965). Beitrag zur Kenntnis der Inversionen, Duplikationen und Strukturtypen von *Drosophila subobscura* Coll. *Z. Vererbungsl.* 96, 285–296. doi: 10.1007/BF00896828

Götz, W. (1967). Untersuchungen über den chromosomalen Strukturpolymorphismus in kleinasiatischen und persischen Populationen von *Drosophila subobscura* Coll. *Mol. Gen. Genet.* 100, 1–38. doi: 10.1007/BF00425773

Grant, C. E., Bailey, T. L., and Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018. doi: 10.1093/bioinformatics/btr064

Greig, D. (2007). A screen for recessive speciation genes expressed in the gametes of F1 hybrid yeast. *PLoS Genet.* 3:e21. doi: 10.1371/journal.pgen.0030021

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximumlikelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Hanson, M. A., and Lemaitre, B. (2020). New insights on Drosophila antimicrobial peptide function in host defense and beyond. *Curr. Opin. Immunol.* 62, 22–30. doi: 10.1016/j.coi.2019.11.008

Hastings, P. J., Ira, G., and Lupski, J. R. (2009). A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* 5:e1000327. doi: 10.1371/journal.pgen.1000327

Hill, T., and Betancourt, A. J. (2018). Extensive exchange of transposable elements in the *Drosophila pseudoobscura* group. *Mob. DNA* 9:20.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi: 10.1093/molbev/msx281

Hoffmann, A. A., and Rieseberg, L. H. (2008). Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annu. Rev. Ecol. Evol. Syst.* 39, 21–42. doi: 10.1146/annurev.ecolsys.39.110707.173532

Holt, C., and Yandell, M. (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491. doi: 10.1186/1471-2105-12-491

Hughes, L. (2000). Biological consequences of global warming: Is the signal already apparent? *Trends Ecol. Evol.* 15, 56–61.

Hughes, S. E., Miller, D. E., Miller, A. L., and Hawley, R. S. (2018). Female meiosis: synapsis, recombination, and segregation in *Drosophila melanogaster*. *Genetics* 208, 875–908. doi: 10.1534/genetics.117.300081

Huson, D. H., and Bryant, I. D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031

Joyce, E. F., Paul, A., Chen, K. E., Tanneti, N., and McKim, K. S. (2012). Multiple barriers to nonhomologous DNA end joining during meiosis in Drosophila. *Genetics* 191, 739–746. doi: 10.1534/genetics.112.140996

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285

Kapun, M., Fabian, D. K., Goudet, J., and Flatt, T. (2016). Genomic evidence for adaptive inversion clines in *Drosophila melanogaster*. *Mol. Biol. Evol.* 33, 1317–1336. doi: 10.1093/molbev/msw016

Kapun, M., and Flatt, T. (2018). The adaptive significance of chromosomal inversion polymorphisms in *Drosophila melanogaster*. *Mol. Ecol.* 28, 1263–1282. doi: 10.1111/mec.14871

Karageorgiou, C., Gámez-Visairas, V., Tarrío, R., and Rodríguez-Trelles, F. (2019). Long-read based assembly and synteny analysis of a reference *Drosophila subobscura* genome reveals signatures of structural evolution driven by inversions recombination-suppression effects. *BMC Genomics* 20:223.

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinformatics* 20, 1160–1166. doi: 10.1093/bib/bbx108

Kaushal, S., and Freudenreich, C. H. (2019). The role of fork stalling and DNA structures in causing chromosome fragility. *Genes Chromosomes Cancer* 58, 270–283. doi: 10.1002/gcc.22721

Kehrer-Sawatzki, H., Sandig, C. A., Goidts, V., and Hameister, H. (2005). Breakpoint analysis of the pericentric inversion between chimpanzee chromosome 10 and the homologous chromosome 12 in humans. *Cytogenet. Genome Res.* 108, 91–97. doi: 10.1159/000080806

Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biol.* 8:e1000501. doi: 10.1371/journal.pbio.1000501

Kirkpatrick, M., and Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics* 173, 419–434. doi: 10.1534/genetics.105.047985

Kolb, J., Chuzhanova, N. A., Högel, J., Vasquez, K. M., Cooper, D. N., Bacolla, A., et al. (2009). Cruciform-forming inverted repeats appear to have mediated many of the microinversions that distinguish the human and chimpanzee genomes. *Chromosome Res.* 17, 469–483.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116

Koske, T., and Maynard-Smith, J. (1954). Genetics and cytology of *Drosophila subobscura*. X. The fifth linkage group. *J. Genet.* 52, 521–541. doi: 10.1007/BF02985076

Kramara, J., Osia, B., and Malkova, A. (2018). Break-induced replication: the where, the why, and the how. *Trends Genet.* 34, 518–531. doi: 10.1016/j.tig.2018.04.002

Krimbas, C. B. (1992). "The inversion polymorphism of *Drosophila subobscura*," in *Drosophila Inversion Polymorphism*, eds C. B. Krimbas and J. R. Powell (Boca Raton, FL: CRC Press), 127–220.

Kulakovskiy, I. V., and Makeev, V. J. (2009). Discovery of DNA motifs recognized by transcription factors through integration of different experimental sources. *Biophysics* 54, 667–674. doi: 10.1134/S0006350909060013

Kunze-Mühl, E., and Müller, E. (1958). Weitere Untersuchungen uber die chromosomale Struktur und die natürlichen Strukturtypen von *Drosophila subobscura* Coll. *Chromosoma* 9, 559–570. doi: 10.1007/BF02568093

Kunze-Mühl, E., and Sperlich, D. (1955). Inversionen und chromosomale Strukturtypen bei *Drosophila subobscura* Coll. *Z. Vererbungsl.* 87, 65–84. doi: 10.1007/BF00308333

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12

Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H. H., Rognes, T., and Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. doi: 10.1093/nar/gkm160

Lazzaro, A., Hilfiker, D., and Zeyer, J. (2015). Structures of microbial communities in alpine soils: seasonal and elevational effects. *Front. Microbiol.* 6:1330. doi: 10.3389/fmicb.2015.01330

Lazzaro, B. P., and Clark, A. G. (2001). Evidence for recurrent paralogous gene conversion and exceptional allelic divergence in the Attacin genes of *Drosophila melanogaster*. *Genetics* 159, 659–671.

Lazzaro, B. P., and Clark, A. G. (2003). Molecular population genetics of inducible antibacterial peptide genes in *Drosophila melanogaster*. *Mol. Biol. Evol.* 20, 914–923. doi: 10.1093/molbev/msg109

Lee, J. A., Carvalho, C. M., and Lupski, J. R. (2007). A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* 131, 1235–1247. doi: 10.1016/j.cell.2007.11.037

Leigh Brown, A. J., and Ish-Horowicz, D. (1981). Evolution of the 87A and 87C heat-shock loci in Drosophila. *Nature* 290, 677–682.

Lobachev, K. S., Rattray, A., and Narayanan, V. (2007). Hairpin- and cruciform-mediated chromosome breakage: causes and consequences in eukaryotic cells. *Front. Biosci.* 12, 4208–4220. doi: 10.2741/2381

Lopes, M., Foiani, M., and Sogo, J. M. (2006). Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol. Cell* 21, 15–27. doi: 10.1016/j.molcel.2005.11.015

Loukas, M., Krimbas, C. B., and Vergini, Y. (1979). The genetics of *Drosophila subobscura* populations. IX. Studies on linkage disequilibrium in four natural populations. *Genetics* 93, 497–523.

Lowe, T. M., and Chan, P. P. (2016). tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 44, W54–W57. doi: 10.1093/nar/gkw413

Lowry, D. B., Popovic, D., Brennan, D. J., and Holeski, L. M. (2019). Mechanisms of a locally adaptive shift in allocation among growth, reproduction, and herbivore resistance in Mimulus guttatus. *Evolution* 73, 1168–1181. doi: 10.1111/evo.13699

Lu, S., Wang, G., Bacolla, A., Zhao, J., Spitser, S., and Vasquez, K. M. (2015). Short inverted repeats are hotspots for genetic instability: relevance to cancer genomes. *Cell Rep.* 10, 1674–1680. doi: 10.1016/j.celrep.2015.02.039

Mani, R. S., and Chinnaiyan, A. M. (2010). Triggers for genomic rearrangements: insights into genomic, cellular and environmental influences. *Nat. Rev. Genet.* 11, 819–829. doi: 10.1038/nrg2883

Matzkin, L. M., Merritt, T. J., Zhu, C. T., and Eanes, W. F. (2005). The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion In(3R)Payne in *Drosophila melanogaster*. *Genetics* 170, 1143–1152. doi: 10.1534/genetics.104.038810

Maynard-Smith, J., and Maynard-Smith, S. (1954). Genetics and cytology of *Drosophila subobscura*. VIII. Heterozygosity, viability and rate of development. *J. Genet.* 52, 152–164. doi: 10.1007/BF02981496

McBroome, J., Liang, D., and Corbett-Detig, R. (2020). Fine-scale position effects shape the distribution of inversion breakpoints in *Drosophila melanogaster*. *Genome Biol. Evol*. doi: 10.1093/gbe/evaa103 [Epub ahead of print].

McKinney, J. A., Wang, G., Mukherjee, A., Christensen, L., Subramanian, S. H. S., Zhao, J., et al. (2020). Distinct DNA repair pathways cause genomic instability at alternative DNA structures. *Nat. Commun.* 11:236.

Menozzi, P., and Krimbas, C. B. (1992). The inversion polymorphism of *D. subobscura* revisited: synthetic maps of gene arrangement frequencies and their interpretation. *J. Evol. Biol.* 5, 625–641. doi: 10.1046/j.1420-9101.1992.5040625.x

Messer, P. W., Ellner, S. P., and Hairston, N. G. Jr. (2016). Can population genetics adapt to rapid evolution? *Trends Genet.* 32, 408–418. doi: 10.1016/j.tig.2016.04.005

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Orengo, D. J., Puerma, E., Cereijo, U., and Aguadé, M. (2019). The molecular genealogy of sequential overlapping inversions implies both homologous chromosomes of a heterokaryotype in an inversion origin. *Sci. Rep.* 9:17009.

Orengo, D. J., Puerma, E., Papaceit, M., Segarra, C., and Aguadé, M. A. (2015). Molecular perspective on a complex polymorphic inversion system with cytological evidence of multiply reused breakpoints. *Heredity* 114, 610–618. doi: 10.1038/hdy.2015

Paaby, A. B., Bergland, A. O., Behrman, E. L., and Schmidt, P. S. (2014). A highly pleiotropic amino acid polymorphism in the Drosophila insulin receptor contributes to life-history adaptation. *Evolution* 68, 3395–3409. doi: 10.1111/evo.12546

Papaceit, M., Segarra, C., and Aguadé, M. (2012). Structure and population genetics of the breakpoints of a polymorphic inversion in *Drosophila subobscura*. *Evolution* 67, 66–79.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* 37, 637–669. doi: 10.1146/annurev.ecolsys.37.091305.110100

Pegueroles, C., Aquadro, C. F., Mestres, F., and Pascual, M. (2013). Gene flow and gene flux shape evolutionary patterns of variation in *Drosophila subobscura*. *Heredity* 110, 520–529. doi: 10.1038/hdy.2012.118

Pegueroles, C., Araúz, P. A., Pascual, M., and Mestres, F. (2010a). A recombination survey using microsatellites: the O chromosome of *Drosophila subobscura*. *Genetica* 138, 795–804.

Pegueroles, C., Ferrés-Coy, A., Martí-Solano, M., Aquadro, C. F., Pascual, M., and Mestres, F. (2016). Inversions and adaptation to the plant toxin ouabain shape DNA sequence variation within and between chromosomal inversions of *Drosophila subobscura*. *Sci. Rep.* 6:23754. doi: 10.1038/srep23754

Pegueroles, C., Ordoñez, V., Mestres, F., and Pascual, M. (2010b). Recombination and selection in the maintenance of the adaptive value of inversions. *J. Evol. Biol.* 23, 2709–2717. doi: 10.1111/j.1420-9101.2010.02136.x

Powell, J. R. (1997). *Progress and Prospects in Evolutionary Biology: The Drosophila Model*. Oxford: Oxford University Press.

Prazeres da Costa, O., González, J., and Ruiz, A. (2009). Cloning and sequencing of the breakpoint regions of inversion 5g fixed in Drosophila buzzatii. *Chromosoma* 118, 349–360. doi: 10.1007/s00412-008-0201-5

Prevosti, A., Ribo, G., Serra, L., Aguadé, M., Balaña, J., Monclus, M., et al. (1988). Colonization of America by *Drosophila subobscura*: experiment in natural populations that supports the adaptive role of chromosomal-inversion polymorphism. *Proc. Natl. Acad. Sci. U.S.A.* 85, 5597–5600. doi: 10.1073/pnas.85.15.5597

Puerma, E., Orengo, D. J., and Aguadé, M. (2016a). Multiple and diverse structural changes affect the breakpoint regions of polymorphic inversions across the Drosophila genus. *Sci. Rep.* 6:36248. doi: 10.1038/srep36248

Puerma, E., Orengo, D. J., and Aguadé, M. (2016b). The origin of chromosomal inversions as a source of segmental duplications in the Sophophora subgenus of Drosophila. *Sci. Rep.* 6:30715. doi: 10.1038/srep30715

Puerma, E., Orengo, D. J., and Aguadé, M. (2017). Inversion evolutionary rates might limit the experimental identification of inversion breakpoints in non-model species. *Sci. Rep*. 7:17281.

Puerma, E., Orengo, D. J., Cruz, F., Gómez-Garrido, J., Librado, P., Salguero, D., et al. (2018). The high-quality genome sequence of the oceanic island endemic species Drosophila guanche reveals signals of adaptive evolution in genes related to flight and genome stability. *Genome Biol. Evol.* 10, 1956–1969. doi: 10.1093/gbe/evy135

Puerma, E., Orengo, D. J., Salguero, D., Papaceit, M., Segarra, C., and Aguadé, M. (2014). Characterization of the breakpoints of a polymorphic inversion complex detects strict and broad breakpoint reuse at the molecular level. *Mol. Biol. Evol.* 31, 2331–2341. doi: 10.1093/molbev/msu177

Puig-Giribets, M., Guerreiro, M. P. G., Santos, M., Ayala, F. J., Tarrío, R., and Rodríguez-Trelles, F. (2019). Chromosomal inversions promote genomic islands of concerted evolution of Hsp70 genes in the *Drosophila subobscura* species subgroup. *Mol. Ecol.* 28, 1316–1332. doi: 10.1111/mec.14511

Ranz, J. M., Maurin, D., Chan, Y. S., von Grotthuss, M., Hillier, L. W., Roote, J., et al. (2007). Principles of genome evolution in the *Drosophila melanogaster* species group. *PLoS Biol.* 5:e152. doi: 10.1371/journal.pbio.0050152

Reese, M. G. (2001). Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput. Chem.* 26, 51–56.

Regan, J. C., Froy, H., Walling, C. A., Moatt, J. P., and Nussey, D. H. (2020). Dietary restriction and insulin-like signalling pathways as adaptive plasticity: a synthesis and re-evaluation. *Funct. Ecol.* 34, 107–128. doi: 10.1111/1365-2435.13418

Rego, C., Balanyà, J., Fragata, I., Matos, M., Rezende, E. L., and Santos, M. (2010). Clinal patterns of chromosomal inversion polymorphisms in *Drosophila subobscura* are partly associated with thermal preferences and heat stress resistance. *Evolution* 64, 385–397. doi: 10.1111/j.1558-5646.2009.00835.x

Rezende, E. L., Balanyà, J., Rodríguez-Trelles, F., Rego, C., Fragata, I., Matos, M., et al. (2010). Climate change and chromosomal inversions in *Drosophila subobscura*. *Clim. Res.* 43, 103–114. doi: 10.1007/s10709-018-0035-x

Richards, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S., and Nielsen, R. (2005). Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution. *Genome Res.* 15, 1–18. doi: 10.1101/gr.3059305

Rivera, J., Keränen, S. V. E., Gallo, S. M., and Halfon, M. S. (2019). REDfly: the transcriptional regulatory element database for Drosophila. *Nucleic Acids Res.* 47, D828–D834. doi: 10.1093/nar/gky957

Rodríguez-Trelles, F. (2003). Seasonal cycles of allozyme-by-chromosomal-inversion gametic disequilibrium in *Drosophila subobscura*. *Evolution* 57, 839–848. doi: 10.1111/j.0014-3820.2003.tb00295.x

Rodríguez-Trelles, F., Alvarez, G., and Zapata, C. (1996). Time-series analysis of seasonal changes of the O inversion polymorphism of *Drosophila subobscura*. *Genetics* 142, 179–187.

Rodríguez-Trelles, F., and Rodríguez, M. A. (1998). Rapid micro-evolution and loss of chromosomal diversity in Drosophila in response to climate warming. *Evol. Ecol.* 12, 829–838. doi: 10.1023/A:1006546616462

Rodríguez-Trelles, F., and Rodríguez, M. A. (2007). Comment on 'Global genetic change tracks global climate warming in *Drosophila subobscura*'. *Science* 315:1497. doi: 10.1126/science.1136298

Rodríguez-Trelles, F., and Rodríguez, M. A. (2010). Measuring evolutionary responses to global warming: cautionary lessons from *Drosophila*. *Insect Conserv. Divers.* 3, 44–50. doi: 10.1111/j.1752-4598.2009.00071.x

Rodríguez-Trelles, F., Tarrío, R., and Santos, M. (2013). Genome-wide evolutionary response to a heat wave in Drosophila. *Biol. Lett.* 9:20130228. doi: 10.1098/rsbl.2013.0228

Romero-Olivares, A. L., Allison, S. D., and Treseder, K. K. (2017). Soil microbes and their response to experimental warming over time: a meta-analysis of field studies. *Soil Biol. Biochem.* 107, 32–40. doi: 10.1016/j.soilbio.2016.12.026

Said, I., Byrne, A., Serrano, V., Cardeno, C., Vollmers, C., and Corbett-Detig, R. B. (2018). Linked genetic variation and not genome structure causes widespread differential expression associated with chromosomal inversions. *Proc. Natl. Acad. Sci. U.S.A.* 115, 5492–5497. doi: 10.1073/pnas.1721275115

Sakofsky, C. J., Ayyar, S., Deem, A. K., Chung, W. H., Ira, G., and Malkova, A. (2015). Translesion polymerases drive microhomology-mediated break-induced replication leading to complex chromosomal rearrangements. *Mol. Cell* 60, 860–872. doi: 10.1016/j.molcel.2015.10.041

Santos, J., Serra, L., Solé, E., and Pascual, M. (2010). FISH mapping of microsatellite loci from *Drosophila subobscura* and its comparison to related species. *Chromosome Res.* 18, 213–226.

Santos, M., Céspedes, W., Balanyà, J., Trotta, V., Calboli, F. C., Fontdevila, A., et al. (2005). Temperature-related genetic changes in laboratory populations of *Drosophila subobscura*: evidence against simple climatic-based explanations for latitudinal clines. *Am. Nat.* 165, 258–273. doi: 10.1086/427093

Savage, J. R. (1976). Classification and relationships of induced chromosomal structual changes. *J. Med. Genet.* 13, 103–122. doi: 10.1136/jmg.13.2.103

Schwenke, R. A., Lazzaro, B. P., and Wolfner, M. F. (2016). Reproduction-immunity trade-offs in insects. *Annu. Rev. Entomol.* 61, 239–256.

Scully, R., Panday, A., Elango, R., and Willis, N. A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell Biol.* 20, 698–714.

Sebastián, A., and Contreras-Moreira, B. (2014). FootprintDB: a database of transcription factors with annotated cis elements and binding interfaces. *Bioinformatics* 30, 258–265. doi: 10.1093/bioinformatics/btt663

Senger, K., Armstrong, G. W., Rowell, W. J., Kwan, J. M., Markstein, M., and Levine, M. (2004). Immunity regulatory DNAs share common organizational features in Drosophila. *Mol. Cell* 13, 19–32.

Seppey, M., Manni, M., and Zdobnov, E. M. (2019). BUSCO: assessing genome assembly and annotation completeness. *Methods Mol. Biol.* 1962, 227–245. doi: 10.1007/978-1-4939-9173-0_14

Shigyo, N., Umeki, K., and Hirao, T. (2019). Seasonal dynamics of soil fungal and bacterial communities in cool-temperate montane forests. *Front. Microbiol.* 10:1944. doi: 10.3389/fmicb.2019.01944

Shtraizent, N., DeRossi, C., Nayar, S., Sachidanandam, R., Katz, L. S., Prince, A., et al. (2017). MPI depletion enhances O-GlcNAcylation of p53 and suppresses the Warburg effect. *eLife* 6:e22477. doi: 10.7554/eLife.22477

Slade, J. D., and Staveley, B. E. (2016). Enhanced survival of Drosophila Akt1 hypomorphs during amino-acid starvation requires foxo. *Genome* 59, 87–93.

Smit, A. F. A., Hubley, R., and Green, P. (2013/2015). *RepeatMasker Open-4.0*. Available online at: http://www.repeatmasker.org (accessed November 15, 2019).

Smith, C., Llorente, B., and Symington, L. (2007). Template switching during break-induced replication. *Nature* 447, 102–105. doi: 10.1038/nature05723

Soderlund, C., Bomhoff, M., and Nelson, W. M. (2011). SyMAP v3.4: a turnkey synteny system with application to plant genomes. *Nucleic Acids Res.* 39:e68. doi: 10.1093/nar/gkr123

Sogo, J. M., Lopes, M., and Foiani, M. (2002). Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. *Science* 297, 599–602. doi: 10.1126/science.1074023

Solé, E., Balanyà, J., Sperlich, D., and Serra, L. (2002). Long-term changes in the chromosomal inversion polymorphism of *Drosophila subobscura*. I. Mediterranean populations from southwestern Europe. *Evolution* 56, 830–835. doi: 10.1111/j.0014-3820.2002.tb01393.x

Sperlich, D., Feuerbach-Mravlag, H., Lange, P., Michaelidis, A., and Pentzos-Daponte, A. (1977). Genetic load and viability distribution in central and marginal populations of *Drosophila subobscura*. *Genetics* 86, 835–848.

Sun, X., Qun, Y., and Xia, X. (2013). An improved implementation of effective number of codons (Nc). *Mol. Biol. Evol.* 30, 191–196. doi: 10.1093/molbev/mss201

Sunder, S., and Wilson, E. T. (2019). Frequency of DNA end joining in trans is not determined by the predamage spatial proximity of double-strand breaks in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 116, 9481–9490. doi: 10.1073/pnas.1818595116

Tajima, F. (1993). Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* 135, 599–607.

Teleman, A. A. (2010). Molecular mechanisms of metabolic regulation by insulin in Drosophila. *Biochem. J.* 425, 13–26.

The Gene Ontology Consortium. (2017). Expansion of the gene ontology knowledgebase and resources. *Nucleic Acids Res.* 45, D331–D338. doi: 10.1093/nar/gkw1108

Tremblay-Belzile, S., Lepage, É., Zampini, É., and Brisson, N. (2015). Short-range inversions: rethinking organelle genome stability: template switching events during DNA replication destabilize organelle genomes. *Bioessays* 37, 1086–1094. doi: 10.1002/bies.201500064

Voineagu, I., Narayanan, V., Lobachev, K. S., and Mirkin, S. M. (2008). Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9936–9941. doi: 10.1073/pnas.0804510105

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963

Wang, G., Christensen, L. A., and Vasquez, K. M. (2006). Z-DNA-forming sequences generate large-scale deletions in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2677–2682. doi: 10.1073/pnas.0511084103

Wang, G., and Vasquez, K. M. (2006). Non-B DNA structure-induced genetic instability. *Mutat. Res.* 598, 103–119.

Wasserman, M. (1968). Recombination-induced chromosomal heterosis. *Genetics* 58, 125–139.

Weirauch, M. T., Yang, A., Albu, M., Cote, A. G., Montenegro-Montero, A., Drewe, P., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* 158, 1431–1443. doi: 10.1016/j.cell.2014.08.009

Wellenreuther, M., and Bernatchez, L. (2018). Eco-evolutionary genomics of chromosomal inversions. *Trends Ecol. Evol.* 33, 427–440. doi: 10.1016/j.tree.2018.04.002

Wesley, C. S., and Eanes, W. F. (1994). Isolation and analysis of the breakpoint sequences of chromosome inversion In(3L)Payne in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 91, 3132–3136. doi: 10.1073/pnas.91.8.3132

Zhang, F., Khajavi, M., Connolly, A. M., Towne, C. F., Batish, S. D., and Lupski, J. R. (2009). The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans. *Nat. Genet.* 41, 849–853. doi: 10.1038/ng.399

Zhao, J., Bacolla, A., Wang, G., and Vasquez, K. M. (2010). Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* 67, 43–62.

Zollinger, E. (1950). Ein strukturell homozygoter Stamm von *Drosophila subobscura* aus einer Wild-population. *Arch. Klaus-Stiftg* 25, 33–35.

Zouros, E., Krimbas, C. B., Tsakas, S., and Loukas, M. (1974). Genic versus chromosomal variation in natural populations of *D. subobscura*. *Genetics* 78, 1223–1244.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 28 October 2020
doi: 10.3389/fgene.2020.555843

Pronounced Plastic and Evolutionary Responses to Unpredictable Thermal Fluctuations in *Drosophila simulans*

Jesper G. Sørensen[1], Tommaso Manenti[1†], Jesper S. Bechsgaard[1], Mads F. Schou[2], Torsten N. Kristensen[3] and Volker Loeschcke[1]*

[1] Department of Biology, Aarhus University, Aarhus, Denmark, [2] Department of Biology, Lund University, Lund, Sweden, [3] Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark

OPEN ACCESS

Edited by:
David Vieites, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Reviewed by:
C.-Elisa Schaum, University of Hamburg, Germany
Norah Paul Saarman, Yale University, United States
Zhongqi Chen, University of Idaho, United States

***Correspondence:**
Jesper G. Sørensen
jesper.soerensen@bio.au.dk

†Present address:
Tommaso Manenti, Laboratori Biokyma Srl, Anghiari, Italy

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *04 May 2020*
Accepted: *06 October 2020*
Published: *28 October 2020*

Citation:
Sørensen JG, Manenti T, Bechsgaard JS, Schou MF, Kristensen TN and Loeschcke V (2020) Pronounced Plastic and Evolutionary Responses to Unpredictable Thermal Fluctuations in Drosophila simulans. Front. Genet. 11:555843. doi: 10.3389/fgene.2020.555843

Organisms are exposed to temperatures that vary, for example on diurnal and seasonal time scales. Thus, the ability to behaviorally and/or physiologically respond to variation in temperatures is a fundamental requirement for long-term persistence. Studies on thermal biology in ectotherms are typically performed under constant laboratory conditions, which differ markedly from the variation in temperature across time and space in nature. Here, we investigate evolutionary adaptation and environmentally induced plastic responses of *Drosophila simulans* to no fluctuations (constant), predictable fluctuations or unpredictable fluctuations in temperature. We whole-genome sequenced populations exposed to 20 generations of experimental evolution under the three thermal regimes and examined the proteome after short-term exposure to the same three regimes. We find that unpredictable fluctuations cause the strongest response at both genome and proteome levels. The loci showing evolutionary responses were generally unique to each thermal regime, but a minor overlap suggests either common laboratory adaptation or that some loci were involved in the adaptation to multiple thermal regimes. The evolutionary response, i.e., loci under selection, did not coincide with induced responses of the proteome. Thus, genes under selection in fluctuating thermal environments are distinct from genes important for the adaptive plastic response observed within a generation. This information is key to obtain a better understanding and prediction of the effects of future increases in both mean and variability of temperatures.

Keywords: heat tolerance, genomics, proteomics, thermal fluctuations, *Drosophila simulans*

INTRODUCTION

It is well known that different adaptive responses for coping with stressful temperature conditions exist. Within generations, organisms can respond plastically to environmental changes (Pigliucci, 1996, 2001; Lande, 2009), while evolutionary responses may occur through both changes in trait means and in the level of plasticity (Hoffmann and Parsons, 1989; Williams et al., 2008; Lande, 2009; Kristensen et al., 2018). It is debated whether plasticity or evolutionary responses constitute the main contributor to temperature adaptation in small ectothermic animals (Gunderson and Stillman, 2015; Sgrò et al., 2016; Sørensen et al., 2016a). Evolutionary change in trait means is

better understood and does provide evidence for local adaptation (e.g., Hoffmann et al., 2002; Kellermann et al., 2012). However, upper thermal limits seem to be evolutionary constrained in some small ectothermic insects (particularly studied in *Drosophila*) (Kellermann et al., 2012; Schou et al., 2014; but see discussion in Logan and Cox, 2020), while not in some species of phytoplankton (Kontopoulos et al., 2020). The constraint among species of *Drosophila* is not founded in an apparent lack of additive genetic variation, as significant levels of genetic variation for heat tolerance in the same species of *Drosophila* have been documented (Williams et al., 2012; Castaneda et al., 2019). Theory predicts that evolution of plasticity should be favored in predictably variable environments (Lande, 2009; Ashander et al., 2016). However, for plasticity in upper (and lower) thermal tolerance empirical evidence supporting this hypothesis is scarce (Gunderson and Stillman, 2015). Thus, it is not clear how and how much small ectothermic animals can and will respond to a warming climate.

Under natural conditions, organisms are exposed to temperatures that vary on diurnal (multiple exposures within a generation for most insect species) and seasonal (often across multiple generations in insects) scales. Diurnal fluctuations are often comparable to seasonal variation in magnitude, but characterized by temperature changes which occur much faster than temperature change across seasons. For example, according to the Danish Meteorological Institute, the difference in mean temperature among seasons in Denmark is ∼15°C (based on monthly mean temperatures 2006–2015[1]). Diurnal temperatures can attain a similar range between a cold night and a warm summer day, where 10 and 25°C, respectively, can be found within the same or a few days. The variability in temperature is expected to further increase with increasing heat waves under climate change (Perkins-Kirkpatrick and Lewis, 2020). Under these conditions, it can be questioned whether evolutionary change in trait means is adequate to maintain fitness sufficient high for a species to persist (Araujo et al., 2013; Radchuk et al., 2019). Phenotypic plasticity has been suggested to be more likely to accommodate rapidly changing temperatures and extreme events (Gienapp et al., 2008; Merilä and Hendry, 2014). However, while large differences among *Drosophila* species in plasticity of different traits exist, within species plasticity in heat tolerance (thermal acclimation capacity) is evolving slowly; plasticity does not to differ among geographically distinct populations of the same species, and is not lost when natural populations are kept at constant temperature in the laboratory (up to 28 generations) (Overgaard et al., 2011; van Heerwaarden et al., 2014; Fragata et al., 2016; Manenti et al., 2017). Furthermore, plasticity does not change under experimental evolution (for 20 generations) in thermal regimes with different variability and predictability (Manenti et al., 2015). An explanation for this could be constraints bounded in the genetic architecture of basal and acclimation trait values (Gerken et al., 2015; see also Lecheta et al., 2020). Temperature fluctuations seem to induce heat tolerance through mechanisms different from mechanisms induced by constant temperatures. For example, thermal fluctuations induce a transcriptomic response that is different from the response induced by differences in mean temperature and from the classic heat stress response induced by increasing thermal stress (Sørensen et al., 2016b; Manenti et al., 2018). However, we have little knowledge on the molecular responses to fluctuating environments, and this limits our understanding of acclimation capacity, evolutionary constraints and trade-offs as well as the costs of induced plastic responses.

A number of evolutionary and ecological studies have recently focused on fitness consequences of short- or long-term variation of temperatures (Ketola et al., 2004; Kingsolver et al., 2009; Hallsson and Bjorklund, 2012; Manenti et al., 2014, 2015; Kellermann et al., 2015; Simões et al., 2020). The use of temperature fluctuations in laboratory experiments has been argued to have a greater ecological relevance compared to constant ones, as they are a better proxy of a natural environment (Boyce et al., 2006; Schreiber, 2010). This study aimed to investigate the molecular responses to environments that differ in amplitude and predictability of daily temperature within the life span of the organism studied. We did this using a well-known insect model system (*Drosophila simulans*), which can easily be collected and manipulated (e.g., applying laboratory natural selection in replicated lines) in the laboratory. To investigate plastic responses to thermal fluctuations we compared the induced proteomic expression profiles among thermal regimes prior to experimental evolution. To investigate evolutionary responses to the thermal regimes we applied full genome sequencing. Specifically, we aimed to identify candidate mechanisms for plastic responses to temperature fluctuations in the proteome, and the genomic patterns of selection responses, of replicate lines exposed to twenty generations of selection in constant, predictable or unpredictable fluctuating thermal regimes. We expect that evolutionary responses act on cis-regulatory elements and will be detected in regions of the genome that encodes proteins inducible by temperature fluctuations (plasticity genes), if inducible and evolved mechanisms of heat tolerance are shared (as predicted if plasticity evolves by genetic assimilation, Pigliucci et al., 2006). Alternatively, evolutionary responses may occur in trans-regulatory elements and observed at genome level will be independent of the proteins induced by temperature fluctuations. Furthermore, we expect that evolutionary responses will be most pronounced in the predictably fluctuating environment if amplitude drives evolution. Alternatively, evolutionary responses will be most pronounced in the unpredictably fluctuating if predictability drives evolution.

MATERIALS AND METHODS

Experimental Animals

Two populations of *D. simulans* both collected at the same field site close to Bologna, Italy, were used in this study. The first was collected in August 2012 (referred to as the collection of '*D. simulans* 2012') and the second was collected at the same field site in August 2014 (referred to as the collection

[1] www.dmi.dk

of '*D. simulans* 2014'). Flies were throughout maintained in plastic bottles containing 50 mL of standard oatmeal-sugar-yeast-agar *Drosophila* medium at 23°C and a 16:8 h light:dark cycle. All experimental flies were generated using density control by transferring 40–45 eggs to plastic shell-vials with 7 mL medium. The first population (*D. simulans* 2012) was used to investigate the evolutionary response to selection in different thermal regimes, while the second population (*D. simulans* 2014) was used to investigate the plasticity induced by the same regimes, respectively (see **Figure 1**). The three regimes all had a mean temperature of 23°C, and were either Constant (C), Predictably fluctuating (PF), or Unpredictably fluctuating (UF) in temperature. Custom build programmable thermal cabinets maintained a 16:8 h light:dark cycle, with the C regime maintained at 23°C throughout, while the PF regime followed a 23–28–23°C sine curve during the light phase and a 23–13–23°C sine curve during the dark phase. The UF regime followed the same sine functions, but with a randomly sampled high temperature point between 23 and 28°C during the light phase and a randomly sampled low temperature point between 23 and 13°C during the dark phase (see **Figure 2**). The thermal regimes are described further in Manenti et al. (2015).

The *D. simulans* 2012 population was used to establish a mass population based on around 350 field caught inseminated females, where after flies were randomly divided into three selection regimes (C, PF, and UF). Each selection regime had three independent biological replicates, each based on three bottles (mixed within replicates each generation) with a combined population size of >500 flies. The selection regimes and maintenance procedures are described further in Manenti et al. (2015). After 20 generations of laboratory natural selection, we froze 250 density controlled females from each biological replicate within each rearing regime, resulting in a total number of 9 samples. These pooled samples were used for full genome sequencing to investigate genomic differences in the three regimes after laboratory experimental evolution.

The *D. simulans* 2014 collection was used to establish a mass population based on 12 bottles of larvae that were offspring of the field collected flies (>500 adult females). This mass population was maintained for two generations in the laboratory at 23°C, before vials with 40 ± 3 eggs were collected and distributed among the three thermal regimes where they developed. Upon emergence, the flies were transferred to fresh food vials and allowed to mature in their respective thermal regimes. When flies were 3–5 days old, flies from 2–3 vials were combined and flash frozen in liquid nitrogen (without anesthesia) to be assayed for the induced response of the proteome. Sampling was performed at 9 am, when the temperature of all three regimes was around 23°C. We sorted females from these samples using a stereomicroscope placed in a 5°C room, with the flies lying on a thin sheet of plastic on top of dry ice to keep the flies as cold as possible. We collected 3 samples of 50 females from each regime (C, PF, and UF, respectively) of *D. simulans* 2014, which were used to investigate the inducible proteomic response to the fluctuating regimes. Furthermore, as the original samples of the founder population before selection collected from the 2012 population were lost (in the initial batch sent for sequencing), we included unselected, density controlled flies collected from *D. simulans* 2014 as an alternative selection control in the full genome sequencing study. For this purpose three replicates of 150 individual females were collected as described above.

Genome Sequencing, Mapping, and SNP Calling

We extracted DNA from 150 females pooled from each of the 12 lines (three replicates of the founder *D. simulans* 2014 collection, and one sample of each of the three replicated selection lines (C, PF, and UF regimes, respectively) from the *D. simulans* 2012 collection). DNA was Illumina sequenced (100 bp PE) by BGI Hong Kong Co., Limited. The raw reads were quality filtered using TrimGalore[2] (parameters; –quality 20, –length 75). The filtered reads were mapped to the *D. simulans* reference genome (ASM75419v3[3]) using bwa (version 0.7.5a, mem algorithm, default parameters). Samtools (version 1.6.0) (Li et al., 2009) was used to convert sam to bam files, to sort bam files, to remove duplicates, and to make mpileup files (mpileup −6 –q 20 –d 100). In total 93,821,525 sites were analyzed. The following analyses were done using PoPoolation2 (Kofler et al., 2011); (1) indels were identified and removed (identify-indel-region.pl –min-count 10 –indel-window 5, filter-sync-by-gft.pl), (2) resulting mpileup files were converted to synchronized (sync) files (mpileup2sync.pl –min-qual 20), and (3) downsampled to coverage 40 (downsample-synchronized.pl –target-coverage 40 –max-coverage 10000 –method fraction).

Genomic Change (Divergence From Base Population)

Using the samples from the unselected '*D. simulans* 2014' as a base population, we estimated consistent allele frequency changes that have occurred during the 20 generations of experimental evolution in the three replicate populations from each of three different thermal selection regimes; constant (C), predictable fluctuating (PF) and unpredictable fluctuating (UF) temperatures. We expect genetic drift to have an equally strong impact across the three selection regimes given that population sizes were kept constant across selection regimes and replicates. Therefore consistent differences in the amount of genomic change among selection regimes will reflect differences in the strength of selection pressure. We performed the Cochran-Mantel-Haenszel (CMH) test (Agresti, 2002) (popoolation2: cmh-test.pl) to identify consistent changes (relative to the base population) in allele frequencies at all polymorphic sites (n = 2,332,305) across the entire genome (the very small chromosome 4 was not analyzed). We note that these SNPs do not represent independent loci due to linkage. To obtain an overall indication of the strength of selection pressure in the three selection regimes, we counted the number of SNPs

[2]https://github.com/FelixKrueger/TrimGalore

[3]https://ensembl.org

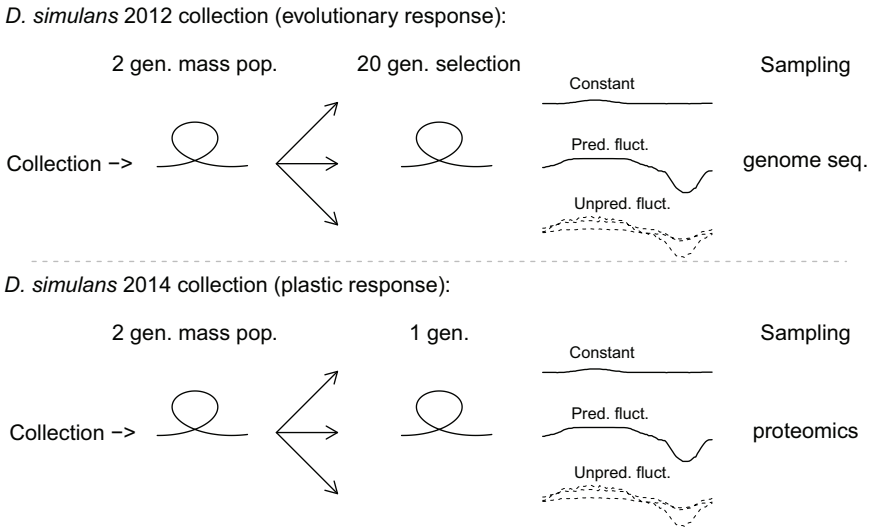


FIGURE 1 | Infographic outlining the design of the study of laboratory responses of *D. simulans* to constant, predictably and unpredictable fluctuating thermal environments. Top part of the figure shows the design of the selection experiment used to evaluate evolutionary responses by genome sequencing. The lower part shows the design of the phenotypic plasticity experiment used to evaluate inducible responses by proteomics. In both experiments the thermal regimes contained independent biological controls. The plotted temperature profiles of the constant, the predictable fluctuating and the unpredictable fluctuating thermal regimes represent the realized cabinet temperatures (temperature data and the thermal regimes are described in more detail in **Figure 2**).

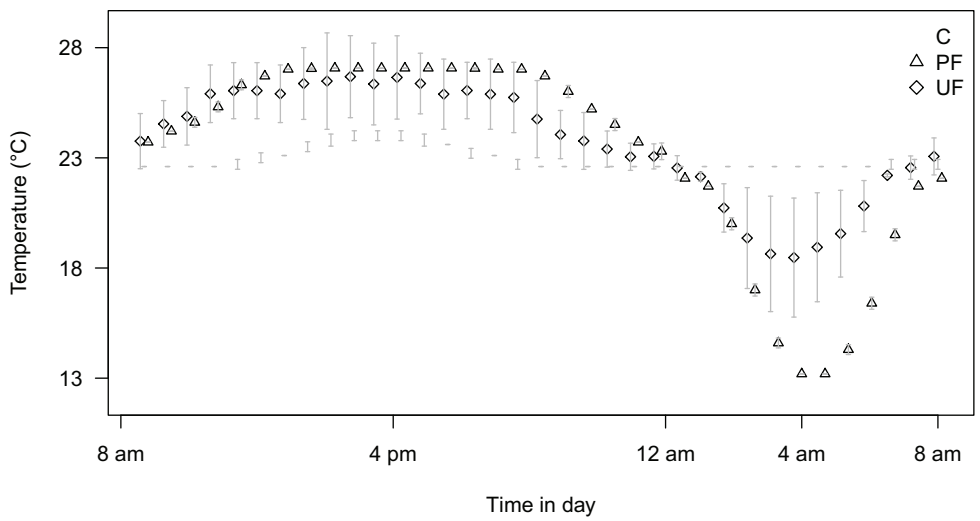


FIGURE 2 | Temperature profile of the three thermal regimes: Constant temperature (C), Predictable fluctuating (PF), and Unpredictable fluctuating (UF). The profiles represent average ± SD for 5 days of recording (points are jittered for better evaluation of error bars). Data and figure modified from Manenti et al. (2014). The average temperature of all regimes is 23°C and flies are exposed to 16:8 h light:dark cycles (note that the light generate a small increase in temperature in the constant regime). The predictable fluctuating regime was programmed to reach 27°C during the light period and 13°C during the dark period, respectively. Low error bars indicate that this was achieved. The unpredictable fluctuating regime was programmed to reach a randomly determined setpoint between 23 and 27°C during the light period and between 23 and 13°C during the dark period, respectively. The average temperatures closer to 23°C and higher SD, respectively, indicate that the fluctuations were on average smaller in amplitude, but unpredictable among days.

that had higher than an arbitrarily chosen threshold *p*-value [–log10(p) > 7] for each selection regime and chromosome separately. Using R (vs. 3.6.1) (R Core Team, 2019) we made Manhattan plots for each chromosome and selection line by plotting the negative log10-transformed *p*-values as a function of chromosome position.

Loci Under Selection (Divergence Among Selection Lines)

Consistent differences in SNP frequencies among replicates of each pair of selection regimes (C-PF, C-UP, and PF-UF) were identified using the CMH test (popoolation2: cmh-test.pl). As such differences can be the product of both random genetic drift

and selection, and the identification of loci under selection is challenging. We quantified genetic drift by performing CMH tests between the first two replicates from within each of the three thermal regimes. The resulting distribution of *p*-values across the genome is a good representation of the pattern that we expect due to drift and other sources of structure in our data. Using this distribution we selected two thresholds, more or less conservative, to detect segments of the genome where frequency changes between two selection regimes are consistent enough that we can interpret it as a product of adaptation (0.001 and 0.0001% percentile). Quantification of drift was also done using the other pairs of replicates within thermal regimes, which produced similar patterns (see **Supplementary Table 1**). Convergent evolution across selection regimes may occur as a consequence of both laboratory adaptation and other similarities among the three selection regimes. We quantified the pairwise overlap of significant SNPs between two selection regimes by creating 30 bins of SNP significance in one regime, and estimating the proportion of SNPs in each bin which has been under selection (according to the most conservative threshold) in another regime. In the case of independent evolution in the two regimes of a pair, we expected the proportion of SNPs in one regime to be independent of the significance level of the bins in the other regime. Alternatively, an overlap would result in an increased proportion of SNPs under selection in one regime as the significance level of bins increased (*p*-values decrease) in the other regime. We performed the same analyses with randomized *p*-values of one of the regimes as a point of reference under the null-expectation. Finally, we used the create-genewise-sync.pl script in PoPoolation2 (Kofler et al., 2011) to only analyze SNPs located in genes. This was also done for each pair of selection lines (C-PF, C-UP, and PF-UF). Similarly, we used the CMH test for each SNP independently, but averaged *p*-values across each gene. We only considered genes with more than 20 SNPs.

Proteomic Protocol

The proteomic investigation was generally performed as described in Sørensen et al. (2017). Briefly, proteins were extracted (lysis buffer: 100 mM Triethylammonium bicarbonate (TEAB) with 0.5% sodium dodecyl sulfate (SDS) and cOmplete ULTRA Tablets (Roche Diagnostics) protease inhibitor), mechanically homogenized (Bio101, Thermo Savant FastPrep FP120 cell disruptor) on ice between cycles to keep samples cold. Protein concentration of the homogenate was determined using a Qubit™ Fluorometer (Invitrogen, Life Technologies) and the Qubit Protein assay kit after centrifugation. Each sample (200 μg protein) was precipitated, dried and re-dissolved in 40 μL Dissolution Buffer with 2 μL Denaturant (iTRAQ® Reagents, AB Sciex). Proteins were reduced (Reducing Reagent, iTRAQ® Reagents, AB Sciex) and subsequently alkylated (Cysteine Blocking Reagent, iTRAQ® Reagents, AB Sciex). Proteins were then enzymatically digested over-night to peptides using a 1:50 trypsin:protein ratio (Sequencing Grade Modified Trypsin, Promega Biotech AB). Each sample was labeled with isobaric tags for relative and absolute quantitation (iTRAQ) (Pottiez et al., 2012). Hereafter, all three samples within one replicate and a common reference were pooled for fractionation and purification. Protein samples were fractioned on columns and fractions were subsequently eluted by increasing pH. Mass spectrometry analysis was performed by high-resolution electrospray ionization tandem mass spectrometry (ESI-MS/MS) (Köcher et al., 2009). Reverse phase nanoLC separation (Dionex UltiMate RSLCnano System) was performed online coupled to the mass spectrometer (LTQ Orbitrap Velos). Three technical replicates of each pooled sample replicate were analyzed, each using a slightly different gradient to increase protein coverage. Tandem mass spectrometry parameters were: positive mode, MS scan range 300–1600 with resolution at 30,000, MS/MS fragmentation was performed using HCD (higher-energy collisional dissociation) on the 20 most intense ions with a normalized collision energy of 40, dynamic exclusion of 90 s and a minimum signal threshold of 10,000.

Proteome Analyses

All raw data from triplicate injections of each fraction were searched and identified (Proteome Discoverer 1.4, Reporter Ions Quantifier, Percolator validator, Mascot 2.4 Search Engine and Swiss-Prot database restricted to taxonomy '*Drosophila*'), with the Percolator algorithm using semi-supervised machine learning and a target-decoy search strategy with reversed sequences to identify correct peptide sequence matches (Käll et al., 2007; Spivak et al., 2009). Search parameters were precursor mass tolerance 10 ppm and fragment mass tolerance 0.1 Da, maximum two missed cleavages, quantification method iTRAQ 4-plex. The strict target FDR was 0.01 for high and the relaxed target FDR was 0.05 for medium confident peptide matches. Protein quantification was based exclusively on unique peptides and among replicate experimental bias correction (Latosinska et al., 2015). The proteomics analysis identified 1319 unique protein IDs in the total data (dataset available as **Supplementary Data**: 'Table 1.xlsx'). Of these, 1001 unique protein IDs were detected in at least two out of three replicate samples for all selection regimes and were retained for analysis. We only accepted proteins that were detected in two of the three samples for all regimes, to avoid a strong bias from non-detected proteins. We performed ANOVA using regimes as a categorical variable. Pairwise (*post hoc* ANOVA) comparisons between selection lines maintained in the constant, the predictably and the unpredictably fluctuating regimes were performed on the resulting dataset. All analyses were performed using the statistical software R (vs. 3.6.1) (R Core Team, 2019) (R code available upon request to the authors).

Location of SNPs Under Selection Relative to Genes Responding Evolutionarily (Transcripts) and Plastically (Proteome) in Their Expression

We investigated whether SNPs under selection were physically linked to genes encoding proteins and transcripts responding to the thermal regimes. To visualize the locations, the genomic location of 204 transcripts that were previously found to show an evolutionary response to the fluctuating temperatures in their expression level (Manenti et al., 2018) were indicated on the

Manhattan plots showing the divergence among selection lines. The genomic location of 34 loci that in this study were found to show a plastic response in protein expression after exposure to the three thermal regimes (see below) were also indicated in the Manhattan plots. To test if the overlap visualized on Manhattan plots was different from the null expectation, we estimated the observed distance from SNPs under selection to (1) genes with protein expression responding to the thermal regimes (plastic responses), and (2) genes with RNA expression responding to selection in the thermal regimes. These distances were compared to distances to a null expectation (distances to random genes). We estimated confidence intervals of the deviation from the null expectation by producing 10,000 random gene sets, each corresponding to the number of significant genes, and for each gene set estimating the average distance to nearest SNP under selection. We then subtracted these 10,000 estimates of distance to random gene sets, from the observed distance to the significant gene set. Distances between SNPs under selection and significant genes being smaller than distances between SNPs under selection and random genes would indicate linkage.

RESULTS

Genomic Change (Divergence From Base Population)

Populations exposed to unpredictable fluctuations showed the largest number of consistently differentiated SNPs as compared to the base population (**Table 1**). This was consistent across chromosome arms, with the exception of chromosome 2L. The unpredictable fluctuations therefore likely expose the flies to a stronger selection pressure than the constant and predictably fluctuating temperatures, which both showed a lower number of consistently differentiated SNPs (**Table 1**). However, consistent differentiation to the base population was found among lines from all three thermal regimes suggesting that all thermal regimes imposed selection. The results of the CMH analyses were visualized as Manhattan plots for each chromosome and thermal regime (**Supplementary Figure 1**).

Loci Under Selection (Divergence Among Selection Lines)

The number of SNPs under selection according to the drift thresholds (see **Supplementary Figure 2**) was highest in populations exposed to the unpredictable fluctuating temperatures (**Table 2**). This result was consistent when selection thresholds were estimated using the other possible within-regime population pairs to quantify genetic drift, although as expected the exact threshold and therefore absolute number of SNPs inferred to be under selection varied (**Supplementary Table 1**). There was only a small overlap among the SNPs under selection between regimes. The magnitude of this overlap was similar in all pairwise comparisons and across all chromosomes (**Figure 3**: chromosome 2L, **Supplementary Figure 3**: all chromosomal segments). From visual inspection of the Manhattan plots comparing pairs of selection regimes (**Figure 4** and **Supplementary Figure 4**), it is evident that the selection responses, resulting in divergence between the three selection regimes, were not limited to a single or few loci with large consistent allele frequency changes. Rather it seems that several loci, showing smaller but consistent allele frequency changes, are involved.

TABLE 1 | Number of SNPs that show consistent allele frequency changes in each selection regime compared to the base population as estimated by CMH test [–log10(p) > 7] at the different chromosomes.

	2L	2R	3L	3R	X	Total
C	394	254	399	371	639	2057
PF	245	352	367	463	582	2009
UF	309	402	568	510	1188	2977

TABLE 2 | Number of SNPs that show consistent allele frequency changes among thermal regimes and have a p-value lower than the 0.001 and 0.0001% (0.001%/0.0001%, respectively) percentile of the drift analysis at the different chromosomes and selection regimes as estimated by CMH test.

	2L	2R	3L	3R	X	Total
C vs. PF	1233/220	2895/448	1447/232	2731/548	1121/176	9427/1624
C vs. UF	1654/241	2128/306	2046/335	2677/540	2436/540	10941/1962
PF vs. UF	1426/207	2846/534	2639/493	3667/790	1972/357	12550/2381

Based on CMH analyses on SNPs located in genes, we estimated an average p-value per gene by averaging over all SNPs in each gene. We only analyzed genes having 20 SNPs or more. We used a similar approach to quantify drift as outlined above, but only using the 0.001% percentile. From each comparison among selection regimes, we only found a few genes having lower p-values than expected due to drift (**Table 3**). The comparisons including unpredictable fluctuations showed the highest number of genes suggested to be under selection (C-PF: 17, C-UF: 30, and PF-UF: 40). Note that 12 genes in the drift analysis have p-values lower than the threshold.

The functional annotation of the identified genes was only available for about half the genes (**Table 3**). Attempting to identify likely *D. melanogaster* orthologs resulted in 'uncharacterized proteins.' Among the characterized genes, many were involved in regulation of expression of the genome, i.e., processes related to transcription, translation and post-transcriptional regulation or regulation (e.g., mRNA, splicing, tRNAs, chaperonins involved in protein folding). Additionally, the desaturase gene Desat 2, involved in the modifications of fatty acids, was identified.

Plastic Proteomic Responses

Of the 1001 analyzed proteins, 34 proteins showed differences in their expression levels among the three thermal regimes. *Post hoc* pairwise comparisons showed that none of the proteins were differentially expressed among all three regimes. Most of protein expression differences were found between the unpredictably fluctuating as compared to either the constant (24 proteins) or the predictably fluctuating (14 proteins) regime, respectively (**Table 4**). Eight of these proteins were shared between the two contrasts and thus showed unique expression

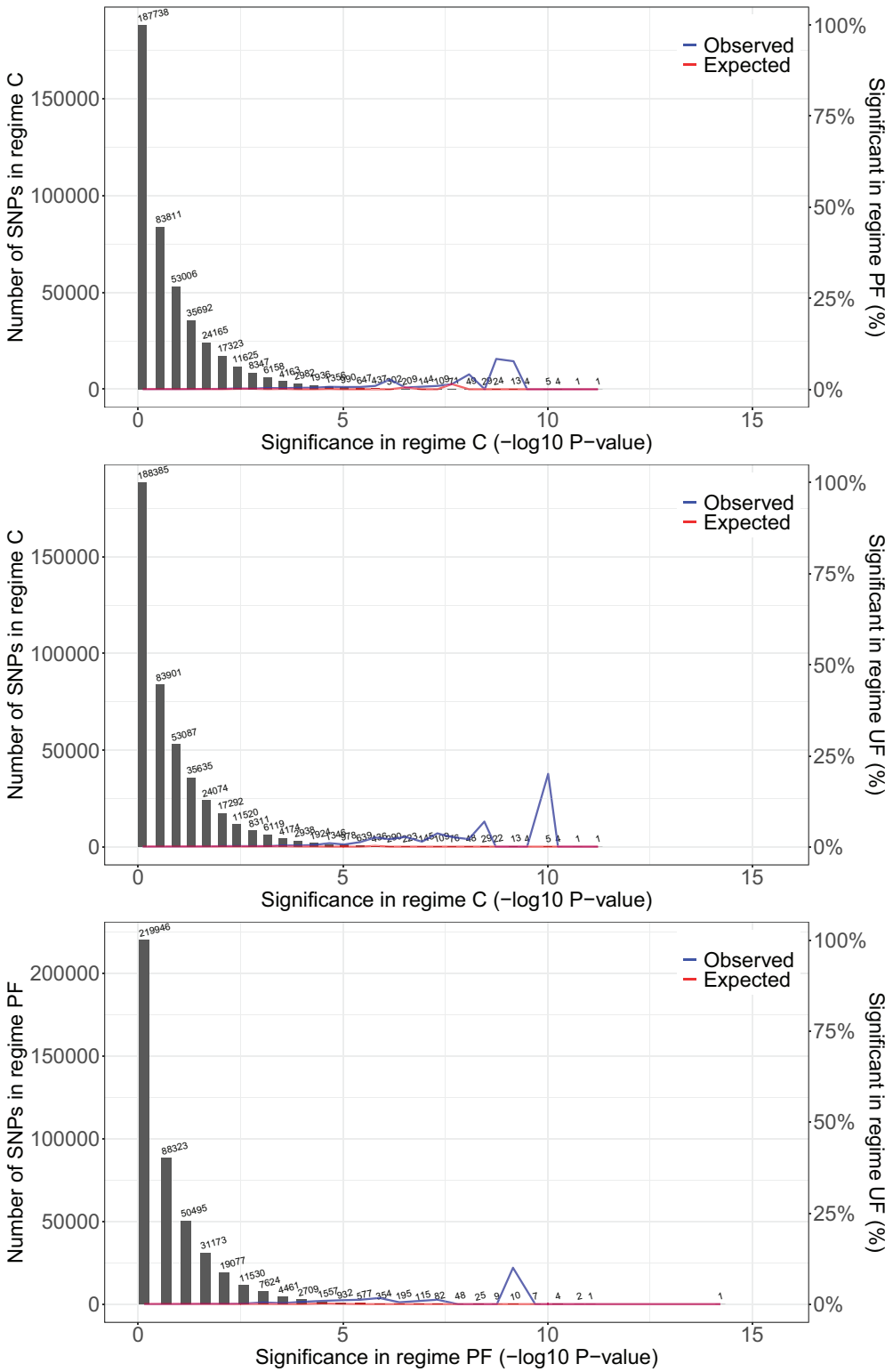


FIGURE 3 | Pairwise overlap of significant SNPs among thermal regimes. Plot shows the distribution of SNP significance [–log10(*p*-values)] in one regime (gray bars), the observed proportion of SNPs that overlap with significant SNPs in the second regime, and the proportion of SNPs with randomized significance in the second regime that overlap (null-expectation). There is a consistent signal of overlap of SNPs with low *p*-values (in a small number of SNPs) in all pairwise comparisons. Thus, of the few SNPs with low *p*-values [high –log10(*p*)] in one selection regime a larger proportion than expected by chance also have low *p*-values in the other selection regimes. Only SNPs from chromosome arm 2L are presented here. Results are consistent across chromosomes (see **Supplementary Figure 3**).

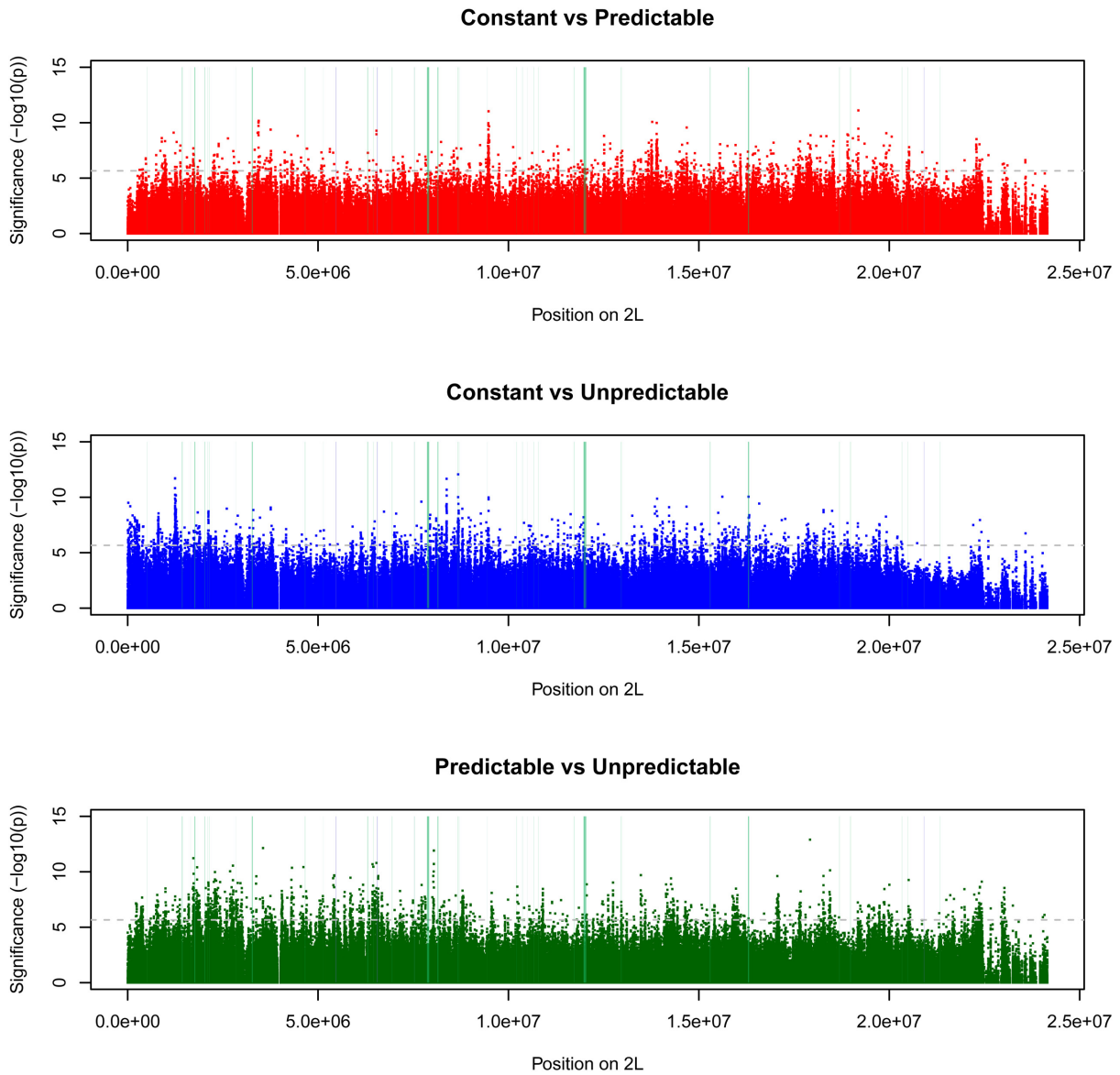


FIGURE 4 | Manhattan plots comparing pairs of selection regimes. Horizontal lines indicate two thresholds used to infer patterns of differentiation due to evolutionary changes across selection regimes; 0.001 and 0.0001% percentiles. SNPs with [–log10(*p*-values)] above these thresholds can be considered evidence of divergence between selection regimes due to adaptation. Vertical lines indicate positions of loci with diverging gene expression levels (Manenti et al., 2018) (blue) or diverging protein expression levels (green), with width of line representing length of gene. Only SNPs from chromosome arm 2L are presented here. Results are consistent across chromosomes (**Supplementary Figure 4**).

in the unpredictable regime. The proteins showing differential expression among flies exposed to the different thermal regimes covered a range of biological functions. Most notable are proteins known to respond to environmental cues, including several heat shock proteins (Hsp60, Hsc70-5, and Hsp83), turandot proteins (Turandot A, Turandot X), a metallothionein and two odorant-binding proteins. We also found several signaling proteins (Calmodulin, Pellino) as well as proteins involved in the transcriptional and translational machinery (ribosomal proteins, elongation and translation initiation factors) and turnover (Ubiquitin-conjugating enzymes).

SNPs Under Selection Are Not Closer to Genes That Show Expression Responses (Evolutionary or Plastic) Than Random Genes

The loci that have shown evolutionary responses in their level of transcription (Manenti et al., 2018) or proteome expression (this study) did not show obvious signs of physical linkage to SNPs under selection in the three thermal regimes. The average distance from these SNPs to nearest genes showing expression responses were in no comparison shorter than their

TABLE 3 | Loci showing average *p*-values below what was estimated for the 99.99% drift values in the contrasts of Constant vs. Predictable fluctuating temperature (C vs. PF), Predictable fluctuating vs. Unpredictable fluctuating temperature (PF vs. UF) or Constant vs. Unpredictable fluctuating temperature (C vs. UF), respectively.

Contrast C vs. PF				
ID	***#SNPs***	***p***	***Gene name***	***GOterm***
FBgn0182628	98	0.0725	GD10865	GO:0016021∼integral component of membrane
FBgn0184839	29	0.0555	GD13117	GO:0050909∼sensory perception of taste, GO:0005886∼plasma membrane, GO:0016021∼integral component of membrane
FBgn0187787	26	0.0694	GD16154	
FBgn0188937	29	0.0111	GD17379	GO:0001952∼regulation of cell-matrix adhesion, GO:0051492∼regulation of stress fiber assembly
FBgn0188963	63	0.0193	GD17407	
FBgn0189060	30	0.0168	GD17510	GO:0016021∼integral component of membrane
FBgn0190058	23	0.0701	GD18535	
FBgn0193190	92	0.0754	GD21769	
FBgn0195290	62	0.0756	GD23922	GO:0003677∼DNA binding
FBgn0195511	25	0.0587	GD24156	
FBgn0195773	39	0.0575	GD24437	GO:0046872∼metal ion binding
FBgn0195794	26	0.0546	GD24459	GO:0003676∼nucleic acid binding, GO:0005524∼ATP binding
FBgn0187743	21	0.0866	Heterochromatin Protein 1D2	GO:0005634∼nucleus
FBgn0268974	58	0.0794	GD27684	
FBgn0270901	22	0.0749	GD29611	
FBgn0270917	25	0.0587	GD29627	
FBgn0268873	23	0.0438	GD27583	
Contrast PF vs. UF				
FBgn0187821	151	0.0587	GD16192	
FBgn0187837	47	0.0531	GD16209	GO:0006807∼nitrogen compound metabolic process, GO:0016811∼hydrolase activity, acting on carbon-nitrogen. . .bonds, in linear amides
FBgn0190058	22	0.0628	GD18535	
FBgn0195794	66	0.0671	GD24459	GO:0003676∼nucleic acid binding, GO:0005524∼ATP binding
FBgn0185758	27	0.0585	GD14065	
FBgn0193177	28	0.0781	GD21755	GO:0016787∼hydrolase activity
FBgn0181897	22	0.0887	GD10122	
FBgn0188378	73	0.0853	GD16791	GO:0006357∼reg. transcript. from RNA polymerase II promoter, GO:0032784∼reg. DNA-templated transcript., elongation
FBgn0188837	34	0.0513	GD17275	GO:0016021∼integral component of membrane
FBgn0188389	42	0.0484	GD16802	GO:0008270∼zinc ion binding
FBgn0187064	22	0.0268	GD15396	
FBgn0187066	86	0.0785	GD15398	GO:0016021∼integral component of membrane, GO:0008173∼RNA methyltransferase activity
FBgn0191046	27	0.0840	GD19551	GO:0008380∼RNA splicing, GO:0030532∼small nuclear ribonucleoprotein complex
FBgn0188859	22	0.0201	GD17298	GO:0004252∼serine-type endopeptidase activity
FBgn0184980	37	0.0608	GD13260	
FBgn0188002	28	0.0499	GD16383	GO:0003676∼nucleic acid binding, GO:0005524∼ATP binding, GO:0008026∼ATP-dependent helicase activity
FBgn0188033	32	0.0861	GD16416	GO:0016012∼sarcoglycan complex, GO:0016021∼integral component of membrane
FBgn0270141	26	0.0599	GD28851	GO:0006457∼protein folding, GO:0005737∼cytoplasm, GO:0005524∼ATP binding
FBgn0193396	25	0.0704	GD21981	
FBgn0184684	21	0.0864	GD12960	
FBgn0187264	36	0.0782	GD15599	GO:0016021∼integral component of membrane
FBgn0187272	51	0.0067	GD15607	
FBgn0187770	21	0.0591	GD16135	

(Continued)

TABLE 3 | Continued

Contrast C vs. PF				
FBgn0186916	52	0.0899	GD15248	GO:0004672~protein kinase activity, GO:0005524~ATP binding
FBgn0186918	99	0.0097	GD15250	GO:0000398~mRNA splicing, via spliceosome, GO:0017070~U6 snRNA binding, GO:0030623~U5 snRNA binding
FBgn0184839	30	0.0753	GD13117	GO:0050909~sensory perception of taste, GO:0005886~plasma membrane,GO:0016021~integral component of membrane
FBgn0188744	28	0.0729	GD17176	GO:0016491~oxidoreductase activity
FBgn0192776	72	0.0800	GD21339	GO:0050909~sensory perception of taste, GO:0016021~integral component of membrane
FBgn0182763	23	0.0604	GD11003	GO:0016021~integral component of membrane
FBgn0197138	44	0.0272	GD25858	GO:0008073~ornithine decarboxylase inhibitor activity
FBgn0197187	31	0.0814	GD25909	GO:0004252~serine-type endopeptidase activity
FBgn0188910	32	0.0751	GD17350	GO:0005634~nucleus
FBgn0188937	24	0.0425	GD17379	GO:0030335~positive regulation of cell migration, GO:0016021~integral component of membrane
FBgn0187661	21	0.0757	GD16020	
FBgn0195562	25	0.0487	GD24212	GO:0016787~hydrolase activity
FBgn0194793	38	0.0796	GD23408	
FBgn0270901	22	0.0810	GD29611	
FBgn0268296	21	0.0789	GD27006	
FBgn0270617	26	0.0674	GD29327	
FBgn0270329	55	0.0831	GD29039	
Contrast C vs. UF				
FBgn0187885	44	0.0877	GD16258	
FBgn0195794	67	0.0471	GD24459	GO:0003676~nucleic acid binding, GO:0005524~ATP binding
FBgn0184467	23	0.0631	GD12740	GO:0016021~integral component of membrane, GO:0004252~serine-type endopeptidase activity
FBgn0188360	28	0.0320	GD16772	GO:0016021~integral component of membrane, GO:0004252~serine-type endopeptidase activity
FBgn0188361	49	0.0320	GD16773	GO:0016021~integral component of membrane, GO:0004252~serine-type endopeptidase activity
FBgn0261743	46	0.0864	Desaturase 2	GO:0006633~fatty acid biosynthetic process, GO:0016021~integral component of membrane
FBgn0188033	34	0.0037	GD16416	GO:0016012~sarcoglycan complex, GO:0016021~integral component of membrane
FBgn0270141	30	0.0750	GD28851	GO:0006457~protein folding, GO:0005737~cytoplasm, GO:0005524~ATP binding
FBgn0194698	29	0.0859	GD23312	GO:0006351~transcription, DNA-templated, GO:0003899~DNA-directed RNA polymerase activity
FBgn0196911	24	0.0324	GD25625	GO:0051539~4 iron, 4 sulfur cluster binding
FBgn0187272	50	0.0898	GD15607	
FBgn0187291	53	0.0644	GD15627	
FBgn0188770	26	0.0878	GD17206	GO:0008270~zinc ion binding
FBgn0182763	21	0.0234	GD11003	GO:0016021~integral component of membrane
FBgn0197138	30	0.0618	GD25858	GO:0008073~ornithine decarboxylase inhibitor activity
FBgn0188402	25	0.0365	GD16815	GO:0006744~ubiquinone biosynthetic process, GO:0055114~oxidation-reduction process
FBgn0186312	37	0.0725	GD14632	GO:0002949~tRNA threonylcarbamoyladenosine modification
FBgn0195553	103	0.0865	GD24203	GO:0016021~integral component of membrane, GO:0016791~phosphatase activity
FBgn0195555	42	0.0421	GD24205	GO:0072669~tRNA-splicing ligase complex, GO:0046872~metal ion binding
FBgn0187310	43	0.0550	GD15648	GO:0001700~embryonic development via the syncytial blastoderm, GO:0007259~JAK-STAT cascade, GO:0005622~intracellular
FBgn0184291	24	0.0721	GD12564	GO:0006886~intracellular protein transport, GO:0006913~nucleocytoplasmic transport, GO:0007264~small GTPase mediated signal transduction

(Continued)

TABLE 3 | Continued

Contrast C vs. PF				
FBgn0193948	80	0.0771	GD22547	GO:0006464~cellular protein modification process
FBgn0270798	21	0.0480	GD29508	
FBgn0268292	25	0.0457	GD27002	
FBgn0268296	21	0.0473	GD27006	
FBgn0268783	39	0.0615	GD27493	
FBgn0268450	33	0.0818	GD27160	
FBgn0268614	47	0.0401	GD27324	
FBgn0268873	28	0.0448	GD27583	
FBgn0269334	32	0.0309	GD28044	

Only loci with at least 20 SNPs were included. The FlyBase accession ID, number of SNPs for each locus (#SNPs), the average p-value (p), Gene name and associated functional annotation (GOterm) are given in the table. Gene name and GOterms (truncated for some redundant terms) were obtained by the Functional Annotation Table of DAVID (The Database for Annotation, Visualization and Integrated Discovery v6.8) (Huang et al., 2009) and FlyBase (flybase.org).

TABLE 4 | Proteomics results.

Accession	**Description**	**Contrast**	**PF-C**	**Contrast**	**UF-C**	**Contrast**	**UF-PF**
		FC	**P**	**FC**	**P**	**FC**	**P**
P48375	12 kDa FK506-binding protein	1.01	0.91	1.10	***0.03***	1.08	***0.05***
O02649	60 kDa heat shock protein, mitochondrial	1.05	0.14	1.13	***0.00***	1.07	***0.04***
Q9VA91	40S ribosomal protein S7	0.97	***0.03***	1.00	0.96	1.03	***0.04***
B4IL76	40S ribosomal protein S3	0.96	0.30	0.92	***0.03***	0.96	0.26
B4II57	Protein Turandot C	1.12	0.67	0.49	***0.02***	0.43	***0.01***
P35128	Ubiquitin-conjugating enzyme E2 N	1.03	0.86	1.16	***0.03***	1.14	0.06
P29845	Heat shock 70 kDa protein cognate 5	1.05	0.32	1.11	***0.03***	1.06	0.18
B4II58	Protein Turandot A1/2	1.28	0.49	0.47	0.13	0.37	***0.03***
P02828	Heat shock protein 83	1.03	0.24	1.09	***0.01***	1.06	***0.04***
P62152	Calmodulin	0.63	***0.05***	0.60	***0.04***	0.95	0.96
P13060	Elongation factor 2	0.97	0.25	0.93	***0.03***	0.97	0.24
Q8I1F4	rRNA 2'-O-methyltransferase fibrillarin	0.92	***0.01***	0.96	0.10	1.04	0.15
Q8MMC4	Protein CDV3 homolog	0.81	0.15	0.68	***0.05***	0.84	0.33
Q24046	Sodium/potassium-transporting ATPase subunit beta-1	0.98	0.34	1.03	0.25	1.05	***0.04***
Q24388	Larval serum protein 2	1.22	0.28	1.42	***0.04***	1.17	0.32
Q8SY61	General odorant-binding protein 56d	1.21	0.08	1.27	***0.03***	1.05	0.75
P07182	Chorion protein S36	0.97	0.82	0.77	***0.01***	0.80	***0.03***
Q9VAI6	General odorant-binding protein 99b	1.23	0.14	1.78	***0.00***	1.44	***0.00***
P24511	Chorion protein S16	1.01	0.98	0.84	0.06	0.84	***0.05***
Q9V7N5	V-type proton ATPase subunit C	0.95	0.22	0.91	***0.03***	0.96	0.35
P04357	Metallothionein-1	1.09	0.77	1.44	***0.03***	1.33	0.07
P48598	Eukaryotic translation initiation factor 4E	1.04	0.42	0.94	0.17	0.90	***0.03***
P41073	Zinc finger protein on ecdysone puffs	0.95	0.19	0.92	***0.04***	0.97	0.50
Q9VPH7	Eukaryotic peptide chain release factor subunit 1	1.18	***0.05***	1.04	0.77	0.88	0.12
Q7KML2	Probable peroxisomal acyl-coenzyme A oxidase 1	1.11	***0.05***	1.09	0.08	0.99	0.94
Q27237	Protein tumorous imaginal disks, mitochondrial	1.15	***0.02***	1.14	***0.03***	0.99	0.97
O77237	Protein pellino	0.76	***0.00***	0.78	***0.01***	1.03	0.90
P16163	Uricase	1.02	0.99	0.53	***0.01***	0.52	***0.01***
Q9VAI1	Probable complex I intermediate-associated protein 30, mitochondrial	0.82	***0.02***	0.80	***0.02***	0.98	0.88
Q9VTY6	Ubiquitin-conjugating enzyme E2 C	1.19	***0.02***	1.10	0.21	0.92	0.22
Q9VVW8	ATP-dependent (S)-NAD(P)H-hydrate dehydratase	1.13	0.29	1.26	***0.04***	1.11	0.33
P42787	Carboxypeptidase D	1.12	***0.01***	1.01	0.86	0.91	***0.02***
O02437	Protein yellow	0.99	0.97	0.90	***0.04***	0.91	0.06
Q9I7T7	La-related protein CG11505	0.94	0.11	0.85	***0.01***	0.91	***0.05***

UniProt accession ID and protein name for Fold change (log2 FC) and p-value for each pairwise contrast is given. P-values in italics and bold are considered as significant ($p < 0.05$).

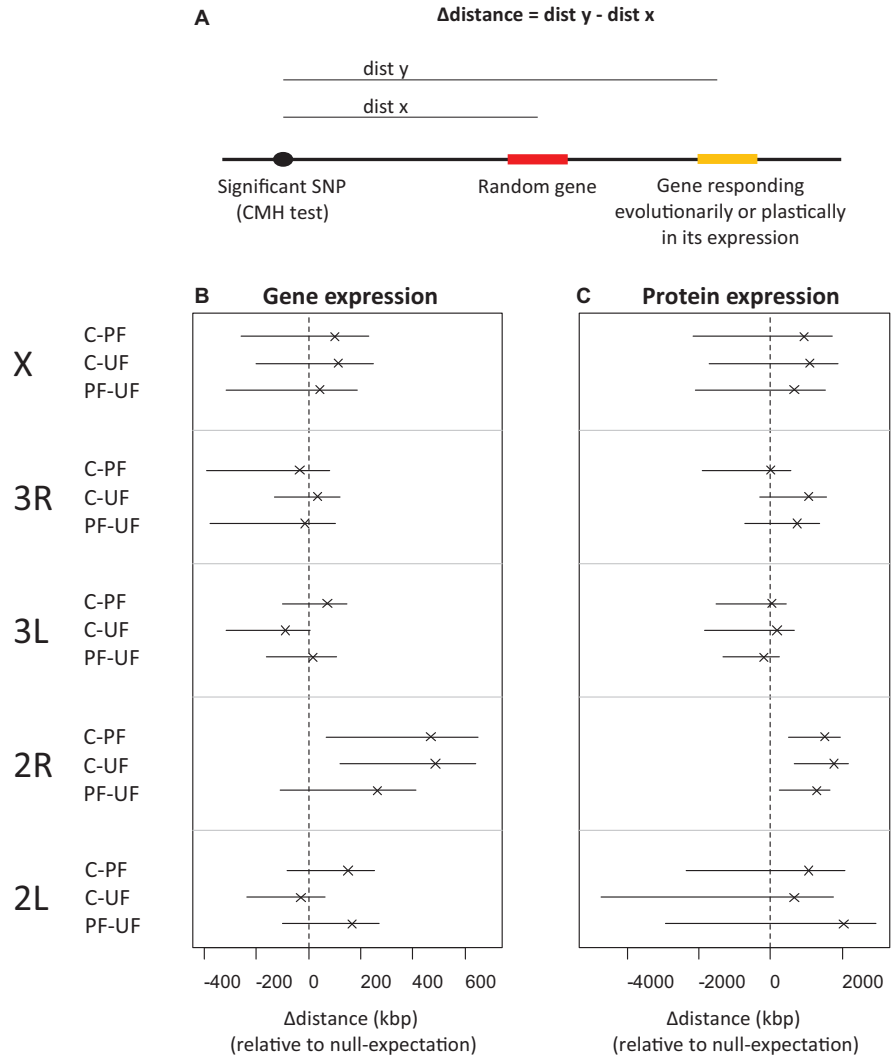


FIGURE 5 | Represents an analysis of distance between SNPs detected to be under selection in the thermal regimes detected by the CMH test and genes that showed an evolutionary response in their gene expression and genes that show a plastic response in their proteome expression. **(A)** illustration of the metric used: Δdistance (the average distance between SNPs under selection and nearest gene showing expression response – the average distance between SNPs under selection and nearest gene out of a random set of 204 genes in gene expression and 34 genes in protein expression). **(B,C)** Δdistance calculated for each chromosome and thermal regime comparison. Crosses (x) are medians and horizontal lines (–) are 95% confidence intervals both based on 10000 random sets of genes. **(B)** Shows results of genes that show an evolutionary response in gene expression and **(C)** the results of genes that show a plastic response in proteome expression.

average distance to random genes (**Figure 5**) suggesting that (1) evolutionary responses in gene expression to the thermal regimes are not caused by selection on cis-elements, and (2) genes that respond plastically to the thermal regimes were not under strong selection.

DISCUSSION

In this study, we applied thermal regimes with temperature variation on replicated populations of *D. simulans* to mimic naturally occurring daily variation (Kingsolver et al., 2009; Bozinovic et al., 2011; Manenti et al., 2014). We found a signal of the 20 generations of experimental evolution as flies maintained in constant, predictable or unpredictable fluctuating environments showed independent genomic differentiation. Natural populations are exposed, and likely adapted, to thermal fluctuations. Thus, as temperatures in natural habitats are not predictable, the unpredictable regime could be argued to represent the most natural condition (Colinet et al., 2015). Our genomic data suggests that the selection pressure in the unpredictable fluctuating regime is distinct from the constant and predictable fluctuating thermal environments. This result corroborates the induced response (i.e., results from the proteomic analysis presented here) and results from an earlier study where the transcriptome responses of the same

selection lines showed similar patterns of a stronger response to unpredictable temperature fluctuations (Manenti et al., 2018). Furthermore, previous studies have documented that exposure to unpredictable temperature fluctuation was more stressful compared to a predictable fluctuating temperature regime with the same mean temperature (as measured by decreased performance in several life history and stress tolerance traits) (Manenti et al., 2014), and that performance of the lines selected in the unpredictable fluctuating regime had evolved enhanced stress tolerance (Manenti et al., 2015).

The finding that the unpredictable fluctuations impose a distinct selection pressure (even if the amplitude was smaller than that of the predictable fluctuating regime) suggests that the unpredictable selection regime does not resemble natural thermal profiles. If our unpredictably fluctuating regime did represent the natural environment, it could be hypothesized to represent the least novel environment and show the least change from the base population. It is often hypothesized that keeping up with environmental change by induction of plastic responses in variable environments can be costly (Dewitt et al., 1998; for an empirical example, see Kristensen et al., 2008). In this study (and in Manenti et al., 2018), predictability rather than amplitude seems to impose the strongest effect on both plastic and evolutionary responses. This points to that the cost of thermal fluctuations (and maybe plasticity generally) is related to the costs of evaluating the environment rather than the amount of regulation needed (Dewitt et al., 1998).

Distinct Selection Responses Among Thermal Regimes

We looked for SNPs that show similar selective responses across selection regimes by comparing *p*-values of each SNP between each pair of the three selection regimes, and for each chromosome separately (**Figure 3** and **Supplementary Figure 3**). There was not a strong pattern that SNPs with low *p*-values [high –log10(p)] in one selection regime also had low *p*-values in the other selection regimes, suggesting that in general the loci under selection are unique to each selection regime. We do note, however, that we did identify an overlap in a small number of SNPs across pairwise comparisons of the three selection regimes. This could represent a weak signal of a common evolutionary response to thermal regimes or shared laboratory conditions, not including temperature, that affect the same loci. While laboratory adaptation can be prominent, for thermal tolerance it does not seem to be a general concern (Maclean et al., 2018), and in this study did not overshadow the effect of the individual regimes.

Our findings support the conclusion that independent selection has occurred in response to the three different selection regimes. This underlines that fluctuating thermal regimes are affecting organism in a very complex way. Thus, the effect of natural fluctuations can be proposed to be associated both with amplitude (e.g., Terblanche et al., 2010; Bozinovic et al., 2011; Ma et al., 2015), but independently also by a marked effect of predictability, possibly mediated by the cost of continuously evaluating and adjusting to the present conditions (Dewitt et al., 1998).

Proteomic Response to Thermal Fluctuation Regimes

Among the significantly differentially expressed proteins several heat shock and Turandot genes were found. Qualitative comparison to the study of the evolved transcriptome by Manenti et al. (2018) suggests a substantial overlap in functional groups, e.g., ribosomal genes, ubiquitin conjugating enzymes, Turandot genes, odorant binding proteins, but surprisingly no overlap with the heat shock genes found in this study. The Turandot and heat shock genes expressed in flies exposed to fluctuations (and particular to unpredictable fluctuations) indicate a heat stress response (Ekengren and Hultmark, 2001; Sørensen et al., 2003). The modest expression levels and a previous transcriptomic study suggest that the up-regulation due to temperature fluctuations are mainly constitutive, rather than tracking the temperature fluctuations (Manenti et al., 2018). The three heat shock proteins found here (Hsp83, Hsc70-5, and Hsp60) support this notion, as they have relatively high constitutive expression and are not expressed at much higher levels by stress (Feder and Hofmann, 1999; Sørensen et al., 2005). The other proteins detected as differentially expressed have no known connection to thermal tolerance, but include proteins important for efficient transcription and translation which might serve to maintain functioning of the cellular machinery allowing animals to survive and reproduce. The low number of proteins found to be differentially expressed is likely due to the relative benign extreme temperatures and the resulting low fold-change induced, but the low number of differentially expressed proteins does also indicate that regulation of quite few proteins is adequate for maintaining cellular function at variable temperatures. A recent study compared heat tolerance and the transcriptomic response to developmental acclimation at different mean temperatures (15 or 25°C), and either fluctuating or constant temperatures (Sørensen et al., 2016b). They found that fluctuations affected heat tolerance markedly, despite of a low number of transcripts (a few hundred) being differentially expressed between the constant and fluctuating regimes differentially expressed. In contrast, a large number of transcripts responded differences in mean temperatures (6000–8000 different expressed) (Sørensen et al., 2016b). This suggests that while some stress response proteins were induced by fluctuations, the molecular underpinnings of benign fluctuations might differ substantially from the pathways known to be involved with more extreme thermal acclimation or hardening.

Evolutionary Adaptation to Fluctuating Temperatures

The loci affected by selection were (for those that could be functionally annotated) related to regulation of the expression of the genome, rather than to genes with known functions in thermal responses or tolerance. This suggests that the evolutionary adaption to fluctuations within benign temperatures is to a large extent achieved by trans-regulation of the genome, rather than by increasing thermal stability and adaptation of proteins seen in extremophiles (Jollivet et al., 2012).

The distribution of distances between SNPs seemingly under selection and the transcriptome differences imposed by selection by the thermal regimes also showed no sign of cis-regulation (i.e., short distances). For the transcripts showing a change in expression following selection this might indicate some degree of trans-regulation. For the proteins responding to our thermal regimes this might suggest that evolved differences and the plastic response are not generally achieved by common genes. However, we acknowledge that linkage and linkage disequilibrium might contribute to our results. Unfortunately, lack of information on linkage in *D. simulans* and the pooled sequencing data prevents us from a more dedicated analysis of this effect. One potential cis-regulated mechanism is related to the Desat2 gene (a lipid desaturase), which could be related to membrane homeoviscous adaptation (Hazel, 1995). However, this remains a hypothesis for future testing.

Fluctuating temperature regimes can vary in terms of period, amplitude, rate of temperature change and predictability in addition to the mean temperature. Thus, even if studies with thermal fluctuations can be considered more ecologically relevant than those with constant temperatures, it is by no means simple to compare results across laboratory studies and to extrapolate findings to field conditions (Ketola and Kristensen, 2017). We found no marked overlap between loci being affected by selection and the proteins induced by the regimes suggesting that the evolutionary and plastic responses are achieved by distinct genes. The genes associated by many significant SNPs suggest that transcription, translation and post-translational modification are targets of evolutionary change. Increased expression of Turandot proteins under fluctuations support this group of genes/proteins as emerging candidates for mediating thermal acclimation to fluctuations. Based on our findings we argue that it is important to acknowledge that predictable and unpredictable (and constant) thermal environments have different impacts on fitness. While other studies have shown that constant and fluctuating temperatures induce different plastic responses and evolutionary pressures (Botero et al., 2015; Dey et al., 2016), this aspect is often ignored in functional and evolutionary studies on thermal adaptation. This is critical because climate models predict more variable and less predictable climates in the future (IPCC, 2014).

DATA AVAILABILITY STATEMENT

The raw genome sequencing data was uploaded to the NCBI SRA (short read archive) BioProject accession: PRJNA629467.

AUTHOR CONTRIBUTIONS

JS and VL acquired the financial support for the project leading to this publication. JS, TM, and VL conceived and designed the study. JS and TM performed experiments and data collection. JS, JB, and MS analyzed the data and wrote first draft of the manuscript. All authors contributed to the final draft and approved the final version of the manuscript.

FUNDING

The study was supported by the Graduate School of Science and Technology (GSST) at Science and Technology, Aarhus University (TM), and by a Sapere Aude: DFF-Starting Grant (JS) and a large frame grant (VL) from the Independent Research Fund Denmark, Natural Sciences.

ACKNOWLEDGMENTS

The authors would like to thank Annemarie Højmark and Trine Bech Søgaard for their help with the laboratory work and Sandro Cavicchi for proving the flies.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.555843/full#supplementary-material

REFERENCES

Agresti, A. (2002). *Categorical Data Analysis*. New York, NY: John Wiley & Sons.

Araujo, M. B., Ferri-Yanez, F., Bozinovic, F., Marquet, P. A., Valladares, F., and Chown, S. L. (2013). Heat freezes niche evolution. *Ecol. Lett.* 16, 1206–1219.

Ashander, J., Chevin, L.-M., and Baskett, M. L. (2016). Predicting evolutionary rescue via evolving plasticity in stochastic environments. *Proc. R. Soc. B Biol. Sci.* 283:20161690. doi: 10.1098/rspb.2016.1690

Botero, C. A., Weissing, F. J., Wright, J., and Rubenstein, D. R. (2015). Evolutionary tipping points in the capacity to adapt to environmental change. *Proc. Natl. Acad. Sci. U.S.A.* 112, 184–189. doi: 10.1073/pnas.1408589111

Boyce, M. S., Haridas, C. V., Lee, C. T., and The Nceas Stochastic Demography Working Groups (2006). Demography in an increasingly variable world. *Trends Ecol. Evol.* 21, 141–148. doi: 10.1016/j.tree.2005.11.018

Bozinovic, F., Bastias, D. A., Boher, F., Clavijo-Baquet, S., Estay, S. A., and Angilletta, M. J. (2011). The mean and variance of environmental temperature interact to determine physiological tolerance and fitness. *Physiol. Biochem. Zool.* 84, 543–552.

Castaneda, L. E., Romero-Soriano, V., Mesas, A., Roff, D. A., and Santos, M. (2019). Evolutionary potential of thermal preference and heat tolerance in *Drosophila subobscura*. *J. Evol. Biol.* 32, 818–824. doi: 10.1111/jeb.13483

Colinet, H., Sinclair, B. J., Vernon, P., and Renault, D. (2015). Insects in fluctuating thermal environments. *Annu. Rev. Entomol.* 60, 123–140. doi: 10.1146/annurev-ento-010814-021017

Dewitt, T. J., Sih, A., and Wilson, D. S. (1998). Costs and limits of phenotypic plasticity. *Trends Ecol. Evol.* 13, 77–81.

Dey, S., Proulx, S. R., and Teotonio, H. (2016). Adaptation to temporally fluctuating environments by the evolution of maternal effects. *PLoS Biol.* 14:e1002388. doi: 10.1371/journal.pbio.1002388

Ekengren, S., and Hultmark, D. (2001). A family of Turandot-related genes in the humoral stress response of *Drosophila*. *Biochem. Biophys. Res. Commun.* 284, 998–1003.

Feder, M. E., and Hofmann, G. E. (1999). Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. *Annu. Rev. Physiol.* 61, 243–282. doi: 10.1146/annurev.physiol.61.1.243

Fragata, I., Lopes-Cunha, M., Barbaro, M., Kellen, B., Lima, M., Faria, G. S., et al. (2016). Keeping your options open: Maintenance of thermal plasticity during adaptation to a stable environment. *Evolution* 70, 195–206.

Gerken, A. R., Eller, O. C., Hahn, D. A., and Morgan, T. J. (2015). Constraints, independence, and evolution of thermal plasticity: probing genetic architecture of long- and short-term thermal acclimation. *Proc. Natl. Acad. Sci. U. S. A.* 112, 4399–4404. doi: 10.1073/pnas.1503456112

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A., and Merila, J. (2008). Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.* 17, 167–178. doi: 10.1111/j.1365-294X.2007.03413.x

Gunderson, A. R., and Stillman, J. H. (2015). Plasticity in thermal tolerance has limited potential to buffer ectotherms from global warming. *Proc. R. Soc. B Biol. Sci.* 282:20150401. doi: 10.1098/rspb.2015.0401

Hallsson, L. R., and Bjorklund, M. (2012). Selection in a fluctuating environment leads to decreased genetic variation and facilitates the evolution of phenotypic plasticity. *J. Evol. Biol.* 25, 1275–1290. doi: 10.1111/j.1420-9101.2012.02512.x

Hazel, J. R. (1995). Thermal adaptation in biological membranes: is homeoviscous adaptation the explanation? *Annu. Rev. Physiol.* 57, 19–42.

Hoffmann, A. A., Anderson, A., and Hallas, R. (2002). Opposing clines for high and low temperature resistance in *Drosophila melanogaster*. *Ecol. Lett.* 5, 614–618.

Hoffmann, A. A., and Parsons, P. A. (1989). An integrated approach to environmental stress tolerance and life-history variation: desiccation tolerance in *Drosophila*. *Biol. J. Linn. Soc.* 37, 117–136.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.

IPCC (2014). "Summary for Policymakers," In *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Eds C.B. Field et al. (Cambridge, NY: Cambridge University Press), 1–32.

Jollivet, D., Mary, J., Gagniere, N., Tanguy, A., Fontanillas, E., Boutet, I., et al. (2012). Proteome adaptation to high temperatures in the ectothermic hydrothermal vent Pompeii worm. *PLoS One* 7:e31150. doi: 10.1371/journal.pone.0031150

Käll, L., Canterbury, J. D., Weston, J., Noble, W. S., and MacCoss, M. J. (2007). Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* 4, 923–925.

Kellermann, V., Overgaard, J., Hoffmann, A. A., Fløjgaard, C., Svenning, J. C., and Loeschcke, V. (2012). Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. *Proc. Natl. Acad. Sci. U. S. A* 109, 16228–16233. doi: 10.1073/pnas.1207553109

Kellermann, V. M., Hoffmann, A. A., Kristensen, T. N., Moghadam, N. N., and Loeschcke, V. (2015). Experimental evolution under fluctuating thermal conditions does not reproduce patterns of adaptive clinal differentation in *Drosophila melanogaster*. *Am. Nat.* 186, 582–593.

Ketola, T., and Kristensen, T. N. (2017). Experimental approaches for testing if tolerance curves are useful for predicting fitness in fluctuating environments. *Front. Ecol. Evol.* 5:129. doi: 10.3389/fevo.2017.00129

Ketola, T., Laakso, J., Kaitala, V., and Airaksinen, S. (2004). Evolution of Hsp90 expression in *Tetrahymena thermophila* (Protozoa, Ciliata) populations exposed to thermally variable environments. *Evolution* 58, 741–748. doi: 10.1111/j.0014-3820.2004.tb00407.x

Kingsolver, J. G., Ragland, G. J., and Diamond, S. E. (2009). Evolution in a constant environment: Thermal fluctuations and thermal sensitivity of laboratory and field populations of *Manduca sexta*. *Evolution* 63, 537–541. doi: 10.1111/j.1558-5646.2008.00568.x

Köcher, T., Pichler, P., Schutzbier, M., Stingl, C., Kaul, A., Teucher, N., et al. (2009). High precision quantitative proteomics using iTRAQ on an LTQ Orbitrap: A new mass spectrometric method combining the benefits of all. *J. Proteome Res.* 8, 4743–4752. doi: 10.1021/pr900451u

Kofler, R., Pandey, R. V., and Schlotterer, C. (2011). PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 27, 3435–3436. doi: 10.1093/bioinformatics/btr589

Kontopoulos, D. G., van Sebille, E., Lange, M., Yvon-Durocher, G., Barraclough, T. G., and Pawar, S. (2020). Phytoplankton thermal responses adapt in the absence of hard thermodynamic constraints. *Evolution* 74, 775–790. doi: 10.1111/evo.13946

Kristensen, T. N., Hoffmann, A. A., Overgaard, J., Sørensen, J. G., Hallas, R., and Loeschcke, V. (2008). Costs and benefits of cold acclimation in field-released *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 216–221.

Kristensen, T. N., Ketola, T., and Kronholm, I. (2018). Adaptation to environmental stress at different timescales. *Ann. N. Y. Acad. Sci.* 1476, 5–22. doi: 10.1111/nyas.13974

Lande, R. (2009). Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J. Evol. Biol.* 22, 1435–1446. doi: 10.1111/j.1420-9101.2009.01754.x

Latosinska, A., Vougas, K., Makridakis, M., Klein, J., Mullen, W., Abbas, M., et al. (2015). Comparative analysis of label-free and 8-Plex iTRAQ approach for quantitative tissue proteomic analysis. *PLoS One* 10:e0137048. doi: 10.1371/journal.pone.0137048

Lecheta, M. C., Awde, D. N., O'Leary, T. S., Unfried, L. N., Jacobs, N. A., Whitlock, M. H., et al. (2020). Integrating GWAS and transcriptomics to identify the molecular underpinnings of thermal stress responses in *Drosophila melanogaster*. *Front. Genet.* 11:658. doi: 10.3389/fgene.2020.00658

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Logan, M. L., and Cox, C. L. (2020). Genetic Constraints, Transcriptome Plasticity, and the Evolutionary Response to Climate Change. *Front. Genet.* 11:538226. doi: 10.3389/fgene.2020.538226

Ma, G., Hoffmann, A. A., and Ma, C. S. (2015). Daily temperature extremes play an important role in predicting thermal effects. *J. Exp. Biol.* 218(Pt 14), 2289–2296. doi: 10.1242/jeb.122127

Maclean, H. J., Kristensen, T. N., Sørensen, J. G., and Overgaard, J. (2018). Laboratory maintenance does not alter ecological and physiological patterns among species: a *Drosophila* case study. *J. Evol. Biol.* 31, 530–542. doi: 10.1111/jeb.13241

Manenti, T., Loeschcke, V., Moghadam, N. N., and Sørensen, J. G. (2015). Phenotypic plasticity is not affected by experimental evolution in constant, predictable or unpredictable fluctuating thermal environments. *J. Evol. Biol.* 28, 2078–2087. doi: 10.1111/jeb.12735

Manenti, T., Loeschcke, V., and Sørensen, J. G. (2018). Constitutive up-regulation of Turandot genes rather than changes in acclimation ability is associated with the evolutionary adaptation to temperature fluctuations in Drosophila simulans. *J. Insect. Physiol.* 104, 40–47. doi: 10.1016/j.jinsphys.2017.11.008

Manenti, T., Sørensen, J. G., and Loeschcke, V. (2017). Environmental heterogeneity does not affect levels of phenotypic plasticity in natural populations of three *Drosophila* species. *Ecol. Evol.* 7, 2716–2724. doi: 10.1002/ece3.2904

Manenti, T., Sørensen, J. G., Moghadam, N. N., and Loeschcke, V. (2014). Predictability rather than amplitude of temperature fluctuations determines stress resistance in a natural population of *Drosophila simulans*. *J. Evol. Biol.* 27, 2113–2122. doi: 10.1111/jeb.12463

Merilä, J., and Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol. Appl.* 7, 1–14. doi: 10.1111/eva.12137

Overgaard, J., Kristensen, T. N., Mitchell, K. A., and Hoffmann, A. A. (2011). Thermal tolerance in widespread and tropical *Drosophila* species: Does phenotypic plasticity increase with latitude? *Am. Nat.* 178, S80–S96.

Perkins-Kirkpatrick, S. E., and Lewis, S. C. (2020). Increasing trends in regional heatwaves. *Nat. Commun.* 11:3357. doi: 10.1038/s41467-020-16970-7

Pigliucci, M. (1996). How organisms respond to environmental changes: From phenotypes to molecules (and vice versa). *Trends Ecol. Evol.* 11, 168–173. doi: 10.1016/0169-5347(96)10008-2

Pigliucci, M. (2001). *Phenotypic Plasticity: Beyond Nature and Nurture*. Baltimore: The John Hopkins University Press.

Pigliucci, M., Murren, C. J., and Schlichting, C. D. (2006). Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Biol.* 209, 2362–2367. doi: 10.1242/jeb.02070

Pottiez, G., Wiederin, J., Fox, H. S., and Ciborowski, P. (2012). Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples. *J. Proteome Res.* 11, 3774–3781. doi: 10.1021/pr300414z

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Radchuk, V., Reed, T., Teplitsky, C., van de Pol, M., Charmantier, A., Hassall, C., et al. (2019). Adaptive responses of animals to climate change are most likely insufficient. *Nat. Commun.* 10:3109. doi: 10.1038/s41467-019-10924-4

Schou, M. F., Kristensen, T. N., Kellermann, V., Schlötterer, C., and Loeschcke, V. (2014). A *Drosophila* laboratory evolution experiment points to low evolutionary potential under increased temperatures likely to be experienced in the future. *J. Evol. Biol.* 27, 1859–1868. doi: 10.1111/jeb.12436

Schreiber, S. J. (2010). Interactive effects of temporal correlations, spatial heterogeneity and dispersal on population persistence. *Proc. Biol. Sci.* 277, 1907–1914. doi: 10.1098/rspb.2009.2006

Sgrò, C., Terblanche, J. S., and Hoffmann, A. A. (2016). What can plasticity contribute to insect responses to climate change? *Annu. Rev. Entomol.* 61, 433–451.

Simões, P., Santos, M. A., Carromeu-Santos, A., Quina, A. S., Santos, M., and Matos, M. (2020). Beneficial developmental acclimation in reproductive performance under cold but not heat stress. *J. Therm. Biol.* 90:102580. doi: 10.1016/j.jtherbio.2020.102580

Sørensen, J. G., Kristensen, T. N., and Loeschcke, V. (2003). The evolutionary and ecological role of heat shock proteins. *Ecol. Lett.* 6, 1025–1037.

Sørensen, J. G., Kristensen, T. N., and Overgaard, J. (2016a). Evolutionary and ecological patterns of thermal acclimation capacity in *Drosophila*: is it important for keeping up with climate change? *Curr. Opin. Insect Sci.* 17, 98–104. doi: 10.1016/j.cois.2016.08.003

Sørensen, J. G., Schou, M. F., Kristensen, T. N., and Loeschcke, V. (2016b). Thermal fluctuations affect the transcriptome through mechanisms independent of average temperature. *Sci. Rep.* 6:30975. doi: 10.1038/srep30975

Sørensen, J. G., Nielsen, M. M., Kruhoffer, M., Justesen, J., and Loeschcke, V. (2005). Full genome gene expression analysis of the heat stress response, in *Drosophila melanogaster*. *Cell Stress Chaperones* 10, 312–328.

Sørensen, J. G., Schou, M. F., and Loeschcke, V. (2017). Evolutionary adaptation to environmental stressors: a common response at the proteomic level. *Evolution* 71, 1627–1642.

Spivak, M., Weston, J., Bottou, L., Käll, L., and Noble, W. S. (2009). Improvements to the percolator algorithm for Peptide identification from shotgun proteomics data sets. *J. Proteome Res.* 8, 3737–3745. doi: 10.1021/pr80 1109k

Terblanche, J. S., Nyamukondiwa, C., and Kleynhans, E. (2010). Thermal variability alters climatic stress resistance and plastic responses in a globally invasive pest, the Mediterranean fruit fly (*Ceratitis capitata*). *Entomol. Exp. Appl.* 137, 304–315.

van Heerwaarden, B., Lee, R. F. H., Overgaard, J., and Sgro, C. M. (2014). No patterns in thermal plasticity along a latitudinal gradient in *Drosophila simulans* from eastern Australia. *J. Evol. Biol.* 27, 2541–2553. doi: 10.1111/jeb. 12510

Williams, B. R., van Heerwaarden, B., Dowling, D. K., and Sgro, C. M. (2012). A multivariate test of evolutionary constraints for thermal tolerance in *Drosophila melanogaster*. *J. Evol. Biol.* 25, 1415–1426. doi: 10.1111/j.1420-9101.2012.02 536.x

Williams, S. E., Shoo, L. P., Isaac, J. L., Hoffmann, A. A., and Langham, G. (2008). Towards an integrated framework for assessing the vulnerability of species to climate change. *PLoS Biol.* 6:2621. doi: 10.1371/journal.pbio.00 60325

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 05 November 2020
doi: 10.3389/fgene.2020.568223

Taxonomic Novelty and Distinctive Genomic Features of Hot Spring Cyanobacteria

*Jaime Alcorta[1], Tomás Alarcón-Schumacher[1,2], Oscar Salgado[1,2,3] and Beatriz Díez[1,4]**

[1] Department of Molecular Genetics and Microbiology, Biological Sciences Faculty, Pontifical Catholic University of Chile, Santiago, Chile, [2] Max Planck Institute for Marine Microbiology, Bremen, Germany, [3] Laboratorio de Bioinformática, Facultad de Educación, Universidad Adventista de Chile, Chillán, Chile, [4] Center for Climate and Resilience Research (CR)2, University of Chile, Santiago, Chile

Several cyanobacterial species are dominant primary producers in hot spring microbial mats. To date, hot spring cyanobacterial taxonomy, as well as the evolution of their genomic adaptations to high temperatures, are poorly understood, with genomic information currently available for only a few dominant genera, including *Fischerella* and *Synechococcus*. To address this knowledge gap, the present study expands the genomic landscape of hot spring cyanobacteria and traces the phylum-wide genomic consequences of evolution in high temperature environments. From 21 globally distributed hot spring metagenomes, with temperatures between 32 and 75°C, 57 medium- and high-quality cyanobacterial metagenome-assembled genomes were recovered, representing taxonomic novelty for 1 order, 3 families, 15 genera and 36 species. Comparative genomics of 93 hot spring genomes (including the 57 metagenome-assembled genomes) and 66 non-thermal genomes, showed that the former have smaller genomes and a higher GC content, as well as shorter proteins that are more hydrophilic and basic, when compared to the non-thermal genomes. Additionally, the core accessory orthogroups from the hot spring genomes of some genera had a greater abundance of functional categories, such as inorganic ion metabolism, translation and post-translational modifications. Moreover, hot spring genomes showed increased abundances of inorganic ion transport and amino acid metabolism, as well as less replication and transcription functions in the protein coding sequences. Furthermore, they showed a higher dependence on the CRISPR-Cas defense system against exogenous nucleic acids, and a reduction in secondary metabolism biosynthetic gene clusters. This suggests differences in the cyanobacterial response to environment-specific microbial communities. This phylum-wide study provides new insights into cyanobacterial genomic adaptations to a specific niche where they are dominant, which could be essential to trace bacterial evolution pathways in a warmer world, such as the current global warming scenario.

Keywords: cyanobacteria, hot springs, metagenomes, thermophiles, taxonomy, MAGs

OPEN ACCESS

Edited by:
Pedro Simões, University of Lisbon, Portugal

Reviewed by:
Francisco (Paco) Barona-Gomez, National Polytechnic Institute of Mexico (CINVESTAV), Mexico
David William Waite, Ministry for Primary Industries, New Zealand
Felipe Hernandes Coutinho, Miguel Hernández University of Elche, Spain

***Correspondence:**
Beatriz Díez
bdiez@bio.puc.cl

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *31 May 2020*
Accepted: *15 October 2020*
Published: *05 November 2020*

Citation:
Alcorta J, Alarcón-Schumacher T, Salgado O and Diez B (2020) Taxonomic Novelty and Distinctive Genomic Features of Hot Spring Cyanobacteria. Front. Genet. 11:568223. doi: 10.3389/fgene.2020.568223

INTRODUCTION

Cyanobacteria are photosynthetic microorganisms that shaped the earth's atmosphere during the Great Oxidation Event 2.6 billion years ago (Castenholz et al., 2001; Schirrmeister et al., 2015). They are morphologically diverse and thrive in most environments exposed to light, such as the ocean, lakes, soils, deserts and hot springs (Castenholz et al., 2001). Indeed, they are members of microbial mat communities from non-acidic hot springs, leading primary production and nitrogen fixation (Castenholz, 1969; Ward et al., 1998; Alcamán et al., 2015). Within these thermal microbial mats, which are found across all continents (Castenholz, 1969; Ionescu et al., 2010), the temperature boundary for cyanobacterial survival is approximately 73°C (Cox et al., 2011). In these environments, the temperature gradient shapes the microbial community (Klatt et al., 2011; Mackenzie et al., 2013), with some thermophilic cyanobacteria living at higher temperatures, and thermotolerant or mesophilic members living in the hot spring outflow or borders (Finsinger et al., 2008; Ward et al., 2012).

Over the last two centuries, several hot spring cyanobacterial species have been described (for descriptions, see Komárek, 1999, 2013; Komárek and Anagnostidis, 2005; Ward et al., 2012). Later, phylogenetic analyses revealed a wide biogeographical distribution of unicellular and filamentous thermal members (Papke et al., 2003; Miller et al., 2007; Ionescu et al., 2010; Sciuto and Moro, 2016), demonstrating that the ability to survive at high temperatures is polyphyletic within the phylum (Sanchez-Baracaldo et al., 2005; Uyeda et al., 2016). Nevertheless, the taxonomic information remains incomplete because most known thermophilic strains have only been assigned at the family or genus level, with newly obtained environmental 16S rRNA gene sequences classified based solely on these references (Ionescu et al., 2010; Mackenzie et al., 2013). This, together with the currently persisting conflict in cyanobacterial taxonomy and nomenclature (see Oren and Ventura, 2017, and references therein), prevents the high-resolution description of new cyanobacterial species and genera from environmental samples.

Three major morphological groups of early-described thermophilic cyanobacteria (Schwabe, 1837; Copeland, 1936) have been widely studied. The first group is the unicellular cyanobacteria that diverged and now specializes along the temperature gradient and vertical layers of hot spring microbial mats (Ward et al., 1998, 2006; Olsen et al., 2015). They survive up to the oxygenic photosynthesis temperature limit (Meeks and Castenholz, 1971; Cox et al., 2011), and are most represented by the genera *Synechococcus* and *Thermosynechococcus*, which comprise two very deep branches near the base of the phylum Cyanobacteria (Shih et al., 2013; Dagan et al., 2013). The second group is represented by filamentous non-heterocystous cyanobacteria, which are morphologically diverse as the *Spirulina*, *Leptolyngbya* and *Phormidium* genera (see Copeland, 1936). Some members have the potential to perform both oxygenic and anoxygenic photosynthesis (Momper et al., 2019). Furthermore, they also are globally distributed and dominant in hot springs (Sciuto and Moro, 2016; Yoon et al., 2017). The third and most studied hot spring morphological group is the filamentous heterocystous cyanobacteria, which includes the true-branching species *Fischerella thermalis*. This group has become a high-temperature model for different research topics, such as photosynthesis, nitrogen fixation, multicellularity and biogeography (Alcorta et al. (2019), and references therein).

Genomes from various hot spring strains have shown diverse adaptations to the thermal environment, such as different strategies for phosphate and nitrogen uptake; light responses at different temperatures and depths; and heterocyst envelope composition (Bhaya et al., 2007; Olsen et al., 2015; Alcorta et al., 2018; Sano et al., 2018). Genome reduction is a main evolutionary trend related to the hot spring environment, such as that observed for *Synechococcus* sp. OS-A and OS-B', and *Thermosynechococcus* (Larsson et al., 2011). The negative correlation of genome size and protein length with increasing temperature is also a major evolutionary trend related to thermophilic bacteria (Sabath et al., 2013), as is differentiated nucleotide content, codon usage, amino acid composition (Vieille and Zeikus, 2001; Singer and Hickey, 2003; Zeldovich et al., 2007) and the prevalence of the CRISPR-Cas system (Weinberger et al., 2012). However, cyanobacterial genomes are not well represented in these studies; thus, the genomic features and functional categories that differentiate thermophilic cyanobacteria are still unknown.

To obtain new information on hot spring cyanobacterial genomic features, metagenomic reconstruction of complete or partial genomes, known as metagenome-assembled genomes (MAGs), was used to uncover an unprecedent diversity. From 21 globally distributed neutral-pH hot spring microbial mats, with temperatures between 32 and 75°C, 57 new cyanobacterial MAGs were obtained, some of which were classified into well-known genera and species, while others represent new taxa at the order to species levels. Comparative genomics corroborates thermophilic features prevalent in other bacteria, and also reveals new trends related to exclusive orthologs, abundances of protein functional categories and adaptative genes involved in the response to the microbial and viral hot spring community. These results highlight various consequences of the ecological speciation process on thermophilic cyanobacterial genomes. However, more studies are now required to reveal the initial colonization mechanism of these organisms to this extreme habitat.

MATERIALS AND METHODS

Study Sites and Sample Collection

Phototrophic microbial mat samples were taken from El Tatio geyser field (Atacama, Chile), Cahuelmó hot spring (Northern Patagonia, Chile) and Kroner Lake (Deception Island, Antarctica). The sample location, collection year and physicochemical parameters are listed in **Supplementary Table 1**. Temperature and pH were determined using a multiparameter instrument (model 35607-85; Oakton, Des Plaines, IL, United Stattes). For molecular analysis, 2 cm core samples were collected and frozen at −80°C until subsequent procedures. DNA was isolated using a bead-beating protocol with xanthogenate lysis buffer and a phenol–chloroform extraction, according to Alcorta et al. (2018). The quality and quantity of nucleic

acids were checked using the Qubit (LifeTechnologies, Carlsbad, CA, United States) and Nanodrop (Thermo Fisher Scientific, Waltham, MA, United States) systems.

Sequencing and Read Quality Assessment

For metagenomic analysis, DNA samples were sequenced on the Illumina HiSeq platform (Illumina, San Diego, CA, United States) at the Research and Testing Laboratory (Lubbock, TX, United States). Briefly, the DNA was fragmented using the NEBNext dsDNA Fragmentase kit (New England Biolabs, Ipswich, MA, United States), followed by DNA clean up via column purification and library construction with the NEBUltra DNA Library Prep kit for Illumina (New England Biolabs). Methodology from Guajardo-Leiva et al. (2018) was followed for quality filtering using Cutadapt (Martin, 2011) with the parameters: paired-end mode, a perfect match of at least 10 bp (-O 10) against the standard Illumina adaptors, hard clipping of the first five leftmost bases (-u 5), 3′ end trimming for bases with a quality score below 28 (-q 28) and retaining only sequences longer than 30 bp (-m 30).

Assembly and Metagenomic Binning

De novo assemblies of trimmed reads were generated using SPAdes v3.10.1 (Bankevich et al., 2012) with the –meta option. Contigs longer than 1,000 bp were grouped into MAGs using the metaWRAP binning module (Uritskiy et al., 2018), which incorporates the following three binning software: metaBAT 2 v2.12.1 (Kang et al., 2019), MaxBin2 (Wu et al., 2016) and CONCOCT v1.1.0 (Alneberg et al., 2014) with default parameters. Next, the bin_refinement module of metaWRAP (Uritskiy et al., 2018) was used to consolidate results from the three methods using the -c 50 and -x 10 options to obtain bins with over 50% completeness and less than 10% contamination according to the CheckM tool v1.0.18 (Parks et al., 2015), which also uses HMMER v3.2.1[1] (Eddy, 1998) as a third-party software. The refineM tool (Parks et al., 2017) was used to clean potential contig contamination with different genomic properties (tetranucleotide signature and coverage) via the scaffold_stats, outliers and filter_bin modules. The refineM tool (Parks et al., 2017) was also used to clean potential contamination based on taxonomic assignment with the following modules: call_genes, which uses Prodigal v.2.6.3 (Hyatt et al., 2010); taxon_profile, using the GTDB R80 custom protein database from the Genome Taxonomy Database (GTDB; Parks et al., 2018; available at[2]); and taxon_filter and ssu_erroneous, using the GTDB R80 custom SSU database[2]. The obtained MAGs were reassessed with the CheckM tool v1.0.18 (Parks et al., 2015), their tRNAs were predicted with the ARAGORN webserver (Laslett and Canback, 2004), and their ribosomal subunit sequences were searched with Barrnap v0.9[3]. According to the Genomic Standards Consortium parameters, this information allowed us to classify the MAGs as high-, medium- or low-quality (Bowers et al., 2017). Taxonomic assignment was performed using GTDB-tk v0.3.2 software with version R89 (Chaumeil et al., 2020), which also uses pplacer as a third-party software (Matsen et al., 2010). Genomes belonging to the phylum Cyanobacteria were used for further analyses. The similarity between MAGs was assessed through all-vs-all comparisons of the average nucleotide identity (ANI) using fastANI software (Jain et al., 2018) and the average amino acid identity (AAI) using compareM software[4]. MAGs identified as possible new taxa were grouped according to thresholds of similarity for the ANI and AAI values stated by Konstantinidis et al. (2017).

Additionally, a locally built database of public metagenomes from hot springs was analyzed. SRA files were first downloaded from the NCBI database, and subsequently quality trimmed, assembled, binned and taxonomically classified as described above. The only procedural difference was that the SRR5581334, SRR7905023 and ERR372908 metagenomes were assembled using MEGAHIT 1.2.9 (Li et al., 2015), due to memory requirements. Details of these metagenomes are listed in **Supplementary Table 1**. High- and medium-quality MAGs assigned to the phylum Cyanobacteria were used for further analyses. All obtained cyanobacterial MAGs were submitted in FASTA format to the Figshare repository[5] under DOI: 10.6084/m9.figshare.12400979. Meanwhile, 34 high- and medium-quality MAGs with > 95% completeness were deposited under NCBI BioProject numbers PRJNA635751 and PRJNA645256. The MAGs obtained using primary data from this study and from Alcamán-Arias et al. (2018) were submitted to the NCBI WGS database, while those obtained from Chan et al. (2015), Kaushal et al. (2018), Ward et al. (2019) and Roy et al. (2020) were submitted to the Third Party Annotation database[6].

Abundance of MAGs in Metagenomes

The abundance of the recovered MAGs was assessed through read mapping. Briefly, quality trimmed reads of each sample were mapped using BBMap v38.71[7] with a minimum identity of 99% (minid = 0.99 and idfilter = 0.97). A MAG was considered present in a sample when it had a coverage > 1x across 75% of its genome; otherwise, the abundance was considered zero. Absolute read counts of selected MAGs were normalized as the number of reads recruited per kilobase of MAG and gigabase of metagenome (RPKG), which allowed the direct comparison of genome abundances between metagenomes of different depths (Reji and Francis, 2020). Normalized read counts were used to calculate diversity metrics, as well as the similarity matrix for multivariate analyses with the "Vegan" package in R.

Phylogenomics and Taxonomy

All genomes classified as Cyanobacteria (1,626 genomes as of September 2019) were downloaded from the NCBI database, and the GTDB taxonomy was assigned to each

[1]http://hmmer.org/

[2]https://data.ace.uq.edu.au/public/misc_downloads/refinem/

[3]https://github.com/tseemann/barrnap

[4]https://github.com/dparks1134/CompareM

[5]https://knowledge.figshare.com/

[6]http://www.insdc.org/tpa.html

[7]http://bbtools.jgi.doe.gov

using GTDB-tk software (Chaumeil et al., 2020). The dRep v2.3.2 software (Olm et al., 2017) was used to dereplicate the entire set due to overrepresented species and low-quality genomes (e.g., only 1136 exhibited > 75% completeness), thus obtaining a final subset of 800 genomes (**Supplementary Table 2**). A total of 120 concatenated single-copy bacterial genes were recovered from the intermediate files of the GTDB-tk analysis (gtdbtk.bac120.msa.fasta files), and their sequences were subsequently aligned using MUSCLE v6.0.98 software (Edgar, 2004). Maximum-likelihood trees were generated using IQtree v.1.5.5 software with the TESTNEW option to choose the best substitution models, after which a non-parametric ultrafast bootstrap (-bb) support of 10,000 replicates was applied (Hoang et al., 2017). Node collapse and rooting of phylogenetic reconstructions were managed using the iTOL web server (Letunic and Bork, 2019).

Comparative Genomics

Genomic features of the 57 MAGs obtained in the present study, as well as for the 800 cyanobacterial genomes from the NCBI database, were extracted from the CheckM summary results (see above, **Supplementary Table 2**). As the MAGs and NCBI genomes have different completeness levels, the expected genome size was calculated as EGS = (genome size * 100)/(completeness). In conjunction with the MAGs, a subset of NCBI genomes that were taxonomically close to the MAGs (according to GTDB-tk), was used to compare the genome size, GC percentage and coding density between hot spring cyanobacteria (all 57 MAGs and 36 NCBI genomes) and non-thermal cyanobacteria (66 NCBI genomes). Together, these 159 genomes are hereafter referred to as the 159-subset. Additionally, the hydrophobicity, protein length, protein molecular weight, isoelectric point (pI) and amino acid usage were compared for the coding DNA sequences (CDSs) calculated with ProPAS v1.03 software (Wu and Zhu, 2012). Furthermore, genomes from both environmental groups were compared for the Thermosynechococcaceae, *Elainellaceae* and Oscillatoriaceae families, and for the *Fischerella* and *Geminocystis* genera ($\geq$ 3 genomes for each environment) to identify differences at these specific taxonomic levels. The isolation environments and associated references for each cyanobacterial genome of the 159-subset are listed in **Supplementary Table 3**.

Additionally, the 159-subset orthogroups (orthologous gene clusters, see **Supplementary Table 4**) obtained with Orthofinder v2.3.3 software (Emms and Kelly, 2015), were used to identify the core and accessory orthogroups. Orthogroups present in > 97% of the subset (153 genomes) were considered the core genome (based on their distribution in the total 47,328 obtained orthogroups). Orthogroups sporadically present in both environmental groups were considered the phylum accessory set, while those exclusive to hot spring or non-thermal genomes were considered as the specific accessory orthogroups. This includes the singletons as well as the specific core accessory orthogroups if they have presence in all genomes an environmental group. This same classification was also performed for seven genera with genomes from both environmental groups (*Calothrix*, *Cyanobacterium*, *Elainella*, *Fischerella*, *Geminocystis*, *Rivularia* and *Trichormus*) and a heatmap of the relative abundance of COG categories was created with R software to make a hierarchical clustering (hclust) of the core and accessory categories. Orthogroups distributed in more than one order, family or genus in each environmental group, were further explored by searching for homologous sequences in the NCBI non-redundant protein database using BLASTP (Altschul et al., 1990). From the retrieved sequences, phylogenetic reconstructions were generated as explained above.

Functional annotation of the 159-subset genomes was done with eggnog-mapper software (Huerta-Cepas et al., 2017, 2019). First, the genomic percentage of COG categories in each genome was compared between hot spring and non-thermal cyanobacteria using STAMP v2.1.3 software (Parks et al., 2014) via the Welch's *t*-test with Bonferroni multiple test correction. Next, specific annotations of complementary metabolism, defense systems and secondary metabolites were searched. Because many proteins are annotated with putative functions or as hypothetical proteins, the following specific terms were used to search within the annotated 159-subset: "nitrogenase," "nitrate," "nitrite," "restriction," "modification," "capsid," "dehydrogenase," "CRISPR," "virus," "viral" and "capsid." Orthogroups harboring protein sequences whose annotations indeed corresponded to these functions were then used for further analyses. Some orthogroups with ambiguous functions (different annotations within the orthogroup) were ignored. For the sulfide-quinone reductase protein, which is characterized as a possible switch between anoxygenic and oxygenic photosynthesis, BLASTP was used with the reference NCBI sequence (AF242368.1 or WP_071516517.1) to search the corresponding orthogroup. Secondary metabolite biosynthetic gene clusters (BGCs) were searched with antiSMASH v4.2.0 software (Blin et al., 2017) using the clusterblast and borderpredict options.

Statistical Analyses

Correlation analyses between MAGs abundances and environmental parameters were calculated with the Mantel test implemented in the R package "Vegan," only considering the variables temperature, pH and geographical location. Additional parameters that were not available for all samples, were excluded from the analyses. Wilcoxon's paired test was used because the comparison of genome features was not balanced between hot spring and non-thermal genome groups. Correlation analyses were performed by determining the adjustment to the "normal" distribution of each variable with the Kolmogorov-Smirnov test and by using parametric Pearson's or non-parametric Spearman's tests, depending on the distribution of data. Correlation indexes were compared using the r-to-z Fischer transformation. For multiple analyses over the same dataset, Bonferroni and FDR corrections were applied to the obtained p-values. The R packages "Tidyverse," "ggpubr," "ggplot2," "cocor," and "maps" (R Core Team, 2017) were used for all statistical analyses and corrections, as well as most of the plots.

RESULTS AND DISCUSSION

Hot Spring Cyanobacterial MAG Recovery

In this study, four new metagenomes from hot spring phototrophic microbial mats in Chile and Antarctica were analyzed along with 79 metagenomes already available in the NCBI database. The four new metagenomes altogether added up to 60.75 Gbp of total trimmed reads and 292,512 assembled contigs (> 1,000 bp), totaling 1.36 Gbp. For the locally built database, cyanobacterial sequences were identified in only 17 of the 79 hot spring metagenomes, which then were used for further analyses. Altogether these 21 cyanobacteria-containing metagenomes were distributed in North and South America, Asia and Antarctica (**Figure 1**), representing a temperature range between 32 and 75°C, and a pH range of 5.8 to 9.2. These comprised 359.8 Gbp of trimmed reads and 5.19 Gbp of assembled contigs (representing 1,393,425 contigs > 1,000 bp). A total of 1,152 medium- or high-quality MAGs were recovered from these 21 metagenomes. According to the GTDB-tk taxonomic assignment, most of these MAGs belong to the phyla Proteobacteria (16.1%), Chloroflexota (13.8%), Bacteroidota (12.9%), Planctomycetota (6.3%) and Cyanobacteria (4.9%) (**Supplementary Table 5**).

A total of 20 cyanobacterial MAGs (phylum Cyanobacteria GTDB R89) were recovered from the four new metagenomes reported in the present study. Overall, 57 cyanobacterial MAGs with an average completeness of 88.5% (SD ± 14.7), were obtained from the entire set of 21 hot spring metagenomes. The average contamination of these MAGs was 1.2% (SD ± 1.1), with 4.8% being the highest value (**Table 1**). According to the GSC quality parameters (Bowers et al., 2017), only two MAGs were classified as high-quality (M3746_SRR7905025_W2019_013 and M46_SRR2626160_R2017_013); while 55 MAGs were medium-quality, from which 36 could be categorized as high-quality if their rRNA operon sequences were binned.

Several strategies have been used to recover partial genomes of hot spring cyanobacteria from metagenomes, including mapping contigs against reference genomes (Bhaya et al., 2007) and the use of binning tools (Thiel et al., 2017; Alcorta et al., 2018; Ward et al., 2019). Metagenomes from the binning-based studies were reanalyzed here, allowing for better genome recovery in terms of contig numbers and completeness. Furthermore, the ssu_erroneous module from refineM enabled removal of partial 16S rRNA gene sequences with incongruent taxonomy from other phyla. Reanalysis of previously published data is useful for maintaining confidence in public repositories of genomes (Shaiber and Eren, 2019) and to obtain genomes that fulfill acceptable standards (Bowers et al., 2017).

Hot Spring Cyanobacterial Taxonomy

Cyanobacterial taxonomy and nomenclature have always been controversial (see Oren and Ventura, 2017, and references therein). Nowadays, applying a valid taxonomy based on

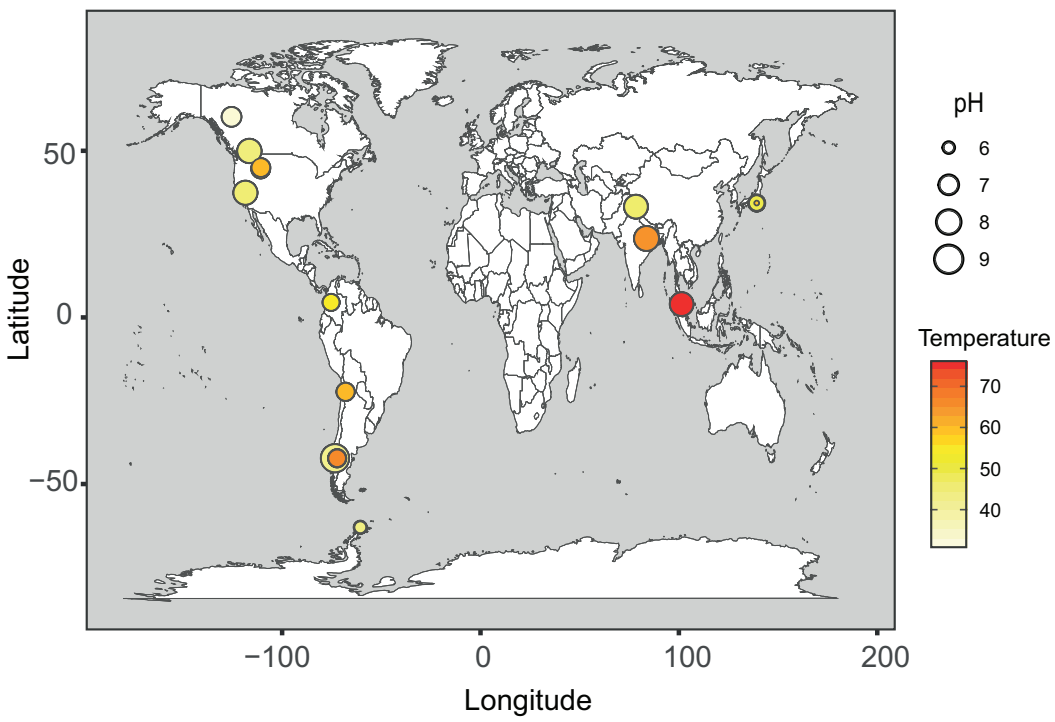


FIGURE 1 | Geographical distribution of phototrophic hot spring metagenomes. Distribution map of the 21 hot spring microbial mat metagenomes used in the present study. Each circle represents one sample. The color scale for circles represents the temperature range from 32 to 75°C, and the circle size represents the pH range (from 5.8 to 9.2) of each sample.

TABLE 1 | Genome features of hot spring cyanobacterial MAGs.

Genome ID	GTDB Classification	NCBI WGS Accession Number	Completeness (%)	Contamination (%)	Genome size (bp)	# Predicted CDS	Ribosomal RNAs
T60_TAT2020_004	c_Sericytochromatia;o_UBA7694;f_;g_;s_	JACYMC000000000	83.3	2.1	4167372	3865	nf
M55_SRR7473442_K2018_030	o_;f_;g_;s_	-	76.3	1.2	3104692	3433	nf
C42_CAH2020_026	o_Cyanobacteriales;f_Chroococcidiopsidaceae; g_Chroogloeocystis;s_Chroogloeocystis siderophila	-	87.8	0.7	4073818	4184	nf
T60_TAT2020_053	o_Cyanobacteriales;f_Cyanobacteriaceae; g_Cyanobacterium;s_	JACYMF000000000	99.1	0.0	3360103	3188	nf
M7585_ERR372908_C2015_104	o_Cyanobacteriales;f_Cyanobacteriaceae;g_Geminocystis;s_	DVEF00000000	98.7	0.2	2823523	2645	p
M4454_SRR7905024_W2019_049	o_Cyanobacteriales;f_Geitlerinemaceae;g_1;s_1	DVDY00000000	99.6	0.7	5103222	4498	nf
M59_SRR7905023_W2019_021	o_Cyanobacteriales;f_Geitlerinemaceae;g_1;s_1	DVED00000000	98.0	0.4	4985496	4422	p
M55_SRR7473442_K2018_032	o_Cyanobacteriales;f_Microcystaceae;g_;s_	-	60.0	1.1	2298071	3211	p
C42_CAH2020_068	o_Cyanobacteriales;f_Microcystaceae;g_Hydrococcus; s_Hydrococcus minor	JACYLS000000000	99.3	0.2	4917303	4538	nf
C42_CAH2020_038	o_Cyanobacteriales;f_Nostocaceae;g_Calothrix;s_	JACYLQ000000000	99.8	0.5	6532080	5465	nf
C42_CAH2020_084	o_Cyanobacteriales;f_Nostocaceae;g_Chlorogloeopsis; s_Chlorogloeopsis fritschii	JACYLT000000000	99.3	1.1	7052188	6270	nf
M7585_ERR372908_C2015_036	o_Cyanobacteriales;f_Nostocaceae;g_Fischerella;s_	-	80.2	2.4	4599829	4836	nf
C42_CAH2020_099	o_Cyanobacteriales;f_Nostocaceae;g_Fischerella;s_Fischerella thermalis	-	51.4	3.7	3171434	3795	nf
M33_SRR5581334_DOE_064	o_Cyanobacteriales;f_Nostocaceae;g_Fischerella;s_Fischerella thermalis	-	76.7	1.2	4199934	4157	p
M48_SRR5451033_A2018_028	o_Cyanobacteriales;f_Nostocaceae;g_Fischerella;s_Fischerella thermalis	JACYLX000000000	98.6	0.2	5141023	4455	nf
M58_SRR5451032_A2018_009	o_Cyanobacteriales;f_Nostocaceae;g_Fischerella;s_Fischerella thermalis	JACYLY000000000	97.5	0.2	5203860	4497	nf
M66_SRR5451031_A2018_004	o_Cyanobacteriales;f_Nostocaceae;g_Fischerella;s_Fischerella thermalis	JACYMA000000000	98.8	0.0	5271540	4589	nf
T60_TAT2020_040	o_Cyanobacteriales;f_Nostocaceae;g_Rivularia;s_	JACYMD000000000	99.1	0.7	6018129	5143	p
M33_SRR5581334_DOE_039	o_Cyanobacteriales;f_Nostocaceae;g_Trichormus;s_	DVDT00000000	99.3	0.4	7018248	5764	nf
M33_SRR5581334_DOE_052	o_Cyanobacteriales;f_Oscillatoriaceae;g_1;s_1	DVDU00000000	99.3	0.8	6239312	5233	p
M59_SRR7905023_W2019_015	o_Cyanobacteriales;f_Oscillatoriaceae;g_2;s_1	-	98.6	1.4	5787682	4574	nf
M7585_ERR372908_C2015_266	o_Cyanobacteriales;f_Oscillatoriaceae;g_3;s_1	DVEG00000000	99.6	0.0	3979146	3395	p
K32_KRO2020_035	o_Elainellales;f_1;g_1;s_1	JACYLV000000000	98.6	1.1	4371702	3742	nf
K44_KRO2020_017	o_Elainellales;f_1;g_1;s_1	JACYLW000000000	99.1	1.1	4432392	3767	nf
M4564_SRR6941191_B2018_003	o_Elainellales;f_1;g_1;s_2	-	54.1	2.3	2340930	3827	p
T60_TAT2020_003	o_Elainellales;f_1;g_1;s_3	JACYMB000000000	95.2	0.8	4020528	3982	nf
C42_CAH2020_086	o_Elainellales;f_Elainellaceae;g_1;s_1	JACYLU000000000	99.5	1.2	5928601	5059	nf
M58_SRR5451032_A2018_015	o_Elainellales;f_Elainellaceae;g_1;s_1	JACYLZ000000000	98.9	1.9	5813440	5026	nf
M66_SRR5451031_A2018_013	o_Elainellales;f_Elainellaceae;g_1;s_1	-	66.7	1.4	3168891	3064	nf

(Continued)

TABLE 1 | Continued

Genome ID	GTDB Classification	NCBI WGS Accession Number	Completeness (%)	Contamination (%)	Genome size (bp)	# Predicted CDS	Ribosomal RNAs
M55_SRR7473442_K2018_004	o_Elainellales;f_Elainellaceae;g_2;s_1	DVEB00000000	97.9	1.4	5013995	4164	nf
M4564_SRR6941191_B2018_001	o_Elainellales;f_Elainellaceae;g_CCP2;s_	-	87.6	2.7	6826097	8104	c
C42_CAH2020_014	o_Elainellales;f_Elainellaceae;g_Elainella;s_1	-	92.7	0.9	6368923	5661	nf
C42_CAH2020_052	o_Elainellales;f_Elainellaceae;g_Elainella;s_2	-	82.6	1.3	6281096	5592	nf
M66_SRR5451031_A2018_017	o_Elainellales;f_Elainellaceae;g_Elainella;s_2	-	76.4	4.8	6099463	5906	nf
C42_CAH2020_010	o_Elainellales;f_Elainellaceae;g_Elainella;s_Elainella sp000733415	JACYLO000000000	99.5	1.0	7459088	6354	p
M55_SRR7473442_K2018_002	o_Elainellales;f_Elainellaceae;g_O-77;s_	DVEA00000000	97.6	1.2	5022360	4498	nf
C42_CAH2020_037	o_Elainellales;f_Elainellaceae;g_O-77;s_O-77 sp001548395	JACYLP000000000	99.1	1.2	5291374	4455	nf
M33_SRR5581334_DOE_111	o_Gloeobacterales;f_Gloeobacteraceae;g_;s_	-	57.4	0.0	3328352	4052	nf
C42_CAH2020_066	o_Gloeomargaritales;f_Gloeomargaritaceae;g_;s_	JACYLR000000000	99.1	2.6	2730382	2813	nf
M33_SRR5581334_DOE_097	o_Leptolyngbyales;f_Leptolyngbyaceae;g_;s_	DVDV00000000	99.5	0.7	6960501	6277	nf
T60_TAT2020_044	o_Leptolyngbyales;f_Leptolyngbyaceae;g_Alkalinema;s_	-	86.3	2.0	3677935	4081	nf
C42_CAH2020_001	o_Leptolyngbyales;f_Leptolyngbyaceae;g_JSC-12;s_1	JACYLN000000000	99.5	0.0	5550306	5066	nf
M33_SRR5581334_DOE_055	o_Leptolyngbyales;f_Leptolyngbyaceae;g_JSC-12;s_2	-	68.1	4.8	4528759	5838	nf
M65_SRR7473443_K2018_014	o_Leptolyngbyales;f_Leptolyngbyaceae;g_JSC-12;s_3	-	62.3	2.2	3370404	4476	nf
M44_SRR5580903_DOE_062	o_PCC-7336;f_JA-3-3Ab;g_JA-3-3Ab;s_1	DVDX00000000	100.0	0.0	2925579	2737	nf
M46_SRR5216251_DOE_021	o_PCC-7336;f_JA-3-3Ab;g_JA-3-3Ab;s_2	-	71.2	1.9	1764837	2077	nf
M60_SRR5451356_T2016_033	o_PCC-7336;f_JA-3-3Ab;g_JA-3-3Ab;s_3	-	56.4	1.4	1200192	1708	nf
M60_SRR5248366_T2016_029	o_PCC-7336;f_JA-3-3Ab;g_JA-3-3Ab;s_JA-3-3Ab sp000013205	-	77.6	0.9	1957034	2127	nf
M60_SRR5451356_T2016_037	o_PCC-7336;f_JA-3-3Ab;g_JA-3-3Ab;s_JA-3-3Ab sp000013205	-	92.1	0.9	2379147	2435	nf
M60_SRR5248366_T2016_047	o_PCC-7336;f_JA-3-3Ab;g_JA-3-3Ab;s_JA-3-3Ab sp000013225	-	91.8	0.4	2358112	2481	nf
M60_SRR5451356_T2016_004	o_PCC-7336;f_JA-3-3Ab;g_JA-3-3Ab;s_JA-3-3Ab sp000013225	-	88.6	1.2	2201305	2331	nf
M65_SRR7473443_K2018_010	o_Phormidesmiales;f_Phormidesmiaceae;g_1;s_1	DVEE00000000	95.4	3.0	3894922	3941	nf
T60_TAT2020_046	o_Phormidesmiales;f_Phormidesmiaceae;g_2;s_1	JACYME000000000	97.8	0.5	4202921	3943	nf
M46_SRR2625865_R2017_001	o_Thermosynechococcales;f_Thermosynechococcaceae; g_Thermosynechococcus;s_1	-	63.1	1.4	1882589	3438	p
M46_SRR2626160_R2017_013	o_Thermosynechococcales;f_Thermosynechococcaceae; g_Thermosynechococcus;s_1	DVDZ00000000	96.9	1.1	2396365	2707	c
M55_SRR7473442_K2018_012	o_Thermosynechococcales;f_Thermosynechococcaceae; g_Thermosynechococcus;s_2	DVEC00000000	97.8	0.2	2399209	2384	p
M3746_SRR7905025_W2019_013	o_Thermosynechococcales;f_Thermosynechococcaceae; g_Thermosynechococcus;s_Thermosynechococcus sp000505665	DVDW00000000	98.1	0.0	2387168	2307	c

Features were determined with the CheckM tool (completeness, contamination, genome size and number of CDSs), barrnap (ribosomal RNAs: nf, not found; p, partial and c, complete 5S-16S-23S subunits) and GTDB-tk R89 (GTDB classification). MAGs are ordered according to the GTDB-tk classification and their NCBI WGS Accession numbers are given for the 34 MAGs submitted to the NCBI.

the genomes of uncultured organisms is impossible (e.g., the rejection of the Whitman (2016) proposal), even when following good practice strategies (Konstantinidis et al., 2017; Chun et al., 2018). This study increased the percentage of available genomes within cyanobacterial taxa that have few cultured representatives that hopefully can be cultivated in the future.

As seen in **Figure 2A**, GTDB classification distributed the new 57 MAGs into 2 classes and 10 orders within the phylum Cyanobacteria (from 13 total orders in the class Cyanobacteriia and 2 orders in the class Sericytochromatia in GTDB R89). This dataset includes one MAG (medium-quality) classified as a new order in the Cyanobacteriia class (M55_SRR7473442_K2018_030: 76.3% completeness and 1.2% contamination). Only T60_TAT2020_004 was assigned to the non-photosynthetic Sericytochromatia class. ANI values between all MAGs distinguished 44 different species (**Table 1**), while GTDB-tk classified 14 MAGs to species with available genomes in the NCBI database (gANI > 95%; 9 different species when the MAGs were dereplicated; **Table 1** and **Figure 2B**). Additionally, 21 MAGs were classified by topology or RED values as new species belonging to established GTDB genera. In summary, 43 MAGs represent taxonomic novelty as 1 new order, 3 new families, 15 new genera and 36 new species (**Table 1**).

This MAG classification is supported by a maximum likelihood phylogenomic reconstruction that incorporated 120 bacterial marker genes (**Figure 2A**) and included 800 NCBI cyanobacterial genomes (> 75% completeness; dereplicated from 1,626 genomes available in the NCBI assembly database as of September 2019). Distribution of the 57 MAGs in the GTDB R89 orders is as follows: o__ (new, 1), Cyanobacteriales

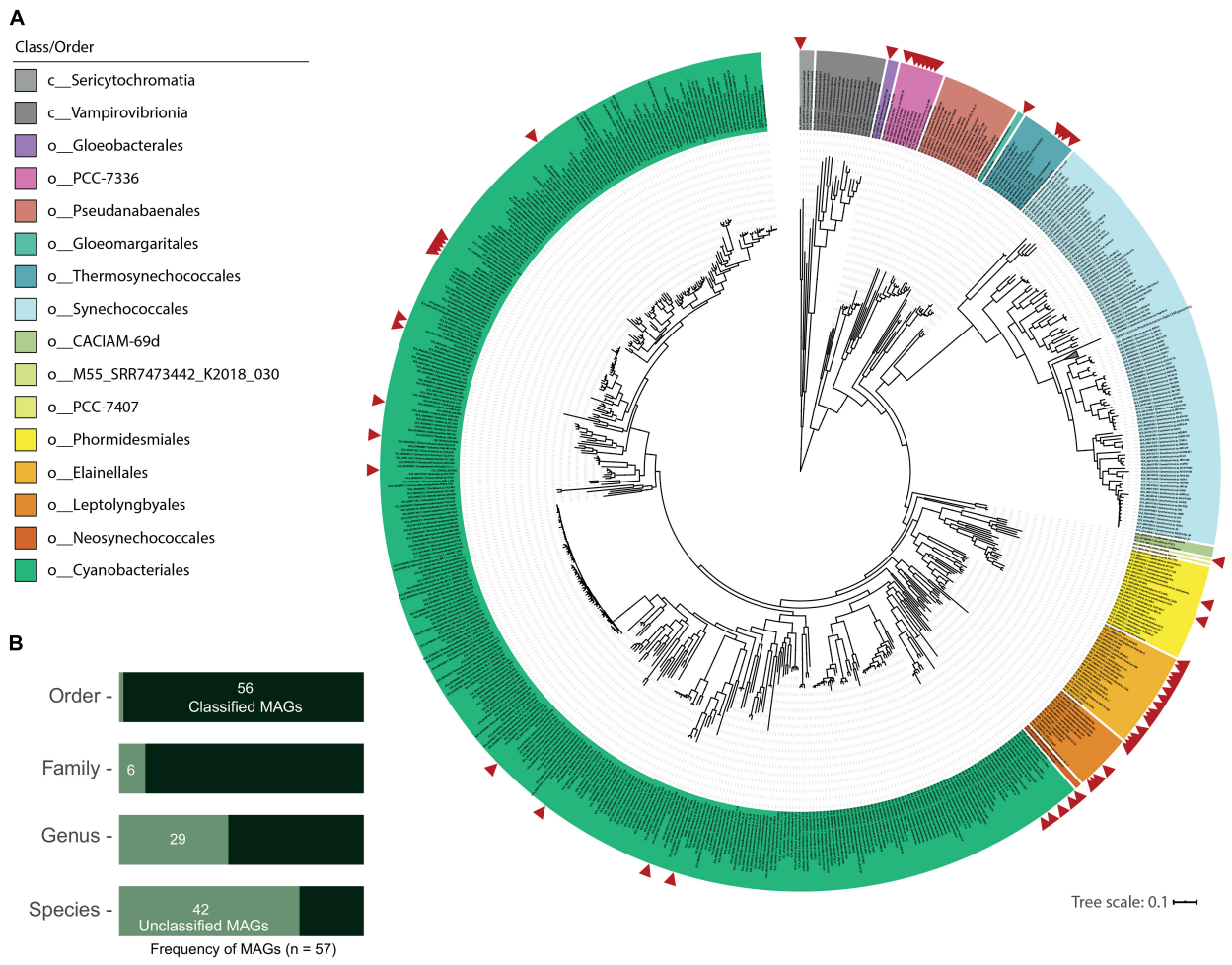


FIGURE 2 | Novel hot spring cyanobacterial MAG phylogeny and taxonomic classification. **(A)** Phylogenomic reconstruction from multiple sequence alignment of 120 bacterial marker genes recovered from 800 NCBI cyanobacterial genomes and 57 new cyanobacterial MAGs from 21 hot spring metagenomes. Maximum likelihood tree reconstruction was performed with IQtree software using the LG + R10 model and a non-parametric UF-bootstrap support of 1,000 replicates. The labels are colored according to the GTDB order classification for c_Cyanobacteriia genomes, and class for Sericytochromatia and Vampirovibrionia members. The tree was rooted in the node between Sericytochromatia/Vambirovibrionia and Cyanobacteriia classes. 305 sequences in the Syn/Pro clade are collapsed in the tree. Red arrows indicate the position of hot spring MAGs in the tree that were distributed into 10 orders from a total of 15 in the cyanobacterial phylum. The new order retrieved in the present study is called M55_SRR7473442_K2018_030 in the legend. **(B)** Classification of MAGs according to the GTDB-tk. Unclassified MAGs correspond to those with no associated genome in the database at the specific taxonomic level.

(20), Elainellales (15), Gloeobacterales (1), Gloeomargaritales (1), Leptolyngbyales (5), PCC-7336 (7), Phormidesmiales (2), Thermosynechococcales (4) and UBA7694 (1, from the Sericytochromatia class). Some of these genomes form clades with known thermophilic cyanobacterial species from the following genera: *Fischerella*, *Thermosynechococcus*, *Synechococcus* (JA-3-3Ab), *Chroogloeocystis*, *Chlorogloeopsis*, *Calothrix*, *Hydrococcus* and O-77 (*Thermoleptolyngbya*). Others represent the first hot spring-associated genome reported within the given genus, specifically *Rivularia* (3 marine genomes), *Alkalinema* (1 freshwater genome), CCP2 (1 saltern genome) and *Cyanobacterium* (3 freshwater/saline lake genomes), as seen in **Supplementary Table 3**. The latter supports the concept that the ability to live at higher temperatures is a secondary specialization from primordial environments like marine or fresh water (Hammerschmidt et al., 2020), even in early-branching cyanobacteria, such as the Thermosynechococcales and PCC7336 orders, that now include non-thermal cyanobacterial members as revealed by GTDB classification (**Supplementary Figure 1**).

Cyanobacterial clades with few genomes are complemented by the group of MAGs that comprises new taxa at the genus or family levels, according to GTDB (see above). For instance, some of these new genomes are related to the recently described Gloeomargaritales order, whose members can accumulate intracellular carbonates (Moreira et al., 2017; Ponce-Toledo et al., 2017), as well as the early diverging non-thylakoid Gloeobacterales (Nakamura et al., 2003). Furthermore, the MAG from the non-photosynthetic Sericytochromatia class (UBA7694 order) further expands the described habitat of this taxon (Monchamp et al., 2019). This MAG also shows other features that are not currently associated with UBA7694 family genomes, such as the potential for H_2 metabolism and a higher number of CRISPR-associated proteins (see below). Altogether, this data provides new insight into the diversity of thermophilic, thermotolerant ($< 40°C$) and hot spring-associated cyanobacteria, and hints at the yet-to-be-discovered taxonomic novelty in these extreme environments (e.g., the large numbers of unassigned cyanobacterial sequences from hot springs reported in Roeselers et al., 2007; Ionescu et al., 2010; Mackenzie et al., 2013; Momper et al., 2019; Uribe-Lorío et al., 2019, among others).

Although GTDB taxonomy solves several of the problems and rearrangements within the phylum, there are still unsolved issues. One of them is the use of placeholder names, such as the O-77 genus that should correspond to *Thermoleptolyngbya* (Sciuto and Moro, 2016; Yoon et al., 2017); the JSC-12 genus that has no generic assignment; and the JA-3-3Ab genus that is mostly known as *Synechococcus* OS-A and OS-B' (Steunou et al., 2006; Bhaya et al., 2007). Moreover, the present phylogenomic analyses using GTDB R89 reproduced the polyphyly within the genera *Chroogloeocystis* (Brown et al., 2005), *Geminocystis* (Korelusova et al., 2009), *Calothrix* (Berrendero et al., 2008) and *Synechococcus* (Walter et al., 2017). Further actions should now focus on selecting and correctly updating the type species and type material to improve cyanobacterial taxonomy (Ramos et al., 2017). This will enable more taxonomically driven sequencing efforts like cyanoGEBA (Shih et al., 2013), as well as the inclusion of genomic type material from uncultured taxa (Chuvochina et al., 2019), such as the MAGs recovered in this study.

Cyanobacterial MAG Abundances in Hot Spring Metagenomes

Taxonomic annotation is relevant for all cyanobacterial MAGs recovered here, but especially for those that may have a global or dominant distribution in hot springs. Accordingly, the proportion of MAG reads mapped against the entire set of metagenomes was determined and ranged from 0.1 to 60% (**Figure 3A**). The most represented orders (35 to 99% of the total cyanobacterial community) were Cyanobacteriales, Elainellales, Thermosynechococcales and PCC-7336 (**Figure 3B** and **Supplementary Figure 3**). The abundance of these MAGs in each hot spring was negatively correlated with the overall cyanobacterial alpha diversity in the respective metagenome (-0.490 rho Spearman's correlation, p-value < 0.05; **Figure 3A** and **Supplementary Figure 2A**). The latter suggests that in low diversity systems, where cyanobacteria represent a major proportion of the microbial community, only a few MAGs outcompete other cyanobacterial species and successfully establish themselves as the dominant producers in the system. Whereas, in samples where non-cyanobacterial taxa dominate the community, a broader range of niches might be available for different groups of cyanobacteria to colonize.

Interestingly, multivariate analyses, which were used to investigate the relationship between abundances, physico-chemical and geographic features, revealed that the cyanobacterial community composition was positively correlated with the hot spring location (Mantel statistic = 0.19, p-value = 9×10^{-4}), but not with physicochemical parameters, such as temperature or pH (Mantel test, p-value > 0.01). For example, MAGs from the PCC-7336 order were more abundant in North American samples, while Thermosynechococcales MAGs were found mainly in Asian samples independent of hot spring conditions (**Supplementary Figure 3**), thereby corroborating the proposed biogeographical islands for these unicellular cyanobacteria (Papke et al., 2003; Cheng et al., 2020). The correlation between community composition and geographical location was also observed through non-metric multidimensional scaling and cluster analyses (**Supplementary Figures 2B,C**). Furthermore, the cyanobacterial community composition presented a similar pattern even within a specific hot spring, independent of the temperature and pH (**Supplementary Figure 2C**).

Another interesting result is that a single clade commonly dominates the cyanobacterial community within each hot spring, where a single order can make up to 99% of all cyanobacterial sequences (**Figure 3B**). Altogether, this information notices the possibility that different hot spring cyanobacteria occupy a similar niche in non-acidic thermal environments and that competitive exclusion between different cyanobacterial clades is a significant force driving the cyanobacterial community composition, as seen in saline freshwater systems (Roney et al., 2009). Accordingly, some hot springs have zones with different

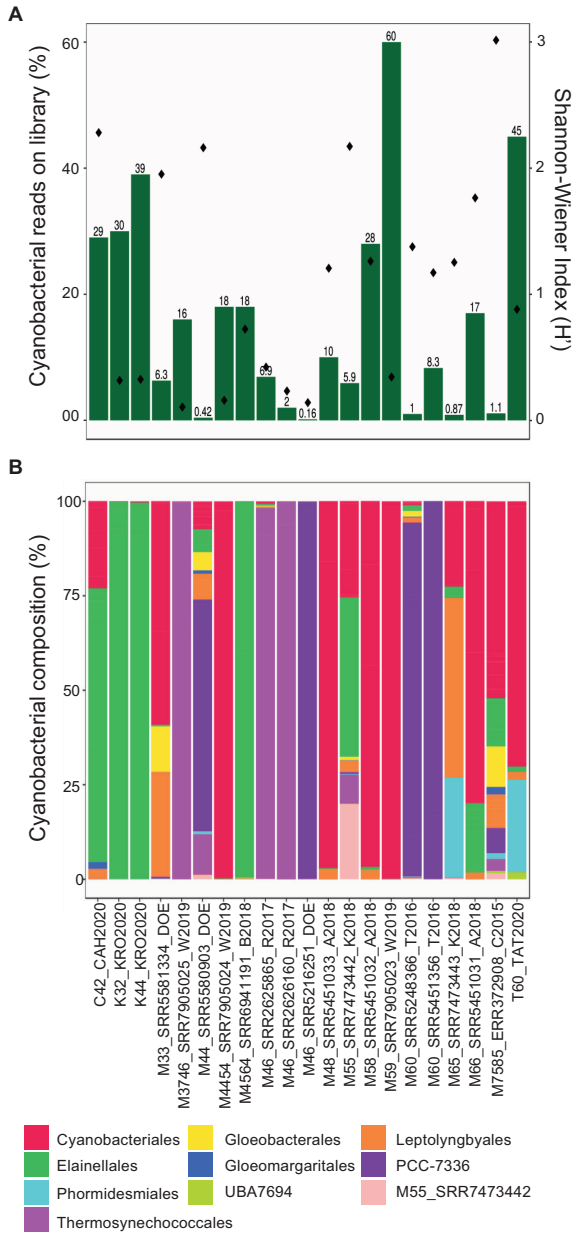


FIGURE 3 | Diversity of hot spring cyanobacterial MAGs in metagenomes. **(A)** Left *y*-axis corresponds to the percentage of total mapped reads from each hot spring to a cyanobacterial MAG (green bars). Right *y*-axis corresponds to alpha-diversity values (dots) for each metagenome. **(B)** Normalized (relative) abundances of the reads assigned to each MAG at the order level in each sample. The possible new order is shown as M55_SRR7473442.

dominant cyanobacteria, sugesting successful competitive exclusion depending on the nitrogen compound availability, sulfide concentration, temperature resistance of each clade, and light adaptations in the thermal gradient (Ward et al., 2012). However, more studies are needed to test the competitive exclusion in these different cyanobacterial communities.

Comparative Genomics: General Features of Genomes and Proteins

Large-scale comparisons have shown that there are common distinctive genomic features that differentiate thermophilic archaea and bacteria from their mesophilic counterparts; however, thermophilic cyanobacteria were underrepresented in these analyses (Singer and Hickey, 2003; Sabath et al., 2013). The cyanobacterial comparative genomics carried out in the present study, both at a general level (the 159-subset) and at specific taxonomic levels (families and genera), was useful to confirm previously detected genome trends of some cyanobacteria, such as genomic streamlining toward extreme environments (Larsson et al., 2011).

For the present study, the trend of an increasing optimum growth temperature with a decreasing genome size (Sabath et al., 2013) was also confirmed as a negative correlation between the source temperature of the metagenome and the expected genome size of the cyanobacterial MAGs (−0.355 Pearson's correlation, p-value < 0.05). Differences in genome size and other features were compared between hot spring and non-thermal cyanobacterial genomes of the 159-subset, which comprised the 57 MAGs, 36 NCBI hot spring genomes and 66 NCBI non-thermal genomes, see section "MATERIALS AND METHODS" and **Supplementary Table 3**. Cyanobacterial MAGs showed a wide variation in genome size from 2.1 to 12.1 Mbp across the 10 different orders (**Figure 4**). However, hot spring cyanobacterial genomes (n = 93) were smaller and exhibited a higher GC percentage (Wilcoxon's paired test, p-value < 0.05 for both analyses) than the non-thermal genomes (n = 66), while the coding density was similar between both groups (Wilcoxon's paired test, p-value = 0.15) (**Figure 5**).

Additionally, whether or not some of these genome changes were reflected in the divergence of cyanobacterial genera and families with $\geq$ 3 genomes in each environmental group was investigated. For the genus *Fischerella*, hot spring genomes (n = 16) were 1.6 ($\pm$ 0.7) Mbp smaller than the non-thermal genomes (n = 6) and had a higher GC content (**Supplementary Figures 4**, **5**, Wilcoxon's paired test, p-value < 0.05). Similarly, the families *Elainellaceae* and *Thermosynechococcaceae* exhibited smaller hot spring genomes (Wilcoxon's paired test, p-value < 0.05; **Supplementary Figure 4**), with an increase in both the GC content and coding density for *Thermosynechococcaceae* (Wilcoxon's paired test, p-value < 0.05; **Supplementary Figures 5**, **6**). However, the low number of genomes for other genera and families in either environmental group allowed for only some trends to be observed. *Alkalinema* and *Rivularia* genomes from hot springs were > 2 Mbp smaller than those from non-thermal environments, and a similar trend was observed for the JA-3-3Ab order. Specifically, genomes of the JA-3-3Ab family (*Synechococcus* sp. OS-A and OS-B' clade) were 2.4 ($\pm$ 0.3) Mbp smaller than the marine *Synechococcus* sp. PCC 7336 genome from the same order.

Furthermore, differences in nucleotide content and protein properties that have been seen in other bacteria (Singer and Hickey, 2003; Sabath et al., 2013) were also found here. Analysis

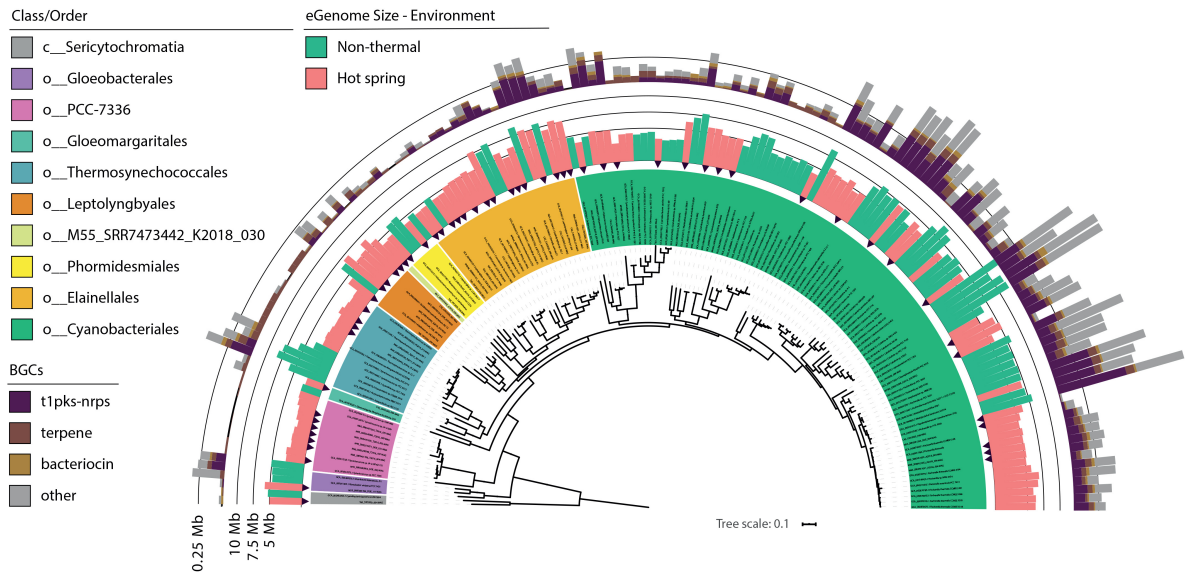


FIGURE 4 | Genome size and secondary metabolite clusters of hot spring and close non-thermal cyanobacteria. Phylogenomic reconstruction from multiple sequence alignment of 120 bacterial marker genes recovered from 102 NCBI cyanobacterial genomes and 57 cyanobacterial MAGs from 21 hot spring metagenomes. Maximum likelihood tree reconstruction was done with IQtree software using the LG + R10 model and a non-parametric UF-bootstrap support of 1,000 replicates. The labels are colored according to the GTDB order classification for c_Cyanobacteriia genomes and class for Sericytochromatia members. The tree was rooted in the node between the Sericytochromatia and Cyanobacteriia classes. Black arrows indicate the position of hot spring MAGs in the tree. The inner ring indicates the estimated genome size (Mbp) for all 159 genomes and is colored according to the respective environmental origin (see **Supplementary Table 3**). The outer ring depicts the total size (Mbp) of secondary metabolite biosynthetic gene clusters (BGCs) detected for each genome using antiSMASH software.

of the coding DNA sequences (**Figure 5**, Wilcoxon's paired test, p-values < 0.05) showed that hot spring cyanobacterial proteins (n = 369,249) are not only shorter than those from non-thermal environments, but also lighter in molecular weight, more hydrophilic, and contain more basic isoelectric points. Indeed, amino acid composition analysis showed that hot spring cyanobacterial proteins have more basic amino acids, such as histidine and arginine, and less acidic amino acids, such as aspartate and glutamate. The aromatic tryptophan and the non-polar residues alanine, leucine, proline and valine were also more abundant in hot spring genomes. The increased alanine abundance (0.81%) and decreased asparagine (0.63%) and lysine (0.62%) abundances were the most notable differences (Wilcoxon's paired test p-values < 0.05, **Supplementary Figure 7**), explaining the tendency toward more hydrophobic proteins in hot spring genomes. These amino acid frequencies agree with previous predictions of increased arginine and valine, and decreased serine and aspartate abundances in other thermophiles and hyperthermophiles (Vieille and Zeikus, 2001; Singer and Hickey, 2003), but disagree for the other amino acid frequencies observed here.

Comparative Genomics: Orthologous Sequences

Analysis of orthologous CDSs was conducted in the 159-subset to determine if there were genes exclusively related to hot spring genomes. Of the 719,564 analyzed protein sequences, 691,501 (96.1%) were assigned into 47,328 orthogroups, 59.2% of which were singletons. Only 183 (0.4%) orthogroups had at least one ortholog sequence in $\geq$ 97% of the analyzed genomes, while 12,657 (26.7%) were sporadically present in both environments. The number of orthogroups with only orthologs ($>$ 1, not singletons) in non-thermal genomes was 3,434 (7.3%, 66 genomes), while that from hot springs was 3,179 (6.7%, 93 genomes). Since environment-specific orthogroups with CDSs across taxa point to potential niche adaptation, CDSs present in more than one class, family or genus were searched. Several orthogroups were shared within 6 different orders or 7 families in hot spring genomes, while non-thermal genomes shared fewer orthogroups at these taxonomic levels (**Supplementary Figure 8**). Most of these widely distributed orthogroups from hot spring genomes were annotated as hypothetical proteins, but other annotated functions included a DsrE family protein and an SDR family oxidoreductase (**Supplementary Table 6**). Phylogenetic analysis of hypothetical protein orthogroups OG0008066 (shared between 5 families) and OG0006223 (shared between 7 genera) not only corroborates the close relationship between hot spring cyanobacteria, but also demonstrates an association with other common thermophilic bacteria, such as Chloroflexota, Deinococcota and Actinobacteriota (**Supplementary Figures 9A,B**). The distribution of these genes across diverse hot spring bacteria could be explained by horizontal gene transfer, as also found between other organisms living at high temperatures (Fuchsman et al., 2017).

Changes within the hot spring core and accessory genomes were determined for 7 genera (*Calothrix*, *Cyanobacterium*, *Elainella*, *Fischerella*, *Geminocystis*, *Rivularia,* and *Trichormus*)

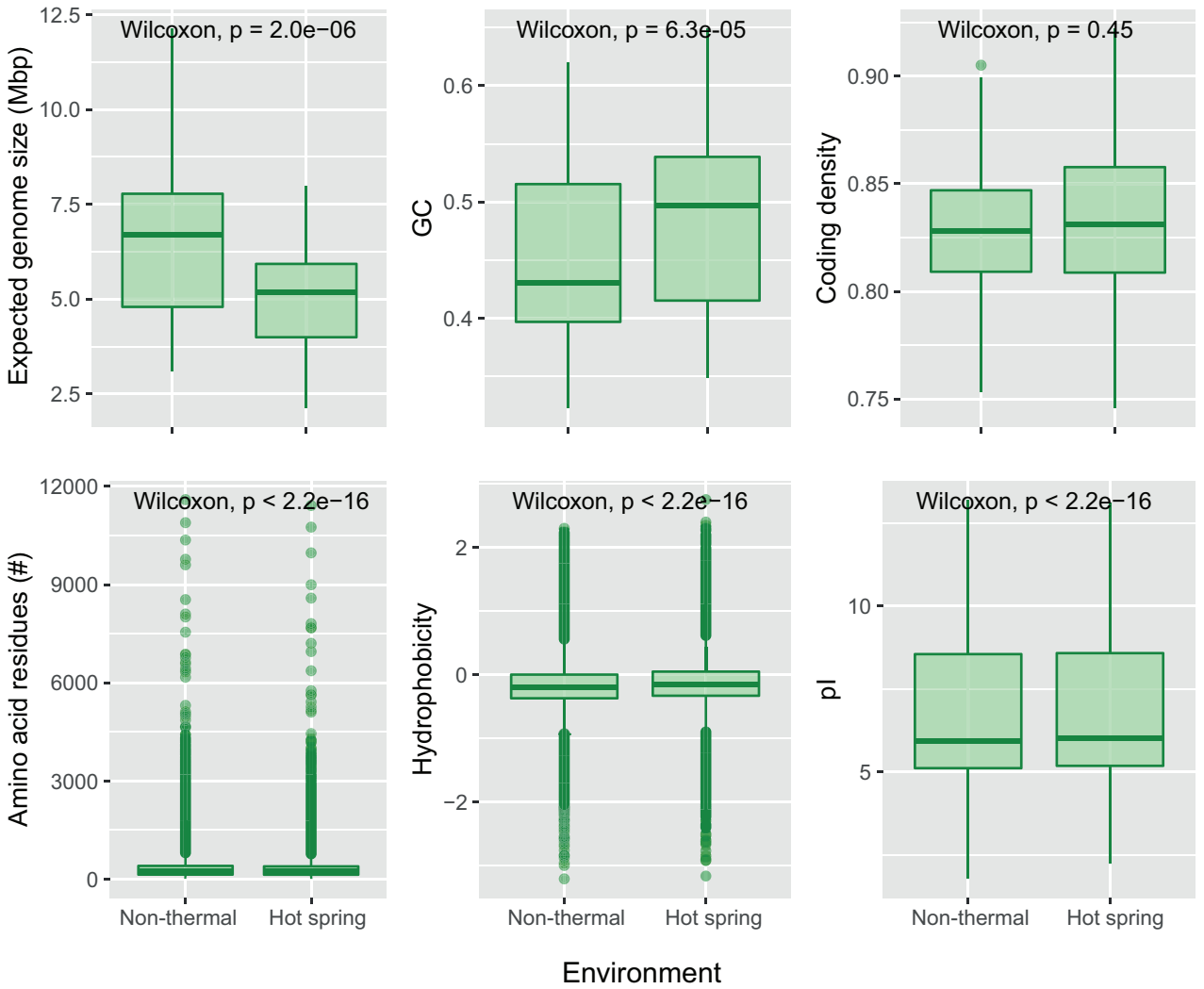


FIGURE 5 | Comparison of genomic and protein properties between hot spring and non-thermal genomes. Boxplots for the "expected genomes size," GC content and coding density were compared for the 159-subset genomes (93 from hot springs and 66 from non-thermal environments). The number of amino acid residues, hydrophobicity and isoelectric point (pI) were compared for the predicted CDSs from the 159-subset genomes from which 369,249 were from hot spring genomes and 348,018 from non-thermal genomes. Wilcoxon's paired test was performed for each dataset, and *p*-values with Bonferroni's multiple test correction are reported.

with both hot spring and non-thermal representatives. The core genome of each genus ranged from 2,007 to 3,488 CDSs, representing 24% (*Calothrix*) to 75% (*Geminocystis* and *Cyanobacterium*) of the CDSs in a genome. The orthogroups were classified according to whether they were present in all genomes (core), exclusively in hot spring or non-thermal genomes (specific accessory and specifc core accessory sets), or genus accessory if they were sporadically present in both environmental groups (**Figure 6A** and **Supplementary Table 7**). The COG distribution (**Supplementary Figure 10**) shows core genomes clustering, along with some of the accessory orthogroups from hot spring and non-thermal genomes. Core accessory orthogroups for all hot spring and non-thermal genera (except for *Cyanobacterium*) clustered together, suggesting common functions associated with these groups.

Regarding gene annotation (**Figure 6B**), non-thermal core accessory orthogroups had more COG function abundances in all genera that were classified as unknown (31, Poorly Char.); replication and repair (24, Information Stor.); energy production and metabolism (22, Metabolism); cell wall/membrane biogenesis (17, Cell Proc.); transcription (17, Information Stor.); and coenzyme metabolism (14, Metabolism) (except for the genus *Rivularia*, whose most represented function was transcription with 54 core accessory orthogroups in the non-thermal genomes). Nevertheless, core accessory genes of hot spring genomes did not exceed 42 orthogroups for each genus, with functions mostly classified in all genera as unknown (50, Poorly Char.); inorganic ion transport (8, Metabolism); amino acid metabolism and transport (6, Metabolism); and cell wall/membrane biogenesis (5, Cell Proc.). For the genus

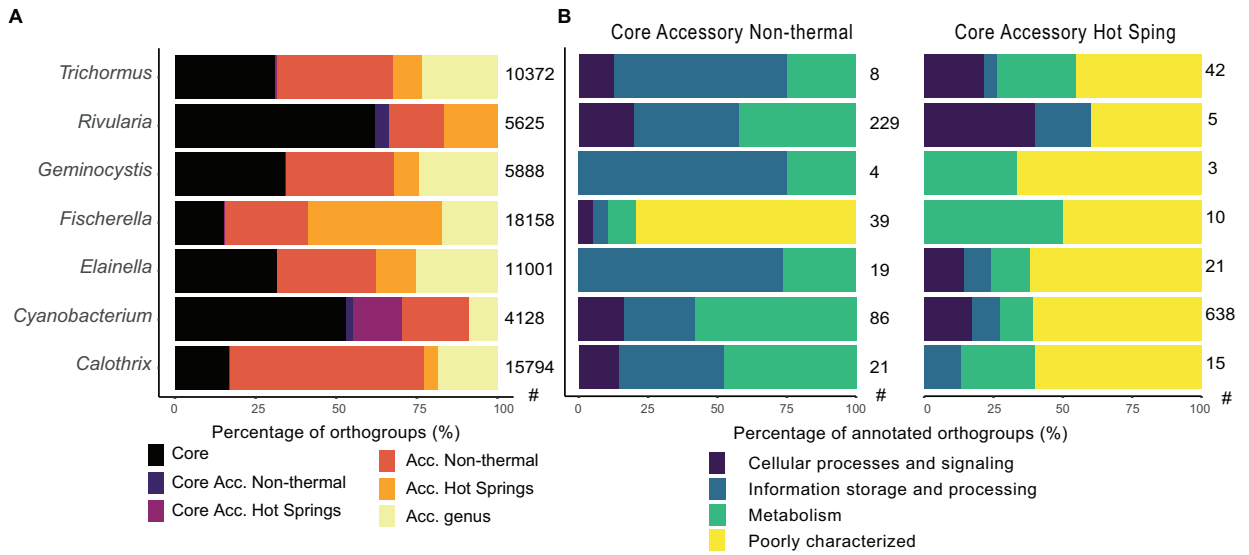


FIGURE 6 | Comparison of core and accessory orthogroups from hot spring and non-thermal cyanobacterial genera. **(A)** The classification of orthogroup fractions according to their presence in all (core), in only hot spring or non-thermal genomes (specific accessory), or in both groups at the genus level (accessory genus). **(B)** COG characterization by eggnog mapper software of exclusive core accessory genes of hot spring and non-thermal environments for each compared genus. The categories were joined in the four main COG functional groups. The number of hot spring (HS) and non-thermal (NT) genomes used for these analyses is as follows: *Trichormus* (2 HS, 6 NT), *Rivularia* (1 HS, 3 NT), *Geminocystis* (3 HS, 3 NT), *Fischerella* (16 HS, 6 NT), *Elainella* (5 HS, 2 NT), *Cyanobacterium* (1 HS, 3NT) and *Calothrix* (2 HS, 8 NT). The number of orthogroups for each analysis is shown in the right side of each bar plot.

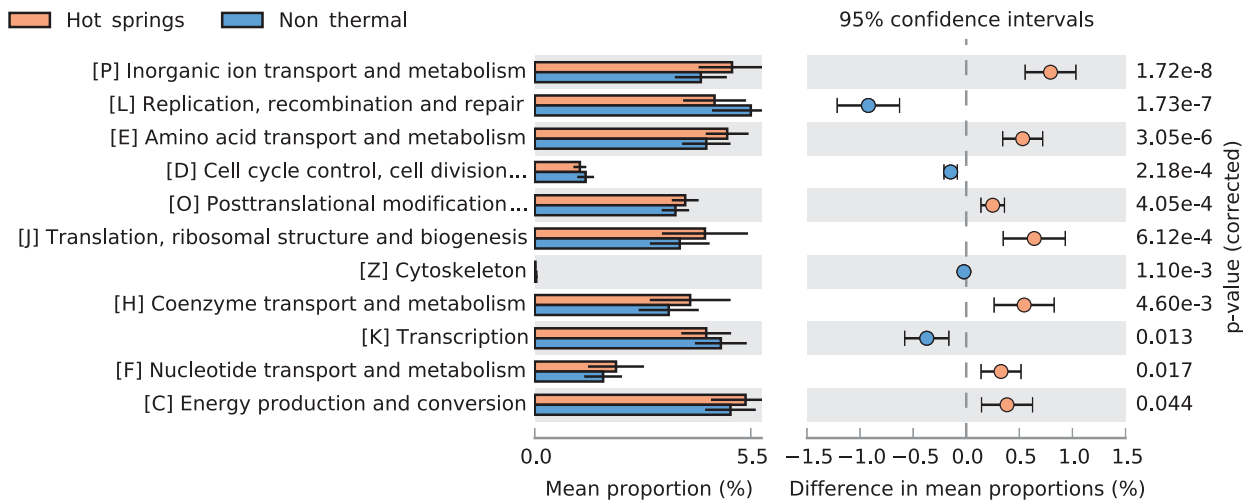


FIGURE 7 | Comparison of COG categories between hot spring and non-thermal genomes. Extended error bar plot showing significantly different predicted COG categories between hot spring (n = 93) and non-thermal (n = 66) cyanobacterial genomes. Only corrected p-values < 0.05 are displayed.

Cyanobacterium, which had only one available hot spring genome, most represented COGs in the hot spring core accessory orthogroups were related to unknown functions (388) (**Figure 6B**).

In general, similar patterns were found for the functional distribution of core genomes between cyanobacterial genera in both the hot spring and non-thermal environments, thereby corroborating previous analyses (Beck et al., 2012). However, these results also show a higher proportion of genes with unknown function in the core accessory genomes of hot spring cyanobacteria. The low number of identified core accessory genes might suggest a reductionist point of view for adaptation to high temperatures, as previously seen for the thermophilic unicellular red algae *Galdieria sulphuraria*, whose adaptation to this environment was mediated by horizontal gene transfer from bacteria and archaea (Schönknecht et al., 2013).

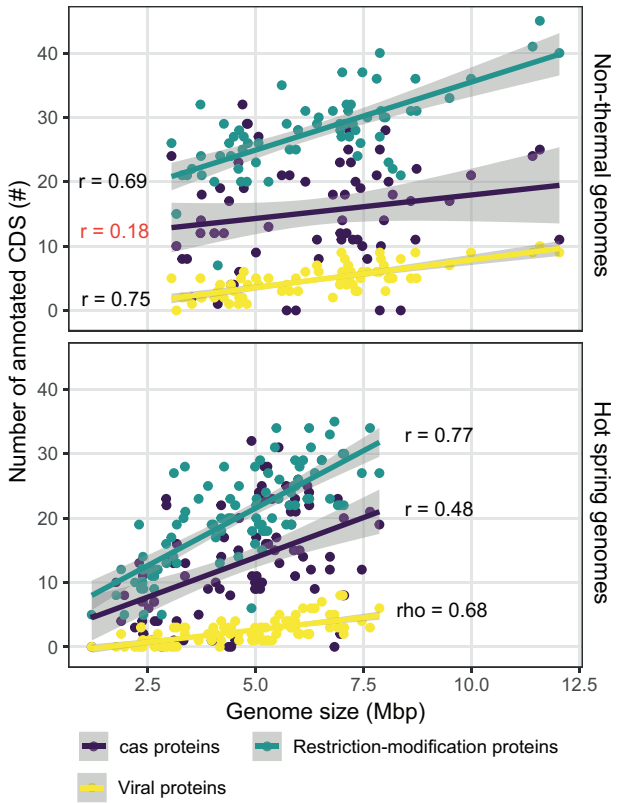


FIGURE 8 | Comparison of defense mechanism-associated genes from hot spring and non-thermal genomes. Relationship between the number of CDSs determined for Cas proteins (purple), viral proteins (yellow) and restriction modification proteins (light blue), and the size of the 159 genomes used for hot spring/non-thermal comparisons. Upper panel shows non-thermal genomes (n = 66), while the bottom panel shows hot spring genomes (n = 93). Smooth linear regression was determined using R software with the ggplot package. Pearson's r and Spearman's rho correlations are provided for each protein category in the plot. Red represents correlations with a p-value > 0.05 and black represents those with p-values < 0.05.

Comparative Genomics: Differences in Functional Categories

Several differences in COG categories were found between the 159-subset of hot spring and non-thermal genomes (Wilcoxon's paired test, p-value < 0.05). The greatest difference was a decrease (∼1% relative abundance) in hot spring genomes for the L COG category (replication, recombination and repair). Additionally, higher abundances were detected in categories P (inorganic ion transport and metabolism), J (translation, ribosomal structure and biogenesis), H (coenzyme transport and metabolism), E (amino acid transport and metabolism) and O (post-translational modification, protein turnover and chaperones). Conversely, other COG categories, such as K (transcription), D (cell cycle control, cell division and chromosome partitioning) and Z (cytoskeleton) decreased in hot spring genomes (**Figure 7**).

An additional comparison of specific functions related to complementary metabolisms and genes involved in microbial community interactions was performed. The results show no differences for the sulfide-quinone reductase gene (related to anoxygenic photosynthesis; Chi-square test p-value = 0.930) or for the *nif* HDK complex and accessory genes (related to nitrogen fixation; Wilcoxon's paired test, p-value = 0.061). However, the potential acquisition of nitrate and its subsequent reduction to ammonium through the *nar*B and *nir*A genes were less represented in hot spring genomes (Wilcoxon's paired test, p-value < 0.05), as were the *hox*, *hup*, *hyp,* and *hyb* genes related to hydrogen metabolism (Wilcoxon's paired test, p-value < 0.05), supporting the competitive exclusion trend by specific nutritional adaptations in the thermal gradient (Ward et al., 2012).

Genes related to the defense mechanisms against foreign nucleic acids and integrated viral proteins were also compared. The number of annotated viral protein orthogroups was 18, showing broad integration of phage CDSs, except for 15 of the hot spring genomes (including all JA-3-3Ab members). Restriction-modification CDSs were classified into 131 different orthogroups, while CRISPR-associated proteins (Cas) were classified into 73 clusters. Considering that genome streamlining was a primary difference between the environmental groups, the correlation between genome size and the number of CDSs annotated with these functions was analyzed. The correlation was higher for restriction-modification and viral proteins (0.778 Pearson's correlation, p-value < 0.05; and 0.747 rho Spearman's correlation, p-value < 0.05, respectively; **Figure 8**), when compared to the Cas proteins (0.360 Pearson's correlation, p-value < 0.05). This scenario varies for Cas proteins when the genomes are separated into hot spring and non-thermal environments (**Figure 8**), in that the Pearson's correlation for hot spring genomes was 0.477 (p-value < 0.05), while the non-thermal genomes showed a lower and non-significant Pearson's correlation index of 0.182 (p-value = 0.142). Conversely, restriction-modification and viral proteins presented similar Pearson's correlation indexes between both environmental groups (Fischer r-to-z transformation p-values > 0.05).

Additionally, BGCs were analyzed because they have been found to be gained and lost during niche transitions (Kurmayer et al., 2015). A total of 6,773,751 bp representing secondary metabolite biosynthetic regions were identified for the 57 MAGs (varying from 0.08 to 11.15% of the total genome). The NRPS-PKS, bacteriocin and terpenes clusters were the most abundant (**Figure 4**), as has been previously seen for some hot spring cyanobacteria (Micallef et al., 2015). BGCs represented a smaller percentage in hot spring genomes (3.07 ± 2.6%, Wilcoxon's paired test, p-value < 0.05) than in non-thermal genomes (5.18 ± 3.39%). For instance, *Fischerella* members show the greatest difference (0.5 ± 0.28 Mb; Wilcoxon's paired test, p-value < 0.05; **Supplementary Figure 11**) in BGC genome percent between genomes from both environments.

For the 159-subset, the correlation between the number of BGC-dedicated base pairs per genome and genome size was higher (0.821 rho Spearman's correlation, p-value < 0.05) than that predicted by Shih et al. (2013) (R^2 = 0.3, p-value < 0.0001), but similar for the hot spring and non-thermal genomes (0.799 and 0.761, respectively). The latter suggests that the streamlining of hot spring genomes also involves a reduction in BGCs. An ecological explanation for BGC-loss in hot spring genomes

could be the lower diversity of these microbial communities compared to other environments (Li et al., 2014). In hot spring communities, the role of secondary metabolites as weapons of inter-microbial warfare (O'Brien and Wright, 2011) could be diminished and susceptible to loss. Furthermore, viral and exogenous nucleic acids are less diverse in hot springs (Parmar et al., 2018), and cyanobacteria are subjected to strong viral predation and potential coevolution (Guajardo-Leiva et al., 2018), as also seen for other thermophilic organisms (Breitbart et al., 2004). Therefore, the prevalence of the CRISPR-Cas system in their genomes could be more important than in non-thermal genomes as seen in other phyla (Weinberger et al., 2012).

The ability to survive at higher temperatures has been gained and lost across bacteria and archaea (Pollo et al., 2015), and all genomic features common to thermophiles are not mandatory for all high-temperature organisms (Puigbò et al., 2008). This was also observed in this study, suggesting once again that different adaptation strategies exist. The polyphyly of this feature in cyanobacteria, seen here widespread in almost all orders, is explained as secondary adaptation (not basal) during the niche expansion stage of trait evolution for this phylum and is intimately related with the ability to form microbial mats (Hammerschmidt et al., 2020). Furthermore, the number of transitions to and from thermophily in cyanobacteria has shown a strong reduction between 0.9 and 0.8 Gya (during the cold temperatures of the Neoproterozoic Oxidation Event), and a slight increase in transitions over the last 0.3 Gya (warmer temperatures). The growing numbers of genera with both environmental groups here described increase the evidence of these recent transitions to thermophily. Altogether, it demonstrates the effect of global temperature changes during this niche expansion (Uyeda et al., 2016) and in the worldwide dispersal of thermophilic cyanobacteria, like *Fischerella*, during the global rise in air temperature 74 mya (Miller et al., 2007). Studying the pathways by which different groups of hot spring cyanobacteria became the autotrophic base for microbial mats in these environments will be the foundation for understanding how different organisms cope and proliferate at higher temperatures in a currently changing world.

CONCLUSION

Cyanobacteria are essential primary producers in hot spring phototrophic microbial mats. In this study, metagenomic binning was used to obtain 57 new cyanobacterial MAGs, subsequently revealing a wide distribution of new thermophilic cyanobacterial members across the phylum. How these cyanobacteria began to colonize these environments is still unknown; however, the adaptation to high temperatures have strong genomic consequences for cyanobacteria that currently live in hot springs. The transition from a non-thermal environment to hot springs is reflected in genomic properties that could be more advantageous to survival, such as small genomes and warm-adapted proteins, as well as a higher abundance of specific protein functional categories to cope with mineral water composition and new microbial and viral communities. This study demonstrates how adaptation to hot spring environments can be traced at the genus level, showing that recent divergences could be restricted to a small set of genes that can be shared by diverse members within the same niche. Under the current scenario of global change, it is important to understand the evolution of hot spring genomes as an example of selective pressure by warmer environments. Studying the information within MAGs from lower hot spring temperatures will allow us to predict changes in genomic features that many species may face in both present and future scenarios on Earth.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, (NCBI BioProject numbers PRJNA635751 and PRJNA645256); https://figshare.com/, 10.6084/m9.figshare.1240097.

AUTHOR CONTRIBUTIONS

JA and BD conceived the study. JA, OS, and TA-S analyzed the metagenomic database and obtained the metagenome-assembled genomes. JA and TA-S performed the evolutionary and bioinformatic analyses under supervision of BD. JA and BD wrote the manuscript. All authors read, commented and approved the drafted manuscript.

FUNDING

This work was financially supported by FONDECYT regular project N° 1150171 and 1190998 (ANID) and ECOS-CONICYT (ANID) project ECOS160025. JA and OS were supported by Doctoral fellowships from CONICYT (ANID) N° 21191763 and 21172022, respectively. JA was also funded by Beca Gastos Operacionales 2020 CONICYT (ANID) N° 24200178. FONDAP-CONICYT (ANID) N°151110009 CR2 is also acknowledged for it support.

ACKNOWLEDGMENTS

The authors would like to thank Ph.D. Roy Mackenzie, Ph.D. Sergio Guajardo, Ph.D. Christina Ridley and Ph.D. (c) Pablo Vergara for their contributions to sampling, sequencing and editing help.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.568223/full#supplementary-material

REFERENCES

Alcamán, M. E., Fernandez, C., Delgado, A., Bergman, B., and Díez, B. (2015). The cyanobacterium Mastigocladus fulfills the nitrogen demand of a terrestrial hot spring microbial mat. *ISME J.* 9, 2290–2303. doi: 10.1038/ismej.2015.63

Alcamán-Arias, M. E., Pedrós-Alió, C., Tamames, J., Fernández, C., Pérez-Pantoja, D., Vásquez, M., et al. (2018). Diurnal changes in active carbon and nitrogen pathways along the temperature gradient in Porcelana hot spring microbial mat. *Front. Microbiol.* 9:2353. doi: 10.3389/fmicb.2018.02353

Alcorta, J., Espinoza, S., Viver, T., Alcamán-Arias, M. E., Trefault, N., Rosselló-Móra, R., et al. (2018). Temperature modulates *Fischerella thermalis* ecotypes in Porcelana hot spring. *Syst. Appl. Microbiol.* 41, 531–543. doi: 10.1016/j.syapm.2018.05.006

Alcorta, J., Vergara-Barros, P., Antonaru, L. A., Alcamán-Arias, M. E., Nürnberg, D. J., and Díez, B. (2019). *Fischerella thermalis*: a model organism to study thermophilic diazotrophy, photosynthesis and multicellularity in cyanobacteria. *Extremophiles* 23, 635–647. doi: 10.1007/s00792-019-01125-4

Alneberg, J., Bjarnason, B. S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nat. Methods* 11, 1144–1146. doi: 10.1038/nmeth.3103

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Beck, C., Knoop, H., Axmann, I. M., and Steuer, R. (2012). The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms. *BMC Genomics* 13:56. doi: 10.1186/1471-2164-13-56

Berrendero, E., Perona, E., and Mateo, P. (2008). Genetic and morphological characterization of Rivularia and Calothrix (Nostocales, Cyanobacteria) from running water. *Int. J. Syst. Evol. Microbiol.* 58, 447–460. doi: 10.1099/ijs.0.65273-0

Bhaya, D., Grossman, A. R., Steunou, A. S., Khuri, N., Cohan, F. M., Hamamura, N., et al. (2007). Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME J.* 1, 703–713. doi: 10.1038/ismej.2007.46

Blin, K., Wolf, T., Chevrette, M. G., Lu, X., Schwalen, C. J., Kautsar, S. A., et al. (2017). antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* 45, W36–W41. doi: 10.1093/nar/gkx319

Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* 35, 725–731. doi: 10.1038/nbt.3893

Breitbart, M., Wegley, L., Leeds, S., Schoenfeld, T., and Rohwer, F. (2004). Phage community dynamics in hot springs. *Appl. Environ. Microbiol.* 70, 1633–1640. doi: 10.1128/AEM.70.3.1633-1640.2004

Brown, I. I., Mummey, D., and Cooksey, K. E. (2005). A novel cyanobacterium exhibiting an elevated tolerance for iron. *FEMS Microbiol. Ecol.* 52, 307–314. doi: 10.1016/j.femsec.2004.11.020

Castenholz, R. W. (1969). Thermophilic blue-green algae and the thermal environment. *Bacteriol. Rev.* 33, 476–504. doi: 10.1128/mmbr.33.4.476-504.1969

Castenholz, R. W., Wilmotte, A., Herdman, M., Rippka, R., Waterbury, J. B., Iteman, I., et al. (2001). "Phylum BX. cyanobacteria," in *Bergey's Manual§of Systematic Bacteriology*, eds D. R. Boone, R. W. Castenholz, and G. M. Garrity (New York, NY: Springer), 473–599. doi: 10.1007/978-0-387-21609-6_27

Chan, C. S., Chan, K. G., Tay, Y. L., Chua, Y. H., and Goh, K. M. (2015). Diversity of thermophiles in a Malaysian hot spring determined using 16S rRNA and shotgun metagenome sequencing. *Front. Microbiol.* 6:177. doi: 10.3389/fmicb.2015.00177

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2020). GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848

Cheng, Y. I., Chou, L., Chiu, Y. F., Hsueh, H. T., Kuo, C. H., and Chu, H. A. (2020). Comparative genomic analysis of a novel strain of Taiwan hot-spring cyanobacterium *Thermosynechococcus* sp. CL-1. *Front. Microbiol.* 11:82. doi: 10.3389/fmicb.2020.00082

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., da Costa, M. S., et al. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68, 461–466. doi: 10.1099/ijsem.0.002516

Chuvochina, M., Rinke, C., Parks, D. H., Rappé, M. S., Tyson, G. W., Yilmaz, P., et al. (2019). The importance of designating type material for uncultured taxa. *Syst. Appl. Microbiol.* 42, 15–21. doi: 10.1016/j.syapm.2018.07.003

Copeland, J. J. (1936). Yellowstone thermal myxophyceae. *Ann. N. Y. Acad. Sci.* 36, 4–223. doi: 10.1111/j.1749-6632.1936.tb56976.x

Cox, A., Shock, E. L., and Havig, J. R. (2011). The transition to microbial photosynthesis in hot spring ecosystems. *Chem. Geol.* 280, 344–351. doi: 10.1016/j.chemgeo.2010.11.022

Dagan, T., Roettger, M., Stucken, K., Landan, G., Koch, R., Major, P., et al. (2013). Genomes of Stigonematalean cyanobacteria (Subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. *Genome. Biol. Evol.* 5, 31–44. doi: 10.1093/gbe/evs117

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics* 14, 755–763. doi: 10.1093/bioinformatics/14.9.755

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Emms, D. M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16:157. doi: 10.1186/s13059-015-0721-2

Finsinger, K., Scholz, I., Serrano, A., Morales, S., Uribe-Lorio, L., Mora, M., et al. (2008). Characterization of true−branching cyanobacteria from geothermal sites and hot springs of Costa Rica. *Environ. Microbiol.* 10, 460–473. doi: 10.1111/j.1462-2920.2007.01467.x

Fuchsman, C. A., Collins, R. E., Rocap, G., and Brazelton, W. J. (2017). Effect of the environment on horizontal gene transfer between bacteria and archaea. *PeerJ* 5:e3865. doi: 10.7717/peerj.3865

Guajardo-Leiva, S., Pedrós-Alió, C., Salgado, O., Pinto, F., and Díez, B. (2018). Active crossfire between Cyanobacteria and Cyanophages in phototrophic mat communities within hot springs. *Front. Microbiol.* 9:2039. doi: 10.3389/fmicb.2018.02039

Hammerschmidt, K., Landan, G., Tria, F. D. K., Alcorta, J., and Dagan, T. (2020). The order of trait emergence in the evolution of cyanobacterial multicellularity. *bioRxiv* [Preprint]. doi: 10.1101/570788

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2017). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi: 10.1093/molbev/msx281

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., Von Mering, C., et al. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* 34, 2115–2122. doi: 10.1093/molbev/msx148

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi: 10.1093/nar/gky1085

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119

Ionescu, D., Hindiyeh, M., Malkawi, H., and Oren, A. (2010). Biogeography of thermophilic cyanobacteria: insights from the Zerka Ma'in hot springs (Jordan). *FEMS Microbiol. Ecol.* 72, 103–113. doi: 10.1111/j.1574-6941.2010.00835.x

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9, 1–8. doi: 10.1038/s41467-018-07641-9

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. doi: 10.7717/peerj.7359

Kaushal, G., Kumar, J., Sangwan, R. S., and Singh, S. P. (2018). Metagenomic analysis of geothermal water reservoir sites exploring carbohydrate-related thermozymes. *Int. J. Biol. Macromol.* 119, 882–895. doi: 10.1016/j.ijbiomac.2018.07.196

Klatt, C. G., Wood, J. M., Rusch, D. B., Bateson, M. M., Hamamura, N., Heidelberg, J. F., et al. (2011). Community ecology of hot spring cyanobacterial mats:

predominant populations and their functional potential. *ISME J.* 5, 1262–1278. doi: 10.1038/ismej.2011.73

Komárek, J. (1999). Cyanoprokaryota 1. Teil: Chroococcales. *Subwasserflora von Mitteleuropa* 19, 1–548. doi: 10.1127/algol_stud/103/2001/1

Komárek, J. (2013). *Süsswasserflora Von Mitteleuropa: Cyanoprokaryota. Heterocytous Genera*. Berlin: Springer.

Komárek, J., and Anagnostidis, K. (2005). *Süßwasserflora von Mitteleuropa, bd. 19/2: Cyanoprokaryota: Oscillatoriales*, Vol. 19. Heidelberg: Spektrum Akademischer Verlag.

Konstantinidis, K. T., Rosselló-Móra, R., and Amann, R. (2017). Uncultivated microbes in need of their own taxonomy. *ISME J.* 11, 2399–2406. doi: 10.1038/ismej.2017.113

Korelusova, J., Kaštovský, J., and Komárek, J. (2009). Heterogeneity of the cyanobacterial genus synechocystis and description of a new genus, geminocystis. *J. Phycol.* 45, 928–937. doi: 10.1111/j.1529-8817.2009.00701.x

Kurmayer, R., Blom, J. F., Deng, L., and Pernthaler, J. (2015). Integrating phylogeny, geographic niche partitioning and secondary metabolite synthesis in bloom-forming Planktothrix. *ISME J.* 9:909. doi: 10.1038/ismej.2014.189

Larsson, J., Nylander, J. A., and Bergman, B. (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evol. Biol.* 11:187. doi: 10.1186/1471-2148-11-187

Laslett, D., and Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32, 11–16. doi: 10.1093/nar/gkh152

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Li, D., Liu, C. M., Luo, R., Sadakane, K., and Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033

Li, S. J., Hua, Z. S., Huang, L. N., Li, J., Shi, S. H., Chen, L. X., et al. (2014). Microbial communities evolve faster in extreme environments. *Sci. Rep.* 4:6205. doi: 10.1038/srep06205

Mackenzie, R., Pedrós-Alió, C., and Díez, B. (2013). Bacterial composition of microbial mats in hot springs in Northern Patagonia: variations with seasons and temperature. *Extremophiles* 17, 123–136. doi: 10.1007/s00792-012-0499-z

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Matsen, F. A., Kodner, R. B., and Armbrust, E. V. (2010). pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11:538. doi: 10.1186/1471-2105-11-538

Meeks, J. C., and Castenholz, R. W. (1971). Growth and photosynthesis in an extreme thermophile, *Synechococcus lividus* (Cyanophyta). *Arch. Mikrobiol.* 78, 25–41. doi: 10.1007/BF00409086

Micallef, M. L., D'Agostino, P. M., Sharma, D., Viswanathan, R., and Moffitt, M. C. (2015). Genome mining for natural product biosynthetic gene clusters in the Subsection V cyanobacteria. *BMC Genomics* 16:669. doi: 10.1186/s12864-015-1855-z

Miller, S., Castenholz, R., and Pedersen, D. (2007). Phylogeography of the thermophilic cyanobacterium *Mastigocladus laminosus*. *Appl. Environ. Microbiol.* 73, 4751–4759. doi: 10.1128/AEM.02945-06

Momper, L., Hu, E., Moore, K. R., Skoog, E. J., Tyler, M., Evans, A. J., et al. (2019). Metabolic versatility in a modern lineage of cyanobacteria from terrestrial hot springs. *Free Radical Biol. Med.* 140, 224–232. doi: 10.1016/j.freeradbiomed.2019.05.036

Monchamp, M. E., Spaak, P., and Pomati, F. (2019). Long term diversity and distribution of non-photosynthetic cyanobacteria in peri-alpine lakes. *Front. Microbiol.* 9:3344. doi: 10.3389/fmicb.2018.03344

Moreira, D., Tavera, R., Benzerara, K., Skouri-Panet, F., Couradeau, E., Gérard, E., et al. (2017). Description of *Gloeomargarita lithophora* gen. nov., sp. nov., a thylakoid-bearing basal-branching cyanobacterium with intracellular carbonates, and proposal for Gloeomargaritales ord. nov. *Int. J. Syst. Evol. Microbiol.* 67:653. doi: 10.1099/ijsem.0.001679

Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., et al. (2003). Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. *DNA Res.* 10, 137–145. doi: 10.1093/dnares/10.4.137

O'Brien, J., and Wright, G. D. (2011). An ecological perspective of microbial secondary metabolism. *Curr. Opin. Biotechnol.* 22, 552–558. doi: 10.1016/j.copbio.2011.03.010

Olm, M. R., Brown, C. T., Brooks, B., and Banfield, J. F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* 11, 2864–2868. doi: 10.1038/ismej.2017.126

Olsen, M. T., Nowack, S., Wood, J. M., Becraft, E. D., LaButti, K., Lipzen, A., et al. (2015). The molecular dimension of microbial species: 3. Comparative genomics of Synechococcus strains with different light responses and in situ diel transcription patterns of associated putative ecotypes in the Mushroom Spring microbial mat. *Front. Microbiol.* 6:604. doi: 10.3389/fmicb.2015.00604

Oren, A., and Ventura, S. (2017). The current status of cyanobacterial nomenclature under the "prokaryotic" and the "botanical" code. *Antonie Leeuwenhoek* 110, 1257–1269. doi: 10.1007/s10482-017-0848-0

Papke, R. T., Ramsing, N. B., Bateson, M. M., and Ward, D. M. (2003). Geographical isolation in hot spring cyanobacteria. *Environ. Microbiol.* 5, 650–659. doi: 10.1046/j.1462-2920.2003.00460.x

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumel, P., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36, 996–1004. doi: 10.1038/nbt.4229

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P. A., Woodcroft, B. J., Evans, P. N., et al. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2, 1533–1542. doi: 10.1038/s41564-017-0012-7

Parks, D. H., Tyson, G. W., Hugenholtz, P., and Beiko, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30, 3123–3124. doi: 10.1093/bioinformatics/btu494

Parmar, K., Dafale, N., Pal, R., Tikariha, H., and Purohit, H. (2018). An insight into phage diversity at environmental habitats using comparative metagenomics approach. *Curr. Microbiol.* 75, 132–141. doi: 10.1007/s00284-017-1357-0

Pollo, S. M., Zhaxybayeva, O., and Nesbø, C. L. (2015). Insights into thermoadaptation and the evolution of mesophily from the bacterial phylum Thermotogae. *Can. J. Microbiol.* 61, 655–670. doi: 10.1139/cjm-2015-0073

Ponce-Toledo, R. I., Deschamps, P., López-García, P., Zivanovic, Y., Benzerara, K., and Moreira, D. (2017). An early-branching freshwater cyanobacterium at the origin of plastids. *Curr. Biol.* 27, 386–391. doi: 10.1016/j.cub.2016.11.056

Puigbò, P., Pasamontes, A., and Garcia-Vallve, S. (2008). Gaining and losing the thermophilic adaptation in prokaryotes. *Trends Genet.* 24, 10–14. doi: 10.1016/j.tig.2007.10.005

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ramos, V., Morais, J., and Vasconcelos, V. M. (2017). A curated database of cyanobacterial strains relevant for modern taxonomy and phylogenetic studies. *Sci. Data* 4, 1–8. doi: 10.1038/sdata.2017.54

Reji, L., and Francis, C. A. (2020). Metagenome-assembled genomes reveal unique metabolic adaptations of a basal marine Thaumarchaeota lineage. *ISME J.* 14, 2105–2115. doi: 10.1038/s41396-020-0675-6

Roeselers, G., Norris, T. B., Castenholz, R. W., Rysgaard, S., Glud, R. N., Kühl, M., et al. (2007). Diversity of phototrophic bacteria in microbial mats from Arctic hot springs (Greenland). *Environ. Microbiol.* 9, 26–38. doi: 10.1111/j.1462-2920.2006.01103.x

Roney, H. C., Booth, G. M., and Cox, P. A. (2009). Competitive exclusion of cyanobacterial species in the Great Salt Lake. *Extremophiles* 13, 355–361. doi: 10.1007/s00792-008-0223-1

Roy, C., Rameez, M. J., Haldar, P. K., Peketi, A., Mondal, N., Bakshi, U., et al. (2020). Microbiome and ecology of a hot spring-microbialite system on the Trans-Himalayan Plateau. *Sci. Rep.* 10, 1–22. doi: 10.1038/s41598-020-62797-z

Sabath, N., Ferrada, E., Barve, A., and Wagner, A. (2013). Growth temperature and genome size in bacteria are negatively correlated, suggesting genomic

streamlining during thermal adaptation. *Genome Biol. Evol.* 5, 966–977. doi: 10.1093/gbe/evt050

Sanchez-Baracaldo, P., Hayes, P. K., and Blank, C. E. (2005). Morphological and habitat evolution in the Cyanobacteria using a compartmentalization approach. *Geobiology* 3, 145–165. doi: 10.1111/j.1472-4669.2005.00050.x

Sano, E. B., Wall, C. A., Hutchins, P. R., and Miller, S. R. (2018). Ancient balancing selection on heterocyst function in a cosmopolitan cyanobacterium. *Nat. Ecol. Evol.* 2, 510–519. doi: 10.1038/s41559-017-0435-9

Schirrmeister, B. E., Gugger, M., and Donoghue, P. C. (2015). Cyanobacteria and the great oxidation event: evidence from genes and fossils. *Palaeontology* 58, 769–785. doi: 10.1111/pala.12178

Schönknecht, G., Chen, W. H., Ternes, C. M., Barbier, G. G., Shrestha, R. P., Stanke, M., et al. (2013). Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* 339, 1207–1210. doi: 10.1126/science.1231707

Schwabe, H. (1837). Über die Algen der Karlsbader warmen Quellen. *Linnaea* 11, 109–127.

Sciuto, K., and Moro, I. (2016). Detection of the new cosmopolitan genus Thermoleptolyngbya (Cyanobacteria, Leptolyngbyaceae) using the 16S rRNA gene and 16S–23S ITS region. *ıMol. Phylogenet. Evol.* 105, 15–35. doi: 10.1016/j.ympev.2016.08.010

Shaiber, A., and Eren, A. M. (2019). Composite metagenome-assembled genomes reduce the quality of public genome repositories. *mBio* 10:e00725-19. doi: 10.1128/mBio.00725-19

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., et al. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1053–1058. doi: 10.1073/pnas.1217107110

Singer, G. A., and Hickey, D. A. (2003). Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content. *Gene* 317, 39–47. doi: 10.1016/S0378-1119(03)00660-7

Steunou, A. S., Bhaya, D., Bateson, M. M., Melendrez, M. C., Ward, D. M., Brecht, E., et al. (2006). In situ analysis of nitrogen fixation and metabolic switching in unicellular thermophilic cyanobacteria inhabiting hot spring microbial mats. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2398–2403. doi: 10.1073/pnas.0507513103

Thiel, V., Hügler, M., Ward, D. M., and Bryant, D. A. (2017). The dark side of the mushroom spring microbial mat: life in the shadow of chlorophototrophs. II. Metabolic functions of abundant community members predicted from metagenomic analyses. *Front. Microbiol.* 8:943. doi: 10.3389/fmicb.2017.00943

Uribe-Lorío, L., Brenes-Guillén, L., Hernández-Ascencio, W., Mora-Amador, R., González, G., Ramírez-Umaña, C. J., et al. (2019). The influence of temperature and pH on bacterial community composition of microbial mats in hot springs from Costa Rica. *MicrobiologyOpen* 8:e893. doi: 10.1002/mbo3.893

Uritskiy, G. V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6, 1–13. doi: 10.1186/s40168-018-0541-1

Uyeda, J. C., Harmon, L. J., and Blank, C. E. (2016). A comprehensive study of cyanobacterial morphological and ecological evolutionary dynamics through deep geologic time. *PLoS One* 11:e0162539. doi: 10.1371/journal.pone.0162539

Vieille, C., and Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* 65, 1–43. doi: 10.1128/MMBR.65.1.1-43.2001

Walter, J. M., Coutinho, F. H., Dutilh, B. E., Swings, J., Thompson, F. L., and Thompson, C. C. (2017). Ecogenomics and taxonomy of *Cyanobacteria phylum*. *Front. Microbiol.* 8:2132. doi: 10.3389/fmicb.2017.02132

Ward, D. M., Bateson, M. M., Ferris, M. J., Kühl, M., Wieland, A., Koeppel, A., et al. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos. Trans. R. Soc. B.* 361, 1997–2008. doi: 10.1098/rstb.2006.1919

Ward, D. M., Castenholz, R. W., and Miller, S. R. (2012). *Cyanobacteria in Geothermal Habitats. In Ecology of Cyanobacteria II*. Dordrecht: Springer, 39–63. doi: 10.1007/978-94-007-3855-3_3

Ward, D. M., Ferris, M. J., Nold, S. C., and Bateson, M. M. (1998). A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol. Mol. Biol. Rev.* 62, 1353–1370. doi: 10.1128/mmbr.62.4.1353-1370.1998

Ward, L. M., Idei, A., Nakagawa, M., Ueno, Y., Fischer, W. W., and McGlynn, S. E. (2019). Geochemical and metagenomic characterization of Jinata Onsen, a Proterozoic-analog hot spring, reveals novel microbial diversity including iron-tolerant phototrophs and thermophilic lithotrophs. *Microbes Environ.* 34, 278–292. doi: 10.1264/jsme2.ME19017

Weinberger, A. D., Wolf, Y. I., Lobkovsky, A. E., Gilmore, M. S., and Koonin, E. V. (2012). Viral diversity threshold for adaptive immunity in prokaryotes. *mBio* 3:e00456-12. doi: 10.1128/mBio.00456-12

Whitman, W. B. (2016). Modest proposals to expand the type material for naming of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 66, 2108–2112. doi: 10.1099/ijsem.0.000980

Wu, S., and Zhu, Y. (2012). ProPAS: standalone software to analyze protein properties. *Bioinformation* 8:167. doi: 10.6026/97320630008167

Wu, Y. W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32, 605–607. doi: 10.1093/bioinformatics/btv638

Yoon, K. S., Nguyen, N. T., Tran, K. T., Tsuji, K., and Ogo, S. (2017). Nitrogen fixation genes and nitrogenase activity of the non-heterocystous cyanobacterium *Thermoleptolyngbya* sp. O-77. *Microbes Environ.* 32, 324–329. doi: 10.1264/jsme2.ME17015

Zeldovich, K. B., Berezovsky, I. N., and Shakhnovich, E. I. (2007). Protein and DNA sequence determinants of thermophilic adaptation. *PLoS Comput. Biol.* 3:e5. doi: 10.1371/journal.pcbi.0030005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership