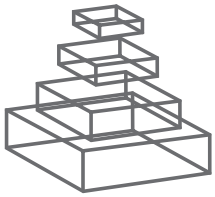
frontiers
RESEARCH TOPICS

TOXICOGENOMICS IN NON-MAMMALIAN SPECIES

Hosted by
Stephen Sturzenbaum, Michael Aschner
and Jonathan Freedman

frontiers in
GENETICS



frontiers

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-064-5
DOI 10.3389/978-2-88919-064-5

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

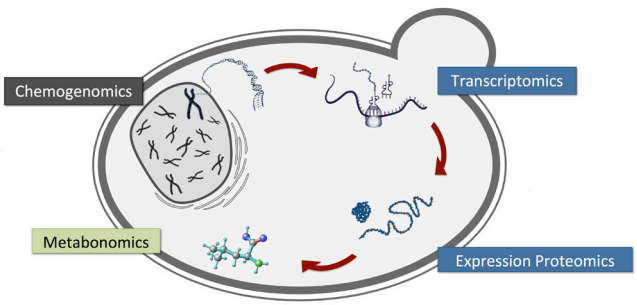
Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - 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: researchtopics@frontiersin.org

TOXICOGENOMICS IN NON-MAMMALIAN SPECIES

Hosted By:
Stephen Sturzenbaum, King's College London, UK
Michael Aschner, Vanderbilt University Medical Center, USA
Jonathan Freedman, National Institutes of Health, USA



Some molecules or conditions are exclusively toxic to biological systems and classified as being non-essential; others are essential for life. Nevertheless, above certain threshold even the essential will become toxic. Tightly controlled homeostatic control mechanisms are thus vital drivers of well being, longevity and survival. The identification and characterization of these intricate pathways form the foundations of Toxicogenomics.

The initiation, and indeed completion, of numerous non-mammalian genome-sequencing projects, has driven the exponential growth of available genetic sequences. Collating this vast amount of data into functional and mechanistically meaningful units will provide novel insights into pathogenesis, new methods of risk assessment, genetic risk-modifications in preventative medicine and new therapeutic targets for pharmaceutical and biological medicines.

This Research Topic issue will explore the current knowledgebase pertaining to the multitude of genomic and toxicological tools within non-mammalian organisms. The encyclopaedic coverage will span the full taxonomic breadth ranging from simple unicellular bacteria and yeast to complex creatures such as birds and fish. The resulting collection of unique, complimentary or indeed contrasting approaches, tools and technologies (which are defined by the availability and feasibility for each organism to study genomics of xenobiotic or stress biology) will not only foster cross-phyla awareness but expand the horizon of Toxicogenomics.

We invite submissions of article types including Original Research, Methods, Hypothesis & Theory, Reviews and Perspectives.

Table of Contents

frontiers in
GENETICS

EDITORIAL
published: 16 October 2012
doi: 10.3389/fgene.2012.00216

Toxicogenomics in non-mammalian species—Editorial

Stephen Stürzenbaum[1]*, Michael Aschner[2] and Jonathan H. Freedman[3]

[1] Analytical and Environmental Science Division, School of Biomedical Sciences, King's College London, London, UK
[2] Vanderbilt University Medical Center, Nashville, TN, USA
[3] Laboratory of Toxicology and Pharmacology, National Institute of Environmental Health Science, National Institute of Health, Research Triangle Park, NC, USA
*Correspondence: stephen.sturzenbaum@kcl.ac.uk

Edited by:
Douglas M. Ruden, Wayne State University, USA

Reviewed by:
Douglas M. Ruden, Wayne State University, USA

In October 2004, the Organisation for Economic Co-operation and Development (OECD) and the International Programme on Chemical Safety (IPCS) organized a workshop in Kyoto Japan to outline the fundamental aspects and future directions of the then emerging science of "Toxicogenomics." The output of this three-day meeting was summarized in the Report of the OECD/IPCS workshop on Toxicogenomics (OECD Series on Testing and Assessment, Number 50, http://www.who.int/ipcs/methods/oecd_report.pdf). In essence, Toxicogenomics was defined as any study that investigates the response of a genome to hazardous substances by means of (1) genomic-scale mRNA expression analyses (Transcriptomics), (2) cell and tissue wide protein expression techniques (Proteomics), or (3) cell and tissue wide metabolite profiling (Metabolomics). These "omic" datasets typically require intricate *in silico* analyses (Bioinformatics) to integrate the results and by doing so provide insights into mechanistic toxicology and biomarkers of exposure.

Since 2004, the initiation, and indeed completion, of numerous whole genome-sequencing projects has driven the exponential growth of available genetic information. Collating this vast amount of data into functional and mechanistically meaningful units is providing novel insights into pathogenesis, new methods of risk assessment, genetic risk-modifications in preventative medicine, and new therapeutic targets for pharmaceutical and biological medicines. Most toxicogenomic responses, however, are multi-dimensional due to the facts that toxicants usual affect multiple intra-and extra-cellular targets and occur as complex mixtures, which are inherently difficult to decipher. Some molecules or conditions are exclusively toxic to biological systems and classified as being non-essential, while others are essential for life. According to Paracelsus, above certain threshold even the essential will become toxic. Thus, tightly controlled homeostatic control mechanisms are vital drivers of well-being, longevity, and survival. The identification and characterization of these mechanisms and the cognate regulatory pathways form the foundations of Toxicogenomics.

This Research Topic issue explores our current knowledge pertaining to the multitude of genomic and toxicological tools within non-mammalian organisms, arguably an underdeveloped niche. This e-book begins with five seminal reviews on (1) the yeast system, which focusses on genome-wide responses to chemical stressors linked to Environmental Health, Pharmacology, and Biotechnology (dos Santos et al., 2012); (2) the fruit fly *Drosophila melanogaster*, as a model for lead neurotoxicology and Toxicogenomics (Hirsch et al., 2012); (3) the nematode *Caenorhabditis elegans*, to investigate the genome-wide response to metal exposure (Caito et al., 2012); (4) next generation sequencing approaches in fish Toxicogenomics (Mehinto et al., 2012); and (5) the application of Toxicogenomics in amphibians (Helbing, 2012). These reviews set the stage for five research articles that describe how genomic tools have aided in uncovering target mechanisms of (eco)toxicological importance. These comprise transcript meta-analyses in *C. elegans* to pinpoint humic acid, quercetin, and tannic acid-mediated effectors of stress and aging (Pietsch et al., 2012 and Menzel et al., 2012); the assessment of a municipal landfill soil via a functional environmental genomic approach using the springtail *Folsomia candida* (Roelofs et al., 2012); a bioinformatic approach to model genotoxic chemical mutations in *Drosophila melanogaster* (Cingolani et al., 2012); and a study that investigates the influence of nitrate and nitrite on targeted gene expression in the tadpole *Rana catesbeiana* (Hinther et al., 2012).

By spanning a wide taxonomic breadth, this encyclopaedic coverage results in a collection of unique approaches, tools and technologies, which are currently defined by the availability and feasibility for each organism to study the genomics of xenobiotic or stress biology. We anticipate that this information will not only foster cross-phyla awareness, but also expand the horizon of Toxicogenomics. We hope you enjoy this eclectic mix of papers.

REFERENCES

Caito, S., Fretham, S., Martinez-Finley, E., Chakraborty, S., Avila, D., Chen, P., et al. (2012). Genome-wide analyses of metal responsive genes in *Caenorhabditis elegans*. *Front. Gene.* 3:52. doi: 10.3389/fgene.2012.00052

Cingolani, P., Patel, V. M., Coon, M., Nguyen, T., Land, S. J., Ruden, D. M., et al. (2012). Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front. Gene.* 3:35. doi: 10.3389/fgene.2012.00035

dos Santos, S. C., Teixeira, M. C., Cabrito, T. R., and Sá-Correia, I. (2012). Yeast toxicogenomics: genome-wide responses to chemical stresses with impact in environmental health, pharmacology, and biotechnology. *Front. Gene.* 3:63. doi: 10.3389/fgene.2012.00063

Helbing, C. C. (2012). The metamorphosis of amphibian toxicogenomics. *Front. Gene.* 3:37. doi: 10.3389/fgene.2012.00037

Hinther, A., Edwards, T. M., Guillette, Jr. L. J., and Helbing, C. C. (2012). Influence of nitrate and nitrite on thyroid hormone responsive and stress-associated gene expression in cultured *Rana catesbeiana* tadpole tail fin tissue. *Front. Gene.* 3:51. doi: 10.3389/fgene.2012.00051

Hirsch, H. V. B., Lnenicka, G., Possidente, D., Possidente, B., Garfinkel, M. D., Wang, L., et al. (2012). *Drosophila melanogaster* as a model for lead neurotoxicology and toxicogenomics research. *Front. Gene.* 3:68. doi: 10.3389/fgene.2012.00068

Mehinto, A. C., Martyniuk, C. J., Spade, D. J., and Denslow, N. D. (2012). Applications of next-generation sequencing in fish ecotoxicogenomics. *Front. Gene.* 3:62. doi: 10.3389/fgene.2012.00062

Menzel, R., Menzel, S., Swain, S. C., Pietsch, K., Tiedt, S., Witczak, J., et al. (2012). The nematode *Caenorhabditis elegans*, stress and aging: identifying the complex interplay of genetic pathways following the treatment with humic substances. *Front. Gene.* 3:50. doi: 10.3389/fgene.2012.00050

Pietsch, K., Saul, N., Swain, S. C., Menzel, R., Steinberg, C. E. W., and Stürzenbaum, S. R. (2012). Meta-analysis of global transcriptomics suggests that conserved genetic pathways are responsible for Quercetin and Tannic acid mediated longevity in *C. elegans*. *Front. Gene.* 3:48. doi: 10.3389/fgene.2012.00048

Roelofs, D., de Boer, M., Agamennone, V., Bouchier, P., Legler, J., and van Straalen, N. (2012). Functional environmental genomics of a municipal landfill soil. *Front. Gene.* 3:85. doi: 10.3389/fgene.2012.00085

Received: 28 September 2012; accepted: 28 September 2012; published online: 16 October 2012.

Citation: Stürzenbaum S, Aschner M and Freedman JH (2012) Toxicogenomics in non-mammalian species—Editorial. Front. Gene. ***3****:216. doi: 10.3389/fgene.2012.00216*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in
GENETICS

REVIEW ARTICLE
published: 19 April 2012
doi: 10.3389/fgene.2012.00063

Yeast toxicogenomics: genome-wide responses to chemical stresses with impact in environmental health, pharmacology, and biotechnology

Sandra C. dos Santos[1], Miguel Cacho Teixeira[1,2], Tânia R. Cabrito[1] and Isabel Sá-Correia[1,2]*

[1] *Institute for Biotechnology and Bioengineering, Centre for Biological and Chemical Engineering, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal*
[2] *Department of Bioengineering, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal*

Edited by:
Stephen Sturzenbaum, King's College London, UK

Reviewed by:
Ralph Menzel, Humboldt-Universität zu Berlin, Germany
Barry Panaretou, King's College London, UK

****Correspondence:***
Isabel Sá-Correia, Institute for Biotechnology and Bioengineering, Centre for Biological and Chemical Engineering, Department of Bioengineering, Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001 Lisbon, Portugal.
e-mail: isacorreia@ist.utl.pt

The emerging transdisciplinary field of Toxicogenomics aims to study the cell response to a given toxicant at the genome, transcriptome, proteome, and metabolome levels. This approach is expected to provide earlier and more sensitive biomarkers of toxicological responses and help in the delineation of regulatory risk assessment. The use of model organisms to gather such genomic information, through the exploitation of Omics and Bioinformatics approaches and tools, together with more focused molecular and cellular biology studies are rapidly increasing our understanding and providing an integrative view on how cells interact with their environment. The use of the model eukaryote *Saccharomyces cerevisiae* in the field of Toxicogenomics is discussed in this review. Despite the limitations intrinsic to the use of such a simple single cell experimental model, *S. cerevisiae* appears to be very useful as a first screening tool, limiting the use of animal models. Moreover, it is also one of the most interesting systems to obtain a truly global understanding of the toxicological response and resistance mechanisms, being in the frontline of systems biology research and developments. The impact of the knowledge gathered in the yeast model, through the use of Toxicogenomics approaches, is highlighted here by its use in prediction of toxicological outcomes of exposure to pesticides and pharmaceutical drugs, but also by its impact in biotechnology, namely in the development of more robust crops and in the improvement of yeast strains as cell factories.

Keywords: yeast model, toxicogenomics, molecular systems biology, genome-wide approaches, response to stress, toxicity mechanisms, predictive toxicology

YEAST TOXICOGENOMICS: GENOME-WIDE APPROACHES TO ELUCIDATE TOXICITY MECHANISMS AND GLOBAL STRESS RESPONSES IN YEAST

The transdisciplinary field of Toxicogenomics is defined as the merging of Omics approaches with toxicology to elucidate the response at the genome level to environmental stressors, drugs, and other toxicants (Hamadeh et al., 2002; Gomase and Tagore, 2008; North and Vulpe, 2010). The coupling of such approaches with classical toxicology studies combined with bioinformatics has the potential to provide a more comprehensive knowledge of the molecular and cellular effects of chemicals in biological systems than more traditional approaches. Understanding these complex responses is of paramount importance in fields ranging from Environmental Health to Pharmacology and drug development and to Biotechnology in general (Hamadeh et al., 2002; Guerreiro et al., 2003; Teixeira et al., 2007). Toxicity testing using animal models has a number of limitations that make it difficult to deal with the increasingly large number of potentially toxic compounds found in the environment and in the pharmaceutical industry for which toxicity data available is scarce (North and Vulpe, 2010). Mechanism-centered analysis represents an alternative approach to animal testing and, in this context, the yeast *Saccharomyces cerevisiae* can be an invaluable asset.

SACCHAROMYCES CEREVISIAE AS AN EXPERIMENTAL MODEL IN TOXICOGENOMICS

S. cerevisiae is a thoroughly established and widely used eukaryotic model for molecular and cellular biology studies. Yeast also plays a significant role in biotechnology, where it is used as a cell factory with diverse applications (Botstein and Fink, 2011). There are several inherent features that make yeast such a proficient model system: (1) it is a unicellular non-pathogenic microorganism with rapid and inexpensive growth, (2) it is amenable to genetic manipulation, (3) genome-wide analyses are easily implemented, with a vast array of experimental tools and biological material readily available, (4) it possesses a strikingly high-level of functional conservation within the human genome and other higher eukaryotes, and (5) it has the unique advantage of possessing functional information available for nearly every gene. The paradigm of research using *S. cerevisiae* changed with the publication of its genome sequence more than 15 years ago (Goffeau et al., 1996). A wealth of biological information has been

gathered over several years of post-genomic research facilitated by easy access to public databases (*Saccharomyces* Genome Database, SGD[1]; and YEASTRACT[2], among others). Research on *S. cerevisiae* pioneered the development of several post-genomic experimental approaches and computational tools, changing the field of yeast research with the application of innovative methodologies in functional genomics and proteomics (Mager and Winderickx, 2005; Smith et al., 2010; Botstein and Fink, 2011). In summary, yeast is a robust and inexpensive experimental platform where molecular studies difficult to carry out in more complex and less accessible eukaryotes are deeply facilitated (Wuster and Madan Babu, 2008; Smith et al., 2010; Botstein and Fink, 2011). Moreover, although many cytotoxic compounds act on their target organisms via physiological mechanisms that do not exist in yeast, many of the basic mechanisms underlying toxicity, adaptation, and resistance to chemical and environmental stresses are apparently conserved between yeast and phylogenetically distant organisms (Foury, 1997; Hohmann and Mager, 1997; Parsons et al., 2003; Mager and Winderickx, 2005).

In this review we will focus on the field of yeast toxicogenomics and how it can be exploited to obtain mechanistic insights into the action of drugs or toxicants with relevance in environmental health and agriculture (pesticides and environmental pollutants), medicinal and biomedical research (drugs used in the treatment of cancer, malaria, bacterial infections, etc.), and biotechnology (wine, beer, and other alcoholic fermentation processes, including the production of bio-ethanol, etc.; **Figure 1**), with a particular emphasis on the contribution of our research group to the field.

OMICS TOOLS APPLIED TO YEAST TOXICOGENOMICS

Variations in gene and protein expression or metabolite levels following exposure to a toxicant can contribute to identify the cellular components and pathways that are most relevant to a

[1]www.yeastgenome.com
[2]http://www.yeastract.com

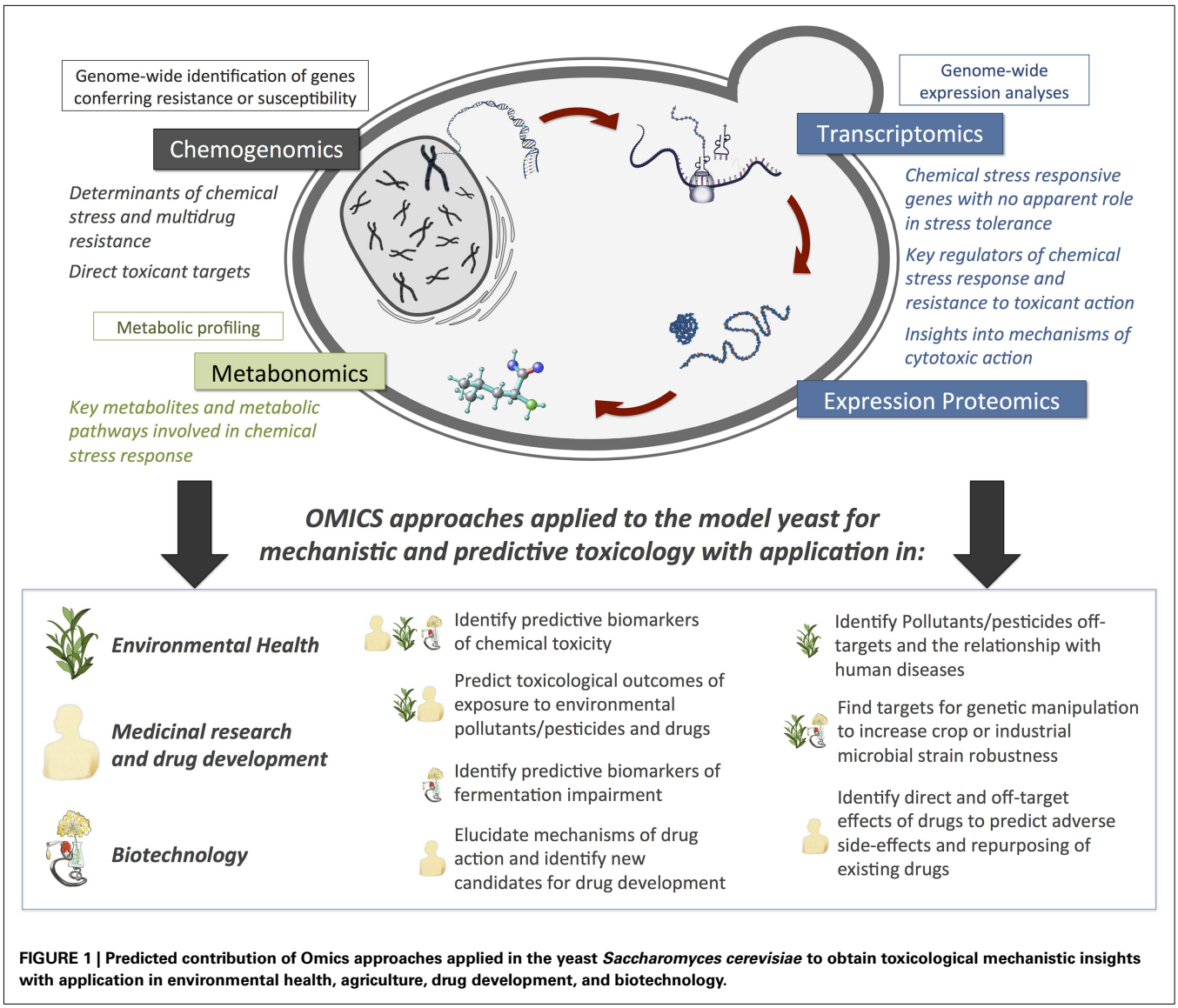


FIGURE 1 | Predicted contribution of Omics approaches applied in the yeast *Saccharomyces cerevisiae* to obtain toxicological mechanistic insights with application in environmental health, agriculture, drug development, and biotechnology.

toxicological response (Hamadeh et al., 2002; North and Vulpe, 2010). *S. cerevisiae* provides a privileged experimental system for obtaining an integrated assessment and genome-wide perspective of toxicity mechanisms (**Figure 1**), through the combination of transcriptomics and quantitative proteomics, for the evaluation of genome-wide expression changes occurring as part of the yeast response to environmental toxicants, metabonomics, for the study of the cell's small-molecule metabolite profile as the ultimate response to the toxicant, and chemogenomics, for the identification of cellular toxicity molecular targets (**Figure 1**). It is noteworthy that, as stand-alone, these analyses are not expected to provide decisive insights into the role of genes or proteins in a toxicological response, and should instead be integrated and coupled with suitable bioinformatics tools. Future research in the field will require the development of computational tools aiming the integration of high-throughput data and the collaborative activity of multidisciplinary teams with expertise in biological sciences, functional and comparative genomics, and bioinformatics.

A considerable amount of toxicity assessment data is available from yeast DNA microarrays. Upon toxicant exposure, yeast cells reprogram mRNA expression in order to adapt to the new environmental conditions, which produces a gene expression pattern (or "signature") characteristic of the compound (Simmons and Portier, 2002; Schwartz et al., 2004). These transcriptional signatures are shared by compounds with similar modes of action and can be used to infer mechanisms of action and predict toxicological outcomes of uncharacterized toxicants (Marton et al., 1998; Lamb et al., 2006; Teixeira et al., 2007). In fact, alterations of global gene expression can occur almost immediately after exposure, and the assessment of these changes could potentially provide an earlier and more sensitive biomarker of a toxic response than traditional toxicological methods (Hamadeh et al., 2002; Simmons and Portier, 2002). However, gene expression alone is not adequate to fully understand a toxicant's action and the resulting outcome (Hamadeh et al., 2002), since abnormalities in protein production and/or function are also expected to occur. As such, proteomics approaches are employed to identify the protein alterations associated with toxicant exposure (Teixeira et al., 2005, 2009b; Santos et al., 2009; Sa-Correia and Teixeira, 2010). Finally, genomic and proteomics methods do not address how the cell's dynamic metabolic status is affected by exposure to a toxicant. Metabonomics is an approach that allows the study of metabolic profiles based on the premise that toxicant-induced alterations will provide information on chemical toxicity (Nicholson et al., 2002). Data in yeast are still scarce, however it has already contributed to increase the current understanding of weak acid toxicity (Hasunuma et al., 2011; Lourenço et al., 2011).

FUNCTIONAL TOXICOGENOMICS USING YEAST GENE DELETION COLLECTIONS

Functional toxicogenomics is defined as the global study of the biological function of genes regarding the toxic effect of a compound or environmental stress, providing a direct link between gene and toxicant (Hamadeh et al., 2002; North and Vulpe, 2010). In this context, a major breakthrough with respect to studies on the mechanism of action and one of the main advantages of using yeast bioassays for toxicity assessment was the generation of heterozygous/homozygous diploid and haploid gene deletion collections (**Figure 2**; Scherens and Goffeau, 2004). Fitness is the primary phenotypic descriptor for yeast studies, and these collections enable comprehensive and systematic genetic screens that provide direct links between a specific gene and the requirement for that gene product function in the cellular response to a particular condition (Auerbach et al., 2005; Hoon et al., 2008b; Wuster and Madan Babu, 2008). The different types of fitness-based assays that can be used to identify toxicant-induced phenotypes include homozygous (knock-out deletion, gene dosage = 0%), haploinsufficiency (heterozygous deletion strains, gene dosage = 50%) and multicopy- and overexpression (gene dosage > 100%) screens (**Figure 2**; Hoon et al., 2008b). In homozygous profiling, each non-essential gene is knocked-out leading to complete loss-of-function. These screens are often used to identify genes that are important for conferring stress resistance, for example by genetic interaction with the toxicant's target. It is also applied to identify biological functions that are affected by a given stress, and infer from those the mechanisms of toxicological action. Finally, it is possible to screen for deletions that confer resistance to a compound in conditions that are deleterious to the wild-type strain. The deleted genes can be direct targets, or be involved in modifications or pathways that enable the compound's cytotoxic action. Naturally, one obvious downside of homozygous or haploid gene deletion strains is that only non-essential genes can be deleted, but the mechanism of action of many toxicants might target essential gene products. On the other hand, haploinsufficiency screens are based on the premise that lowering the gene dosage of the molecular target increases susceptibility, and thus enables direct identification of a toxicant's cellular target (see Hoon et al., 2008b; Wuster and Madan Babu, 2008). Libraries of double mutants have also been generated to uncover interactions between genes through synthetic lethality, which can be integrated with functional screens data to elucidate toxicity mechanisms and modes of action (Costanzo et al., 2010). Many bioinformatics tools are available to facilitate interpretation of the results. For example, Gene Ontology (GO) annotation and GO-based resources such as GOToolBox[3] (Martin et al., 2004) allow the identification of biological functions that are enriched within datasets, thus creating a "functional fingerprint" comparable to transcriptional signatures by expression profiling. These can be used to identify responses and pathways that are common to different classes of toxicants.

Our current literature survey found at least 80 publications that have used yeast deletion collections in the context of toxicity testing, in which hundreds of different stresses were explored and phenotypes were attributed to over 90% of all yeast genes (see **Table 1** for an overview). Remarkably, although many of these conditions had been previously scrutinized using classical methodologies or even DNA microarrays, many of the genes identified using deletion collections had not been known to be involved in the toxicological pathways investigated. One major contribution arising from the use of homozygous and heterozygous yeast deletion collections was the so-called "chemical genomic portrait of yeast" (Hillenmeyer et al., 2008). The authors carried out over 1100

[3] http://genome.crg.es/GOToolBox/

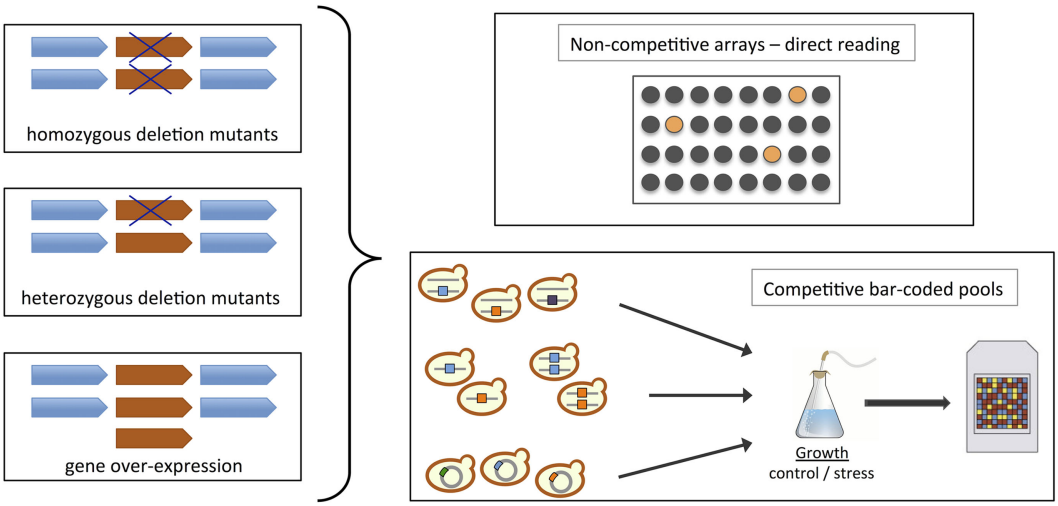


FIGURE 2 | Construction and screening of yeast collections. Schematic representation of methodologies and cell libraries available for chemogenomics testing in *S. cerevisiae* (homozygous or haploid deletion – gene dosage 0%, heterozygous deletion – gene dosage 50%, and overexpression – gene dosage > 100%; see Section "Functional Toxicogenomics using Yeast Gene Deletion Collections"; Auerbach et al., 2005; Hoon et al., 2008b; Wuster and Madan Babu, 2008; North and Vulpe, 2010; Smith et al., 2010). The fitness of strains upon chemical treatment is usually assessed in non-competitive arrays or in competitive bar-coded pools. In the first case, the toxicant can be added to a well plate and each mutant occupies a separate well; the effects are observed directly by comparison with wild-type strain fitness. In the second case, the screen is executed in a pooled format where uniquely tagged ("bar-coded") strains are grown together in the presence of a toxicant. Fitness is assessed by determining the abundance of the different mutant strains using microarrays coupled with a PCR strategy that amplifies the molecular bar-codes associated with each mutant. Strain depletion in the toxicant-treated pool indicates chemical hypersensitivity.

assays in the presence of chemical or environmental stress conditions, ranging from approved therapeutic drugs to compounds with uncertain activity. Functional enrichment analysis led to the identification of several biological functions that were required for growth in at least 20% of the assays performed, in particular endosome transport, vacuolar degradation, and transcription, a coordinated system that is conserved from yeast to humans (Hillenmeyer et al., 2008). The remarkable amount of data generated can be applied, for example, to predict toxicological modes of action for the compounds tested and other structurally related, or to predict the effects of synergies between some of these stresses. Other important contributions using yeast deletion collections will be described throughout this review.

INTEGRATION OF GENOME-WIDE DATA TO IDENTIFY MECHANISMS OF TOXICITY

To achieve a more complete understanding of the mechanisms of action and toxicological response, it is necessary to intersect and integrate the genome-wide data coming from the different approaches (**Figure 1**). Interestingly, several studies have now shown that the previous belief that genes that are up-regulated under a given stress are also required for growth under those conditions is often false (Giaever et al., 2002; Mettetal et al., 2008; Batova et al., 2010; Landstetter et al., 2010; dos Santos and Sá-Correia, 2011). Although this seems to suggest that deletion collections provide a better assessment of genes required for the response to a toxicant treatment, an integrated approach can lead to other important insights, for example the identification of key regulators of stress response. In yeast, the identification of the transcription factors predicted to underlie the transcriptomic response to stress is facilitated by freely accessible databases and computational tools such as YEASTRACT, a database focused on the delineation of yeast transcription regulatory networks, at a genomic scale (Teixeira et al., 2006b; Abdulrehman et al., 2011). Comprising more than 48,000 documented regulatory associations between yeast transcription factors and target genes (Abdulrehman et al., 2011), YEASTRACT offers one of the best platforms for the study and understanding of genomic regulation in an integrative perspective. The growing use of mathematical descriptions of the dynamical behavior of such networks (Teixeira et al., 2010a) is now allowing the prediction of the possible behaviors of biological systems under the action of pollutants, pesticides, drugs, and other chemical stresses, emphasizing the outstanding position of the yeast model in the field of toxicogenomics. Moreover, yeast toxicogenomics data can be integrated with studies from other sources, such as toxicological results obtained in the toxicant's target organism when appropriate, to obtain a view of a compound's toxicity at the systems level. This raises the potential of yeast toxicogenomics very high but poses additional challenges to the tasks of data integration and usage.

YEAST TOXICOGENOMICS APPLIED TO ENVIRONMENTAL POLLUTANTS AND XENOBIOTIC COMPOUNDS WIDELY USED IN AGRICULTURE

GENOME-WIDE RESPONSES TO ENVIRONMENTAL POLLUTANTS

The toxicological outcome of sudden or chronic exposure to environmental pollutants (e.g., metal ions or organic solvents, including benzene, or phenol derived compounds), is scarcely understood at the molecular and cellular levels. However the genome-wide yeast response to toxic concentrations of metal ions,

Table 1 | Selected publications in yeast toxicogenomics studies using deletion mutant collections.

Assay	Result	Reference
Quinine	Identification of 279 mutants that display hypersensitivity and 62 mutants that display resistance to quinine	dos Santos and Sá-Correia (2011)
	Identification of 43 quinine-sensitive strains and tryptophan uptake as a target of quinine toxicity	Khozoie et al. (2009)
214 psychoactive drugs	Identification of off-target effects	Ericson et al. (2008)
78 compounds with therapeutic activity	Identification of lanosterol synthase as a target of the antianginal drug molsidomine, and identification of rRNA processing exosome was identified as a potential target of the growth inhibitor 5-fluorouracil	Lum et al. (2004)
Imatinib mesylate	Identification of V-ATPase activity and vacuolar function as potential new imatinib targets	dos Santos and Sá-Correia (2009)
Antifungal agents	Identification of 20 strains displaying increased caspofungin sensitivity	Markovich et al. (2004)
12 bioactive compounds	Identification of multidrug sensitivity in yeast mutants lacking a functional V-ATPase	Parsons et al. (2004)
DNA-damaging anticancer agents	Identification of 231 mutants that display hypersensitivity and five mutants that display resistance to bleomycin	Aouida et al. (2004)
	Role of V-ATPase and cytosolic acidification in sensitivity to DNA-damaging agents such as cisplatin	Liao et al. (2007)
	Identification of 117 and 73 genes whose deletion results in increased or decreased resistance to tirapazamine	Hellauer et al. (2005)
	Identification of gene ERK5 as susceptible to cisplatin, methyl methane sulfonate and 5-fluorouracil, confirmed in human studies	Sletta et al. (2011)
Antimicrobials	No deletion strains are sensitive to amoxicillin, penicillin G, rifampin, or vancomycin. Two strains are sensitive to tetracycline sensitive and four to oxytetracycline	Blackburn and Avery (2003)
	Dermaseptin induces programmed cell death	Morton et al. (2007)
10 small therapeutic molecules	Identification of a chemical core structure shared among three compounds that inhibit the *ERG24* deletion strain	Giaever et al. (2004)
Nitrogen-containing bisphosphonates	Identification of tubulin cofactor B as a new target and *DBF4* as a key player in cytotoxicity	Bivi et al. (2009)
Introduction of human Huntingtin or α-synuclein fragments	Identification of 52 strains sensitive to mutant Huntingtin, 86 that are sensitive to α-synuclein, and one mutant sensitive to both	Willingham et al. (2003)
Library of 188 novel synthetic chemical compounds	Identification of potential targets and structure–activity relationships	Hoon et al. (2008a)
Endoplasmic reticulum stress	Identification of MAPK signaling pathways	Chen et al. (2005)
Fitness profiling under non-optimal growth conditions	Identification of genes required for growth in the presence of high salt or sorbitol or [60] galactose, or at pH8, or in minimal medium, or following nystatin treatment	Giaever et al. (2002)
High glucose	Identification of 44 susceptible strains	Teixeira et al. (2010b)
Ethanol	Identification of 250 determinants of resistance to ethanol and of gene *FPS1*	Teixeira et al. (2009a)
Weak acids	Identification of 650 determinants of resistance to acetic acid	Mira et al. (2010b)
	Identification of vacuolar function and of the RIM101 pathway in propionic acid resistance	Mira et al. (2009)
Oxidative stress	Identification of 394 strains sensitive to hydrogen peroxide and/or menadione	Tucker and Fields (2004)
	Identification of 456 mutants sensitive to at least one of five different types of oxidant	Thorpe et al. (2004)
Multiple environmental stresses and small molecules (1154 assays)	"A chemical genomic portrait of yeast: uncovering a phenotype for all genes"	Hillenmeyer et al. (2008)
Benzene	Confirmation by RNAi in human cells	Zhang et al. (2010)
Metals	Identification of determinants of resistance to cadmium, nickel, mercury, zinc, cobalt, and iron	Ruotolo et al. (2008)
	Identification of a regulatory crosstalk of iron and zinc regulons	Landstetter et al. (2010)
	Identification of mRNA mistranslation as a primary cause of cellular chromium toxicity	Holland et al. (2007)

(Continued)

Table 1 | Continued

Assay	Result	Reference
Fungicides	Identification of 286 determinants of resistance to mancozeb	Dias et al. (2010)
	Identification of intracellular superoxide production and oxidative stress as a mode of action of CTBT	Batova et al. (2010)
Killer toxin HM1	Identification of eight resistant strains including high-osmolarity glycerol pathways HOG1 and FPS1	Miyamoto et al. (2011)
Toxicants inducing Parkinson's disease	Identification of the multivesicular body pathway as an element of toxicity induced by MPP and paraquat	Doostzadeh et al. (2007)

such as nickel, cadmium, copper, chromium, arsenic, cobalt, manganese, and zinc has been studied by exploring chemogenomics and transcriptomics approaches. These studies led to the identification of several functional groups that are important in the yeast response to all or to part of the metal ions tested, mostly involved in sulfur amino acid and iron metabolism, oxidative stress response, vacuolar function, protein modification, transport and degradation, enzyme inactivation, cation and transition metal transport, mRNA decay, and DNA metabolism (Momose and Iwahashi, 2001; Jin et al., 2008; Ruotolo et al., 2008; Serero et al., 2008; Yasokawa et al., 2008; Takumi et al., 2010; Bleackley et al., 2011). The toxicological outcome of the exposure to agrochemicals, including herbicides and agricultural fungicides, is difficult to predict, since many times it takes years to develop. Genome-wide analyses in yeast have been successfully used to identify the genes responsible for response and resistance to stresses induced by pesticides of agricultural interest (Cabrito et al., 2009, 2011). These pesticides include the herbicide sulfometuron methyl (Jia et al., 2000), the dithiocarbamate fungicides mancozeb (Santos et al., 2009; Dias et al., 2010), thiuram, zineb and maneb (Kitagawa et al., 2003), the benzimidazole fungicide benomyl (Lucau-Danila et al., 2005), the pesticide lindane (Parveen et al., 2003), the herbicide 2,4-D (Teixeira et al., 2005, 2006a, 2007), and the herbicides paraquat and cyperquat (1-methyl-4-phenylpyridinium – MPP+; Doostzadeh et al., 2007). Toxicogenomics approaches have been applied to define and predict new toxicological outcomes of exposure to pesticides, including the agricultural fungicide mancozeb (Santos et al., 2009; Dias et al., 2010) and the herbicide 2,4-D (Teixeira et al., 2005, 2006a). A review of the main results of these two case studies follows.

TOXICOGENOMIC STUDIES FOCUSED ON THE AGRICULTURAL FUNGICIDE MANCOZEB

Mancozeb, a mixture of manganese- and zinc-ethylene-*bis*-dithiocarbamate (Mn:Zn, 9:1), is an agricultural fungicide with a broad spectrum of action and multiple cell targets, widely used against phytopathogenic fungi in several crops and vineyards (Maroni et al., 2000; Ballantyne, 2004). This compound displays low acute toxicity, however, in recent years, mounting evidence suggests that chronic exposure to this fungicide increases the probability of developing Parkinson's disease and certain forms of cancer (Belpoggi et al., 2002; Zhou et al., 2004; Calviello et al., 2006).

The early global response to mancozeb and the genome-wide resistance mechanisms established by *S. cerevisiae* were analyzed by expression proteomics (Santos et al., 2009) and chemogenomics (Dias et al., 2010; see **Figure 3** for a schematic representation of the main findings obtained in these studies with mancozeb). Interestingly, 70% of the proteins differently expressed in cells exposed to mancozeb (Santos et al., 2009) and 53% of the determinants of yeast resistance to the fungicide (Dias et al., 2010) have human orthologs. This is the case for proteins involved in V-ATPase function (Vma4 and Vma13), protein synthesis, folding (e.g., Kar2), protein degradation/proteasome sub-units (e.g., Pre3, Pre7, Pre8, Pre9, Nas2, Sem1, and Ubp6), and in the oxidative stress and anti-oxidant response (e.g., Tsa1, Tsa2, Glr1, Gsh1, Gsh2, Sod1, Sod2, and Yap1). Interestingly, V-ATPases are overexpressed in several metastatic cancers (Sennoune et al., 2004) and the overexpression of proteasome sub-units leads to increased survival rate of human cell lines following oxidative stress, due to a higher proteasome degradation of oxidized modified protein (Chondrogianni et al., 2005). Other human orthologs related with the oxidative stress response (*PRDX2*, *PRDX3*, *GSR*, *GCLC*, *GSS*, *SOD1*, and *SOD2*), have been involved in the onset and progression of neurodegenerative diseases, namely Parkinson's disease, by protecting the cell, acting as anti-oxidant agents and neuroprotectors (Chang et al., 2004; Ihara et al., 2005; Cumming et al., 2007). Besides, *PAK1*, *MAP2K1*, and *LCP1*, three other genes described as being involved in tumor development and invasiveness (Wang et al., 2006a,b), are orthologs of yeast determinants of mancozeb resistance *CLA4*, *PBS2*, and *SAC6* (Dias et al., 2010).

Using the YEASTRACT database, more than 90% of the genes that encode proteins up-regulated under mancozeb imposed stress were found to be known targets of Yap1 (Santos et al., 2009), the major oxidative stress regulator in yeast. Yap1 was also found, based on a chemogenomics study (Dias et al., 2010), to be a determinant of yeast resistance to the fungicide, and also to control the regulatory network underlying the up-regulation of the multidrug transporter encoding gene *FLR1* (Teixeira et al., 2010a; Monteiro et al., 2011). Remarkably, the human orthologs of Yap1, Jun, and Jdp2, are activated during acute and chronic phases of several neurodegenerative diseases (Shaulian and Karin, 2002), establishing a possible link between the predicted response to mancozeb toxicity and neurodegenerative disease progression through the Yap1/Jun/Jdp2 regulators. Although mancozeb was reported to induce reactive oxygen species (ROS) production as a consequence of mitochondrial dysfunction in mesencephalic cells (Domico et al., 2007), Dias et al. (2010) registered no increase in ROS production in yeast in response to the fungicide, probably due to the fact that in glucose fermenting yeast the level of mitochondrial

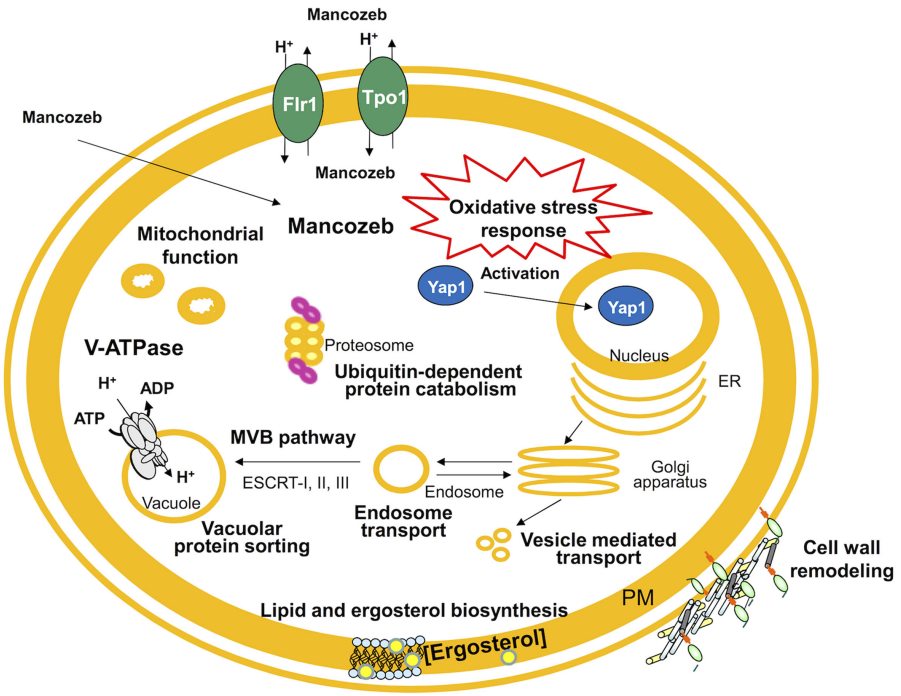


FIGURE 3 | Proposed model for the action of mancozeb in *S. cerevisiae* cells. This model results from the integration of yeast chemogenomics (Dias et al., 2010) and proteomics (Santos et al., 2009) approaches. The complex mancozeb-induced expression changes and mancozeb determinants of yeast resistance, were found to be related to oxidative stress, V-ATPase function, protein translation initiation and protein folding, disassembling of protein aggregates and degradation of damaged proteins, lipid and ergosterol biosynthesis, mitochondrial function, cell wall remodeling, and multidrug resistance transporters.

respiration is reduced. In the absence of mitochondrial electron leakage, which masks other aspects of mancozeb toxicity, a direct role of mancozeb in protein damage, as a thiol-reactive compound, was thus identified in the yeast model, with a possible parallel in human cells (Dias et al., 2010).

TOXICOGENOMIC STUDIES FOCUSED ON THE HERBICIDE 2,4-D

The herbicide 2,4-D is the most commonly used member of the auxin-like herbicide family. Although being considered relatively safe, exposure to 2,4-D has been linked to the development of non-Hodgkin lymphoma and sarcoma (Ibrahim et al., 1991). Furthermore, several cases of 2,4-D resistant weeds have been described[4], raising the need to increase herbicidal application rates, with the risk of reaching environmental toxic levels.

S. cerevisiae genome-wide approaches, including transcriptomics (Teixeira et al., 2006a) and expression proteomics (Teixeira et al., 2005), were used to gain insights into the mechanisms of response and resistance to 2,4-D. The results obtained in yeast have been used to guide studies on the molecular mechanisms underlying 2,4-D toxicity and response in plants and other higher eukaryotes. Interestingly, the quantification of the relative toxicity of 2,4-D compared to other herbicides measuring electrophysiological parameters and vitality of an animal nervous system (the frog's sciatic nerve) was comparable to the results obtained when yeast growth inhibition due the same herbicides was tested (Papaefthimiou et al., 2004). The early transcriptional response of yeast to 2,4-D includes the up-regulation of several genes involved in oxidative stress and anti-oxidant response (Teixeira et al., 2006a), which correlates with the increase in hydroxyl radicals and lipid peroxidation levels registered as a consequence of acute 2,4-D stress in yeast (Teixeira et al., 2004). Significantly, 2,4-D was found to induce fatty acid β-oxidation and also electron leakage from the mitochondrial respiration and catalase activity in rat cells (Bradberry et al., 2000, 2004). Given the fact that oxidative stress is associated with neurological diseases, aging, and cancer, these results can give clues on the effect of massive or repeated human exposure to the herbicide. In parallel to the implications of these studies in environmental health, results of herbicide resistance obtained in the yeast model have also proven useful to study herbicide toxicity mechanisms in plants, with expected impact in agriculture and plant biotechnology (Cabrito et al., 2009; Cabrito et al., unpublished results).

In acidified growth medium, yeast cells challenged with the herbicide 2,4-D suffer a strong reduction in their cytosolic and vacuolar pH (Fernandes et al., 2003; Simões et al., 2003), which is counteracted by the activation of the plasma and vacuolar membrane H^+-ATPases (Fernandes et al., 2003; Teixeira et al., 2005). In fact, auxins, similar to 2,4-D, were also shown to induce the activity of the model plant *Arabidopsis* plasma membrane H^+-ATPase, contributing to maintain the intracellular pH in plant roots (Shen et al., 2006).

[4]www.weedscience.org

Both the microarray and the proteomic analyses carried out indicated the up-regulation of a large number of genes involved in carbon and nutrient source metabolism and amino acid homeostasis in response to 2,4-D (Teixeira et al., 2005, 2006a), which can be a response to the damaging effects of the herbicide on the plasma membrane lipid and sterol organization and permeability, with consequences in terms of nutrient uptake. Teixeira et al. (2007) suggested that the auxin-like herbicide 2,4-D exerts a repressing effect over the TOR (target of rapamycin) pathway, recognized as a central controller of cell growth in all eukaryotes, which control the balance between protein synthesis and degradation in response to nutrient availability and quality (Crespo and Hall, 2002). Interestingly, in plants, namely *Arabidopsis thaliana*, the existence of a functional TOR kinase pathway has been reported and linked to cell proliferation and growth regulation, auxin being suggested as a possible signaling molecule in the pathway (Berkowitz et al., 2008).

Additionally, several other mechanisms of yeast resistance to the herbicide are similar to processes described in plants, namely those involved in cytosolic detoxification (Smart and Fleming, 1996; Teixeira and Sá-Correia, 2002; Ito and Gray, 2006; Teixeira et al., 2006a; Cabrito et al., 2009). Exposure to the herbicide leads to the up-regulation of several genes encoding MDR transporters in yeast (Teixeira et al., 2006a). The MDR transporters Tpo1, Pdr5, and Pdr18 were confirmed as determinants of yeast resistance to 2,4-D (Teixeira et al., 2004; Cabrito et al., 2011). 2,4-D induces the expression of *PDR5* orthologs, *SpTUR2*, in the aquaphyte *Spirodela polyrrhiza*, and *AtPDR9*, in the model plant *A. thaliana*, the latter catalyzing 2,4-D extrusion from root cells (Smart and Fleming, 1996; Ito and Gray, 2006). Recently, Tpo1 homologs encoding putative plasma membrane MFS transporters from *A. thaliana* were analyzed by Cabrito et al. (2009) for a possible role in 2,4-D resistance. *At5g13750*/*ZIFL1* transcript levels were found to increase in 2,4-D stressed plants. The functional heterologous expression of *AtZIFL1* in yeast was found to confer increased resistance to the herbicide in wild-type and Δ*tpo1* cells, through the reduction of the intracellular concentration of 2,4-D counter-ion (Cabrito et al., 2009). Interestingly, Zifl1 was the first eukaryotic transporter of the MFS identified as a multidrug resistance determinant, opening an entirely new field of research with promising repercussions in medicine, biotechnology, and agriculture.

These case studies highlight the similarities of toxicological effects of these pesticides from yeast to higher eukaryotes, such as humans and plants. Hence, the use of the yeast model system is expected to continue to contribute to the understanding of the molecular mechanisms underlying pesticide toxicity in more complex and less easily accessible eukaryotes.

YEAST TOXICOGENOMICS IN BIOMEDICAL AND MEDICINAL RESEARCH

YEAST IN DRUG DEVELOPMENT AND PHARMACOLOGICAL RESEARCH

The use of yeast as a eukaryotic model is particularly important in the field of medicinal research and drug discovery (Simon and Bedalov, 2004; Mager and Winderickx, 2005; Menacho-Marquez and Murguia, 2007; Hoon et al., 2008b). Approximately 17% of all yeast genes are members of orthologous gene families associated with human disease, and for the majority of these genes their mammalian homolog is functional in yeast and complements the yeast deletion mutant (Foury, 1997; Heinicke et al., 2007). Modern medicine faces the challenge of developing safer and more effective therapies to treat human diseases; toxicogenomics represents a new paradigm in drug development and risk assessment, particularly in mechanistic and predictive toxicology as well as in biomarker discovery (Guerreiro et al., 2003; Gomase and Tagore, 2008). To better evaluate drug-associated adverse effects, the drug's specific mode of action needs to be elucidated first. However, the successful identification of drug targets and mechanisms of action requires a prior understanding of the high-level functional interaction between the key components of cells and systems (Guerreiro et al., 2003; North and Vulpe, 2010).

The primary advantage of yeast in drug discovery is the contribution to identify the mechanisms of action of compounds when they are unknown. Moreover, yeast is currently the only system where it is possible to assess all targets in the cell simultaneously and *in vivo* (Smith et al., 2010). These strategies are important not only to identify new drugs for further development, but also to find new uses for already approved drugs. The contribution of yeast toxicogenomics to this field takes on a pharmacogenomics perspective. This emerging post-genomic discipline is characterized by having a genome-wide perspective on the action of drugs, making use of global approaches to identify candidate drug targets and off-target effects (Swen et al., 2007; Wang and Weinshilboum, 2008; Ruderfer et al., 2009). Off-target effects are more difficult to detect and are often the cause of deleterious side effects. Such effects can arise, for example, when the direct binding interaction between one protein and a target results in an interaction with a second protein (Parsons et al., 2006; Ericson et al., 2008; West et al., 2010). In recent years, yeast fitness screens and other Omics approaches have been employed to search for new targets and elucidate the mode of action of different compounds, including anticancer drugs, antimalarials, antimicrobials, and other bioactive compounds (see **Table 1**). Yeast methodologies have also been used to study human disease genes and model human disorders (Steinmetz et al., 2002; Outeiro and Lindquist, 2003; Willingham et al., 2003; Gammie et al., 2007; Yuen et al., 2007), screen for new drugs to treat cancer, obesity, prion disease, etc. (Hammonds et al., 1998; Bach et al., 2003; Tribouillard et al., 2006; Marjanovic et al., 2010), and predict drug responses in relation to indicators such as genotype and expression levels (Perlstein et al., 2007; Ruderfer et al., 2009; Chen et al., 2010).

In an illustrative example, Ericson et al. (2008) identified 81 psychoactive drugs that affected yeast fitness at the level of evolutionarily conserved cellular processes such as secretion, protein folding, RNA processing, and chromatin structure. These processes might constitute secondary drug targets and point to additional, previously uncharacterized mechanisms of action for these drugs in humans. Information of this nature can be used to guide the rational design of new compound derivatives with fewer side effects and for tailoring drug treatment to individual patient genotypes, in a personalized medicine perspective. In another interesting study, yeast was applied to identify secondary targets of nitrogen-containing bisphosphonates, drugs commonly used to treat bone-related disorders including cancer (Bivi et al., 2009). The only known target of these compounds was farnesyl

pyrophosphate synthase, but the authors showed that the action of this drug in yeast involves additional mechanisms, predominantly at the level of DNA damage, and cytoskeleton dynamics. The dataset obtained from the yeast screen was validated in a mammalian system, and confirmed the involvement of new biological processes and specific genes that represent potential new targets for compounds with antitumor activity (Bivi et al., 2009). Genome-wide expression patterns were applied for target validation and identification of secondary drug targets of FK506, an immunosuppressant drug known to inhibit the protein phosphatase calcineurin (Marton et al., 1998). The authors identified a transcriptional signature of FK506 and found that it closely resembled that of the calcineurin null mutant. However, an increase of the drug dosage resulted in a different expression profile, suggesting that targets other than calcineurin mediated FK506 effects. Using a similar approach, Hughes et al. (2000) compared the transcriptomes of 13 compounds with those of 286 deletion mutants representing a variety of functional classes, identifying novel targets and off-targets for several drugs.

Yeast deletion mutants have also been used to identify targets for 78 compounds with diverse chemical structures and therapeutic relevance (Lum et al., 2004), resulting in identification of a lanosterol synthase in the sterol biosynthetic pathway as a target of the antianginal drug molsidomine, while the rRNA processing exosome was singled-out as a potential target of the cell growth inhibitor 5-fluorouracil. A similar study screened 10 different compounds (including anticancer and antifungal agents) against a deletion collection in 80 competitive pool assays (Giaever et al., 2004) and identified a chemical core structure that is shared by three compounds that inhibit the *ERG24* deletion strain, suggesting that cells respond similarly to compounds of related structure. Several compounds with *in vivo* activity against yeast and mammalian prions were identified in a yeast-based screen (Tribouillard et al., 2006), establishing this method as an economic and efficient high-throughput approach to identify novel prion inhibitors or to carry out comprehensive structure-activity studies for already isolated anti-mammalian prion drugs. The results also seem to highlight the extensive conservation of biochemical pathways controlling prion formation and/or maintenance from yeast to human.

GENOME-WIDE RESPONSES AND DETERMINANTS OF RESISTANCE TO ANTIMALARIAL DRUGS

A comprehensive example of yeast toxicogenomics is its application to the study of the antimalarial quinine. It was the first effective treatment against malaria and is still the drug of choice for chloroquine-resistant or severe malaria (WHO, 2006; Alkadi, 2007), but the molecular mechanism of action and toxicity of quinine in the *Plasmodium* parasite are still a topic of debate (Fitch, 2004). In a clear demonstration of the merging of toxicogenomics with pharmacogenomics, Khozoie et al. (2009) screened a yeast deletion collection to address the mode of action of quinine and its adverse side effects. The authors observed an enrichment of genes involved in tryptophan biosynthesis, and additional assays seemed to demonstrate that quinine causes tryptophan starvation and that dietary tryptophan supplements could help to avert the toxic effects of quinine (Khozoie et al., 2009). Another chemogenomics-based analysis identified for the first time several genes encoding ribosome protein sub-units whose deletion leads to increased quinine resistance (dos Santos and Sá-Correia, 2011). The particular involvement of phosphate signaling and transport in quinine tolerance was also studied, with indications that phosphate-starvation responsive genes are activated in response to quinine. *P. falciparum* homology searches identified several relevant functional homologs in the parasite, suggesting that the quinine targets identified in the yeast model are good candidates to be transposed and explored in a parasitic context (**Figure 4A**). The transcriptomic analysis of the yeast early response to quinine unveiled glucose de-repression reprogramming (**Figure 4A**; dos Santos et al., 2009). Moreover, quinine was shown to inhibit the uptake of glucose into yeast cells following a competitive inhibition kinetic model. These findings have an important parallel in the malaria parasite, where glucose uptake is vital and mediated by PfHT1, a single-copy transporter homologous to yeast's hexose HXT transporters (Woodrow et al., 1999; Saliba et al., 2004). The mechanism by which quinine enters and accumulates in the parasitic cell is not clear, but it is believed that a carrier-mediated import system is involved (Sanchez et al., 2008). The yeast results suggested PfHT1 as a quinine target, possibly involved in quinine uptake into the cell and depriving the parasite of glucose (dos Santos et al., 2009).

Yeast has also been used as model for mechanistic studies with other antimalarial drugs, namely mefloquine (Delling et al., 1998), chloroquine (Emerson et al., 2002), artemisinins (Li et al., 2005; Alenquer et al., 2006), and quinidine (Delling et al., 1998; Nunes et al., 2001; Felder et al., 2002; Tenreiro et al., 2002; Vargas et al., 2004).

GENOME-WIDE RESPONSES AND DETERMINANTS OF RESISTANCE TO ANTICANCER DRUGS

Anticancer drugs have been the subject of several yeast pharmacogenomics studies. For example, at least three genome-wide screens for cisplatin susceptibility have been performed using yeast deletion collections (Wu et al., 2004; Huang et al., 2005; Liao et al., 2007). The vacuolar H^+-ATPase (V-ATPase) and its action in cytoplasmic pH maintenance was shown to have an important role in sensitivity to this drug (Liao et al., 2007), an important result given how cisplatin cytotoxicity is potentiated by synergistic treatment with a V-ATPase inhibitor in human cell lines (Murakami et al., 2001). In a demonstration of how results obtained in the yeast model can be extended to human cells, Schenk et al. (2001, 2002) identified *SKY1* as a cisplatin sensitivity gene whose disruption conferred cisplatin resistance in yeast; later work in human ovarian carcinoma cell lines showed that inactivation of its human homolog, *SRPK1*, induces cisplatin resistance as well. The paradigmatic anticancer drug imatinib mesylate (Glivec, Novartis) was also studied in yeast. Imatinib is a selective tyrosine kinase inhibitor used in chronic myeloid leukemia with outstanding results, but drug resistance is an arising problem (Quintás-Cardama et al., 2009; Volpe et al., 2009). Fifty-one genes emerged as determinants of resistance to imatinib from the screening of a yeast deletion mutant collection, including 83% human homologs (dos Santos and Sá-Correia, 2009). Imatinib

was also shown to act as a potent inhibitor of the highly conserved yeast V-ATPase, both *in vivo* and *in vitro* (dos Santos and Sá-Correia, 2009; dos Santos and Sá-Correia, unpublished results), suggesting that vacuolar function is a novel imatinib target. In fact, V-ATPase activity has been shown to be necessary to limit the deleterious effects of several drugs in yeast, namely the genotoxic tirapazamine and cisplatin, weak acids, mancozeb, toxins, and others (Parsons et al., 2004; Hellauer et al., 2005; Liao et al., 2007; Mira et al., 2009; Dias et al., 2010). However, in none of these studies vacuolar acidification was affected directly, which is in striking contrast with imatinib, where a clear loss of vacuolar acidification in cells treated with this drug was reported (dos Santos and Sá-Correia, 2009). Quantitative- and phosphoproteomic analyses identified 18 proteins altered at the content level or displaying imatinib-repressed phosphorylation (dos Santos and Sá-Correia, unpublished results). All these proteins have human homologs and are mainly involved in glycolytic pathways, translation, and protein folding. A role for HSP70 proteins in the response to imatinib in yeast, as well as decreased glycolysis as a metabolic marker of imatinib action were suggested, consistent with findings from studies in human cell lines (Nowicki et al., 2003; Ferrari et al., 2007; Pocaly et al., 2008; Kominsky et al., 2009). The previously proposed effect of imatinib as an inhibitor of V-ATPase function was supported by the identification of an under-expressed subunit of this complex in imatinib-stressed yeast cells (**Figure 4B**).

YEAST TOXICOGENOMICS TOOLS APPLIED TO OVERCOME FERMENTATION-RELATED STRESSES RELEVANT IN INDUSTRIAL BIOTECHNOLOGY

S. cerevisiae has been used for millennia in fermentation processes behind wine, beer, and spirits production. Its remarkable capability of carrying out alcoholic fermentation very efficiently, leading to the production of very high ethanol titers, has also launched *S. cerevisiae* as a preferable host for the production of bio-ethanol, as a renewable biofuel. Furthermore, genetic engineering and synthetic biology methods have allowed the development of *S. cerevisiae* strains to be used as cell factories for the production of a number of interesting biomolecules, of biotechnological and pharmaceutical interest. In all these industrial processes, yeast cells have to cope with stressful environmental conditions, including chemical stress coming from the raw material composition, and from the accumulation of ethanol, weak acids, and other toxic byproducts of the yeast metabolism (Teixeira et al., 2011).

Toxicogenomics tools have been used with success to characterize the toxicological outcome of yeast exposure to fermentation-related chemical stress inducers. Such an approach has the potential to elucidate the mechanisms of yeast tolerance to fermentation stressors, thus providing clues on how to improve process conditions and to engineer yeast strains to increase fermentation yield. A particularly good example of the use of such an approach can be found in the efforts to improve bio-ethanol production process whose efficiency is compromised by several stress factors throughout fermentation. First generation bio-ethanol production relies on the use of very high gravity (VHG) media, highly enriched in fermentable carbon sources, which induce osmotic stress in the beginning of the fermentation process. In recent years, the interest in the production of bio-ethanol from agricultural lignocellulosic residues, the so-called second-generation bio-ethanol, has gained strength. These residues appear to be preferable for a sustainable large-scale production of bio-ethanol since they are largely available and do not compete with food resources (van Maris et al., 2006). However, in lignocellulosic hydrolysate fermentations, the first phase of the process is hindered by the presence of toxic concentrations of inhibitory side-products of the raw material hydrolysis process, including acetic acid, furfural, and vanillin. During the later stages of alcoholic fermentation for first or second-generation bio-ethanol production, the accumulation of toxic concentrations of ethanol and weak organic acids are responsible for lower fermentation productivity and, eventually, for fermentation arrest, limiting the final ethanol concentration achieved. Having this in mind, transcriptomics, expression proteomics, and metabolomics approaches have been used to study the expression and metabolic profile of yeast cells exposed to sudden stress induced by ethanol (Alexandre et al., 2001; Hirasawa et al., 2007; Stanley et al., 2010), weak acids (Mira et al., 2009, 2010a; Hasunuma et al., 2011), high sugar concentrations (Erasmus et al., 2003; Pham et al., 2006), but also throughout industrial or industrial-like fermentation processes (Devantier et al., 2005; Marks et al., 2008; Ding et al., 2009; Li et al., 2010). Using such toxicogenomics tools, the involvement of three signaling pathways mediated by the transcription factors War1 (Schüller et al., 2004), Haa1 (Mira et al., 2010a), and Rim101 (Mira et al., 2009) in the yeast response to weak acids was recently characterized. Weak acid toxicity mechanisms are additionally interesting in this context, given that they are widely used as food-preservatives against spoilage yeasts and molds and because *S. cerevisiae* is arising as an interesting alternative host for the industrial production of carboxylic acids, being more tolerant to their toxicity than currently used bacterial systems (Abbott et al., 2009). In this context, the use of transcriptomic and chemogenomic (Mollapour et al., 2004; Schüller et al., 2004) screenings focused on the food preservative sorbate has further highlighted the importance of vacuolar acidification and redox homeostasis for weak acid stress resistance. Clues on the mechanisms of weak acid toxicity have also come from metabolomics approaches (Hasunuma et al., 2011; Lourenço et al., 2011). For example, in a *S. cerevisiae* strain, modified through metabolic engineering tools to be able to ferment xylose, metabolomics data revealed that metabolites involved in the pentose phosphate pathway (PPP) were significantly accumulated by the addition of acetate during xylose fermentation, suggesting that acetic acid slows down the flux of the pathway (Hasunuma et al., 2011). Based on this result, a gene encoding a PPP-related enzyme was overexpressed in the xylose-fermenting yeast, conferring increased ethanol productivity in the presence of acetic acid (Hasunuma et al., 2011).

A particularly successful approach, in this context, has proven to be the use of the yeast deletion mutant collections to identify the determinants of yeast resistance to all these stresses, individually, or in combination. This chemogenomics strategy was used to unveil the global mechanisms and determinants of yeast resistance to stresses occurring during alcoholic fermentation, in particular to high ethanol (Fujita et al., 2006; van Voorst et al., 2006; Teixeira

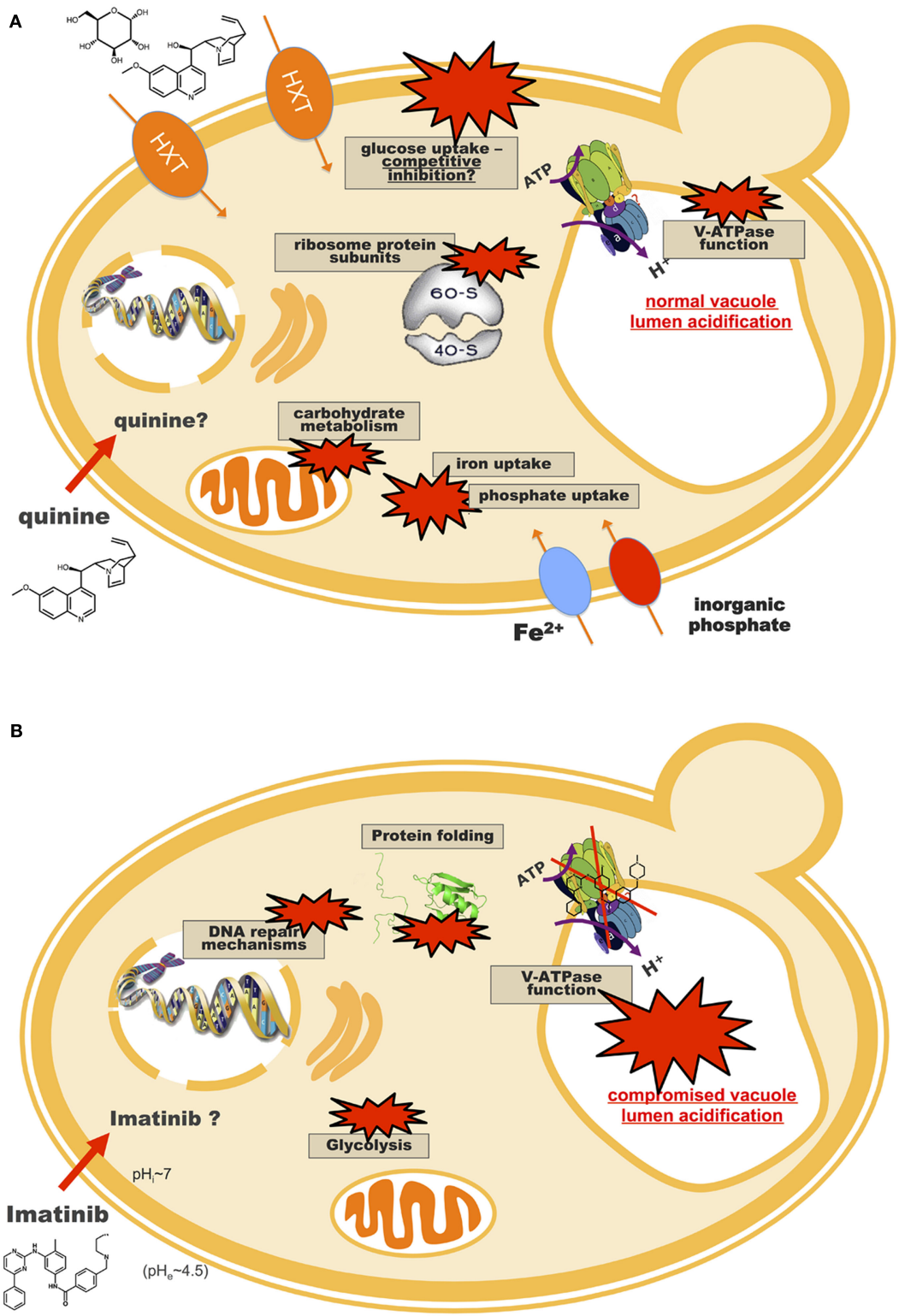


FIGURE 4 | Proposed model for the action of (A) quinine and (B) imatinib in *S. cerevisiae* cells. These models result from the integration of chemogenomics, transcriptomics and proteomics approaches (dos Santos and Sá-Correia, 2009; dos Santos et al., 2009; dos Santos and Sá-Correia, 2011; dos Santos and Sá-Correia, unpublished results), suggesting new targets and modes of action for quinine and imatinib that possess extensive functional conservation in the organisms of interest, *Plasmodium falciparum*, and human cells, respectively. The most important results are the identification of PfHT1 as a potential target of quinine, as well as the vacuolar H^+-ATPase (V-ATPase) as a target of imatinib (see Genome-wide Responses and Determinants of Resistance to Antimalarial Drugs and Genome-wide Responses and Determinants of Resistance to Anticancer Drugs).

et al., 2009a; Yoshikawa et al., 2009), high glucose (Teixeira et al., 2010b), and acetic acid (Mira et al., 2010b) concentrations. Based on these results, the aquaglyceroporin Fps1 was proposed as a major determinant of yeast resistance to ethanol and shown to play a role in reducing intracellular ethanol accumulation. The manipulation of *FPS1* expression levels was found to result in an increase of the final concentration of ethanol produced under conditions close to high gravity industrial fermentation (Teixeira et al., 2009a). The chemogenomics analysis of acetic acid stress resistance has further highlighted, among other things, the importance the potassium concentration in this process, suggesting that the control of potassium levels in the fermentation broth may be crucial to increase fermentation performance (Mira et al., 2010b).

Since it is the combination of all fermentation stresses, and not their individual effect, that affects yeast fermentative capacity, a recent study (Pereira et al., 2011) focused on the integration of several chemogenomics studies to identify the few genes that are able to increase yeast tolerance to: (1) ethanol (Fujita et al., 2006; van Voorst et al., 2006; Teixeira et al., 2009a; Yoshikawa et al., 2009), acetate (Mira et al., 2010b), and high glucose (Teixeira et al., 2010b) concentrations, (2) ethanol (Fujita et al., 2006; van Voorst et al., 2006; Teixeira et al., 2009a; Yoshikawa et al., 2009), acetate (Mira et al., 2010b), and vanillin (Endo et al., 2008), and (3) ethanol (Fujita et al., 2006; van Voorst et al., 2006; Teixeira et al., 2009a; Yoshikawa et al., 2009), acetate (Mira et al., 2010b), and furfural (Gorsich et al., 2006). The effect of the deletion of these genes in VHG or biomass fermentation performance, respectively, was evaluated. The identified genes, including *BUD31*, *HPR1*, *PHO85*, *VRP1*, and *YGL024w*, found to contribute to improved performances in VHG, and *ERG2*, *PRS3*, *RAV1*, *RPB4*, and *VMA8*, required for improved performance in wheat straw hydrolysate fermentations, stand as preferential targets for genetic engineering in order to generate more robust industrial yeast strains, better suited for industrial bio-ethanol production (Pereira et al., 2011).

CONCLUDING REMARKS

Functional genomics approaches applied to the simple single cell model *S. cerevisiae* have deeply modified the understanding of drug/chemical stress resistance and response mechanisms. The gathered data reinforces the idea that observations made in yeast are very likely to have a parallel in more complex eukaryotes, validating the use of this model organism in the context of toxicogenomics. The extrapolation of these results to higher eukaryotes, which has been so far carried out with relative success, is obviously of paramount importance.

This article emphasizes how yeast toxicogenomics enables a rapid and reproducible assessment of the mechanisms of toxicity of, and resistance to, many chemicals, requiring small amounts of growth medium and compound under testing, leading to reduced costs and reduced toxic wastes, effectively contributing to reduce, refine, and replace (3R) the use of animals in toxicological testing of pesticides (Gad, 1990). However, and in spite of these numerous advantages, yeast is not without its drawbacks. The simplicity of yeast is a disadvantage, since unicellularity is not conducive to study complex phenomena and cannot provide data on organ or tissue-specific toxicity. Moreover, studies in yeast do not provide accurate indicators for determination of toxic doses of a compound, since *S. cerevisiae* is usually much more tolerant to high doses of toxicants than higher eukaryotic cells. This is likely due to the barrier presented by the cell wall as well as the expression of numerous active efflux pumps and detoxification mechanisms that are highly abundant in yeast cells, who together make it virtually impossible to know the real concentration that is acting on the toxicant targets (Sá-Correia et al., 2009; Smith et al., 2010). Another important limitation is the possible absence of adequate molecular targets in yeast, since many cytotoxic compounds act in their target organisms via physiological mechanisms that do not exist in yeast (Foury, 1997; Hohmann and Mager, 1997; Parsons et al., 2003; Mager and Winderickx, 2005). Furthermore, the finding of homology between yeast and human genes does not necessarily imply that they are orthologs (Foury, 1997; Heinicke et al., 2007), demanding experimental verification of which (if any) is the true homolog of interest.

Nonetheless, the exploitation of high-throughput technologies and the global molecular analyses of the effects of drugs and other chemicals using the yeast model are revealing previously unsuspected on unknown molecular targets or adverse effects. It has proven a valuable first platform for the screening and prediction of the toxicological outcome of new or still unstudied drugs/chemicals and for the study of toxicity mechanisms. These first analyses are very much facilitated by the exploitation of the yeast model, given that signaling and regulatory pathways are highly conserved and may uncover the interactions of a chemical with its expected and unexpected gene/protein/metabolite targets. The use of yeast, together with other model organisms, and cross-species comparison of important genes/proteins in the toxicological response will facilitate the understanding of the response of an organism to toxic insults at a systems level. This strategy is expected to allow the description of all toxicological interactions occurring in a living system under chemical stress and the prediction of action of similar compounds in other species.

In conclusion, and based on the results reviewed in this paper, we believe that yeast stands out as an unavoidable and preferential system for toxicogenomics studies, through a combination of large-scale experimental approaches and expertise with biological and computational tools.

ACKNOWLEDGMENTS

Research carried out at IBB-IST and overviewed in this paper was financially supported by the EU (HERBICBIOREM project QLK3-CT-1999-00041) and by FEDER and "Fundação para a Ciência e a Tecnologia" (FCT, Portugal) – contracts POCTI/AGG/38110/2001, POCTI/AGR/45347/2002, POCTI/BME/46526/2002, POCTI/BIO/56838/2004, PTDC/SAU-FCF/71760/2006, PTDC/BIO/72063/2006, PTDC/AGR-AAM/67858/2006, PTDC/AGR-AAM/102967/2008, PTDC/AGR-ALI/102608/2008, and ERA-IB/0002/2010, and postdoctoral fellowships to Sandra C. dos Santos (SFRH/BPD/75483/2010) and Tânia R. Cabrito (SFRH/BPD/81221/2011).

REFERENCES

Abbott, D. A., Zelle, R. M., Pronk, J. T., and Van Maris, A. J. (2009). Metabolic engineering of *Saccharomyces cerevisiae* for production of carboxylic acids: current status and challenges. *FEMS Yeast Res.* 9, 1123–1136.

Abdulrehman, D., Monteiro, P. T., Teixeira, M. C., Mira, N. P., Lourenco, A. B., dos Santos, S. C., Cabrito, T. R., Francisco, A. P., Madeira, S. C., Aires, R. S., Oliveira, A. L., Sa-Correia, I., and Freitas, A. T. (2011). YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in *Saccharomyces cerevisiae* through a web services interface. *Nucleic Acids Res.* 39, D136–D140.

Alenquer, M., Tenreiro, S., and Sá-Correia, I. (2006). Adaptive response to the antimalarial drug artesunate in yeast involves Pdr1p/Pdr3p-mediated transcriptional activation of the resistance determinants TPO1 and PDR5. *FEMS Yeast Res.* 6, 1130–1139.

Alexandre, H., Ansanay-Galeote, V., Dequin, S., and Blondin, B. (2001). Global gene expression during short-term ethanol stress in *Saccharomyces cerevisiae*. *FEBS Lett.* 498, 98–103.

Alkadi, H. O. (2007). Antimalarial drug toxicity: a review. *Chemotherapy* 53, 385–391.

Aouida, M., Pagé, N., Leduc, A., Peter, M., and Ramotar, D. (2004). A genome-wide screen in *Saccharomyces cerevisiae* reveals altered transport as a mechanism of resistance to the anticancer drug bleomycin. *Cancer Res.* 64, 1102–1109.

Auerbach, D., Arnoldo, A., Bogdan, B., and Fetchko, M. (2005). Drug discovery using yeast as a model system: a functional genomic and proteomic view. *Curr. Proteomics* 2, 1–13.

Bach, S., Talarek, N., Andrieu, T., Vierfond, J. M., Mettey, Y., Galons, H., Dormont, D., Meijer, L., Cullin, C., and Blondel, M. (2003). Isolation of drugs active against mammalian prions using a yeast-based screening assay. *Nat. Biotechnol.* 21, 1075–1081.

Ballantyne, B. (2004). "Toxicology in fungicides," in *Pesticide Toxicology and International Regulation*, eds T. C. Marrs and B. Ballantyne (Chichester: John Wiley and Sons), 191–303.

Batova, M., Klobucnikova, V., Oblasova, Z., Gregan, J., Zahradnik, P., Hapala, I., Subik, J., and Schuller, C. (2010). Chemogenomic and transcriptome analysis identifies mode of action of the chemosensitizing agent CTBT (7-chlorotetrazolo[5,1-c]benzo[1,2,4]triazine). *BMC Genomics* 11, 153. doi:10.1186/1471-2164-11-153

Belpoggi, F., Soffritti, M., Guarino, M., Lambertini, L., Cevolani, D., and Maltoni, C. (2002). Results of long-term experimental studies on the carcinogenicity of ethylene-bis-dithiocarbamate (Mancozeb) in rats. *Ann. N. Y. Acad. Sci.* 982, 123–136.

Berkowitz, O., Jost, R., Pollmann, S., and Masle, J. (2008). Characterization of TCTP, the translationally controlled tumor protein, from *Arabidopsis thaliana*. *Plant Cell* 20, 3430–3447.

Bivi, N., Romanello, M., Harrison, R., Clarke, I., Hoyle, D. C., Moro, L., Ortolani, F., Bonetti, A., Quadrifoglio, F., Tell, G., and Delneri, D. (2009). Identification of secondary targets of N-containing bisphosphonates in mammalian cells via parallel competition analysis of the barcoded yeast deletion collection. *Genome Biol.* 10, R93.

Blackburn, A. S., and Avery, S. V. (2003). Genome-wide screening of *Saccharomyces cerevisiae* to identify genes required for antibiotic insusceptibility of eukaryotes. *Antimicrob. Agents Chemother.* 47, 676–681.

Bleackley, M. R., Young, B. P., Loewen, C. J., and Macgillivray, R. T. (2011). High density array screening to identify the genetic requirements for transition metal tolerance in *Saccharomyces cerevisiae*. *Metallomics* 3, 195–205.

Botstein, D., and Fink, G. R. (2011). Yeast: an experimental organism for 21st century biology. *Genetics* 189, 695–704.

Bradberry, S. M., Proudfoot, A. T., and Vale, J. A. (2004). Poisoning due to chlorophenoxy herbicides. *Toxicol. Rev.* 23, 65–73.

Bradberry, S. M., Watt, B. E., Proudfoot, A. T., and Vale, J. A. (2000). Mechanisms of toxicity, clinical features, and management of acute chlorophenoxy herbicide poisoning: a review. *J. Toxicol. Clin. Toxicol.* 38, 111–122.

Cabrito, T. R., Teixeira, M. C., Duarte, A. A., Duque, P., and Sa-Correia, I. (2009). Heterologous expression of a Tpo1 homolog from *Arabidopsis thaliana* confers resistance to the herbicide 2,4-D and other chemical stresses in yeast. *Appl. Microbiol. Biotechnol.* 84, 927–936.

Cabrito, T. R., Teixeira, M. C., Singh, A., Prasad, R., and Sa-Correia, I. (2011). The yeast ABC transporter Pdr18 (ORF YNR070w) controls plasma membrane sterol composition, playing a role in multidrug resistance. *Biochem. J.* 440, 195–202.

Calviello, G., Piccioni, E., Boninsegna, A., Tedesco, B., Maggiano, N., Serini, S., Wolf, F. I., and Palozza, P. (2006). DNA damage and apoptosis induction by the pesticide Mancozeb in rat cells: involvement of the oxidative mechanism. *Toxicol. Appl. Pharmacol.* 211, 87–96.

Chang, D. J., Lee, S. H., Lim, C. S., Jang, D. H., Lee, C. H., Lee, Y. D., and Kaang, B. K. (2004). Thiol oxidation-mediated cell death in *Aplysia* cultured sensory neurons. *Brain Res.* 1007, 71–77.

Chen, B.-J., Causton, H. C., Mancenido, D., Goddard, N. L., Perlstein, E. O., and Pe'er, D. (2010). Harnessing gene expression to identify the genetic basis of drug resistance. *Mol. Syst. Biol.* 5, 1–14.

Chen, Y., Feldman, D. E., Deng, C., Brown, J. A., De Giacomo, A. F., Gaw, A. F., Shi, G., Le, Q. T., Brown, J. M., and Koong, A. C. (2005). Identification of mitogen-activated protein kinase signaling pathways that confer resistance to endoplasmic reticulum stress in *Saccharomyces cerevisiae*. *Mol. Cancer Res.* 3, 669–677.

Chondrogianni, N., Tzavelas, C., Pemberton, A. J., Nezis, I. P., Rivett, A. J., and Gonos, E. S. (2005). Overexpression of proteasome beta5 assembled subunit increases the amount of proteasome and confers ameliorated response to oxidative stress and higher survival rates. *J. Biol. Chem.* 280, 11840–11850.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L., Toufighi, K., Mostafavi, S., Prinz, J., St Onge, R. P., Vandersluis, B., Makhnevych, T., Vizeacoumar, F. J., Alizadeh, S., Bahr, S., Brost, R. L., Chen, Y., Cokol, M., Deshpande, R., Li, Z., Lin, Z. Y., Liang, W., Marback, M., Paw, J., San Luis, B. J., Shuteriqi, E., Tong, A. H., Van Dyk, N., Wallace, I. M., Whitney, J. A., Weirauch, M. T., Zhong, G., Zhu, H., Houry, W. A., Brudno, M., Ragibizadeh, S., Papp, B., Pal, C., Roth, F. P., Giaever, G., Nislow, C., Troyanskaya, O. G., Bussey, H., Bader, G. D., Gingras, A. C., Morris, Q. D., Kim, P. M., Kaiser, C. A., Myers, C. L., Andrews, B. J., and Boone, C. (2010). The genetic landscape of a cell. *Science* 327, 425–431.

Crespo, J. L., and Hall, M. N. (2002). Elucidating TOR signaling and rapamycin action: lessons from *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* 66, 579–591.

Cumming, R. C., Dargusch, R., Fischer, W. H., and Schubert, D. (2007). Increase in expression levels and resistance to sulfhydryl oxidation of peroxiredoxin isoforms in amyloid beta-resistant nerve cells. *J. Biol. Chem.* 282, 30523–30534.

Delling, U., Raymond, M., and Schurr, E. (1998). Identification of *Saccharomyces cerevisiae* genes conferring resistance to quinoline ring-containing antimalarial drugs. *Antimicrob. Agents Chemother.* 42, 1034–1041.

Devantier, R., Scheithauer, B., Villas-Boas, S. G., Pedersen, S., and Olsson, L. (2005). Metabolite profiling for analysis of yeast stress response during very high gravity ethanol fermentations. *Biotechnol. Bioeng.* 90, 703–714.

Dias, P. J., Teixeira, M. C., Telo, J. P., and Sá-Correia, I. (2010). Insights into the mechanisms of toxicity and tolerance to the agricultural fungicide mancozeb in yeast, as suggested by a chemogenomic approach. *OMICS* 14, 211–227.

Ding, M. Z., Cheng, J. S., Xiao, W. H., Qiao, B., and Yuan, Y. J. (2009). Comparative metabolomic analysis on industrial continuous and batch ethanol fermentation processes by GC-TOF-MS. *Metabolomics* 5, 229–238.

Domico, L. M., Cooper, K. R., Bernard, L. P., and Zeevalk, G. D. (2007). Reactive oxygen species generation by the ethylene-bis-dithiocarbamate (EBDC) fungicide mancozeb and its contribution to neuronal toxicity in mesencephalic cells. *Neurotoxicology* 28, 1079–1091.

Doostzadeh, J., Davis, R. W., Giaever, G. N., Nislow, C., and Langston, J. W. (2007). Chemical genomic profiling for identifying intracellular targets of toxicants producing Parkinson's disease. *Toxicol. Sci.* 95, 182–187.

dos Santos, S. C., and Sá-Correia, I. (2009). Genome-wide identification of genes required for yeast growth under imatinib stress: vacuolar H^+-ATPase function is an important target of this anticancer drug. *OMICS* 13, 185–198.

dos Santos, S. C., and Sá-Correia, I. (2011). A genome-wide screen identifies yeast genes required for protection against or enhanced cytotoxicity of the antimalarial drug quinine. *Mol. Genet. Genomics* 286, 333–346.

dos Santos, S. C., Tenreiro, S., Palma, M., Becker, J., and Sa-Correia, I. (2009). Transcriptomic profiling of the *Saccharomyces cerevisiae* response to quinine reveals a glucose limitation response attributable

to drug-induced inhibition of glucose uptake. *Antimicrob. Agents Chemother.* 53, 5213–5223.

Emerson, L. R., Nau, M. E., Martin, R. K., Kyle, D. E., Vahey, M., and Wirth, D. F. (2002). Relationship between chloroquine toxicity and iron acquisition in *Saccharomyces cerevisiae*. *Antimicrob. Agents Chemother.* 46, 787–796.

Endo, A., Nakamura, T., Ando, A., Tokuyasu, K., and Shima, J. (2008). Genome-wide screening of the genes required for tolerance to vanillin, which is a potential inhibitor of bioethanol fermentation, in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 1, 3.

Erasmus, D. J., Van Der Merwe, G. K., and Van Vuuren, H. J. (2003). Genome-wide expression analyses: metabolic adaptation of *Saccharomyces cerevisiae* to high sugar stress. *FEMS Yeast Res.* 3, 375–399.

Ericson, E., Gebbia, M., Heisler, L. E., Wildenhain, J., Tyers, M., Giaever, G., and Nislow, C. (2008). Off-target effects of psychoactive drugs revealed by genome-wide assays in yeast. *PLoS Genet.* 4, e1000151. doi:10.1371/journal.pgen.1000151

Felder, T., Bogengruber, E., Tenreiro, S., Ellinger, A., Sá-Correia, I., and Briza, P. (2002). Dtrlp, a multidrug resistance transporter of the major facilitator superfamily, plays an essential role in spore wall maturation in *Saccharomyces cerevisiae*. *Eukaryotic Cell* 1, 799–810.

Fernandes, A. R., Durão, P. J., Santos, P. M., and Sá-Correia, I. (2003). Activation and significance of vacuolar H^+-ATPase in *Saccharomyces cerevisiae* adaptation and resistance to the herbicide 2,4-dichlorophenoxyacetic acid. *Biochem. Biophys. Res. Commun.* 312, 1317–1324.

Ferrari, G., Pastorelli, R., Buchi, F., Spinelli, E., Gozzini, A., Bosi, A., and Santini, V. (2007). Comparative proteomic analysis of chronic myelogenous leukemia cells: inside the mechanism of imatinib resistance. *J. Proteome Res.* 6, 367–375.

Fitch, C. D. (2004). Ferriprotoporphyrin IX, phospholipids, and the antimalarial actions of quinoline drugs. *Life Sci.* 74, 1957–1972.

Foury, F. (1997). Human genetic diseases: a cross-talk between man and yeast. *Gene* 195, 1–10.

Fujita, K., Matsuyama, A., Kobayashi, Y., and Iwahashi, H. (2006). The genome-wide screening of yeast deletion mutants to identify the genes required for tolerance to ethanol and other alcohols. *FEMS Yeast Res.* 6, 744–750.

Gad, S. C. (1990). Recent developments in replacing, reducing, and refining animal use in toxicologic research and testing. *Fundam. Appl. Toxicol.* 15, 8–16.

Gammie, A. E., Erdeniz, N., Beaver, J., Devlin, B., Nanji, A., and Rose, M. D. (2007). Functional characterization of pathogenic human MSH2 missense mutations in *Saccharomyces cerevisiae*. *Genetics* 177, 707–721.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K. D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kotter, P., Labonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C. Y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., and Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391.

Giaever, G., Flaherty, P., Kumm, J., Proctor, M., Nislow, C., Jaramillo, D. F., Chu, A. M., Jordan, M. I., Arkin, A. P., and Davis, R. W. (2004). Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 101, 793–798.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., and Oliver, S. G. (1996). Life with 6000 genes. *Science* 274, 546, 563–547.

Gomase, V. S., and Tagore, S. (2008). Toxicogenomics. *Curr. Drug Metab.* 9, 250–254.

Gorsich, S. W., Dien, B. S., Nichols, N. N., Slininger, P. J., Liu, Z. L., and Skory, C. D. (2006). Tolerance to furfural-induced stress is associated with pentose phosphate pathway genes ZWF1, GND1, RPE1, and TKL1 in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 71, 339–349.

Guerreiro, N., Staedtler, F., Grenet, O., Kehren, J., and Chibout, S. D. (2003). Toxicogenomics in drug development. *Toxicol. Pathol.* 31, 471–479.

Hamadeh, H. K., Amin, R. P., Paules, R. S., and Afshari, C. A. (2002). An overview of toxicogenomics. *Curr. Issues Mol. Biol.* 4, 45–56.

Hammonds, T. R., Maxwell, A., and Jenkins, J. R. (1998). Use of a rapid throughput in vivo screen to investigate inhibitors of eukaryotic topoisomerase II enzymes. *Antimicrob. Agents Chemother.* 42, 889–894.

Hasunuma, T., Sanda, T., Yamada, R., Yoshimura, K., Ishii, J., and Kondo, A. (2011). Metabolic pathway engineering based on metabolomics confers acetic and formic acid tolerance to a recombinant xylose-fermenting strain of *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 10, 2.

Heinicke, S., Livstone, M. S., Lu, C., Oughtred, R., Kang, F., Angiuoli, S. V., White, O., Botstein, D., and Dolinski, K. (2007). The Princeton protein orthology database (P-POD): a comparative genomics analysis tool for biologists. *PLoS ONE* 2, e766. doi:10.1371/journal.pone.0000766

Hellauer, K., Lesage, G., Sdicu, A.-M., and Turcotte, B. (2005). Large-scale analysis of genes that alter sensitivity to the anticancer drug tirapazamine in *Saccharomyces cerevisiae*. *Mol. Pharmacol.* 68, 1365–1375.

Hillenmeyer, M. E., Fung, E., Wildenhain, J., Pierce, S. E., Hoon, S., Lee, W., Proctor, M., St Onge, R. P., Tyers, M., Koller, D., Altman, R. B., Davis, R. W., Nislow, C., and Giaever, G. (2008). The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* 320, 362–365.

Hirasawa, T., Yoshikawa, K., Nakakura, Y., Nagahisa, K., Furusawa, C., Katakura, Y., Shimizu, H., and Shioya, S. (2007). Identification of target genes conferring ethanol stress tolerance to *Saccharomyces cerevisiae* based on DNA microarray data analysis. *J. Biotechnol.* 131, 34–44.

Hohmann, S., and Mager, W. H. (1997). "Stress response mechanisms in the yeast *Saccharomyces cerevisiae*," in *Yeast Stress Responses*, eds S. Hohmann and W. H. Mager (Austin, TX: Molecular Biology Intelligence Unit, R.G. Landes Company), 1–5.

Holland, S., Lodwig, E., Sideri, T., Reader, T., Clarke, I., Gkargkas, K., Hoyle, D. C., Delneri, D., Oliver, S. G., and Avery, S. V. (2007). Application of the comprehensive set of heterozygous yeast deletion mutants to elucidate the molecular basis of cellular chromium toxicity. *Genome Biol.* 8, R268.

Hoon, S., Smith, A. M., Wallace, I. M., Suresh, S., Miranda, M., Fung, E., Proctor, M., Shokat, K. M., Zhang, C., Davis, R. W., Giaever, G., St Onge, R. P., and Nislow, C. (2008a). An integrated platform of genomic assays reveals small-molecule bioactivities. *Nat. Chem. Biol.* 4, 498–506.

Hoon, S., St Onge, R. P., Giaever, G., and Nislow, C. (2008b). Yeast chemical genomics and drug discovery: an update. *Trends Pharmacol. Sci.* 29, 499–504.

Huang, R.-Y., Eddy, M., Vujcic, M., and Kowalski, D. (2005). Genome-wide screen identifies genes whose inactivation confer resistance to cisplatin in *Saccharomyces cerevisiae*. *Cancer Res.* 65, 5890–5897.

Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., Bennett, H. A., Coffey, E., Dai, H., He, Y. D., Kidd, M. J., King, A. M., Meyer, M. R., Slade, D., Lum, P. Y., Stepaniants, S. B., Shoemaker, D. D., Gachotte, D., Chakraburtty, K., Simon, J., Bard, M., and Friend, S. H. (2000). Functional discovery via a compendium of expression profiles. *Cell* 102, 109–126.

Ibrahim, M. A., Bond, G. G., Burke, T. A., Cole, P., Dost, F. N., Enterline, P. E., Gough, M., Greenberg, R. S., Halperin, W. E., Mcconnell, E., Munro, I. C., Swenberg, J. A., Zahm, S. H., and Graham, J. D. (1991). Weight of the evidence on the human carcinogenicity of 2,4-D. *Environ. Health Perspect.* 96, 213–222.

Ihara, Y., Nobukuni, K., Takata, H., and Hayabara, T. (2005). Oxidative stress and metal content in blood and cerebrospinal fluid of amyotrophic lateral sclerosis patients with and without a Cu, Zn-superoxide dismutase mutation. *Neurol. Res.* 27, 105–108.

Ito, H., and Gray, W. M. (2006). A gain-of-function mutation in the Arabidopsis pleiotropic drug resistance transporter PDR9 confers resistance to auxinic herbicides. *Plant Physiol.* 142, 63–74.

Jia, M. H., Larossa, R. A., Lee, J. M., Rafalski, A., Derose, E., Gonye, G., and Xue, Z. (2000). Global expression profiling of yeast treated with an inhibitor of amino acid biosynthesis, sulfometuron methyl. *Physiol. Genomics* 3, 83–92.

Jin, Y. H., Dunlap, P. E., Mcbride, S. J., Al-Refai, H., Bushel, P. R.,

and Freedman, J. H. (2008). Global transcriptome and deletome profiles of yeast exposed to transition metals. *PLoS Genet.* 4, e1000053. doi:10.1371/journal.pgen.1000053

Khozoie, C., Pleass, R. J., and Avery, S. V. (2009). The antimalarial drug quinine disrupts Tat2p-mediated tryptophan transport and causes tryptophan starvation. *J. Biol. Chem.* 284, 17968–17974.

Kitagawa, E., Momose, Y., and Iwahashi, H. (2003). Correlation of the structures of agricultural fungicides to gene expression in *Saccharomyces cerevisiae* upon exposure to toxic doses. *Environ. Sci. Technol.* 37, 2788–2793.

Kominsky, D. J., Klawitter, J., Brown, J. L., Boros, L. G., Melo, J. V., Eckhardt, S. G., and Serkova, N. J. (2009). Abnormalities in glucose uptake and metabolism in imatinib-resistant human BCR-ABL-positive cells. *Clin. Cancer Res.* 15, 3442–3450.

Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., Lerner, J., Brunet, J. P., Subramanian, A., Ross, K. N., Reich, M., Hieronymus, H., Wei, G., Armstrong, S. A., Haggarty, S. J., Clemons, P. A., Wei, R., Carr, S. A., Lander, E. S., and Golub, T. R. (2006). The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313, 1929–1935.

Landstetter, N., Glaser, W., Gregori, C., Seipelt, J., and Kuchler, K. (2010). Functional genomics of drug-induced ion homeostasis identifies a novel regulatory crosstalk of iron and zinc regulons in yeast. *OMICS* 14, 651–663.

Li, B. Z., Cheng, J. S., Qiao, B., and Yuan, Y. J. (2010). Genome-wide transcriptional analysis of *Saccharomyces cerevisiae* during industrial bioethanol fermentation. *J. Ind. Microbiol. Biotechnol.* 37, 43–55.

Li, W., Mo, W., Shen, D., Sun, L., Wang, J., Lu, S., Gitschier, J. M., and Zhou, B. (2005). Yeast model uncovers dual roles of mitochondria in action of artemisinin. *PLoS Genet.* 1, e36. doi:10.1371/journal.pgen.0010036

Liao, C., Hu, B., Arno, M. J., and Panaretou, B. (2007). Genomic screening in vivo reveals the role played by vacuolar H^+- ATPase and cytosolic acidification in sensitivity to DNA-damaging agents such as cisplatin. *Mol. Pharmacol.* 71, 416–425.

Lourenço, A. B., Ascenso, J. R., and Sá-Correia, I. (2011). Metabolic insights into the yeast response to propionic acid based on high resolution 1H NMR spectroscopy. *Metabolomics* 7, 457–468.

Lucau-Danila, A., Lelandais, G., Kozovska, Z., Tanty, V., Delaveau, T., Devaux, F., and Jacq, C. (2005). Early expression of yeast genes affected by chemical stress. *Mol. Cell. Biol.* 25, 1860–1868.

Lum, P. Y., Armour, C. D., Stepaniants, S. B., Cavet, G., Wolf, M. K., Butler, J. S., Hinshaw, J. C., Garnier, P., Prestwich, G. D., Leonardson, A., Garrett-Engele, P., Rush, C. M., Bard, M., Schimmack, G., Phillips, J. W., Roberts, C. J., and Shoemaker, D. D. (2004). Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell* 116, 121–137.

Mager, W. H., and Winderickx, J. (2005). Yeast as a model for medical and medicinal research. *Trends Pharmacol. Sci.* 26, 265–273.

Marjanovic, J., Chalupska, D., Patenode, C., Coster, A., Arnold, E., Ye, A., Anesi, G., Lu, Y., Okun, I., Tkachenko, S., Haselkorn, R., and Gornicki, P. (2010). Recombinant yeast screen for new inhibitors of human acetyl-CoA carboxylase 2 identifies potential drugs to treat obesity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9093–9098.

Markovich, S., Yekutiel, A., Shalit, I., Shadkchan, Y., and Osherov, N. (2004). Genomic approach to identification of mutations affecting caspofungin susceptibility in *Saccharomyces cerevisiae*. *Antimicrob. Agents Chemother.* 48, 3871–3876.

Marks, V. D., Ho Sui, S. J., Erasmus, D., Van Der Merwe, G. K., Brumm, J., Wasserman, W. W., Bryan, J., and Van Vuuren, H. J. (2008). Dynamics of the yeast transcriptome during wine fermentation reveals a novel fermentation stress response. *FEMS Yeast Res.* 8, 35–52.

Maroni, M., Colosio, C., Ferioli, A., and Fait, A. (2000). Biological monitoring of pesticide exposure: a review. *Toxicology* 143, 1–118.

Martin, D., Brun, C., Remy, E., Mouren, P., Thieffry, D., and Jacq, B. (2004). GOToolBox: functional analysis of gene datasets based on Gene Ontology. *Genome Biol.* 5, R101.

Marton, M. J., Derisi, J. L., Bennett, H. A., Iyer, V. R., Meyer, M. R., Roberts, C. J., Stoughton, R., Burchard, J., Slade, D., Dai, H., Bassett, D. E. Jr., Hartwell, L. H., Brown, P. O., and Friend, S. H. (1998). Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat. Med.* 4, 1293–1301.

Menacho-Marquez, M., and Murguia, J. R. (2007). Yeast on drugs: *Saccharomyces cerevisiae* as a tool for anticancer drug research. *Clin. Transl. Oncol.* 9, 221–228.

Mettetal, J. T., Muzzey, D., Gomez-Uribe, C., and Van Oudenaarden, A. (2008). The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*. *Science* 319, 482–484.

Mira, N. P., Becker, J. D., and Sa-Correia, I. (2010a). Genomic expression program involving the Haa1p-regulon in *Saccharomyces cerevisiae* response to acetic acid. *OMICS* 14, 587–601.

Mira, N. P., Palma, M., Guerreiro, J. F., and Sa-Correia, I. (2010b). Genome-wide identification of *Saccharomyces cerevisiae* genes required for tolerance to acetic acid. *Microb. Cell Fact.* 9, 79.

Mira, N. P., Lourenço, A. B., Fernandes, A. R., Becker, J. D., and Sá-Correia, I. (2009). The RIM101 pathway has a role in *Saccharomyces cerevisiae* adaptive response and resistance to propionic acid and other weak acids. *FEMS Yeast Res.* 9, 202–216.

Miyamoto, M., Furuichi, Y., and Komiyama, T. (2011). Genome-wide screen of *Saccharomyces cerevisiae* for killer toxin HM-1 resistance. *Yeast* 28, 27–41.

Mollapour, M., Fong, D., Balakrishnan, K., Harris, N., Thompson, S., Schuller, C., Kuchler, K., and Piper, P. W. (2004). Screening the yeast deletant mutant collection for hypersensitivity and hyper-resistance to sorbate, a weak organic acid food preservative. *Yeast* 21, 927–946.

Momose, Y., and Iwahashi, H. (2001). Bioassay of cadmium using a DNA microarray: genome-wide expression patterns of *Saccharomyces cerevisiae* response to cadmium. *Environ. Toxicol. Chem.* 20, 2353–2360.

Monteiro, P. T., Dias, P. J., Ropers, D., Oliveira, A. L., Sá-Correia, I., Teixeira, M. C., and Freitas, A. T. (2011). Qualitative modeling and formal verification of the FLR1 gen mancozeb response in *Saccharomyces cerevisiae*. *IET Syst. Biol.* 5, 308–316.

Morton, C. O., dos Santos, S. C., and Coote, P. (2007). An amphibian-derived, cationic, alpha-helical antimicrobial peptide kills yeast by caspase-independent but AIF-dependent programmed cell death. *Mol. Microbiol.* 65, 494–507.

Murakami, T., Shibuya, I., Ise, T., Chen, Z. S., Akiyama, S., Nakagawa, M., Izumi, H., Nakamura, T., Matsuo, K., Yamada, Y., and Kohno, K. (2001). Elevated expression of vacuolar proton pump genes and cellular PH in cisplatin resistance. *Int. J. Cancer* 93, 869–874.

Nicholson, J. K., Connelly, J., Lindon, J. C., and Holmes, E. (2002). Metabonomics: a platform for studying drug toxicity and gene function. *Nat. Rev. Drug Discov.* 1, 153–161.

North, M., and Vulpe, C. D. (2010). Functional toxicogenomics: mechanism-centered toxicology. *Int. J. Mol. Sci.* 11, 4796–4813.

Nowicki, M. O., Pawlowski, P., Fischer, T., Hess, G., Pawlowski, T., and Skorski, T. (2003). Chronic myelogenous leukemia molecular signature. *Oncogene* 22, 3952–3963.

Nunes, P. A., Tenreiro, S., and Sá-Correia, I. (2001). Resistance and adaptation to quinidine in *Saccharomyces cerevisiae*: role of QDR1 (YIL120w), encoding a plasma membrane transporter of the major facilitator superfamily required for multidrug resistance. *Antimicrob. Agents Chemother.* 45, 1528–1534.

Outeiro, T. F., and Lindquist, S. (2003). Yeast cells provide insight into alpha-synuclein biology and pathobiology. *Science* 302, 1772–1775.

Papaefthimiou, C., Cabral Mde, G., Mixailidou, C., Viegas, C. A., Sá-Correia, I., and Theophilidis, G. (2004). Comparison of two screening bioassays, based on the frog sciatic nerve and yeast cells, for the assessment of herbicide toxicity. *Environ. Toxicol. Chem.* 23, 1211–1218.

Parsons, A. B., Brost, R. L., Ding, H., Li, Z., Zhang, C., Sheikh, B., Brown, G. W., Kane, P. M., Hughes, T. R., and Boone, C. (2004). Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. *Nat. Biotechnol.* 22, 62–69.

Parsons, A. B., Geyer, R., Hughes, T. R., and Boone, C. (2003). Yeast genomics and proteomics in drug discovery and target validation. *Prog. Cell Cycle Res.* 5, 159–166.

Parsons, A. B., Lopez, A., Givoni, I. E., Williams, D. E., Gray, C. A., Porter, J., Chua, G., Sopko, R., Brost, R. L., Ho, C. H., Wang, J., Ketela, T., Brenner, C., Brill, J. A., Fernandez, G. E., Lorenz, T. C., Payne, G. S., Ishihara, S., Ohya, Y., Andrews, B., Hughes, T. R., Frey, B. J., Graham, T. R., Andersen, R. J., and Boone, C. (2006). Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* 126, 611–625.

Parveen, M., Momose, Y., Kitagawa, E., Kurita, S., Kodama, O., and Iwahashi, H. (2003). Bioassay of pesticide Lindane using yeast-DNA microarrays technology. *Chem-Bio Informatics J.* 3, 12–29.

Pereira, F. B., Guimaraes, P. M., Gomes, D. G., Mira, N. P., Teixeira, M. C., Sa-Correia, I., and Domingues, L. (2011). Identification of candidate genes for yeast engineering to improve bioethanol production in very-high-gravity and lignocellulosic biomass industrial fermentations. *Biotechnol. Biofuels* 4, 57.

Perlstein, E. O., Ruderfer, D. M., Roberts, D. C., Schreiber, S. L., and Kruglyak, L. (2007). Genetic basis of individual differences in the response to small-molecule drugs in yeast. *Nat. Genet.* 39, 496–502.

Pham, T. K., Chong, P. K., Gan, C. S., and Wright, P. C. (2006). Proteomic analysis of *Saccharomyces cerevisiae* under high gravity fermentation conditions. *J. Proteome Res.* 5, 3411–3419.

Pocaly, M., Lagarde, V., Etienne, G., Dupouy, M., Lapaillerie, D., Claverol, S., Vilain, S., Bonneu, M., Turcq, B., Mahon, F.-X., and Pasquet, J.-M. (2008). Proteomic analysis of an imatinib-resistant K562 cell line highlights opposing roles of heat shock cognate 70 and heat shock 70 proteins in resistance. *Proteomics* 8, 2394–2406.

Quintás-Cardama, A., Kantarjian, H. M., and Cortes, J. E. (2009). Mechanisms of primary and secondary resistance to imatinib in chronic myeloid leukemia. *Cancer Control* 16, 122–131.

Ruderfer, D. M., Roberts, D. C., Schreiber, S. L., Perlstein, E. O., and Kruglyak, L. (2009). Using expression and genotype to predict drug response in yeast. *PLoS ONE* 4, e6907. doi:10.1371/journal.pone.0006907

Ruotolo, R., Marchini, G., and Ottonello, S. (2008). Membrane transporters and protein traffic networks differentially affecting metal tolerance: a genomic phenotyping study in yeast. *Genome Biol.* 9, R67.

Sá-Correia, I., dos Santos, S. C., Teixeira, M. C., Cabrito, T. R., and Mira, N. P. (2009). Drug:H^+ antiporters in chemical stress response in yeast. *Trends Microbiol.* 17, 22–31.

Sa-Correia, I., and Teixeira, M. C. (2010). 2D electrophoresis-based expression proteomics: a microbiologist's perspective. *Expert Rev. Proteomics* 7, 943–953.

Saliba, K. J., Krishna, S., and Kirk, K. (2004). Inhibition of hexose transport and abrogation of pH homeostasis in the intraerythrocytic malaria parasite by an O-3-hexose derivative. *FEBS Lett.* 570, 93–96.

Sanchez, C. P., Stein, W. D., and Lanzer, M. (2008). Dissecting the components of quinine accumulation in *Plasmodium falciparum*. *Mol. Microbiol.* 67, 1081–1093.

Santos, P. M., Simões, T., and Sá-Correia, I. (2009). Insights into yeast adaptive response to the agricultural fungicide mancozeb: a toxicoproteomics approach. *Proteomics* 9, 657–670.

Schenk, P. W., Boersma, A. W., Brandsma, J. A., Den Dulk, H., Burger, H., Stoter, G., Brouwer, J., and Nooter, K. (2001). SKY1 is involved in cisplatin-induced cell kill in *Saccharomyces cerevisiae*, and inactivation of its human homologue, SRPK1, induces cisplatin resistance in a human ovarian carcinoma cell line. *Cancer Res.* 61, 6982–6986.

Schenk, P. W., Boersma, A. W. M., Brok, M., Burger, H., Stoter, G., and Nooter, K. (2002). Inactivation of the *Saccharomyces cerevisiae* SKY1 gene induces a specific modification of the yeast anticancer drug sensitivity profile accompanied by a mutator phenotype. *Mol. Pharmacol.* 61, 659–666.

Scherens, B., and Goffeau, A. (2004). The uses of genome-wide yeast mutant collections. *Genome Biol.* 5, 229.

Schüller, C., Mamnun, Y. M., Mollapour, M., Krapf, G., Schuster, M., Bauer, B. E., Piper, P. W., and Kuchler, K. (2004). Global phenotypic analysis and transcriptional profiling defines the weak acid stress response regulon in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 15, 706–720.

Schwartz, D. A., Freedman, J. H., and Linney, E. A. (2004). Environmental genomics: a key to understanding biology, pathophysiology and disease. *Hum. Mol. Genet.* 13, R217–R224.

Sennoune, S. R., Bakunts, K., Martínez, G. M., Chua-Tuan, J. L., Kebir, Y., Attaya, M. N., and Martínez-Zaguilán, R. (2004). Vacuolar H^+-ATPase in human breast cancer cells with distinct metastatic potential: distribution and functional activity. *Am. J. Physiol. Cell Physiol.* 286, C1443–C1452.

Serero, A., Lopes, J., Nicolas, A., and Boiteux, S. (2008). Yeast genes involved in cadmium tolerance: identification of DNA replication as a target of cadmium toxicity. *DNA Repair (Amst.)* 7, 1262–1275.

Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. *Nat. Cell Biol.* 4, E131–E136.

Shen, H., Chen, J., Wang, Z., Yang, C., Sasaki, T., Yamamoto, Y., Matsumoto, H., and Yan, X. (2006). Root plasma membrane H^+-ATPase is involved in the adaptation of soybean to phosphorus starvation. *J. Exp. Bot.* 57, 1353–1362.

Simmons, P. T., and Portier, C. J. (2002). Toxicogenomics: the new frontier in risk analysis. *Carcinogenesis* 23, 903–905.

Simões, T., Teixeira, M. C., Fernandes, A. R., and Sá-Correia, I. (2003). Adaptation of *Saccharomyces cerevisiae* to the herbicide 2,4-dichlorophenoxyacetic acid, mediated by Msn2p- and Msn4p-regulated genes: important role of SPI1. *Appl. Environ. Microbiol.* 69, 4019–4028.

Simon, J. A., and Bedalov, A. (2004). Yeast as a model system for anticancer drug discovery. *Nat. Rev. Cancer* 4, 481–492.

Sletta, H., Klinkenberg, G., Winnberg, A., Kvitvang, H. F., Nilsen, M. B., Krokan, H. E., Otterlei, M., and Bruheim, P. (2011). A new high resolution screening method for study of phenotype stress responses of *Saccharomyces cerevisiae* mutants. *J. Microbiol. Methods* 87, 363–367.

Smart, C. C., and Fleming, A. J. (1996). Hormonal and environmental regulation of a plant PDR5-like ABC transporter. *J. Biol. Chem.* 271, 19351–19357.

Smith, A. M., Ammar, R., Nislow, C., and Giaever, G. (2010). A survey of yeast genomic assays for drug and target discovery. *Pharmacol. Ther.* 127, 156–164.

Stanley, D., Chambers, P. J., Stanley, G. A., Borneman, A., and Fraser, S. (2010). Transcriptional changes associated with ethanol tolerance in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 88, 231–239.

Steinmetz, L. M., Scharfe, C., Deutschbauer, A. M., Mokranjac, D., Herman, Z. S., Jones, T., Chu, A. M., Giaever, G., Prokisch, H., Oefner, P. J., and Davis, R. W. (2002). Systematic screen for human disease genes in yeast. *Nat. Genet.* 31, 400–404.

Swen, J. J., Huizinga, T. W., Gelderblom, H., De Vries, E. G. E., Assendelft, W. J. J., Kirchheiner, J., and Guchelaar, H.-J. (2007). Translating pharmacogenomics: challenges on the road to the clinic. *PLoS Med.* 4, e209. doi:10.1371/journal.pmed.0040209

Takumi, S., Kimura, H., Matsusaki, H., Kawazoe, S., Tominaga, N., and Arizono, K. (2010). DNA microarray analysis of genomic responses of yeast *Saccharomyces cerevisiae* to nickel chloride. *J. Toxicol. Sci.* 35, 125–129.

Teixeira, M. C., Dias, P. J., Monteiro, P. T., Sala, A., Oliveira, A. L., Freitas, A. T., and Sá-Correia, I. (2010a). Refining current knowledge on the yeast FLR1 regulatory network by combined experimental and computational approaches. *Mol. Biosyst.* 6, 2471–2481.

Teixeira, M. C., Raposo, L. R., Palma, M., and Sa-Correia, I. (2010b). Identification of genes required for maximal tolerance to high-glucose concentrations, as those present in industrial alcoholic fermentation media, through a chemogenomics approach. *OMICS* 14, 201–210.

Teixeira, M. C., Duque, P., and Sá-Correia, I. (2007). Environmental genomics: mechanistic insights into toxicity of and resistance to the herbicide 2,4-D. *Trends Biotechnol.* 25, 363–370.

Teixeira, M. C., Fernandes, A. R., Mira, N. P., Becker, J. D., and Sá-Correia, I. (2006a). Early transcriptional response of *Saccharomyces cerevisiae* to stress imposed by the herbicide 2,4-dichlorophenoxyacetic acid. *FEMS Yeast Res.* 6, 230–248.

Teixeira, M. C., Monteiro, P., Jain, P., Tenreiro, S., Fernandes, A. R., Mira, N. P., Alenquer, M., Freitas, A. T., Oliveira, A. L., and Sá-Correia, I. (2006b). The YEASTRACT database: a tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 34, D446–D451.

Teixeira, M. C., Mira, N. P., and Sa-Correia, I. (2011). A genome-wide perspective on the response and tolerance to food-relevant stresses in *Saccharomyces cerevisiae*. *Curr. Opin. Biotechnol.* 22, 150–156.

Teixeira, M. C., Raposo, L. R., Mira, N. P., Lourenço, A. B., and Sá-Correia, I. (2009a). Genome-wide identification of *Saccharomyces cerevisiae* genes required for maximal tolerance to ethanol. *Appl. Environ. Microbiol.* 75, 5761–5772.

Teixeira, M. C., Santos, P. M., Rodrigues, C., and Sa-Correia, I. (2009b). Teaching expression proteomics: from the wet-lab to the laptop. *Biochem. Mol. Biol. Educ.* 37, 279–286.

Teixeira, M. C., and Sá-Correia, I. (2002). *Saccharomyces cerevisiae* resistance to chlorinated phenoxyacetic acid herbicides involves Pdr1p-mediated transcriptional activation of TPO1 and PDR5 genes. *Biochem. Biophys. Res. Commun.* 292, 530–537.

Teixeira, M. C., Santos, P. M., Fernandes, A. R., and Sá-Correia, I. (2005). A proteome analysis of the yeast response to the herbicide 2,4-dichlorophenoxyacetic acid. *Proteomics* 5, 1889–1901.

Teixeira, M. C., Telo, J. P., Duarte, N. F., and Sá-Correia, I. (2004). The herbicide 2,4-dichlorophenoxyacetic acid induces the generation of free-radicals and associated oxidative stress responses in yeast. *Biochem. Biophys. Res. Commun.* 324, 1101–1107.

Tenreiro, S., Nunes, P. A., Viegas, C. A., Neves, M. S., Teixeira, M. C., Cabral, M. G., and Sá-Correia, I. (2002). AQR1 gene (ORF YNL065w) encodes a plasma membrane transporter of the major facilitator superfamily that confers resistance to short-chain monocarboxylic acids and quinidine in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* 292, 741–748.

Thorpe, G. W., Fong, C. S., Alic, N., Higgins, V. J., and Dawes, I. W. (2004). Cells have distinct mechanisms to maintain protection against different reactive oxygen species: oxidative-stress-response genes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6564–6569.

Tribouillard, D., Bach, S., Gug, F., Desban, N., Beringue, V., Andrieu, T., Dormont, D., Galons, H., Laude, H., Vilette, D., and Blondel, M. (2006). Using budding yeast to screen for anti-prion drugs. *Biotechnol. J.* 1, 58–67.

Tucker, C. L., and Fields, S. (2004). Quantitative genome-wide analysis of yeast deletion strain sensitivities to oxidative and chemical stress. *Comp. Funct. Genomics* 5, 216–224.

van Maris, A. J., Abbott, D. A., Bellissimi, E., Van Den Brink, J., Kuyper, M., Luttik, M. A., Wisselink, H. W., Scheffers, W. A., Van Dijken, J. P., and Pronk, J. T. (2006). Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie Van Leeuwenhoek* 90, 391–418.

van Voorst, F., Houghton-Larsen, J., Jonson, L., Kielland-Brandt, M. C., and Brandt, A. (2006). Genome-wide identification of genes required for growth of *Saccharomyces cerevisiae* under ethanol stress. *Yeast* 23, 351–359.

Vargas, R. C., Tenreiro, S., Teixeira, M. C., Fernandes, A. R., and Sá-Correia, I. (2004). *Saccharomyces cerevisiae* multidrug transporter Qdr2p (Yil121wp): localization and function as a quinidine resistance determinant. *Antimicrob. Agents Chemother.* 48, 2531–2537.

Volpe, G., Panuzzo, C., Ulisciani, S., and Cilloni, D. (2009). Imatinib resistance in CML. *Cancer Lett.* 274, 1–9.

Wang, H. P., Long, X. H., Sun, Z. Z., Rigaud, O., Xu, Q. Z., Huang, Y. C., Sui, J. L., Bai, B., and Zhou, P. K. (2006a). Identification of differentially transcribed genes in human lymphoblastoid cells irradiated with 0.5 Gy of gamma-ray and the involvement of low dose radiation inducible CHD6 gene in cell proliferation and radiosensitivity. *Int. J. Radiat. Biol.* 82, 181–190.

Wang, R. A., Zhang, H., Balasenthil, S., Medina, D., and Kumar, R. (2006b). PAK1 hyperactivation is sufficient for mammary gland tumor formation. *Oncogene* 25, 2931–2936.

Wang, L., and Weinshilboum, R. M. (2008). Pharmacogenomics: candidate gene identification, functional validation and mechanisms. *Hum. Mol. Genet.* 17, R174–R179.

West, G. M., Tucker, C. L., Xu, T., Park, S. K., Han, X., Yates, J. R. III, and Fitzgerald, M. C. (2010). Quantitative proteomics approach for identifying protein-drug interactions in complex mixtures using protein stability measurements. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9078–9082.

WHO. (2006). *Guidelines for the Treatment of Malaria*. Geneva: World Health Organization.

Willingham, S., Outeiro, T. F., Devit, M. J., Lindquist, S. L., and Muchowski, P. J. (2003). Yeast genes that enhance the toxicity of a mutant huntingtin fragment or alpha-synuclein. *Science* 302, 1769–1772.

Woodrow, C. J., Penny, J. I., and Krishna, S. (1999). Intraerythrocytic *Plasmodium falciparum* expresses a high affinity facilitative hexose transporter. *J. Biol. Chem.* 274, 7272–7277.

Wu, H. I., Brown, J. A., Dorie, M. J., Lazzeroni, L., and Brown, J. M. (2004). Genome-wide identification of genes conferring resistance to the anticancer agents cisplatin, oxaliplatin, and mitomycin C. *Cancer Res.* 64, 3940–3948.

Wuster, A., and Madan Babu, M. (2008). Chemogenomics and biotechnology. *Trends Biotechnol.* 26, 252–258.

Yasokawa, D., Murata, S., Kitagawa, E., Iwahashi, Y., Nakagawa, R., Hashido, T., and Iwahashi, H. (2008). Mechanisms of copper toxicity in *Saccharomyces cerevisiae* determined by microarray analysis. *Environ. Toxicol.* 23, 599–606.

Yoshikawa, K., Tanaka, T., Furusawa, C., Nagahisa, K., Hirasawa, T., and Shimizu, H. (2009). Comprehensive phenotypic analysis for identification of genes affecting growth under ethanol stress in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 9, 32–44.

Yuen, K. W. Y., Warren, C. D., Chen, O., Kwok, T., Hieter, P., and Spencer, F. A. (2007). Systematic genome instability screens in yeast and their potential relevance to cancer. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3925–3930.

Zhang, L., Mchale, C. M., Rothman, N., Li, G., Ji, Z., Vermeulen, R., Hubbard, A. E., Ren, X., Shen, M., Rappaport, S. M., North, M., Skibola, C. F., Yin, S., Vulpe, C., Chanock, S. J., Smith, M. T., and Lan, Q. (2010). Systems biology of human benzene exposure. *Chem. Biol. Interact.* 184, 86–93.

Zhou, Y., Shie, F. S., Piccardo, P., Montine, T. J., and Zhang, J. (2004). Proteasomal inhibition induced by manganese ethylene-bis-dithiocarbamate: relevance to Parkinson's disease. *Neuroscience* 128, 281–291.

Conflict of Interest Statement: 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.

Received: 16 February 2012; paper pending published: 09 March 2012; accepted: 03 April 2012; published online: 19 April 2012.

Citation: dos Santos SC, Teixeira MC, Cabrito TR and Sá-Correia I (2012) Yeast toxicogenomics: genome-wide responses to chemical stresses with impact in environmental health, pharmacology, and biotechnology. Front. Gene. ***3****:63. doi: 10.3389/fgene.2012.00063*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in GENETICS

REVIEW ARTICLE
published: 04 May 2012
doi: 10.3389/fgene.2012.00068

Drosophila melanogaster as a model for lead neurotoxicology and toxicogenomics research

Helmut V. B. Hirsch[1], Gregory Lnenicka[1], Debra Possidente[2], Bernard Possidente[2], Mark D. Garfinkel[3], Luan Wang[4], Xiangyi Lu[4] and Douglas M. Ruden[4,5]*

[1] Department of Biological Sciences, University at Albany, State University of New York, Albany, NY, USA

[2] Department of Biology, Skidmore College, Saratoga Springs, NY, USA

[3] Department of Environmental Health Sciences, University of Alabama at Birmingham, Birmingham, AL, USA

[4] Institute of Environmental Health Sciences, Wayne State University, Detroit, MI, USA

[5] Department of Obstetrics and Gynecology, C. S. Mott Center for Human Growth and Development, Wayne State University, Detroit, MI, USA

Edited by:
Michael Aschner, Vanderbilt University Medical Center, USA

Reviewed by:
Eugenia Xu, Raymond and Beverly Sackler Foundation, USA
Marcos De Donato, Cornell University, USA

***Correspondence:**
Douglas M. Ruden, Institute of Environmental Health Sciences, Wayne State University, Detroit, MI 48201, USA. e-mail: douglasr@wayne.edu

Drosophila melanogaster is an excellent model animal for studying the neurotoxicology of lead. It has been known since ancient Roman times that long-term exposure to low levels of lead results in behavioral abnormalities, such as what is now known as attention deficit hyperactivity disorder (ADHD). Because lead alters mechanisms that underlie developmental neuronal plasticity, chronic exposure of children, even at blood lead levels below the current CDC community action level (10 μg/dl), can result in reduced cognitive ability, increased likelihood of delinquency, behaviors associated with ADHD, changes in activity level, altered sensory function, delayed onset of sexual maturity in girls, and changes in immune function. In order to better understand how lead affects neuronal plasticity, we will describe recent findings from a *Drosophila* behavioral genetics laboratory, a *Drosophila* neurophysiology laboratory, and a *Drosophila* quantitative genetics laboratory who have joined forces to study the effects of lead on the *Drosophila* nervous system. Studying the effects of lead on *Drosophila* nervous system development will give us a better understanding of the mechanisms of Pb neurotoxicity in the developing human nervous system.

Keywords: *Drosophila*, toxicology, toxicogenomics, behavioral toxicology

PART 1: QTL MAPPING: BEHAVIORS AND TOXINS

BEHAVIOR AS A QTL ENDPOINT FOR TOXICOLOGY STUDIES

There are two main reasons why behavioral assays are so useful in detecting effects of toxins. First is their richness. Behavior can be described as an ongoing, generally complex, spatio-temporal pattern; subtle changes in that pattern can signal the action of very low doses of a toxin well before there are dramatic changes in organ systems: in children overt clinical encephalopathy is associated with blood lead levels of 80–100 μg/dl, while changes in intelligence quotient (IQ) and learning occur at or below a tenth of that dose (UNEP, 2010). Second is that even subtle toxin-dependent changes in behavior can have consequences for well-being which makes them meaningful and relevant.

Since chronic developmental exposure to toxins affects assembly of the neuronal and hormonal systems mediating adult behavior, their effects are generally more pronounced than that following acute adult exposure. Developmental exposure to sub-lethal and sub-teratogenic levels of toxins may alter or degrade physiological and behavioral mechanisms in a quantitative manner. Examples include effects of low doses of lead on behavior in fruit flies (Hirsch et al., 2003), cognitive effects of polychlorobiphenyls in people (Faroon et al., 2000), lead effects on cognitive function in people (Counter et al., 1998), and the effects of aluminum on adult behavior and developmental rate of their offspring in mice (Abu-Taweel et al., 2012). Once toxin-dependent behavioral changes are observed, the next steps may include study of their underlying physiological effects, for example, effects of lead on synaptic function in fruit fly larvae (He et al., 2009).

The bulk of what is known about the effects of various toxins is based on studies focusing on a single toxin; combinations of toxins can have additive, protective, or synergistic effects (Rai et al., 2010; Singh et al., 2010).

DROSOPHILA BEHAVIORAL QTLs

Quantitative trait locus (QTL) mapping is used for unbiased genome-wide screens to identify genetic loci causing variation in a trait, but they are most efficient when large numbers of individual genomes can be assayed. The large sample sizes required for precise QTL analysis favors model organisms that can be bred in the lab easily and inexpensively, and traits that are simple to assay yet rich in their information content and thus provide sensitive measures of effects of toxins are especially desirable.

Drosophila is an ideal model organism for both efficiency and genetic analysis, since they are easy and relatively inexpensive to breed and maintain in large numbers, and have a surprising degree of genetic homology to mammals (Mackay and Anholt, 2006). Because of the small physical size of *Drosophila* many phenotypic traits can be difficult to assay rapidly or quantitatively. Exceptions include very simple traits such as bristle number, or

those easy to quantify such as number of offspring; others may be easy to automate given an investment in the required technical support.

Automated phenotypic assays in *Drosophila* tend to favor either gene expression during embryogenesis, where molecular marker extraction and detection can be automated, or locomotor behavior which, by using *Drosophila* Activity Monitors (DAMs), can be recorded relatively easily for large numbers of flies over periods of many days (**Figure 1A**; TriKinetics Inc., Waltham, MA, USA; Rosato and Kyriacou, 2006; Rosato et al., 2006).

The richness of the data gathered using DAMs makes it an ideal automated assay for genetic analysis of toxins. DAM data always provides, at a minimum, an estimate of mean locomotor activity level for the duration of the experiment. However, by measuring locomotor activity under the influence of a controlled light–dark cycle the photosensitivity pattern of activity can be analyzed. Using regulated light–dark cycles allows for determining the distribution of activity between the two phases of the photoperiod; the effect of light–dark cycle length other than 24 h; the amplitude of "daily" activity rhythms, the percentage of flies that show significant daily rhythms; and the timing, length, and distribution of sleep periods (Shaw et al., 2000).

Finally, by replacing light–dark cycles with constant darkness the free-running circadian period can be estimated. This provides a direct measure of the output of the biological circadian pacemaker driving the activity rhythm. Toxins that alter the circadian periodicity pattern of locomotor activity are likely, therefore, to perturb the circadian synchronization of many different biological functions regulated by a common central circadian clock. Temporal analysis of behavioral effects of toxins could also suggest the presence of daily rhythms in sensitivity to specific toxins, or in the manifestation of toxic effects. In humans, certain forms of cancer chemotherapy are more effective when administered at specific times of day (Levi et al., 2010; Li et al., 2010); other toxins may also vary in their effects during the rhythm of daytime sunlight and nighttime darkness.

In short, DAMs allow for analysis of toxic effects on both the level of locomotor activity *per se* and on its temporal pattern of expression (von Mayersbach, 1975; Mayersbach, 1976).

DROSOPHILA ACTIVITY MONITORS AND ACTIVITY ASSAYS

Behavioral toxicogenetic analysis is a relatively new application of DAM technology. DAMs were originally devised for genetic analysis of the circadian oscillator controlling locomotor activity. Locomotor activity in fruit flies, similar to running wheel activity in rodents, is an excellent reporter phenotype for underlying circadian clock function (Takahashi et al., 2008). Since precise analysis of circadian rhythms require frequent, continuous long-term data sampling for time-series estimation of periodic oscillations in the 24-h range, DAMs are designed to produce a data-rich profile of the locomotor activity of individual fruit flies. Flies are housed, usually singly, in a 5 mm × 60 mm transparent tube with enough food to last for approximately 2 weeks. A single monitor holds 32 tubes, and a photobeam in the center of each tube tracks locomotor activity as numbers of beam-crossing per unit of time, usually 10-min intervals.

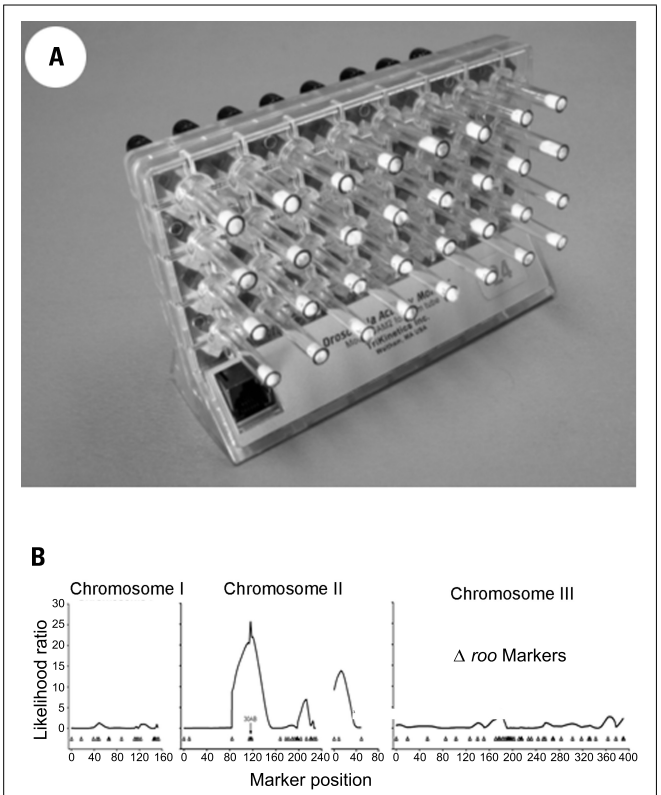


FIGURE 1 | (A) *Drosophila* Activity Monitor with glass tubes containing one adult fly, food at the far end and a cotton plug visible on the other end of each tube. **(B)** A behavioral QTL at region 30AB of *Drosophila melanogaster* chromosome two, for changes in activity level induced by developmental exposure to lead.

Drosophila Activity Monitors permit a variety of experimental designs. It is relatively easy to assay developmental effects of toxins by adding them to the fly culture food during egg, larval, or pupal development, and then monitoring adults. Effects of toxins can also be assayed acutely by adding them to the food in the activity monitor tubes. A single DAM, approximately 5″ long, 4″ high, and 3″ deep holds 32 flies so one refrigerator-sized incubator can house thousands of flies per experiment. The efficiency of DAMs facilitates experimental analysis of sex differences, dose–response curves, sensitive periods during development, interaction among multiple toxins, *trans*-generational effects, and various types of genetic analysis including behavioral QTL. It is also just as easy to monitor survival time of individual flies to the minute in assays of lethal toxic effects and degrees of resistance to them.

BEHAVIORAL QTL ANALYSIS OF DEVELOPMENTAL EXPOSURE TO TOXINS

We have identified a toxin-induced behavioral QTL by assaying locomotor activity levels for a set of recombinant inbred (RI) fly lines raised on medium containing 250 μM lead acetate or control medium made with 250 μM sodium acetate. The variation among the RI fly lines in the difference in mean activity level per line between the two treatments indicated a behavioral QTL in the

30AB region of the second *Drosophila* chromosome (**Figure 1B**; Hirsch et al., 2009). This QTL was independently identified as the site of an expression QTL (eQTL) in response to the same lead treatment in the same set of RI lines (RILs). Using the gene expression assay, we further showed the 30AB eQTL to function as a master modulatory regulator of approximately 70 additional eQTLs throughout the genome in response to developmental lead exposure (Ruden et al., 2009). Current research underway in our labs will map these behavioral QTLs in response to developmental lead exposure more precisely.

PART 2: SYNAPSES AND Ca^{2+} REGULATION IN *DROSOPHILA*

Early findings suggested that chronic Pb^{2+} exposure could produce its behavioral effects by altering synaptic development. Mammalian studies reported a variety of alterations in synaptic morphology and physiology in the brains of animals exposed to Pb^{2+} during development (Petit and LeBoutillier, 1979; Kiraly and Jones, 1982; Altmann et al., 1993). Today, the synapse remains a focus for studying the effects of toxins on brain development. Here the *Drosophila* larval neuromuscular junction (NMJ) offers a distinct advantage since one can compare the same, identified synapse in controls and treated animals (**Figure 2**). Note that this neuromuscular system is unique even amongst invertebrates because there is a stereotypic pattern of neuromuscular connections where both the presynaptic (motor neuron) and postsynaptic cells (muscle fiber) can be uniquely identified (Keshishian et al., 1996). These neuromuscular synapses have been used to demonstrate the effects of second messengers, cell-adhesion molecules and their modulators, and impulse activity on synaptic development (Budnik, 1996; Budnik et al., 1996; Davis et al., 1996).

We found that indeed one could detect synaptic abnormalities resulting from chronic Pb^{2+} exposure at the *Drosophila* larval NMJ. Initial studies found that the motor terminal formed by motoneuron RP3 on Muscle Fiber 6 (MF6) showed greater variability in size for animals raised in media containing Pb^{2+} compared to controls (Morley et al., 2003). One could imagine that increasing the variability of synaptic size could result in a loss of fine-tuning of synaptic function; this effect, if seen in the brain, would likely be detrimental to circuit dynamics and behavior. It is noteworthy that this synaptic change was subtle and it likely would have been difficult to detect when making comparisons among a large population of synapses in the mammalian brain. Given that the behavioral effects of toxins, such as Pb^{2+}, can be subtle, it seems reasonable that the effects on synapses might also be small.

Pb^{2+} affects proteins that bind Ca^{2+} including those that regulate the intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$), such as Ca^{2+} channels and Ca^{2+} pumps. Changes in Ca^{2+} regulation could be particularly important for the synapse since $[Ca^{2+}]_i$ controls multiple steps in synaptic development; e.g., growth cone guidance (Jin et al., 2005), synapse formation (Xu et al., 2009), and synapse elimination and stabilization (Pratt et al., 2003; Lohmann and Bonhoeffer, 2008). In addition, altered Ca^{2+} regulation could influence transmitter release (Zucker, 1996) and both long-term and short-term forms of synaptic plasticity at the mature synapse (Zucker and Regehr, 2002; MacDonald et al., 2006).

Acute Pb^{2+} exposure blocks Ca^{2+} channels. This has been demonstrated for a variety of voltage-dependent Ca^{2+} channels in both invertebrates (Audesirk and Audesirk, 1989) and mammals (Evans et al., 1991). The acute application of micromolar concentrations of Pb^{2+} can reduce the activity of the plasma membrane Ca^{2+} ATPase (PMCA) in humans and rats (Mas-Oliva, 1989; Sandhir and Gill, 1994a,b); however the PMCA can be stimulated by low Pb^{2+} concentrations (Mas-Oliva, 1989; Campagna et al., 2000). In addition, chronic *in vivo* Pb^{2+} exposure has been found to produce a persistent inhibition of PMCA activity in human erythrocytes (Campagna et al., 2000) and rat synaptosomes (Sandhir and Gill, 1994a,b).

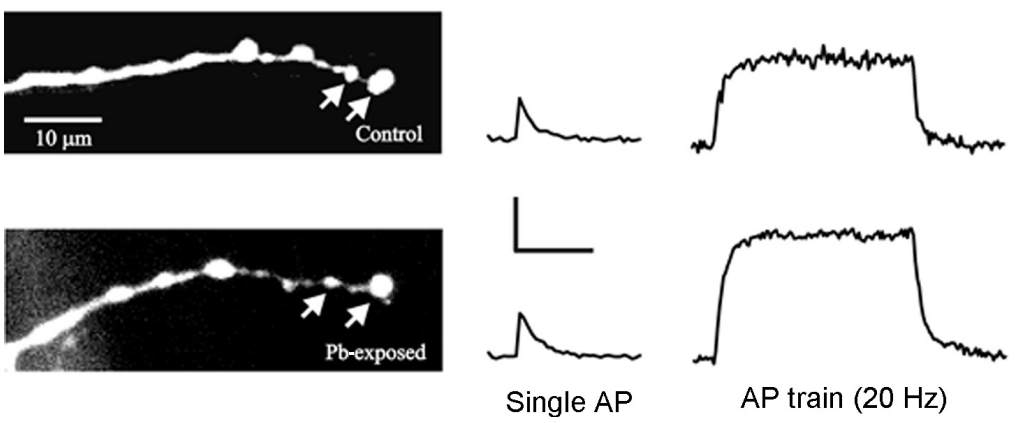


FIGURE 2 | Ca^{2+} transients recorded from identified synaptic terminals in control and Pb^{2+}-exposed animals. (Left) The RP3 motor terminals on MF6 were filled with OGB-1 in control larvae and in those exposed to Pb^{2+}. The arrows point to two typical synaptic boutons where Ca^{2+} transients were measured. (Right) Typical Ca^{2+} transients produced by single action potentials (AP) and AP trains for control and Pb^{2+}-exposed boutons. For single APs, the Ca^{2+} transients were similar for control and Pb^{2+}-exposed synaptic boutons; however, the Ca^{2+} transients produced by AP trains were larger and decayed more slowly in Pb^{2+}-exposed boutons compared to controls. Calibration: single Ap – 20% ΔF/F, 0.4 s; AP train – 20% ΔF/F, 2 s. Adapted from He et al. (2009).

We studied the effect of chronic Pb^{2+} exposure on presynaptic Ca^{2+} regulation at the RP3-to-MF6 synapse in *Drosophila* larvae (He et al., 2009). Ca^{2+} indicators were loaded in these motor terminals and we measured the changes in $[Ca^{2+}]_i$ produced by single action potentials and action potential trains (**Figure 2**). We found that chronic exposure to Pb^{2+} resulted in a greater increase in $[Ca^{2+}]_i$ during trains of action potentials and a slower decay of postsynaptic $[Ca^{2+}]_i$ at the end of the train. This is likely due to a decrease in the activity of the PMCA and it provides an interesting parallel to the effect of Pb^{2+} exposure on the PMCA in mammals. The direct effect of this large increase in $[Ca]_i$ was that the Pb^{2+}-exposed animals showed greater synaptic facilitation. This was consistent with the residual Ca^{2+} model for synaptic facilitation (Zucker and Regehr, 2002) and with mammalian studies showing that knocking down expression of the PMCA resulted in enhanced synaptic facilitation (Empson et al., 2007). In that study, reduced PMCA expression also produced changes in neuronal structure (Empson et al., 2007) and it may be that the Pb^{2+}-induced changes in the structure of the larval NMJ resulted from reduced PMCA activity and altered Ca^{2+} regulation.

In summary, identified NMJs in *Drosophila* larvae allow for the detection of subtle changes in synaptic structure and function resulting from developmental exposure to toxins. Our evidence suggests that chronic Pb^{2+} exposure produces parallel synaptic changes in *Drosophila* and mammals; this would not be surprising given that chemical synapses found in vertebrates and invertebrates are very similar.

PART 3: GENETICAL GENOMICS STUDIES IN *DROSOPHILA*

Developmental neurotoxicology research requires an approach that reduces the candidate toxin-regulated genes to a manageable number. Fortunately, a new multi-dimensional strategy has been developed, called genetical genomics, which identifies master modulatory loci that regulate the expression of hundreds of other genes (Jansen and Nap, 2001; Broman, 2005; de Koning and Haley, 2005; Li and Burmeister, 2005; Rockman and Kruglyak, 2006). Genetical genomics combines two methodologies, microarray-based whole transcriptome analyses and extracting QTLs by using RILs. Global gene expression levels are determined for each RIL, and then QTL mapping software (e.g., R/QTL; Broman et al., 2003) is used to find above-threshold statistical associations between specific chromosomal loci and transcript levels for all the genes measured on the microarrays. With QTL analysis, we can identify the specific chromosomal loci that regulate genes, which are referred to as eQTLs (Mueller et al., 2006; West et al., 2007; Majewski and Pastinen, 2011; Zhang et al., 2012).

Our research is the first that combines genetical genomics with toxicogenomics, an experimental approach we call "genetical toxicogenomics" (Ruden et al., 2009). Li et al. (2008) showed that adding environmental perturbations in genetical genomics studies, as we did in our study (Ruden et al., 2009), allows toxin-response genes to be identified. We believe that this genetical toxicogenomics approach will drive the field of toxicology and aid in understanding the effects of toxins such as lead.

Expression QTL analyses were performed for all ~18,000 genes and other microarray features by treating the expression level of each gene as a quantitative trait. The genetics of gene expression in RILs can be mapped as eQTLs. Flies from each of 75 RILs were fed, from egg to adult, either control food or lead-treated food (made with 250 μM lead acetate). RNA expression analyses of whole adult male flies (5–10 days old) were performed with Affymetrix Dros2 whole genome arrays (18,952 probe sets). Loci that are linked to a gene are called locally acting or *cis*-eQTL, whereas loci that are distantly acting are called *trans*-eQTLs (Broman, 2005).

The first genetical genomics study to identify genes with significant GxE interactions was done in *Caenorhabditis elegans*; the authors identified a group of genes with *trans*-eQTL that are induced by heat shock, which they called plastic QTL (Li et al., 2006). Smith and Kruglyak (2008) recently performed a detailed analysis of GxE-eQTL in yeast (which they call "*gxe*QTL") grown in either glucose or ethanol as the sole carbon source. Others have identified GxE interactions in which the environment is a different tissue (e.g., brain *vs.* liver; Hovatta et al., 2007). In our studies, among the 1,389 genes with *cis*-eQTL, there were 405 genes unique to control flies and 544 genes unique to lead-treated ones (440 genes had the same *cis*-eQTLs in both samples).

There were 2,396 genes with *trans*-eQTL that mapped to 12 major hotspots that met statistical significance ($p < 0.05$, chi-squared test, based on permutation analyses, on 5-cM windows). Unexpectedly, we identified two hotspots, one located on the second chromosome (polytene region 30AB) and one on the third chromosome (polytene region 73D), which co-regulate 33 genes, all of which are induced by lead (**Figure 3**). We have shown by QTL analysis that marker locus 30AB contributes to lead-dependent changes in locomotion, which suggests that the genes in the 30AB hotspot can be used as a functional test to identify both the lead-dependent *trans*-regulatory factor and the common *cis*-regulatory motifs (Hirsch et al., 2009). We propose that a *trans*-regulator located at 73D increases expression of its target genes when it binds lead, regardless of the genotype of the 73D hotspot. In contrast, a second *trans*-regulator located at 30AB increases expression of the co-regulated target genes only when it has the *ORE* genotype (**Figure 3**).

This model can explain the data, but other explanations are possible. For example, the putative *trans*-regulator could increase the stability of the mRNA of the target genes, or there could be an indirect effect on steady-state mRNA levels. MEME (Multiple Em for Motif Elicitation) analyses of the genes regulated by the 30AB and 73D *trans*-regulators identified both conserved proximal promoter and 3′-untranslated region (UTR) sequences (data not shown). Fine-mapping of the genes that underlie the QTLs and molecular and biochemical analyses should enable us to determine the mechanism involved.

FUTURE STUDIES IN *DROSOPHILA* NEUROTOXICOLOGY

In this review, we describe recent results from several laboratories that are collaborating on studying the effects of lead on the developing *Drosophila* nervous system. The Possidente and Hirsch

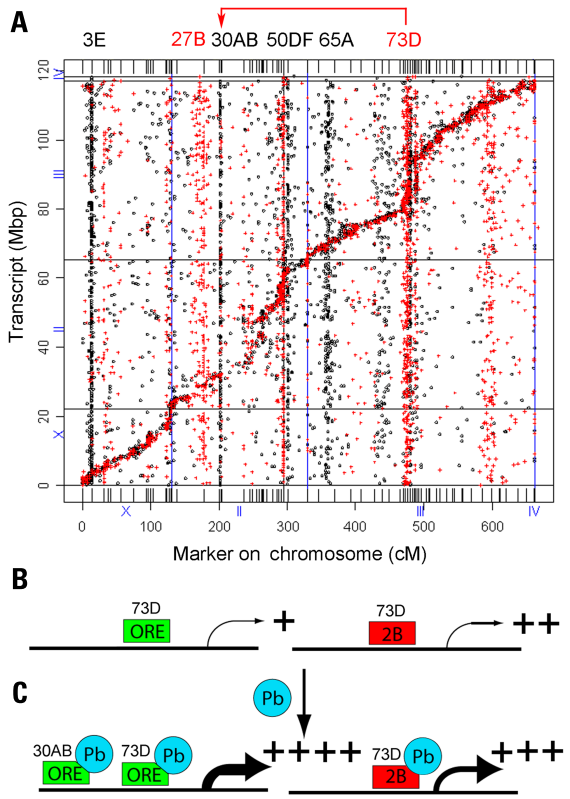


FIGURE 3 | Genetical genomics results of *Drosophila* recombinant inbred lines with developmental lead exposure. (A) The *x*-axis is the chromosome locations of the *cis*- and *trans*-eQTLs in centimorgans (cM). The *y*-axis is the location of the transcript in mega-base pairs (Mbp). The eQTLs in control (red) and the eQTLs treated (black) are indicated as dots and crosses, respectively. The locations of the control *trans*-eQTL are shown on top in red and the lead-treated *trans*-eQTL locations are shown in black letters. Many of the eQTL at 70D in the control flies shift to 30AB in the lead-treated flies (red arrow at top). **(B)** The *trans*-activator encoded by the 73D hotspot is required for basal transcription of the 33 genes co-regulated by both 73D and 30AB master modulatory genes. **(C)** In the presence of lead, the *trans*-regulator encoded by the 73D hotspot further increases steady-state mRNA levels when it has either the ORE or 2B genotype. However, the *trans*-regulator encoded by the 30AB hotspot further increases steady-state mRNA levels for the 33 target genes only when it has the ORE genotype. Adapted from Ruden et al. (2009).

laboratories study the effects of lead on behavioral quantitative traits. The Lnenicka laboratory studies the effects of lead on the NMJ in the larvae. The Ruden, Lu, and Garfinkel laboratories study the effects of lead on gene expression using quantitative genetics techniques.

Further innovations to the field of genetical genomics will make use of next-generation RNA sequencing (RNA-seq) rather than gene expression microarrays. RNA-seq will allow us not only to accurately determine gene expression levels, but also to analyze alternative splicing products of genes, thereby adding a new dimension to these studies. Since over 90% of human genes and a similar percentage of *Drosophila* genes are alternatively spliced, RNA-seq could provide some exciting and novel findings about environmental regulation of alternative mRNA splicing. The potential for RNA-seq to supplant microarrays requires, however, the development of robust statistical tools for analyzing the resulting huge datasets.

Adding metabolomics profiling to these studies will allow us to identify metabolite QTL (mQTL) that are specific for lead. To our knowledge mQTL analyses have never been conducted in any organism, although genome-wide association studies combined with metabolic profiling were conducted in a recent human study (Prakash, 2011; Suhre et al., 2011). The human study identified 37 genetic loci associated with blood metabolite concentrations, of which 25 showed effect sizes of 10–60% which is extraordinarily high for genome-wide association studies (Prakash, 2011; Suhre et al., 2011).

We are now entering the stage of understanding the complex genetic pathways that are affected by developmental lead exposure. The sophisticated genetic analyses that are possible in *Drosophila* will soon allow us to manipulate these pathways to better understand how they are affected by lead. The sequencing of several *Drosophila* strains and studying the effects of lead on neurodevelopment in these strains will allow us to better understand the evolution of lead-sensitive pathways. Together, these studies will provide a better understanding of the health effects of lead in humans.

ACKNOWLEDGMENTS

This work was supported by NIH R01 grants ES012933 to Douglas M. Ruden, and DK071073 to Xiangyi Lu.

REFERENCES

Abu-Taweel, G. M., Ajarem, J. S., and Ahmad, M. (2012). Neurobehavioral toxic effects of perinatal oral exposure to aluminum on the developmental motor reflexes, learning, memory and brain neurotransmitters of mice offspring. *Pharmacol. Biochem. Behav.* 101, 49–56.

Altmann, L., Weinsberg, F., Sveinsson, K., Lilienthal, H., Wiegand, H., and Winneke, G. (1993). Impairment of long-term potentiation and learning following chronic lead exposure. *Toxicol. Lett.* 66, 105–112.

Audesirk, G., and Audesirk, T. (1989). Effects of in vitro lead exposure on voltage-sensitive calcium channels differ among cell types in central neurons of *Lymnaea stagnalis*. *Neurotoxicology* 10, 659–669.

Broman, K. W. (2005). Mapping expression in randomized rodent genomes. *Nat. Genet.* 37, 209–210. [see comments].

Broman, K. W., Wu, H., Sen, S., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19, 889–890.

Budnik, V. (1996). Synapse maturation and structural plasticity at *Drosophila* neuromuscular junctions. *Curr. Opin. Neurobiol.* 6, 858–867.

Budnik, V., Koh, Y. H., Guan, B., Hartmann, B., Hough, C., Woods, D., and Gorczyca, M. (1996). Regulation of synapse structure and function by the *Drosophila* tumor suppressor gene dlg. *Neuron* 17, 627–640.

Campagna, D., Huel, G., Hellier, G., Girard, F., Sahuquillo, J., Fagot-Campagna, A., Godin, J., and Blot, P. (2000). Negative relationships between erythrocyte Ca-pump activity and lead levels in mothers and newborns. *Life Sci.* 68, 203–215.

Counter, S. A., Buchanan, L. H., Rosas, H. D., and Ortega, F. (1998). Neurocognitive effects of chronic lead intoxication in Andean children. *J. Neurol. Sci.* 160, 47–53.

Davis, G. W., Schuster, C. M., and Goodman, C. S. (1996). Genetic dissection of structural and functional components of synaptic plasticity. III. CREB is necessary for presynaptic functional plasticity. *Neuron* 17, 669–679.

de Koning, D. J., and Haley, C. S. (2005). Genetical genomics in humans and model organisms. *Trends Genet.* 21, 377–381.

Empson, R. M., Garside, M. L., and Knopfel, T. (2007). Plasma membrane Ca^{2+} ATPase 2 contributes

to short-term synapse plasticity at the parallel fiber to Purkinje neuron synapse. *J. Neurosci.* 27, 3753–3758.

Evans, M. L., Busselberg, D., and Carpenter, D. O. (1991). Pb^{2+} blocks calcium currents of cultured dorsal root ganglion cells. *Neurosci. Lett.* 129, 103–106.

Faroon, O., Jones, D., and de Rosa, C. (2000). Effects of polychlorinated biphenyls on the nervous system. *Toxicol. Ind. Health* 16, 305–333.

He, T., Hirsch, H. V. B., Ruden, D. M., and Lnenicka, G. A. (2009). Chronic lead exposure alters presynaptic calcium regulation and synaptic facilitation in *Drosophila* larvae. *Neurotoxicology* 30, 777–784.

Hirsch, H. V., Mercer, J., Sambaziotis, H., Huber, M., Stark, D. T., Torno-Morley, T., Hollocher, K., Ghiradella, H., and Ruden, D. M. (2003). Behavioral effects of chronic exposure to low levels of lead in *Drosophila melanogaster*. *Neurotoxicology* 24, 435–442.

Hirsch, H. V. B., Possidente, D., Averill, S., Despain, T. P., Buytkins, J., Thomas, V., Goebel, W. P., Shipp-Hilts, A., Wilson, D., Hollocher, K., Possidente, B., Lnenicka, G., and Ruden, D. M. (2009). Variations at a quantitative trait locus (QTL) affect development of behavior in lead-exposed *Drosophila melanogaster*. *Neurotoxicology* 30, 305–311.

Hovatta, I., Zapala, M. A., Broide, R. S., Schadt, E. E., Libiger, O., Schork, N. J., Lockhart, D. J., and Barlow, C. (2007). DNA variation and brain region-specific expression profiles exhibit different relationships between inbred mouse strains: implications for eQTL mapping studies. *Genome Biol.* 8, R25.

Jansen, R. C., and Nap, J. P. (2001). Genetical genomics: the added value from segregation. *Trends Genet.* 17, 388–391.

Jin, M., Guan, C.-B., Jiang, Y.-A., Chen, G., Zhao, C.-T., Cui, K., Song, Y.-Q., Wu, C.-P., Poo, M.-M., and Yuan, X.-B. (2005). Ca^{2+}-dependent regulation of rho GTPases triggers turning of nerve growth cones. *J. Neurosci.* 25, 2338–2347.

Keshishian, H., Broadie, K., Chiba, A., and Bate, M. (1996). The *Drosophila* neuromuscular junction: a model system for studying synaptic development and function. *Annu. Rev. Neurosci.* 19, 545–575.

Kiraly, E., and Jones, D. G. (1982). Dendritic spine changes in rat hippocampal pyramidal cells after postnatal lead treatment: a Golgi study. *Exp. Neurol.* 77, 236–239.

Levi, F., Okyar, A., Dulong, S., Innominato, P. F., and Clairambault, J. (2010). Circadian timing in cancer treatments. *Annu. Rev. Pharmacol. Toxicol.* 50, 377–421.

Li, J., and Burmeister, M. (2005). Genetical genomics: combining genetics with gene expression analysis. *Hum. Mol. Genet.* 14, R163–R169.

Li, X. M., Delaunay, F., Dulong, S., Claustrat, B., Zampera, S., Fujii, Y., Teboul, M., Beau, J., and Lévi, F. (2010). Cancer inhibition through circadian reprogramming of tumor transcriptome with meal timing. *Cancer Res.* 70, 3351–3360.

Li, Y., Álvarez, O. A., Gutteling, E. W., Tijsterman, M., Fu, J., Riksen, J. A. G., Hazendonk, E., Prins, P., Plasterk, R. H. A., Jansen, R. C., Breitling, R., and Kammenga, J. E. (2006). Mapping determinants of gene expression plasticity by genetical genomics in *C. elegans*. *PLoS Genet.* 2, e222. doi: 10.1371/journal.pgen.0020222

Li, Y., Breitling, R., and Jansen, R. C. (2008). Generalizing genetical genomics: getting added value from environmental perturbation. *Trends Genet.* 24, 518–524.

Lohmann, C., and Bonhoeffer, T. (2008). A role for local calcium signaling in rapid synaptic partner selection by dendritic filopodia. *Neuron* 59, 253–260.

MacDonald, J. F., Jackson, M. F., and Beazely, M. A. (2006). Hippocampal long-term synaptic plasticity and signal amplification of NMDA receptors. *Crit. Rev. Neurobiol.* 18, 71–84.

Mackay, T. F. C., and Anholt, R. R. H. (2006). Of flies and man: *Drosophila* as a model for human complex traits. *Annu. Rev. Genomics Hum. Genet.* 7, 339–367.

Majewski, J., and Pastinen, T. (2011). The study of eQTL variations by RNA-seq: from SNPs to phenotypes. *Trends Genet.* 27, 72–79.

Mas-Oliva, J. (1989). Effect of lead on the erythrocyte (Ca^{2+}, Mg^{2+})-ATPase activity. Calmodulin involvement. *Mol. Cell. Biochem.* 89, 87–93.

Mayersbach, H. (1976). Time – a key in experimental and practical medicine. *Arch. Toxicol.* 36, 185–216.

Morley, E. J., Hirsch, H. V., Hollocher, K., and Lnenicka, G. A. (2003). Effects of chronic lead exposure on the neuromuscular junction in *Drosophila* larvae. *Neurotoxicology* 24, 35–41.

Mueller, M., Goel, A., Thimma, M., Dickens, N. J., Aitman, T. J., and Mangion, J. (2006). eQTL Explorer: integrated mining of combined genetic linkage and expression experiments. *Bioinformatics* 22, 509–511.

Petit, T. L., and LeBoutillier, J. C. (1979). Effects of lead exposure during development on neocortical dendritic and synaptic structure. *Exp. Neurol.* 64, 482–492.

Prakash, S. (2011). Human metabolic individuality in biomedical and pharmaceutical research. *Circ. Cardiovasc. Genet.* 4, 714–715.

Pratt, K. G., Watt, A. J., Griffith, L. C., Nelson, S. B., and Turrigiano, G. G. (2003). Activity-dependent remodeling of presynaptic inputs by postsynaptic expression of activated CaMKII. *Neuron* 39, 269–281.

Rai, A., Maurya, S. K., Khare, P., Srivastava, A., and Bandyopadhyay, S. (2010). Characterization of developmental neurotoxicity of As, Cd, and Pb mixture: synergistic action of metal mixture in glial and neuronal functions. *Toxicol. Sci.* 118, 586–601.

Rockman, M. V., and Kruglyak, L. (2006). Genetics of global gene expression. *Nat. Rev. Genet.* 7, 862–872.

Rosato, E., and Kyriacou, C. P. (2006). Analysis of locomotor activity rhythms in *Drosophila*. *Nat. Protoc.* 1, 559–568.

Rosato, E., Tauber, E., and Kyriacou, C. P. (2006). Molecular genetics of the fruit-fly circadian clock. *Eur. J. Hum. Genet.* 14, 729–738.

Ruden, D. M., Chen, L., Possidente, D., Possidente, B., Rasouli, P., Wang, L., Lu, X., Garfinkel, M. D., Hirsch, H. V., and Page, G. P. (2009). Genetical toxicogenomics in *Drosophila* identifies master-modulatory loci that are regulated by developmental exposure to lead. *Neurotoxicology* 30, 898–914.

Sandhir, R., and Gill, K. D. (1994a). Lead perturbs calmodulin dependent cyclic AMP metabolism in rat central nervous system. *Biochem. Mol. Biol. Int.* 33, 729–742.

Sandhir, R., and Gill, K. D. (1994b). Alterations in calcium homeostasis on lead exposure in rat synaptosomes. *Mol. Cell. Biochem.* 131, 25–33.

Shaw, P. J., Cirelli, C., Greenspan, R. J., and Tononi, G. (2000). Correlates of sleep and waking in *Drosophila melanogaster*. *Science* 287, 1834–1837.

Singh, M. P., Ram, K. R., Mishra, M., Shrivastava, M., Saxena, D. K., and Chowdhuri, D. K. (2010). Effects of co-exposure of benzene, toluene and xylene to *Drosophila melanogaster*: alteration in hsp70, hsp60, hsp83, hsp26, ROS generation and oxidative stress markers. *Chemosphere* 79, 577–587.

Smith, E. N., and Kruglyak, L. (2008). Gene-environment interaction in yeast gene expression. *PLoS Biol.* 6, e83. doi: 10.1371/journal.pbio.0060083

Suhre, K., Shin, S.-Y., Petersen, A.-K., Mohney, R. P., Meredith, D., Wägele, B., A.-K., Altmaier, E., CARDIoGRAM, Deloukas, P., Erdmann, J., Grundberg, E., Hammond, C. J., Hrabé de Angelis, M., Kastenmüller, G., Köttgen, A., Kronenberg, F., Mangino, M., Meisinger, C., Meitinger, T., Mewes, H.-W., Milburn, M. V., Prehn, C., Raffler, J., Ried, J. S., Römisch-Margl, W., Samani, N. J., Small, K. S., Wichmann, H.-E., Zhai, G., Illig, T., Spector, T. D., Adamski, J., Soranzo, N., and Gieger, C. (2011). Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 477, 54–60.

Takahashi, J. S., Shimomura, K., and Kumar, V. (2008). Searching for genes underlying behavior: lessons from circadian rhythms. *Science* 322, 909–912.

UNEP. (2010). *Lead Exposure and Human Health*. Nairobi: Chemicals Branch, Division of Technology, Industry and Economics, United Nations Environment Program. Available at: www.chem.unep.ch/pops/pdf/lead/leadexp.pdf [accessed January 3, 2012].

von Mayersbach, H. (1975). Proceedings: Time – a key in experimental and practical medicine. *Naunyn Schmiedebergs Arch. Pharmacol.* 287(Suppl.), R107.

West, M. A., Kim, K., Kliebenstein, D. J., van Leeuwen, H., Michelmore, R. W., Doerge, R. W., and St. Clair, D. A. (2007). Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. *Genetics* 175, 1441–1450.

Xu, F., Hennessy, D. A., Lee, T. K. M., and Syed, N. I. (2009). Trophic factor-induced intracellular calcium oscillations are required for the expression of postsynaptic acetylcholine receptors during synapse formation between *Lymnaea* neurons. *J. Neurosci.* 29, 2167–2176.

Zhang, X., Huang, S., Sun, W., and Wang, W. (2012). Rapid and robust resampling-based multiple testing correction with application in genome-wide eQTL study. *Genetics*. doi: 10.1534/genetics.111.137737

Zucker, R. S. (1996). Exocytosis: a molecular and physiological perspective. *Neuron* 17, 1049–1055.

Zucker, R. S., and Regehr, W. G. (2002). Short-term synaptic plasticity. *Annu. Rev. Physiol.* 64, 355–405.

Conflict of Interest Statement: 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.

Received: 29 February 2012; paper pending published: 19 March 2012; accepted: 09 April 2012; published online: 04 May 2012.
Citation: Hirsch HVB, Lnenicka G, Possidente D, Possidente B, Garfinkel MD, Wang L, Lu X and Ruden, DM. (2012) Drosophila melanogaster as a model for lead neurotoxicology and toxicogenomics research. Front. Gene. ***3****:68. doi: 10.3389/fgene.2012.00068*
This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in
GENETICS

REVIEW ARTICLE
published: 10 April 2012
doi: 10.3389/fgene.2012.00052

Genome-wide analyses of metal responsive genes in *Caenorhabditis elegans*

Samuel Caito[1,2], Stephanie Fretham[1], Ebany Martinez-Finley[1,2], Sudipta Chakraborty[2,3], Daiana Avila[4], Pan Chen[1] and Michael Aschner[1,2,3,5]*

[1] Division of Clinical Pharmacology and Pediatric Toxicology, Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN, USA
[2] Center in Molecular Toxicology, Vanderbilt University Medical Center, Nashville, TN, USA
[3] Center for Molecular Neuroscience, Vanderbilt University Medical Center, Nashville, TN, USA
[4] Biochemistry Graduate Program, Universidade Federal do Pampa, Uruguaiana, Rio Grande do Sul, Brazil
[5] The Kennedy Center for Research on Human Development, Vanderbilt University Medical Center, Nashville, TN, USA

Edited by:
Stephen Sturzenbaum, King's College London, UK

Reviewed by:
Jonathan Freedman, National Institute of Health, USA
Suresh Chandra Swain, Aberystwyth University, UK

***Correspondence:**
Michael Aschner, Division of Pediatric Toxicology, Vanderbilt University Medical Center, 11425 MRB IV, 2215-B Garland Avenue, Nashville, TN 37232-0414, USA.
e-mail: michael.aschner@vanderbilt.edu

Metals are major contaminants that influence human health. Many metals have physiologic roles, but excessive levels can be harmful. Advances in technology have made toxicogenomic analyses possible to characterize the effects of metal exposure on the entire genome. Much of what is known about cellular responses to metals has come from mammalian systems; however the use of non-mammalian species is gaining wider attention. *Caenorhabditis elegans* is a small round worm whose genome has been fully sequenced and its development from egg to adult is well characterized. It is an attractive model for high throughput screens due to its short lifespan, ease of genetic mutability, low cost, and high homology with humans. Research performed in *C. elegans* has led to insights in apoptosis, gene expression, and neurodegeneration, all of which can be altered by metal exposure. Additionally, by using worms one can potentially study mechanisms that underline differential responses to metals in nematodes and humans, allowing for identification of novel pathways and therapeutic targets. In this review, toxicogenomic studies performed in *C. elegans* exposed to various metals will be discussed, highlighting how this non-mammalian system can be utilized to study cellular processes and pathways induced by metals. Recent work focusing on neurodegeneration in Parkinson's disease will be discussed as an example of the usefulness of genetic screens in *C. elegans* and the novel findings that can be produced.

Keywords: metals, gene expression, *C. elegans*, apoptosis, neurodegeneration

INTRODUCTION

Metals are persistent environmental contaminants that have been associated with developmental and behavioral deficits and neurodegeneration (Kordas, 2010; Koyashiki et al., 2010; Neal and Guilarte, 2010). Several metals, such as Zn, Mn, and Fe, are essential for cell viability and function; however excessive levels of metals can be deleterious. Metals can generate free radicals and reactive oxygen species (ROS) through Fenton chemistry, leading to protein and DNA oxidation and lipid peroxidation. Oxidation of macromolecules can damage cellular organelles and generate additional ROS, thus creating a vicious cycle. Exposure to metals can range in their clinical symptoms, but similar cellular effects are observed, such as oxidative stress. With the advent of complete genome sequencing and genome-wide mapping, it is now possible to identify novel genes involved in metal homeostasis, detoxification, and involved in response to toxic exposures. One important consideration in performing a toxicogenomic screen is the choice of model organism. Traditional mammalian models, such as rats and mice have provided vital information about metal exposures. However rodents are complex organisms where subtle, yet important, changes may not be observed due to the numerous cell types and their relative abundance in the tissue examined. Non-mammalian model organisms provide an alternative, less complex system to perform these screens while maintaining high homology to mammals.

Caenorhabditis elegans is an attractive model for studying effects of metals on gene expression. *C. elegans* are small (∼1.5 mm) round worms with a short lifespan of 3 weeks and a life cycle of 3 days, that can be maintained cheaply in a humid environment containing atmospheric oxygen and bacteria grown on agar as a food source. A single adult hermaphrodite is capable of producing ∼300 progeny. Upon hatching (L1 phase), worms proceed through three subsequent larval stages (L2 after 12 h, L3 after 8 h, and L4 after 10 h). As an adult, an hermaphrodite is able to self-fertilize, or may be used for mating with an adult male, a rare sex with a frequency of 0.5%, allowing for crossing

Abbreviations: AD, Alzheimer's disease; AO, acridine orange; *C. elegans*, *Caenorhabditis elegans*; DAergic, dopaminergic; DTT, dichlorodiphenylchloroethane; eIF2α, eukaryotic translation initiation factor; ER, endoplasmic reticulum; ERAD, ER-associated degradation; HD, Huntington's disease; HSP, heat shock protein; LRRK2, leucine-rich repeat kinase 2; Nrf2, nuclear factor (erythoid-derived 2)-like 2; PD, Parkinson's disease; PINK1, PTEN-induced novel kinase 1; PTEN, phosphatase and tensin homolog; SNpc, substantia nigra pars compacta; ROS, reactive oxygen species; UPR, unfolded protein response.

worms of different genotypes. Importantly, worms contain many of the transporters and stress response genes critical for xenobiotic and metal detoxification, including metallothioneins (MTs), transporters involved in metal homeostasis, heat shock proteins (HSPs), and genes involved in glutathione homeostasis.

Because of the short lifespan and ease of manipulability, *C. elegans* are a perfect model for high throughput toxicogenomic screens. *C. elegans*' body comprises less than 1000 cells, allowing for the mapping of every cell's development and lineage (Sulston and Horvitz, 1977). This allows one to investigate not only general overall toxicity, but neurotoxicity, reproductive toxicity, muscular toxicity, toxicity to the GI track, and developmental toxicity. The *C. elegans* genome is fully sequenced and shows 60–80% homology with mammals (Kaletta and Hengartner, 2006; McDonald et al., 2006). The genetic architecture of *C. elegans* genes is well characterized and there are standardized protocols for knocking down genes through RNAi and introducing DNA through injections, allowing for the creation of numerous strains of knockout alleles and transgenic overexpression alleles. These advantages allow for quick but informative investigations in signaling pathways and gene–environment interactions that may be more difficult in more complex systems. It is possible that genes induced by metal exposure in *C. elegans* may differ from mammals, but this may provide insight into pathways and protective genes that can be investigated as therapeutic targets. In this review, we highlight recent toxicogenomic screens in *C. elegans*, including screens in various worm lines, RNAi feeding screens, and microarray analyses, to characterize genes and pathways altered by metal exposure. Additionally we discuss novel findings from studies in *C. elegans* of metal-induced neurodegeneration in Parkinson's disease (PD) as an application of toxicogenomic screens in studying human disease in a non-mammalian system.

DNA DAMAGE, GENE EXPRESSION, EPIGENETICS

Various metals damage DNA, change gene expression profiles, and epigenetically modify DNA, consequences which can contribute to the toxicity of metal exposure. Many of these alterations have been identified and characterized using mammalian systems (Bal and Kasprzak, 2002; Beyersmann and Hartwig, 2008; Salnikow and Zhitkovich, 2008; Robinson et al., 2010), however there are novel genetic screens in *C. elegans* that can be utilized to examine metal-induced genetic changes.

A survey for naturally occurring single base substitutions assayed the entire genomes of 10, mutation accumulation *C. elegans* strains generated by bottlenecking individual wildtype animals for many generations (Denver et al., 2009). High throughput synthesis- and pyro-DNA sequencing revealed more base substitutions than previous studies and found that G:C → A:T and G:C → T:A mutations occurred at higher than expected rates, indicative of oxidative DNA damage. While this study did not directly examine metal exposure, it suggests that, as in mammalian systems, oxidative stress resulting from metals likely causes DNA damage in *C. elegans* (Bal and Kasprzak, 2002). DNA strand breaks are another form of DNA damage induced by exposure to metals such as Cd, Cu, and Zn in *C. kiiensis* invertebrate larvae (Al-Shami et al., 2012). This single cell assay has not been applied specifically to metal toxicity in *C. elegans*, however it has been used to examine nicotine genotoxicity (Sobkowiak and Lesicki, 2009).

Microarray-technology alone and in combination with other techniques, has demonstrated changes in gene expression following Ag and Cd exposure *C. elegans*. Whole genome microarray identified altered expression of 1632 genes after exposure to Ag nanoparticles and ions, including genes involved in oxidative stress and metabolism (Roh et al., 2009). In addition, this study combined genetic changes with *in vivo* toxicological outcomes of survival, growth, and reproduction. Another study identified altered expression of 290 genes using whole genome microarray following Cd exposure (Cui et al., 2007). Many of the Cd-altered genes had previously known functions including metabolism, proteolysis, fatty acid, and lipid metabolism, although the majority of identified genes were novel. RNAi against 92 of the most significantly altered genes was used to screen for interactions between gene function and growth and reproductive phenotypes associated with Cd exposure (Cui et al., 2007).

The short lifespan of *C. elegans*, presence of epigenetic machinery, and heritability of epigenetic alterations (Greer et al., 2011; Wenzel et al., 2011) make it possible to examine epigenetic contributions to metal toxicity in the worm. One recent study combined several genomic assays including microarray, ChIP-seq with histone modification specific antibodies, and RT-PCR to assess global genomic differences between adults that had experienced Dauer and those that had not (Hall et al., 2010). Such an approach has not yet been applied to the toxicogenomic consequences of metal exposure, but would be highly informative particularly for non-lethal and latent toxicity.

ANTIOXIDANT RESPONSE TO METAL EXPOSURE

Another way in which metals damage cells is through oxidative stress. There are several *C. elegans* mutants that are hypersensitive to oxidative stress including *mev-1*, *mev-3*, *gas-1*, *rad-8*, *skn-1*, and *nnt-1* (Ishii et al., 1990; Yamamoto et al., 1996; Hartman et al., 2001; Kayser et al., 2001; An and Blackwell, 2003; Arkblad et al., 2005), with *nnt-1*, and *skn-1* primarily involved in antioxidant defense. Many genes are involved in generation of ROS, Fujii et al. (2009) screened for *C. elegans* mutants that showed increased sensitivity to oxidative stress, and isolated a novel mutant, *oxy-4(qa5001)* (Fujii et al., 2009). This mutant showed an increased sensitivity to a high concentration of oxygen, and decreased longevity that is temperature-dependent. Genetic analysis showed that *oxy-4* had a mutation in an [FeFe]-hydrogenase-like gene, responsible for catalyzing the formation and the splitting of molecular hydrogen ($H_2 \leftrightarrow 2H^+ + 2e^-$), and functioning in anaerobic respiration but whose properties are less characterized in aerobic eukaryotes. Results from this study suggest that [FeFe]-hydrogenase-like genes are involved in the regulation of sensitivity to oxygen in *C. elegans* (Fujii et al., 2009).

The *daf-2* insulin-like signaling pathway plays a major role in *C. elegans* longevity by controlling the expression of a large number of genes, including free radical detoxifying genes (Honda and Honda, 1999). Samuelson and colleagues performed a genome-wide RNAi screen to identify genes necessary for the extended lifespan of *daf-2* mutants and identified ∼159 gene inactivations responsible for shortening the lifespan. Endocytosis and vesicular trafficking to

lysosomes made up a majority of the genes identified for lifespan extension in the *daf-2* mutant. Decreased *daf-2* signaling causes nuclear localization of DAF-16 and results in upregulation of manganese superoxide dismutase (SOD-3), whose function is to protect cells from oxidative stress. Thirty-four gene inactivations were identified that suppressed the induction of SOD-3 expression in non-neuronal cells (Samuelson et al., 2007).

Toxicogenomic studies have described the utility of *C. elegans* as an environmental monitor and biosensor. Chu et al. (2005) tested a number of mutant strains defective in genes controlling ROS response for enhanced sensitivity to metals. A double mutant *daf-16 unc-75* strain was identified as most sensitive, exhibiting a sixfold increase in sensitivity to Cd, threefold increase in sensitivity to Cu, and twofold increase in sensitivity to Zn compared to wild-type worms. Roh et al. (2006) investigated the toxicity of Cd, Pb, Cr, and As using a transgenic *C. elegans* model. Cd led to a more than threefold increase in HSP 16.2, HSP 70, metallothionein 2, cytochrome P450 family protein 35A2, glutathione-*S*-transferase 4, superoxide dismutase 1, catalase 2, *C. elegans* p53-like protein 1, and apoptosis enhancer 1 genes compared to controls. The Pb-, Cr-, and As-exposed nematodes showed little change in gene expression showing that Cd has a higher tolerance level compared to the other metals tested. Cd- and Cr-exposed worms also exhibited alterations in growth and reproduction. This study illustrated that stress responses must be measured following exposure to several different metals (Roh et al., 2006). The DNA microarray experiments of Cui et al. (2007) examined the toxicogenomic response to Cd and identified 237 up-regulated and 53 down-regulated genes that significantly changed following either 4 or 24 h exposure. Early response genes were those that regulate the localization and transport of various metals and are important for ion homeostasis. Cd exposure resulted in the overexpression of 25 biotransformation genes, proteolysis genes, and expression of four ABC transporters, *pgp-1*, *pgp-8*, *pgp-9*, *and mrp-3*. Fatty acid and cellular lipid metabolism and cell wall catabolism pathways represented a significant portion of the downregulated gene set (Cui et al., 2007).

Caenorhabditis elegans' SKN-1 [mammalian nuclear factor (erythoid-derived 2)-like 2, Nrf2] is critical for oxidative stress resistance. Wang et al. (2009) conducted a genome-scale RNAi screen to identify mechanisms that prevent inappropriate *skn-1* target gene expression under non-stressed conditions (Wang et al., 2009). They identified 41 genes that if knocked down lead to activation of a SKN-1 target gene, *gcs-1*. These genes represent many cellular processes, including mRNA translation. Results suggest that SKN-1 acts as a monitor of many metabolic and regulatory processes.

HEAT SHOCK PROTEINS AND ER STRESS

Cellular functions are greatly dependent upon efficient and accurate cooperation between different proteins, relying on efficient protein synthesis, processing, trafficking, and degradation. Thus, it is not surprising that cells have evolved several defense mechanisms to address misfolded and damaged proteins. Molecular chaperone proteins are highly conserved and ubiquitously expressed in all subcellular compartments and are essential for the stability of the proteome under normal and stressful conditions (Frydman, 2001). The expression of many molecular chaperones is regulated by different forms of environmental and physiological stresses that can interfere with folding stability, thus leading to a flux of misfolded proteins. Stress responsive molecular chaperones are referred to as HSPs and classified by gene families according to their molecular mass. HSPs exert their physiological effect by assisting in the formation of new proteins as well as by preserving existing structures. However, they also display major functions in pathological conditions, especially through structural rectification of denatured proteins and solubilization of protein aggregates carrying them on to the proteasome system (Soti et al., 2005; Powers et al., 2010).

Despite the robust nature of the heat shock response and the capacity of chaperones to recognize misfolded proteins, chronic exposure to metals even at low doses, or acute toxicity at high doses may result in the accumulation of misfolded and damaged proteins. The accumulation of alternate folded states and toxic species overburdens and functionally depletes the proteostasis machinery, which in turn amplifies protein damage (Gidalevitz et al., 2006, 2009). This suggests that the regulation of the protein quality control system is essential for proteostasis to monitor the state of the proteome throughout the lifetime of an organism. Hence, HSP expression has emerged as an indicator of cellular stress in animals exposed to metals at sub-toxic and toxic doses.

Toxicogenomics has helped in elucidating whether HSP levels are altered by metals exposure. HSP levels vary after metal exposure depending on the metal, type of cell or organ studied, and type of chaperone evaluated (Cheng et al., 2003; Qian et al., 2005; Rodella et al., 2008; Escobar Mdel et al., 2009). Some metals such as Pb, Cd, Al, and Hg can directly inhibit protein folding or inhibit their expression, thus causing accumulation of unfolded proteins and endoplasmic reticulum (ER) stress (Rodella et al., 2008; Sharma et al., 2008; Aremu et al., 2011). There are several diseases associated with long-term metal exposure, increased misfolded protein aggregation, and disturbance in the ubiquitin/proteasome system including Alzheimer's disease (AD), PD, and polyglutamine diseases, such as Huntington's disease (HD; Imaizumi et al., 2001; Gitler et al., 2009; Squitti, 2012). Once proteins are misfolded, they can aggregate and damage the cell. Metal-induced oxidative stress has been shown to promote protein aggregation (Santner and Uversky, 2010; Dudzik et al., 2011). Additionally, metals can also perform aberrant interactions with proteins such as beta-amyloid, α-synuclein, and prion proteins (Santner and Uversky, 2010; Hong and Simon, 2011). For instance, Zn and Al strongly stimulate the aggregation/fibrillogenesis of variations of ataxin-3, a protein involved in the polyglutamine disease ataxia (Ricchelli et al., 2007). In addition, a series of mono, di, and trivalent metal ions such as Li, Mn, Zn, Cu, Al can cause significant acceleration of α-synuclein fibril formation *in vitro* (Santner and Uversky, 2010). These processes trigger HSP activation and consequently, a defense response in order to detoxify and protect the cell against the oxidative injury and alterations in the proteostasis caused by the pro-oxidant (Stefani and Dobson, 2003; Dudzik et al., 2011; Hong and Simon, 2011). Several studies using toxicogenomics have demonstrated that HSP levels are elevated following metal exposure in different animal models. In *C. elegans*, Anbalagan et al. demonstrated that Cd, Cu, Zn, and both

organic MeHg and inorganic Hg exposure induced the heat shock genes, HSP-16.1, HSP-16.2, HSP-6, HSP-60 were at least twofold increased, using transgenic GFP strains (Helmcke and Aschner, 2010; Anbalagan et al., 2012). Currently, there is little data unraveling the mechanisms that underlie metal detoxification via HSP, and toxicogenomics will be an important tool in this research.

The ER serves as the major organelle for protein synthesis, modification, folding, and transportation. Disruption of normal ER function results in accumulation of unfolded or misfolded proteins inside ER lumen, known as ER stress, triggering the unfolded protein response (UPR), in order to decrease protein synthesis, enhance protein folding, and degradation. The molecular chaperone GRP78/BiP (immunoglobulin-binding protein) serves as a sensor in this signaling pathway. When ER functions normally, BiP binds to three ER transmembrane proteins, including an endoribonuclease IRE1, protein kinase PERK, and a transcription factor ATF6. Upon ER stress, BiP dissociates from these proteins and binds to unfolded protein to improve folding, thus releasing these three proteins from ER membrane and activating subsequent signaling pathways. These three transmembrane proteins are corresponding to three distinct UPR signaling branches. The transmembrane kinase PERK phosphorylates its downstream target, eukaryotic translation initiation factor (eIF2α), through which the protein synthesis is shut down. IRE1 cleaves a transcription factor xbp-1 pre-mRNA, which then gets translated and translocated into nucleus to activate genes for protein degradation, such as ER-associated degradation (ERAD) complex. ATF6 dissociates from ER membrane, travels to Golgi apparatus and gets C-terminal cleaved, releasing the N-terminal cytosolic fragment to enter the nucleus, and activate targets genes. These genes are mainly ER resident chaperones (such as BiP and protein disulfide isomerase), upregulation of which enhances protein folding in the ER. Activation of three UPR branches would alleviate the stress, thus restore ER protein homeostasis; on the other hand, if the stress remains unsolved, apoptosis is activated to kill these cells. To date, ER stress has been found in many human neurological disorder diseases, including early-onset torsion dystonia (Chen et al., 2010), AD (Hoozemans et al., 2005, 2009), PD (Jiang et al., 2010; Calì et al., 2011), amyotrophic lateral sclerosis (Atkin et al., 2008; Nishitoh et al., 2008; Mori et al., 2011), and HD (Reijonen et al., 2008).

Recently heavy metal exposure has been associated with ER stress. Gardarin et al. (2010) found that ER was the major target of Cd exposure in yeast *Saccharomyces cerevisiae*. Upon 50 μM of Cd treatment, splicing of *HAC1* (yeast *xbp-1*) mRNA was observed, Δ*ire1* and Δ*hac1* strains showed hypersensitive to Cd, similar to ER stress inducer tunicamycin and dichlorodiphenylchloroethane (DTT; Gardarin et al., 2010). In addition, 14 out of 16 mutants in MAPK signaling pathway, which is important for tolerance of ER stress, is also sensitive to Cd treatment (Gardarin et al., 2010). Fauchon et al. (2002) noticed activation ER chaperones PDI, FKB2, LHS1, and JEM1 and UPR components upon Cd treatment. MeHg has been shown to induce hermetic changes in GRP78 levels, with low acute doses increasing GRP78 expression but high doses decreased its expression, suggesting that low level MeHg exposure induces cytoprotective ER stress pathway (Zhang et al., 2011). In mammalian cell culture, Pb (Qian et al., 2001; Shinkai et al., 2010), Mn (Chun et al., 2001), Hg (Qian et al., 2001), Ni (Hiramatsu et al., 2007), Co (Hiramatsu et al., 2007), and Cd (Hiramatsu et al., 2007) also caused ER stress and activated UPR by upregulating BiP expression. In transgenic mice administered with Cd^{2+}, rapid and transient ER stress was induced in liver and kidney predominantly, but not obvious in other tissues (Hiramatsu et al., 2007). Although it remains unclear whether these heavy metals directly or indirectly interfere with ER functions, restoration of ER homeostasis could be a treatment of these metal exposures. Induction of ER stress has not been reported in *C. elegans* in response to metal exposure, however there is potential to investigate ER stress in transgenic strains for genes associated with diseases where ER stress has been observed.

METAL-INDUCED APOPTOSIS

Despite the existence of multiple protective pathways, cellular damage in response to metal exposure may at times be too severe. Exposure to heavy metals has been shown to induce programmed cell death or apoptosis, by influencing various components of several apoptotic pathways (Rana, 2008). However, most of these studies have been *in vitro* and they do not seem to recapitulate the effects in live animals. Consequently, the *C. elegans* model system has become a favorable biosensor in studies on the induction of apoptotic pathways *in vivo* following exposure to various types of toxicants with most studies focusing on germline apoptosis in the nematodes (Wu et al., 2006; Leung et al., 2008). In addition to the physiologically requisite germline apoptosis that occurs in nematodes, stress-induced apoptosis can occur through multiple signaling transduction pathways. Genotoxic insults have been shown to mediate apoptosis through a pathway involving the checkpoint protein HUS-1, the p53 homolog CED-9, and the necessary apoptotic CED-3 caspase (Derry et al., 2001; Hofmann et al., 2002; Conradt and Xue, 2005). However, other stressors can lead to the induction of p53-dependent or independent pathways that require MAPK-related signaling transduction in order for germline apoptosis to occur (Salinas et al., 2006; Rutkowski et al., 2011). The conservation of these pathways in *C. elegans* allows for the validation on their role in mammalian *in vivo* metal-induced apoptotic mechanisms (Caffrey et al., 1999).

The *C. elegans* germline is composed of two U-shaped gonad arms that join proximally at the uterus of the nematode. Both physiologically relevant and stress-induced germline apoptosis occurs at the loop of the gonad (Gartner et al., 2008). A common germline apoptosis assay that can be used in *C. elegans* involves staining worms with the fluorescent dye acridine orange (AO) that labels the characteristic, highly fragmented DNA of apoptotic cells (Kelly et al., 2000). For example, *C. elegans* studies found that Ni exposure can induce germline apoptosis in a dose- and time-dependent manner (Kezhou et al., 2010). This increase in Ni-induced germ cell corpses is independent of the ERK pathway as shown in *C. elegans* knockout strains for the MAPKKK (*lin-45*), MAPKK (*mek-2*), and MAPK (*mpk-1*) homologs. Moreover, mutant worms lacking *cep-1* (homolog for the tumor suppressor p53), *hus-1*, and *egl-1* (involved in DNA damage checkpoints) show increased germline apoptosis upon Ni exposure, and are not required for Ni-induced germline apoptosis when compared to wildtype Ni-exposed worms. However, Ni-induced germline apoptosis is decreased in mutant homologs of the JNK (*jkk-1*,

mek-1, *jnk-1*, and *mkk-4* loss-of-function strains) and p38 MAPK (*nsy-1*, *sek-1*, *pmk-1*, and *pmk-3* loss-of-function strains) signaling cascades, indicating the significance of these two pathways in Ni-induced apoptosis in *C. elegans* (Kezhou et al., 2010). Similarly, Cd exposure in *C. elegans* also induces germline apoptosis in a manner dependent upon the JNK and p38 MAPK pathways, as indicated by the blockage of apoptosis in Cd-exposed JNK and p38 MAPK pathway homolog knockouts (except in loss-of-function *pmk-3* mutants; Wang et al., 2008). Also analogous to Ni-induced apoptotic pathways, both *cep-1* and *hus-1* play non-essential roles in germline apoptosis in warms exposed to 50 μM Cd (Wang et al., 2008), suggesting the lack of involvement of the DNA damage response in mediating metal-induced apoptosis in *C. elegans.* However, another gene expression profiling study in *C. elegans* has found that *cep-1* and *ape-1* (apoptosis enhancer 1) mRNA is up-regulated upon exposure to 0.85 mg/L Cd (Roh et al., 2006). Such a discrepancy illustrates the complex control of these apoptotic pathways that may be highly dose-dependent. Interestingly, a separate study found that cobalt exposure mimicked the same apoptotic pathways delineated in the above studies: DNA damage-independent, but JNK- and p38 MAPK-dependent mechanisms (Chong et al., 2009). These studies indicate consistently shared signaling cascades that mediate apoptosis caused by a variety of metals.

Moreover, metal exposure also mediates apoptosis through increased oxidative stress. For example, Cu induces its toxicity via ROS generation that damage DNA, lipids, and proteins (Jomova and Valko, 2011). This is thought to via a Fenton-type of reaction (Held et al., 1996). However, the exact pathways involved in Cu-induced apoptosis are not clearly understood. Using *C. elegans* model, Wang et al. (2009) established that Cu induces germline apoptosis in analogous fashion to Ni and Cd. Similar to Ni- and Cd-induced apoptosis, Cu-induced germline apoptosis does not rely on DNA checkpoint genes, as *hus-1*, *clk-1*, *ced-9*, and *egl-1* knockouts do not prevent Cu-induced apoptosis. Similar to the other metals, Cu-induced apoptosis does not depend on CEP-1/p53, as these knockouts caused a significant increase in germline apoptosis upon Cu exposure. However, knockouts of the caspase *ced-3* gene, Apaf-1 homolog *ced-4* gene, ERK pathway homologs (*lin-45*, *mek-2*, and *mpk-1*), JNK pathway homologs (*nsy-1*, 3 *mek-1*, *jkk-1*, *mkk-4*, *jnk-1*), and some p38 MAPK homologs (*sek-1*, *pmk-1*, but not *pmk-3*) are all essential for Cu-induced germline apoptosis (Wang et al., 2009). The involvement of multiple MAPK signaling cascades suggests that the cell requires more effort to manage the encompassing effects of Cu toxicity, potentially due to its highly reactive nature in generating ROS in mitochondria to initiate the caspase cascade. However, the DNA damage response remains non-essential in mediating this apoptosis.

APPLICATION OF TOXICOGENOMIC SCREENS IN NEURODEGENERATIVE DISEASES

Caenorhabditis elegans have been used to investigate several neurological disorders primarily due to its simple and well characterized nervous system. With its short lifespan and ease of genetic mutability, toxicogenomic screens may be quickly and easily performed to identify genes involved in metal-induced neurotoxicity. For example, *C. elegans* possess only eight dopaminergic (DAergic) neurons, which are easily counted in transgenic worms expressing GFP driven by the dopamine transporter, *dat-1*, promoter. Dopamine controls specific, measurable behaviors in worms, including food searching behavior, defecation, and egg laying (Sulston et al., 1975; Weinshenker et al., 1995), which are easily measured for loss of DAergic functioning. Herein, we discuss the application of toxigenomic screens in *C. elegans* to examine the role of metals in PD.

PARKINSON'S DISEASE

Parkinson's disease is the second most common neurodegenerative disease, afflicting ∼2% of the US population (Bushnell and Martin, 1999). It is characterized by the selective loss of DAergic cells in the substantia nigra pars compacta (SNpc) and locus coeruleus regions of the brain (Wilson et al., 1996). The mechanisms underlying the selective degeneration of the DAergic neurons are poorly understood; however genetic factors and environmental and endogenous toxins have been implicated (Dauer and Przedborski, 2003). Epidemiological studies suggest that PD is more common in rural areas, where the increased prevalence is associated with the use of pesticides, herbicides, and heavy metals (Gorell et al., 2004). Observable symptoms in patients with PD include emotional and cognitive decline, bradykinesia, rigidity, tremors, and postural instability (Lees et al., 2009). In addition to cell loss, proteinaceous intracellular inclusions called Lewy bodies are observed in postmortem brains. The majority of PD cases are sporadic, while 10–20% of cases have a genetic component. Familial PD may be due to autosomal dominant genes for α-synuclein and leucine-rich repeat kinase 2 (LRRK2) or autosomal recessive genes for parkin, DJ-1, and phosphatase and tensin homolog (PTEN)-induced novel kinase 1 (PINK1), all of which have orthologs in worms, such as *pdr-1* (Parkin), *djr-1* (DJ-1), *lrk-1* (LRRK2), *pink-1* (PINK1), but not α-synuclein. Proteinaceous inclusions can be induced in *C. elegans* by overexpression of either wildtype of mutant α-synuclein under the dopamine transporter (*dat-1*) promoter (Kuwahara et al., 2006). Interestingly, mammalian models of α-synuclein overexpression have not been successful in causing selective DAergic neurodegeneration, which makes *C. elegans* an invaluable tool.

In addition to genetic predisposition to PD, environmental exposures contribute to the disease development as well. Exposure to Mn, either from the environment or in occupational settings, can produce Parkinsonian-like symptoms. These symptoms include rigidity, tremor, gait disturbances, and hypokinesia, and are attributed to selective interaction of Mn with the basal ganglia downstream of the nigrostriatal DAergic projection, which are areas that readily accumulate Mn (Aschner et al., 2007). Mn has been shown oxidize DA to a highly toxic reactive metabolite leukoaminochrome *o*-semiquinone (Graumann et al., 2002). MeHg exposure has also been implicated in PD. MeHg poisoning causes severe neurological deficits due to brain lesions and the disruption of neurotransmitter systems (Aschner and Syversen, 2005). Both PD and MeHg poisoning present resting tremors and alterations in motor functioning (Biernat et al., 1999; Kaur et al., 2007). MeHg exposure occurs through seafood consumption, due to the global cycling, and accumulation of Hg from industries that reaches the aquatic environment (Fitzgerald and Clarkson, 1991).

Several studies have shown an association between PD and fish consumption or occupational exposure to Hg (Biernat et al., 1999; Kirkey et al., 2001; Fabrizio et al., 2007; Petersen et al., 2008a,b). Serum levels of Hg have been shown to increase post-PD diagnosis, and low serum Hg levels have been associated with decreased risk for development of PD (Gellein et al., 2008).

Toxicogenomic screens have been performed in *C. elegans* knockdown strains exposed to Mn to identify genes that are involved in Mn toxicity. Benedetto et al. (2010) found that Mn exposure induces a dose-dependent degeneration in DAergic neurons in *C. elegans*, which required the presence of the reuptake transporter, DAT-1, as neurodegeneration was not observed in *dat-1* knockdown worms. Using various strains of *C. elegans*, Benedetto et al. (2010) also found that toxicity was prevented by the loss of tyrosine hydroxylase (TH)/CAT-2 function in the double knockout strain, *cat-2(e1112);dat-1(ok157)* and knockdown of vesicular monoamine transporter (VMAT2)/CAT-1 in *cat-1(e1111)* mutants, in which DAergic neurons were unable to release DA at the synaptic cleft. As in mammals, Mn enters DAergic neurons in *C. elegans* through NRAMP/divalent metal transporters, as deletion of the NRAMP ortholog *smf-1* or *smf-3* gene attenuated Mn induced DAergic neurodegeneration and increased survival (Au et al., 2009; Settivari et al., 2009).

The use of *C. elegans* has also facilitated identification of genes that are involved in the cellular response to Mn exposure. Metals cause oxidative stress through the generation of ROS through Fenton chemistry. Cells respond to increased ROS through upregulation of antioxidant molecules and enzymes via activation of the Nrf2 transcription factor. Benedetto et al. (2010) found that loss-of-function mutants for *skn-1*, the worm ortholog to Nrf2, had increased sensitivity to Mn toxicity, whereas wildtype worms expressing SKN-1::GFP showed activation of the transcription factor after Mn exposure. Cells can also produce ROS as a defense mechanism using specific enzymatic pathways. Loss-of-function of *bli-3* gene, a dual oxidase involved in pathogen-induced ROS production, caused increased resistance to Mn toxicity and surprisingly had no increase in ROS production from Mn exposures, implying that BLI-3 is required for ROS-mediated effects in Mn exposure (Benedetto et al., 2010). It is unknown if dual oxidases are involved in mammalian responses to Mn.

Although MeHg exposure causes several toxic effects in *C. elegans*, including decreased survival, developmental delay, and decreased pharyngeal pumping, MeHg exposure did not cause neurodegeneration (Helmcke et al., 2009; Helmcke and Aschner, 2010). This suggests the presence of protective mechanisms in *C. elegans*' neurons. MeHg exposure causes oxidative stress in *C. elegans* through alterations in GSH levels, increased expression of HSPs, and glutathione-*S*-transferase (GST; Helmcke et al., 2009; Helmcke and Aschner, 2010). GST levels are controlled by Nrf2, which is activated by MeHg in mammalian glial cells (Ni et al., 2010). Knockdown of the *skn-1* gene led to DAergic neuron degeneration in 30% of the worms exposed to 1 mM MeHg, whereas no degeneration was observed in wildtype animals (Vanduyn et al., 2010). Further studies are necessary to understand how *C. elegans* DAergic neurons are protected from MeHg induced neurodegeneration.

CONCLUSION

Many of the pathways and cellular processes affected by metals in humans are also present in *C. elegans*. This combined with the ease of genetic manipulability and inexpensive cost make toxicogenomic studies in the worm an attractive alternative to mammalian systems. The *C. elegans* model has proved to be an invaluable tool in studying metal-induced gene expression, DNA damage, and apoptosis. The short lifespan has proved useful in performing genetic screens for xenobiotics and the large RNAi libraries available have been utilized to find novel proteins and protein interactions involved in cellular responses to metal exposure. Through the high homology between human and worm genomes novel targets for therapeutic intervention may become apparent through further large scale toxicogenomic studies in the nematode.

ACKNOWLEDGMENTS

We are grateful for support by NIEHS R01ES07331, R01ES10563, the Center in Molecular Toxicology NIH grant P30ES00267 and the Training Program in Environmental Toxicology grant T32ES007028.

REFERENCES

Al-Shami, S. A., Md Rawi, C. S., Ahmad, A. H., and Mohd Nor, S. A. (2012). Genotoxicity of heavy metals to the larvae of Chironomus kiiensis Tokunaga after short-term exposure. *Toxicol. Ind. Health.* PMID: 22025505. [Epub ahead of print].

An, J. H., and Blackwell, T. K. (2003). SKN-1 links *C. elegans* mesendodermal specification to a conserved oxidative stress response. *Genes Dev.* 17, 1882–1893.

Anbalagan, C., Lafayette, I., Antoniou-Kourounioti, M., Haque, M., King, J., Johnsen, B., Baillie, D., Gutierrez, C., Rodriguez Martin, J. A., and de Pomerai, D. (2012). Transgenic nematodes as biosensors for metal stress in soil pore water samples. *Ecotoxicology* 21, 439–455.

Aremu, D. A., Ezomo, O. F., and Meshitsuka, S. (2011). Gene expression in primary cultured astrocytes affected by aluminum: alteration of chaperons involved in protein folding. *Environ. Health Prev. Med.* 16, 16–24.

Arkblad, E. L., Tuck, S., Pestov, N. B., Dmitriev, R. I., Kostina, M. B., Stenvall, J., Tranberg, M., and Rydstrom, J. (2005). A *Caenorhabditis elegans* mutant lacking functional nicotinamide nucleotide transhydrogenase displays increased sensitivity to oxidative stress. *Free Radic. Biol. Med.* 38, 1518–1525.

Aschner, M., Guilarte, T. R., Schneider, J. S., and Zheng, W. (2007). Manganese: recent advances in understanding its transport and neurotoxicity. *Toxicol. Appl. Pharmacol.* 221, 131–147.

Aschner, M., and Syversen, T. (2005). Methylmercury: recent advances in the understanding of its neurotoxicity. *Ther. Drug. Monit.* 27, 278–283.

Atkin, J. D., Farg, M. A., Walker, A. K., McLean, C., Tomas, D., and Horne, M. K. (2008). Endoplasmic reticulum stress and induction of the unfolded protein response in human sporadic amyotrophic lateral sclerosis. *Neurobiol. Dis.* 30, 400–407.

Au, C., Benedetto, A., Anderson, J., Labrousse, A., Erikson, K., Ewbank, J. J., and Aschner, M. (2009). SMF-1, SMF-2 and SMF-3 DMT1 orthologues regulate and are regulated differentially by manganese levels in *C. elegans*. *PLoS ONE* 4, e7792. doi:10.1371/journal.pone.0007792

Bal, W., and Kasprzak, K. S. (2002). Induction of oxidative DNA damage by carcinogenic metals. *Toxicol. Lett.* 127, 55–62.

Benedetto, A., Au, C., Avila, D. S., Milatovic, D., and Aschner, M. (2010). Extracellular dopamine potentiates Mn-induced oxidative stress, lifespan reduction, and dopaminergic neurodegeneration in a BLI-3-dependent manner in *Caenorhabditis elegans*. *PLoS Genet.* 6, e1001084. doi:10.1371/journal.pgen.1001084

Beyersmann, D., and Hartwig, A. (2008). Carcinogenic metal compounds: recent insight into molecular and cellular mechanisms. *Arch. Toxicol.* 82, 493–512.

Biernat, H., Ellias, S. A., Wermuth, L., Cleary, D., de Oliveira Santos, E. C., Jorgensen, P. J., Feldman, R. G., and Grandjean, P. (1999). Tremor frequency patterns in mercury vapor exposure, compared with early Parkinson's disease and essential tremor. *Neurotoxicology* 20, 945–952.

Bushnell, D. M., and Martin, M. L. (1999). Quality of life and Parkinson's disease: translation and validation of the US Parkinson's Disease Questionnaire (PDQ-39). *Qual. Life Res.* 8, 345–350.

Caffrey, D. R., O'Neill, L. A., and Shields, D. C. (1999). The evolution of the MAP kinase pathways: coduplication of interacting proteins leads to new signaling cascades. *J. Mol. Evol.* 49, 567–582.

Calì, T., Ottolini, D., and Brini, M. (2011). Mitochondria, calcium, and endoplasmic reticulum stress in Parkinson's disease. *Biofactors* 37, 228–240.

Chen, P., Burdette, A. J., Porter, J. C., Ricketts, J. C., Fox, S. A., Nery, F. C., Hewett, J. W., Berkowitz, L. A., Breakefield, X. O., Caldwell, K. A., and Caldwell, G. A. (2010). The early-onset torsion dystonia-associated protein, torsinA, is a homeostatic regulator of endoplasmic reticulum stress response. *Hum. Mol. Genet.* 19, 3502–3515.

Cheng, R. Y., Zhao, A., Alvord, W. G., Powell, D. A., Bare, R. M., Masuda, A., Takahashi, T., Anderson, L. M., and Kasprzak, K. S. (2003). Gene expression dose-response changes in microarrays after exposure of human peripheral lung epithelial cells to nickel(II). *Toxicol. Appl. Pharmacol.* 191, 22–39.

Chong, R., Ke-Zhou, C., and Zeng-Liang, Y. (2009). Induction of germline apoptosis by cobalt and relevant signal transduction pathways in *Caenorhabditis elegans*. *Toxicol. Mech. Methods* 19, 541–546.

Chu, K. W., Chan, S. K., and Chow, K. L. (2005). Improvement of heavy metal stress and toxicity assays by coupling a transgenic reporter in a mutant nematode strain. *Aquat. Toxicol.* 74, 320–332.

Chun, H. S., Lee, H., and Son, J. H. (2001). Manganese induces endoplasmic reticulum (ER) stress and activates multiple caspases in nigral dopaminergic neuronal cells, SN4741. *Neurosci. Lett.* 316, 5–8.

Conradt, B., and Xue, D. (2005). "Programmed cell death," in *WormBook*, ed. The *C. elegans* Research Community, 1–13.

Cui, Y., McBride, S. J., Boyd, W. A., Alper, S., and Freedman, J. H. (2007). Toxicogenomic analysis of *Caenorhabditis elegans* reveals novel genes and pathways involved in the resistance to cadmium toxicity. *Genome Biol.* 8, R122.

Dauer, W., and Przedborski, S. (2003). Parkinson's disease: mechanisms and models. *Neuron* 39, 889–909.

Denver, D. R., Dolan, P. C., Wilhelm, L. J., Sung, W., Lucas-Lledo, J. I., Howe, D. K., Lewis, S. C., Okamoto, K., Thomas, W. K., Lynch, M., and Baer, C. F. (2009). A genome-wide view of *Caenorhabditis elegans* base-substitution mutation processes. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16310–16314.

Derry, W. B., Putzke, A. P., and Rothman, J. H. (2001). *Caenorhabditis elegans* p53: role in apoptosis, meiosis, and stress resistance. *Science* 294, 591–595.

Dudzik, C. G., Walter, E. D., and Millhauser, G. L. (2011). Coordination features and affinity of the Cu(2)+ site in the alpha-synuclein protein of Parkinson's disease. *Biochemistry* 50, 1771–1777.

Escobar Mdel, C., Souza, V., Bucio, L., Hernandez, E., Gomez-Quiroz, L. E., and Gutierrez Ruiz, M. C. (2009). MAPK activation is involved in cadmium-induced Hsp70 expression in HepG2 cells. *Toxicol. Mech. Methods* 19, 503–509.

Fabrizio, E., Vanacore, N., Valente, M., Rubino, A., and Meco, G. (2007). High prevalence of extrapyramidal signs and symptoms in a group of Italian dental technicians. *BMC Neurol.* 7, 24. doi:10.1186/1471-2377-7-24

Fauchon, M., Lagniel, G., Aude, J.-C., Lombardia, L., Soularue, P., Petat, C., Marguerie, G., Sentenac, A., Werner, M., and Labarre, J. (2002). Sulfur sparing in the yeast proteome in response to sulfur demand. *Mol. Cell* 9, 713–723.

Fitzgerald, W. F., and Clarkson, T. W. (1991). Mercury and monomethylmercury: present and future concerns. *Environ. Health Perspect.* 96, 159–166.

Frydman, J. (2001). Folding of newly translated proteins in vivo: the role of molecular chaperones. *Annu. Rev. Biochem.* 70, 603–647.

Fujii, M., Adachi, N., Shikatani, K., and Ayusawa, D. (2009). [FeFe]-hydrogenase-like gene is involved in the regulation of sensitivity to oxygen in yeast and nematode. *Genes Cells* 14, 457–468.

Gardarin, A., Chédin, S., Lagniel, G., Aude, J.-C., Godat, E., Catty, P., and Labarre, J. (2010). Endoplasmic reticulum is a major target of cadmium toxicity in yeast. *Mol. Microbiol.* 76, 1034–1048.

Gartner, A., Boag, P. R., and Blackwell, T. K. (2008). Germline survival and apoptosis. *WormBook* 1–20.

Gellein, K., Syversen, T., Steinnes, E., Nilsen, T. I., Dahl, O. P., Mitrovic, S., Duraj, D., and Flaten, T. P. (2008). Trace elements in serum from patients with Parkinson's disease – a prospective case-control study: the Nord-Trondelag Health Study (HUNT). *Brain Res.* 1219, 111–115.

Gidalevitz, T., Ben-Zvi, A., Ho, K. H., Brignull, H. R., and Morimoto, R. I. (2006). Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* 311, 1471–1474.

Gidalevitz, T., Krupinski, T., Garcia, S., and Morimoto, R. I. (2009). Destabilizing protein polymorphisms in the genetic background direct phenotypic expression of mutant SOD1 toxicity. *PLoS Genet.* 5, e1000399. doi:10.1371/journal.pgen.1000399

Gitler, A. D., Chesi, A., Geddie, M. L., Strathearn, K. E., Hamamichi, S., Hill, K. J., Caldwell, K. A., Caldwell, G. A., Cooper, A. A., Rochet, J. C., and Lindquist, S. (2009). Alpha-synuclein is part of a diverse and highly conserved interaction network that includes PARK9 and manganese toxicity. *Nat. Genet.* 41, 308–315.

Gorell, J. M., Peterson, E. L., Rybicki, B. A., and Johnson, C. C. (2004). Multiple risk factors for Parkinson's disease. *J. Neurol. Sci.* 217, 169–174.

Graumann, R., Paris, I., Martinez-Alvarado, P., Rumanque, P., Perez-Pastene, C., Cardenas, S. P., Marin, P., Diaz-Grez, F., Caviedes, R., Caviedes, P., and Segura-Aguilar, J. (2002). Oxidation of dopamine to aminochrome as a mechanism for neurodegeneration of dopaminergic systems in Parkinson's disease. Possible neuroprotective role of DT-diaphorase. *Pol. J. Pharmacol.* 54, 573–579.

Greer, E. L., Maures, T. J., Ucar, D., Hauswirth, A. G., Mancini, E., Lim, J. P., Benayoun, B. A., Shi, Y., and Brunet, A. (2011). Transgenerational epigenetic inheritance of longevity in *Caenorhabditis elegans*. *Nature* 479, 365–371.

Hall, S. E., Beverly, M., Russ, C., Nusbaum, C., and Sengupta, P. (2010). A cellular memory of developmental history generates phenotypic diversity in *C. elegans*. *Curr. Biol.* 20, 149–155.

Hartman, P. S., Ishii, N., Kayser, E. B., Morgan, P. G., and Sedensky, M. M. (2001). Mitochondrial mutations differentially affect aging, mutability and anesthetic sensitivity in *Caenorhabditis elegans*. *Mech. Ageing Dev.* 122, 1187–1201.

Held, K. D., Sylvester, F. C., Hopcia, K. L., and Biaglow, J. E. (1996). Role of Fenton chemistry in thiol-induced toxicity and apoptosis. *Radiat. Res.* 145, 542–553.

Helmcke, K. J., and Aschner, M. (2010). Hormetic effect of methylmercury on *Caenorhabditis elegans*. *Toxicol. Appl. Pharmacol.* 248, 156–164.

Helmcke, K. J., Syversen, T., Miller, D. M. III, and Aschner, M. (2009). Characterization of the effects of methylmercury on *Caenorhabditis elegans*. *Toxicol. Appl. Pharmacol.* 240, 265–272.

Hiramatsu, N., Kasai, A., Du, S., Takeda, M., Hayakawa, K., Okamura, M., Yao, J., and Kitamura, M. (2007). Rapid, transient induction of ER stress in the liver and kidney after acute exposure to heavy metal: evidence from transgenic sensor mice. *FEBS Lett.* 581, 2055–2059.

Hofmann, E. R., Milstein, S., Boulton, S. J., Ye, M., Hofmann, J. J., Stergiou, L., Gartner, A., Vidal, M., and Hengartner, M. O. (2002). *Caenorhabditis elegans* HUS-1 is a DNA damage checkpoint protein required for genome stability and EGL-1-mediated apoptosis. *Curr. Biol.* 12, 1908–1918.

Honda, Y., and Honda, S. (1999). The daf-2 gene network for longevity regulates oxidative stress resistance and Mn-superoxide dismutase gene expression in *Caenorhabditis elegans*. *FASEB J.* 13, 1385–1393.

Hong, L., and Simon, J. D. (2011). Insights into the thermodynamics of copper association with amyloid-beta, alpha-synuclein and prion proteins. *Metallomics* 3, 262–266.

Hoozemans, J. J. M., van Haastert, E. S., Nijholt, D. A. T., Rozemuller, A. J. M., Eikelenboom, P., and Scheper, W. (2009). The unfolded protein response is activated in pretangle neurons in Alzheimer's disease hippocampus. *Am. J. Pathol.* 174, 1241–1251.

Hoozemans, J. J. M., Veerhuis, R., Van Haastert, E. S., Rozemuller, J. M., Baas, F., Eikelenboom, P., and Scheper, W. (2005). The unfolded

protein response is activated in Alzheimer's disease. *Acta Neuropathol.* 110, 165–172.

Imaizumi, K., Miyoshi, K., Katayama, T., Yoneda, T., Taniguchi, M., Kudo, T., and Tohyama, M. (2001). The unfolded protein response and Alzheimer's disease. *Biochim. Biophys. Acta* 1536, 85–96.

Ishii, N., Takahashi, K., Tomita, S., Keino, T., Honda, S., Yoshino, K., and Suzuki, K. (1990). A methyl viologen-sensitive mutant of the nematode *Caenorhabditis elegans*. *Mutat. Res.* 237, 165–171.

Jiang, P., Gan, M., Ebrahim, A., Lin, W.-L., Melrose, H., and Yen, S.-H. (2010). ER stress response plays an important role in aggregation of alpha-synuclein. *Mol. Neurodegener.* 5, 56.

Jomova, K., and Valko, M. (2011). Advances in metal-induced oxidative stress and human disease. *Toxicology* 283, 65–87.

Kaletta, T., and Hengartner, M. O. (2006). Finding function in novel targets: *C. elegans* as a model organism. *Nat. Rev. Drug Discov.* 5, 387–398.

Kaur, P., Aschner, M., and Syversen, T. (2007). Role of glutathione in determining the differential sensitivity between the cortical and cerebellar regions towards mercury-induced oxidative stress. *Toxicology* 230, 164–177.

Kayser, E. B., Morgan, P. G., Hoppel, C. L., and Sedensky, M. M. (2001). Mitochondrial expression and function of GAS-1 in *Caenorhabditis elegans*. *J. Biol. Chem.* 276, 20551–20558.

Kelly, K. O., Dernburg, A. F., Stanfield, G. M., and Villeneuve, A. M. (2000). *Caenorhabditis elegans* msh-5 is required for both normal and radiation-induced meiotic crossing over but not for completion of meiosis. *Genetics* 156, 617–630.

Kezhou, C., Chong, R., and Zengliang, Y. (2010). Nickel-induced apoptosis and relevant signal transduction pathways in *Caenorhabditis elegans*. *Toxicol. Ind. Health* 26, 249–256.

Kirkey, K. L., Johnson, C. C., Rybicki, B. A., Peterson, E. L., Kortsha, G. X., and Gorell, J. M. (2001). Occupational categories at risk for Parkinson's disease. *Am. J. Ind. Med.* 39, 564–571.

Kordas, K. (2010). Iron, lead, and children's behavior and cognition. *Annu. Rev. Nutr.* 30, 123–148.

Koyashiki, G. A., Paoliello, M. M., and Tchounwou, P. B. (2010). Lead levels in human milk and children's health risk: a systematic review. *Rev. Environ. Health* 25, 243–253.

Kuwahara, T., Koyama, A., Gengyo-Ando, K., Masuda, M., Kowa, H., Tsunoda, M., Mitani, S., and Iwatsubo, T. (2006). Familial Parkinson mutant alpha-synuclein causes dopamine neuron dysfunction in transgenic *Caenorhabditis elegans*. *J. Biol. Chem.* 281, 334–340.

Lees, A. J., Hardy, J., and Revesz, T. (2009). Parkinson's disease. *Lancet* 373, 2055–2066.

Leung, M. C., Williams, P. L., Benedetto, A., Au, C., Helmcke, K. J., Aschner, M., and Meyer, J. N. (2008). *Caenorhabditis elegans*: an emerging model in biomedical and environmental toxicology. *Toxicol. Sci.* 106, 5–28.

McDonald, P. W., Jessen, T., Field, J. R., and Blakely, R. D. (2006). Dopamine signaling architecture in *Caenorhabditis elegans*. *Cell. Mol. Neurobiol.* 26, 593–618.

Mori, A., Yamashita, S., Uchino, K., Suga, T., Ikeda, T., Takamatsu, K., Ishizaki, M., Koide, T., Kimura, E., Mita, S., Maeda, Y., Hirano, T., and Uchino, M. (2011). Derlin-1 overexpression ameliorates mutant SOD1-induced endoplasmic reticulum stress by reducing mutant SOD1 accumulation. *Neurochem. Int.* 58, 344–353.

Neal, A. P., and Guilarte, T. R. (2010). Molecular neurobiology of lead (Pb(2+)): effects on synaptic function. *Mol. Neurobiol.* 42, 151–160.

Ni, M., Li, X., Yin, Z., Jiang, H., Sidoryk-Wegrzynowicz, M., Milatovic, D., Cai, J., and Aschner, M. (2010). Methylmercury induces acute oxidative stress, altering Nrf2 protein level in primary microglial cells. *Toxicol. Sci.* 116, 590–603.

Nishitoh, H., Kadowaki, H., Nagai, A., Maruyama, T., Yokota, T., Fukutomi, H., Noguchi, T., Matsuzawa, A., Takeda, K., and Ichijo, H. (2008). ALS-linked mutant SOD1 induces ER stress- and ASK1-dependent motor neuron death by targeting Derlin-1. *Genes Dev.* 22, 1451–1464.

Petersen, M. S., Halling, J., Bech, S., Wermuth, L., Weihe, P., Nielsen, F., Jorgensen, P. J., Budtz-Jorgensen, E., and Grandjean, P. (2008a). Impact of dietary exposure to food contaminants on the risk of Parkinson's disease. *Neurotoxicology* 29, 584–590.

Petersen, M. S., Weihe, P., Choi, A., and Grandjean, P. (2008b). Increased prenatal exposure to methylmercury does not affect the risk of Parkinson's disease. *Neurotoxicology* 29, 591–595.

Powers, M. V., Jones, K., Barillari, C., Westwood, I., van Montfort, R. L., and Workman, P. (2010). Targeting HSP70: the second potentially druggable heat shock protein and molecular chaperone? *Cell Cycle* 9, 1542–1550.

Qian, Y., Falahatpisheh, M., Zheng, Y., Ramos, K., and Tiffany-Castiglioni, E. (2001). Induction of 78 kD glucose-regulated protein (GRP78) expression and redox-regulated transcription factor activity by lead and mercury in C6 rat glioma cells. *Neurotox. Res.* 3, 581–589.

Qian, Y., Zheng, Y., Ramos, K. S., and Tiffany-Castiglioni, E. (2005). GRP78 compartmentalized redistribution in Pb-treated glia: role of GRP78 in lead-induced oxidative stress. *Neurotoxicology* 26, 267–275.

Rana, S. V. (2008). Metals and apoptosis: recent developments. *J. Trace Elem. Med. Biol.* 22, 262–284.

Reijonen, S., Putkonen, N., Nørremølle, A., Lindholm, D., and Korhonen, L. (2008). Inhibition of endoplasmic reticulum stress counteracts neuronal cell death and protein aggregation caused by N-terminal mutant huntingtin proteins. *Exp. Cell Res.* 314, 950–960.

Ricchelli, F., Fusi, P., Tortora, P., Valtorta, M., Riva, M., Tognon, G., Chieregato, K., Bolognin, S., and Zatta, P. (2007). Destabilization of non-pathological variants of ataxin-3 by metal ions results in aggregation/fibrillogenesis. *Int. J. Biochem. Cell Biol.* 39, 966–977.

Robinson, J. F., Port, J. A., Yu, X., and Faustman, E. M. (2010). Integrating genetic and toxicogenomic information for determining underlying susceptibility to developmental disorders. *Birth Defects Res. Part A Clin. Mol. Teratol.* 88, 920–930.

Rodella, L. F., Ricci, F., Borsani, E., Stacchiotti, A., Foglio, E., Favero, G., Rezzani, R., Mariani, C., and Bianchi, R. (2008). Aluminium exposure induces Alzheimer's disease-like histopathological alterations in mouse brain. *Histol. Histopathol.* 23, 433–439.

Roh, J. Y., Lee, J., and Choi, J. (2006). Assessment of stress-related gene expression in the heavy metal-exposed nematode *Caenorhabditis elegans*: a potential biomarker for metal-induced toxicity monitoring and environmental risk assessment. *Environ. Toxicol. Chem.* 25, 2946–2956.

Roh, J. Y., Sim, S. J., Yi, J., Park, K., Chung, K. H., Ryu, D. Y., and Choi, J. (2009). Ecotoxicity of silver nanoparticles on the soil nematode *Caenorhabditis elegans* using functional ecotoxicogenomics. *Environ. Sci. Technol.* 43, 3933–3940.

Rutkowski, R., Dickinson, R., Stewart, G., Craig, A., Schimpl, M., Keyse, S. M., and Gartner, A. (2011). Regulation of *Caenorhabditis elegans* p53/CEP-1-dependent germ cell apoptosis by Ras/MAPK signaling. *PLoS Genet.* 7, e1002238. doi:10.1371/journal.pgen.1002238

Salinas, L. S., Maldonado, E., and Navarro, R. E. (2006). Stress-induced germ cell apoptosis by a p53 independent pathway in *Caenorhabditis elegans*. *Cell Death Differ.* 13, 2129–2139.

Salnikow, K., and Zhitkovich, A. (2008). Genetic and epigenetic mechanisms in metal carcinogenesis and cocarcinogenesis: nickel, arsenic, and chromium. *Chem. Res. Toxicol.* 21, 28–44.

Samuelson, A. V., Carr, C. E., and Ruvkun, G. (2007). Gene activities that mediate increased life span of *C. elegans* insulin-like signaling mutants. *Genes Dev.* 21, 2976–2994.

Santner, A., and Uversky, V. N. (2010). Metalloproteomics and metal toxicology of alpha-synuclein. *Metallomics* 2, 378–392.

Settivari, R., Levora, J., and Nass, R. (2009). The divalent metal transporter homologues SMF-1/2 mediate dopamine neuron sensitivity in *Caenorhabditis elegans* models of manganism and Parkinson disease. *J. Biol. Chem.* 284, 35758–35768.

Sharma, S. K., Goloubinoff, P., and Christen, P. (2008). Heavy metal ions are potent inhibitors of protein folding. *Biochem. Biophys. Res. Commun.* 372, 341–345.

Shinkai, Y., Yamamoto, C., and Kaji, T. (2010). Lead Induces the expression of endoplasmic reticulum chaperones GRP78 and GRP94 in vascular endothelial cells via the JNK-AP-1 pathway. *Toxicol. Sci.* 114, 378–386.

Sobkowiak, R., and Lesicki, A. (2009). Genotoxicity of nicotine in cell culture of *Caenorhabditis elegans* evaluated by the comet assay. *Drug Chem. Toxicol.* 32, 252–257.

Soti, C., Nagy, E., Giricz, Z., Vigh, L., Csermely, P., and Ferdinandy, P. (2005). Heat shock proteins as emerging therapeutic targets. *Br. J. Pharmacol.* 146, 769–780.

Squitti, R. (2012). Metals in Alzheimer's disease: a systemic perspective. *Front. Biosci.* 17, 451–472.

Stefani, M., and Dobson, C. M. (2003). Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and

biological evolution. *J. Mol. Med.* 81, 678–699.

Sulston, J., Dew, M., and Brenner, S. (1975). Dopaminergic neurons in the nematode *Caenorhabditis elegans. J. Comp. Neurol.* 163, 215–226.

Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans. Dev. Biol.* 56, 110–156.

Vanduyn, N., Settivari, R., Wong, G., and Nass, R. (2010). SKN-1/Nrf2 inhibits dopamine neuron degeneration in a *Caenorhabditis elegans* model of methylmercury toxicity. *Toxicol. Sci.* 118, 613–624.

Wang, S., Tang, M., Pei, B., Xiao, X., Wang, J., Hang, H., and Wu, L. (2008). Cadmium-induced germline apoptosis in *Caenorhabditis elegans*: the roles of HUS1, p53, and MAPK signaling pathways. *Toxicol. Sci.* 102, 345–351.

Wang, S., Wu, L., Wang, Y., Luo, X., and Lu, Y. (2009). Copper-induced germline apoptosis in *Caenorhabditis elegans*: the independent roles of DNA damage response signaling and the dependent roles of MAPK cascades. *Chem. Biol. Interact.* 180, 151–157.

Weinshenker, D., Garriga, G., and Thomas, J. H. (1995). Genetic and pharmacological analysis of neurotransmitters controlling egg laying in *C. elegans. J. Neurosci.* 15, 6975–6985.

Wenzel, D., Palladino, F., and Jedrusik-Bode, M. (2011). Epigenetics in *C. elegans*: facts and challenges. *Genesis* 49, 647–661.

Wilson, J. M., Levey, A. I., Rajput, A., Ang, L., Guttman, M., Shannak, K., Niznik, H. B., Hornykiewicz, O., Pifl, C., and Kish, S. J. (1996). Differential changes in neurochemical markers of striatal dopamine nerve terminals in idiopathic Parkinson's disease. *Neurology* 47, 718–726.

Wu, Y., Wu, Z., Butko, P., Christen, Y., Lambert, M. P., Klein, W. L., Link, C. D., and Luo, Y. (2006). Amyloid-beta-induced pathological behaviors are suppressed by Ginkgo biloba extract EGb 761 and ginkgolides in transgenic *Caenorhabditis elegans. J. Neurosci.* 26, 13102–13113.

Yamamoto, K., Honda, S., and Ishii, N. (1996). Properties of an oxygen-sensitive mutant mev-3 of the nematode *Caenorhabditis elegans. Mutat. Res.* 358, 1–6.

Zhang, Y., Yu, F., Lu, R., Liu, Y., Wu, Y., Qian, H., Zhao, X., Wang, S., Xing, G., and Aschner, M. (2011). Hormetic effects of acute methylmercury exposure on GRP78 expression in rat brain cortex. *Dose-Response.* (in press).

Conflict of Interest Statement: 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.

Received: 25 January 2012; paper pending published: 23 February 2012; accepted: 20 March 2012; published online: 10 April 2012.

Citation: Caito S, Fretham S, Martinez-Finley E, Chakraborty S, Avila D, Chen P and Aschner M (2012) Genome-wide analyses of metal responsive genes in Caenorhabditis elegans. Front. Gene. ***3****:52. doi: 10.3389/fgene.2012.00052*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in GENETICS

REVIEW ARTICLE
published: 25 April 2012
doi: 10.3389/fgene.2012.00062

Applications of next-generation sequencing in fish ecotoxicogenomics

Alvine C. Mehinto[1], Christopher J. Martyniuk[2], Daniel J. Spade[1†] and Nancy D. Denslow[1]*

[1] Center for Environmental and Human Toxicology, Department of Physiological Sciences, University of Florida, Gainesville, FL, USA

[2] Canadian Rivers Institute, Department of Biology, University of New Brunswick, Saint John, NB, Canada

Edited by:
Jonathan Freedman, National Institute of Health, USA

Reviewed by:
Jing-Woei Li, The Chinese University of Hong Kong, Hong Kong
Manikandan Jayapal, King Abdulaziz University, Saudi Arabia

***Correspondence:**
Nancy D. Denslow, Center for Environmental and Human Toxicology, Department of Physiological Sciences, University of Florida, 2187 Mowry Road, Building 471, PO Box 110885, Gainesville, FL 32611, USA.
e-mail: ndenslow@ufl.edu

†***Current address:***
Daniel J. Spade, Department of Pathology and Laboratory Medicine, Brown University, Providence, RI, USA

The new technologies for next-generation sequencing (NGS) and global gene expression analyses that are widely used in molecular medicine are increasingly applied to the field of fish biology. This has facilitated new directions to address research areas that could not be previously considered due to the lack of molecular information for ecologically relevant species. Over the past decade, the cost of NGS has decreased significantly, making it possible to use non-model fish species to investigate emerging environmental issues. NGS technologies have permitted researchers to obtain large amounts of raw data in short periods of time. There have also been significant improvements in bioinformatics to assemble the sequences and annotate the genes, thus facilitating the management of these large datasets. The combination of DNA sequencing and bioinformatics has improved our abilities to design custom microarrays and study the genome and transcriptome of a wide variety of organisms. Despite the promising results obtained using these techniques in fish studies, NGS technologies are currently underused in ecotoxicogenomics and few studies have employed these methods. These issues should be addressed in order to exploit the full potential of NGS in ecotoxicological studies and expand our understanding of the biology of non-model organisms.

Keywords: toxicogenomics, non-model fish, next-generation sequencing, bioinformatics, pathway analysis

INTRODUCTION

Research in fish physiology, genetics, evolution, immunology, and endocrinology using non-model species has seen a marked increase in the utilization of genomic information over the last decade. Traditionally, obtaining genomic information was achieved through Sanger sequencing methods which utilizes fluorescent dye-labeled dideoxynucleotide triphosphates as DNA chain terminators. However, Sanger sequencing is limiting because of the high cost and labor intensity. The development of next-generation sequencing (NGS) technologies has facilitated the collection of large amounts of nucleotide information in sequence read-length from 30 to 1,500 nucleotides (nt) for hundreds of thousands to millions of DNA molecules simultaneously. In parallel, the bioinformatics tools required to analyze these large datasets and identify unique gene sequences have also significantly improved. The different steps involved in NGS studies are illustrated in **Figure 1**. NGS technologies are already considered revolutionary tools in the fields of eukaryotic microorganism (Nowrousian, 2010), plant (Bräutigam and Gowik, 2010), animal, and human genomics (Pareek et al., 2011) and their application has demonstrated great potential to study genome evolution (Holt et al., 2008), gene expression profiling (Wang et al., 2008), and gene regulation (e.g., DNA methylation; Pomraning et al., 2009). With regard to fish studies, the number of publications using NGS technologies has increased approximately 10-fold in the last 3 years (**Figure 2**).

Researchers in fish biology stand to gain a great deal of insight using NGS to learn more about genome-wide and transcriptome-wide control of biological processes, discover novel biomarkers for ecotoxicological applications, characterize toxicity pathways, and investigate evolutionary questions to a greater degree of resolution than previously provided by using more traditional population genetic markers such as DNA microsatellites. In ecotoxicogenomics, gene expression profiling using techniques such as microarrays plays a key role for biomarkers characterization and discovery of toxicity pathways (Denslow et al., 2007; Ju et al., 2007). But research in this field often requires the analysis of complex genomic events using extensive time course and dose response studies in multiple tissues of teleost fish, which can be difficult due to logistics and cost. Fortunately the cost of sequencing is now decreasing, permitting the analysis of many biological replicates (i.e., multiple individual genomes) in a single study. The application of NGS technologies will permit to better link knowledge of individual genotype to phenotype and transcriptomic responses under varying environmental conditions and experimental paradigms.

This review describes the latest NGS platforms available and bioinformatics tools that can be employed to examine the transcriptome of non-model fish species. Specifically, we aim to discuss the possible factors involved in platform selection for researchers working with non-model fish species. Studies that have utilized NGS technologies using fish species are also reviewed

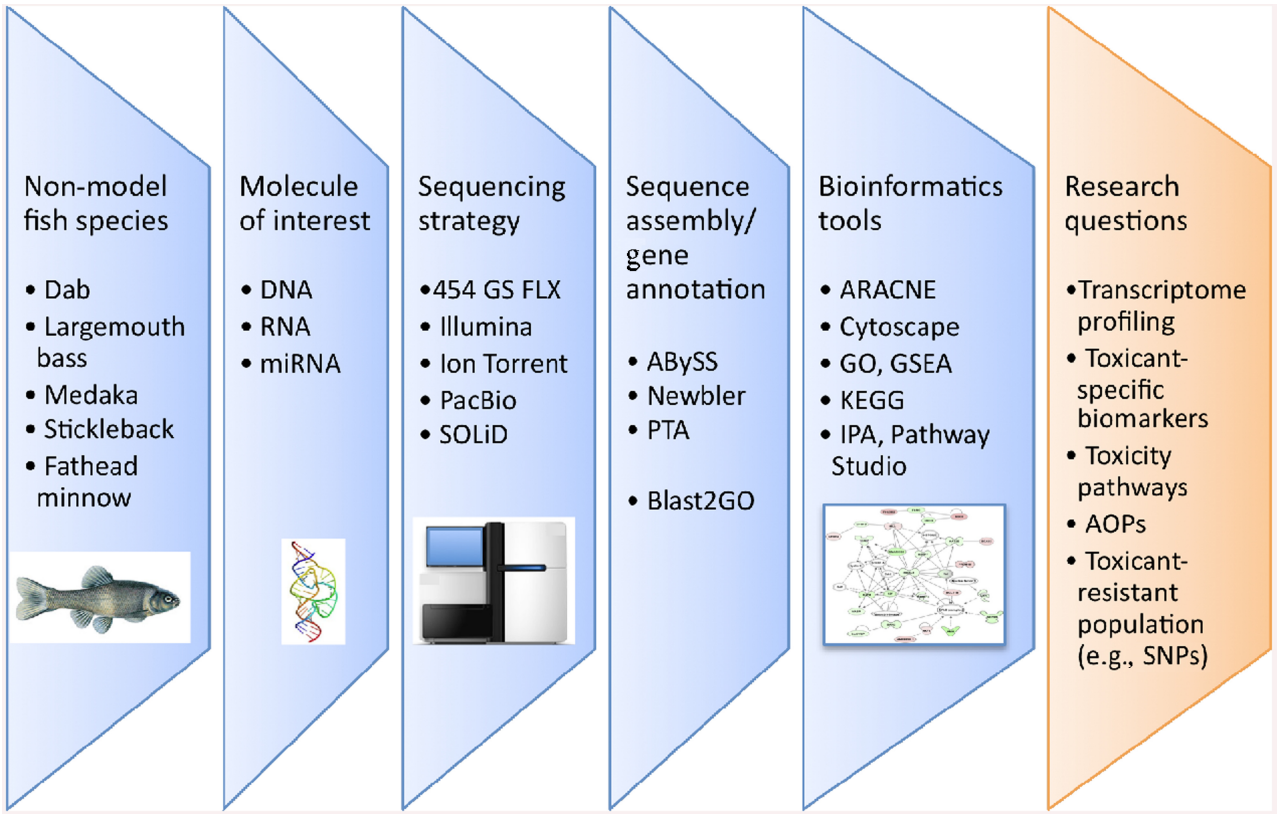


FIGURE 1 | Flowchart of the different steps involved in NGS based studies in fish ecotoxicology. AOPs, adverse outcome pathways; GO, gene ontology; GSEA, gene set enrichment analysis; IPA, ingenuity pathway analysis; PTA, paracel transcript assembler; SNPs, single nucleotide polymorphism.

with the conclusion that NGS data can contribute significantly to our understanding of the detrimental effects of aquatic pollution.

PLATFORMS AND TECHNOLOGY

There are five leading instruments that can be classified as part of the NGS technologies: the 454 GS FLX, the Ion Torrent, the SOLiD, the Illumina, and the more recently released PacBio instrument. These can be distinguished from each other based on the chemistry employed for sequencing, the amount of sequence information produced, the length of each sequence read, and the overall price per nt. While next-generation sequencers are reviewed for general purposes elsewhere (e.g., Mardis, 2008), we provide a brief description of the various technologies followed by a discussion of the relative advantages of each platform for fish toxicogenomics research.

454 GENOME SEQUENCER-FLX^{TM}

The 454 pyrosequencer, manufactured by Roche[1], is the NGS instrument most utilized in fish genomics research (**Table 1**). This platform operates on a principle referred to as "pyrosequencing," a method of detecting single nucleotide addition by capturing the emission of light produced from the release of the by-product pyrophosphate during the polymerization of the DNA molecule (Droege and Hill, 2008; Rothberg and Leamon, 2008). During 454 sequencing, DNA is fragmented and ligated to sepharose beads with one DNA fragment per bead, optimally. This DNA library is then amplified using a process called emulsion PCR (emPCR), producing many copies of a unique single-stranded template on each bead. Following amplification, a single DNA bead and enzyme beads (sulfurylase, luciferase) are deposited in each well of a picotiter plate where as many as one million sequencing reactions – one per bead – occur in parallel. For each nucleotide added during the polymerization reaction, inorganic pyrophosphate and proton by-products are released, which interact with the luciferase to produce a pulse of light that is read by a high-density camera.

ION TORRENT SEMICONDUCTOR SEQUENCER

The Ion Torrent is a modified version of the 454 pyrosequencing approach and operates based on the same sequencing chemistry, except that it makes use of the H^+ that is released with every nucleotide incorporated, instead of the pyrophosphate (Rothberg et al., 2011). To detect the H^+ released, the picotiter plate sits on top of a massively parallel semiconductor-sensing device or ion chip. The integrated circuits take advantage of metal-oxide semiconductor technology, which significantly reduces the cost of sequencing since luciferase and other costly enzymes and scanners

[1]www.454.com

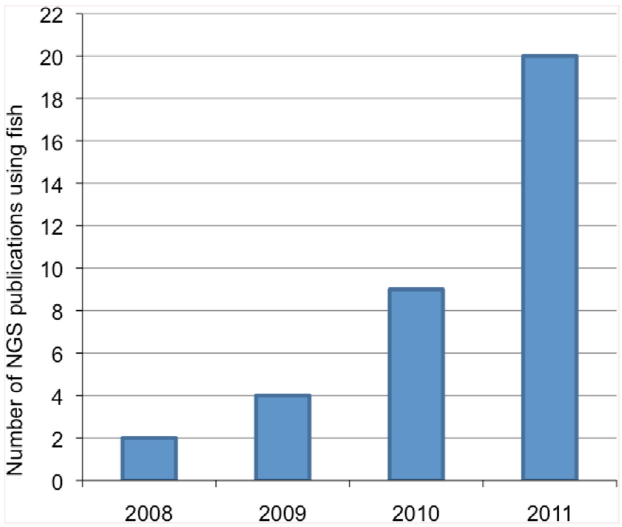


FIGURE 2 | Number of publications using next-generation sequencing (NGS) with non-model fish species in the last 4 years. Keywords used for PubMed query include: fish, next-gen sequencing, high throughput sequencing, and toxicology.

are not needed. To date, this instrument can sequence about 100 nt but it should soon be able to read sequence lengths up to 200 nt (Ion Torrent System, Inc.[2]). Modifications of this technology to increase the length of the sequences produced will likely increase the use of the Ion Torrent in fish ecotoxicology.

SOLiD™ SYSTEM

The SOLiD genome sequencer from Applied Biosystems uses an emPCR process similar to 454, but parallel DNA sequencing is achieved by repeatedly ligating two-nucleotide probes instead of a sequencing reaction catalyzed by DNA polymerase (Morozova and Marra, 2008). The two-nucleotide probes are used to query adjacent bases on the DNA fragment, therefore each nucleotide is actually probed twice. This system is designed to make sequence calls on two signals per base, rather than one, resulting in a lower error rate (for more information on this process, see Morozova and Marra, 2008; Rusk and Kiermer, 2008). Originally, SOLiD technology could only read approximately 35 nucleotides (Morozova and Marra, 2008), but current versions of the instrument have increased the read-length to about 50 nucleotides (Applied Biosystems[3]).

ILLUMINA GENOME ANALYZER

The Illumina/Solexa technology is the second most utilized in fish genomics research (**Table 1**). This sequencing platform differs from 454 and SOLiD in terms of its amplification strategy. Rather than amplifying DNA-covered beads by emPCR, the Illumina technology amplifies clusters of DNA fragments that are affixed to a glass slide using a strategy called bridge amplification. The parallel sequencing process uses dye-labeled nucleotides (one fluorophore per base) that are added simultaneously, rather than sequentially as in the 454 process. The DNA clusters are then subjected to laser excitation that cleaves the dye and permits the addition of the next nucleotide. In 2008, Illumina sequencer projects reported reads of 25–50 nt. Base-calling algorithms have been improving to increase read-length and base-calling confidence (Rougemont et al., 2008; Smith et al., 2008). Currently, the Illumina sequencer can produce longer reads of 100 nt (Illumina, Inc.[4]).

PacBio RS

The PacBio is a single-molecule sequencing approach that has been developed to further reduce the cost and time required to obtain the sequence of a genome or transcriptome. It is thought of as a "third generation" sequencing platform. This instrument has recently become commercially available and only a few institutions have used it. The PacBio works based on a nanophotonic tool called zero-mode waveguide (ZMW; Levene et al., 2003). ZMW technology allows for a DNA polymerase to work in real time using fluorescently labeled nucleotides and tracks synthesis of a single molecule per DNA fragment (Eid et al., 2009). Like the 454 and Illumina instruments, the PacBio sequences by measuring the burst of light produced when the pyrophosphate and fluorescent label are released during the polymerization reaction. This instrument is able to sequence single molecules up to 1500 nt long, but the error rate (around 15%) is still relatively high (Pacific Biosciences[5]). However, pairing this instrument with other more robust sequencers can be a real advantage for non-model species, as one can get a relatively long intact scaffold against which to build and assemble genomes or transcriptomes for species of interest.

ADVANTAGES AND DISADVANTAGES OF SEQUENCING PLATFORMS

The instruments described above use different technologies and each approach has its advantages and disadvantages. Currently, Illumina sequencing produces short reads of about 100 nt in length but has the ability to do this from each end of the DNA molecule when paired ends are used. The SOLiD likewise produces reads of approximately 35–60 nt in length. The short sequences yielded by Illumina and SOLiD platforms have proven useful for the detection of miRNA (small RNA molecules of about 22 nt; Chi et al., 2011; Johansen et al., 2011) and comparative genome analysis of different fish populations (Chi et al., 2011). They could also be useful to design microarrays probes for a variety of non-model fish species. However the use of short sequence reads can be challenging for *de novo* sequencing, sequence assembly and accurate annotation of the genes. It must be noted that the Illumina and SOLiD are working toward increasing the number of base pairs reads and this will improve in the future.

The DNA sequencing techniques employed in Illumina and SOLiD technologies are effective to assess genetic variations in fish at individual (i.e., single nucleotide polymorphisms, SNPs) and population level (Liu et al., 2011). Indeed, while the 454 pyrosequencer determines the length of homopolymers in one step based on the intensity of the light signal (Morozova and Marra, 2008),

[2]www.iontorrent.com

[3]www.appliedbiosystems.com

[4]www.illumina.com

[5]www.pacificbiosciences.com

Table 1 | Some examples of next-generation sequencing studies in fish.

Fish species	Sequencing platform	Source/tissue	Molecule	Discipline	Reference
Atlantic cod (*Gadus morhua*)	Roche GS FLX (454)	Tissue pool	DNA/RNA	Evolution/physiology	Star et al. (2011)
Atlantic killifish (*F. heteroclitus*)	Roche GS FLX (454)	Embryos	RNA	Aquatic toxicology	Oleksiak et al. (2011)
Atlantic salmon (*Salmo salar*)	Roche GS FLX (454)	BAC library	DNA	Evolution/physiology	Quinn et al. (2008)
Brown trout (*Salmo trutta*)	Roche GS FLX (454)	Mitochondrial DNA	DNA	Physiology	Keller et al. (2011)
Carp (*Hypophthalmichthys* spp.)	Illumina GAI	5 tissues: brain, liver, etc.	miRNA	Immunology	Chi et al. (2011)
Catfish (*Ictalurus* spp.)	Illumina GAII	11 tissues: skin, heart, etc.	RNA	Aquaculture	Liu et al. (2011)
Chum salmon (*Oncorhynchus keta*)	Roche GS FLX (454)	Testis	RNA	Evolution	Seeb et al. (2011)
Cichlid fish (*Amphilophus* spp.)	Roche GS FLX (454)	Fry and juvenile	RNA	Evolution	Elmer et al. (2010)
Common dab (*Limanda limanda*)	Illumina GAII	Liver	DNA	Aquatic toxicology	Mirbahai et al. (2011)
European eelpout (*Zoarces viviparus*)	Roche GS FLX (454)	Liver	RNA	Evolution	Kristiansson et al. (2009)
European seabass (*D. labrax*)	Roche GS FLX (454)	BAC library	DNA	Evolution	Kuhl et al. (2011)
Gulf killifish (*Fundulus grandis*)	Illumina GAI	Liver	RNA	Aquatic toxicology	Whitehead et al. (2011)
Japanese medaka (*Oryzias latipes*)	ABI SOLiD	Whole fish	miRNA	Genetics	Li et al. (2010)
Lake sturgeon (*Acipenser fulvescens*)	Roche GS FLX (454)	Ovary and testis	RNA	Evolution	Hale et al. (2010)
Lake trout (*Salvelinus namaycush*)	Roche GS FLX (454)	Liver	RNA	Evolution	Goetz et al. (2010)
Lake whitefish (*Coregonus* spp.)	Roche GS FLX (454)	Brain, liver, white muscle	RNA	Evolution	Renaut et al. (2010)
Largemouth bass (*M. salmoides*)	Roche GS FLX (454)	Brain, liver, gonad	RNA	Aquatic toxicology	Garcia-Reyero et al. (2008)
Mangrove killifish (*K. marmoratus*)	Roche GS FLX (454)	Liver and ovary	RNA	Aquatic toxicology	Rhee et al. (2011)
Orange-spotted grouper (*E. coioides*)	Roche GS FLX (454)	Spleen	RNA	Evolution	Huang et al. (2011)
Pygmy perch (*Nannoperca* spp.)	Roche GS FLX (454)	Muscle (mitochondria)	DNA	Evolution/physiology	Prosdocimi et al. (2011)
Rainbow trout (*Oncorhynchus mykiss*)	Roche GS FLX (454)	Tissue pool	RNA	Evolution	Salem et al. (2010)
Southern platyfish (*X. maculatus*)	Roche GS FLX (454)	Whole fish	RNA	Physiology	Zhang et al. (2011b)
Spotted gar (*Lepisosteus oculatus*)	Illumina GAIIX	Embryo, larvae	DNA/RNA	Evolution	Amores et al. (2011)
Yellow perch (*Perca flavescens*)	Roche GS FLX (454)	Liver	RNA	Aquatic toxicology	Pierron et al. (2011)
Zebrafish (*Danio rerio*)	Roche GS FLX (454)	Embryo, young adult	miRNA	Physiology	Soares et al. (2009)

To date, the large majority of studies in fish biology have utilized the 454 Roche GS FLX system.

the Illumina reads all nucleotides individually. In addition, the SOLiD sequencing system can more reliably distinguish between true sequence polymorphisms and sequencing errors. In SOLiD sequencing, each base is probed twice in two independent ligation reactions, rather than one synthesis reaction. If one of the two ligation reactions gives an unexpected nucleotide, this is recognized as an error. If a consistent result is found for both ligation reactions, it is recognized as a polymorphism (Morozova and Marra, 2008). This distinction is paramount for fish genetics studies due to the increased polymorphic loci resulting from genome duplication events. It should also be mentioned that the newer SOLiD instrument is organized in such a way that individual lanes can be run, without having to fill an entire plate which may improve accessibility for smaller projects in non-model fish.

The new Ion Torrent instrument is relatively inexpensive and will allow individual researchers to have one in their laboratories, much like they do for qPCR. However this instrument also produces relatively short reads. The 454 pyrosequencer and the PacBio are superior in term of read-length and are capable of producing up to 700 and 1,500 nt per read respectively, making them ideal techniques for *de novo* sequencing of fish species as a scaffold is required. The 454 pyrosequencing technology has already shown great potential for whole transcriptome analysis using non-model fish (Garcia-Reyero et al., 2008; Jeukens et al., 2010). We should point out that the PacBio is still in its infancy and the platform still requires a lot of care. With improvements, this instrument will surely become a mainstay for *de novo* sequencing of non-model fish species.

Researchers should consider carefully each sequencing platform based on the aims of the project (i.e., assessment of genetic variation, *de novo* sequencing or transcriptome sequencing). In toxicogenomics studies with non-model fish species, it may be more beneficial to use a hybrid sequencing strategy. For example, combining the short pair-ends reads of the Illumina with the longer single-end reads of the 454 will likely enhance sequence assembly and gene annotation. This was demonstrated recently by Jiang et al. (2011). The authors used Illumina and 454 sequencing to investigate the genome of the channel catfish (*Ictalurus punctatus*), and demonstrated that sequencing data from two NGS platforms improved the sequencing depth and increased the number of contigs assembled.

BIOINFORMATICS: EXTRACTING INFORMATIVE TOXICOLOGICAL INFORMATION FROM NGS BASED STUDIES IN NON-MODEL FISH

Next-generation sequencing technologies produce massive amounts of data that need to be processed, annotated, and aligned to the genome before expression analysis (Garber et al., 2011). This is a significant obstacle for ecotoxicogenomics because many researchers are using non-model fish species to study the impacts of aquatic pollutants. Therefore, the advances in DNA sequencing technologies require corresponding improvements in bioinformatics approaches to better manage and interpret genomic and transcriptomics data. There are new algorithms, such as GENEcounter (Cumbie et al., 2011) that can assist with processing and managing the data but these methods have not been tested with non-model fish species. The process to align reads in NGS will not be covered here and there are a number of pipelines for obtaining meaningful sequencing data in order to quantitate transcriptome data (Goncalves et al., 2011). Although some algorithms incorporate splicing events of transcripts into the analysis, the detection of splice variants could be more challenging in teleosts because there are multiple copies of genes. For example, in some teleost species there are four gene variants of the estrogen receptor that show differences in ontogeny and sex expression (Boyce-Derricott et al., 2010). The ER isoforms show high conservation in the DNA and ligand binding domains and are more variable in other regions. Sequencing a gene with multiples isoforms in the conserved region by chance could make interpretation and quantitation difficult, especially when counting differentially expressed tags in RNA-seq studies.

In recent years, there has been a movement away from single gene characterization and toward the integration and quantification of high-throughput sequencing data in ecotoxicology. To supplement and enhance biologically relevant observations made from gene expression analysis using NGS, bioinformatics algorithms have been developed to consider all affected genes, many of which appear functionally unrelated, and to identify cellular processes and molecular functions perturbed by toxicants. This approach circumvents concerns with multiple hypotheses testing of both microarray and RNA-Seq data which severely restrict expression data because genes are grouped into larger categories resulting in fewer comparisons. Bioinformatics approaches implemented for fish ecotoxicogenomics experiments include functional enrichment, gene set enrichment, pathway analysis, and reverse engineering. We provide some brief examples of their use in ecotoxicology.

GENE ONTOLOGY

Gene ontology (GO) is a manually curated database of genes using a standardized vocabulary that includes biological process, molecular function, and cellular component. Using NCBI PubMed for a literature search, more than 40 scientific publications investigating the impact of aquatic pollutants in fish have characterized differentially expressed transcripts using GO to identify functionally enriched biological processes. As an example, there has been valuable insight obtained into the effects of endocrine disrupting chemicals that mimic estrogens. Many studies with different experimental paradigms reported common biological processes and molecular functions affected by environmental estrogens, despite the variety of genes that were differentially regulated. These include electron transport, amino acid synthesis, primary metabolism, cell communication and signaling, steroid binding, and steroid metabolism (Martyniuk et al., 2007; Benninghoff and Williams, 2008; Hoffmann et al., 2008; Garcia-Reyero et al., 2009).

GENE SET ENRICHMENT ANALYSIS

In contrast to functional enrichment that utilizes a user defined gene list based on predetermined criteria (i.e., fold change or *p*-value cutoff), Gene set enrichment analysis (GSEA) considers the entire list of genes in the analysis (Subramanian et al., 2005) and can be used for microarray and RNA-seq data. GSEA is a computational method that determines whether an *a priori* defined set

of genes shows statistical differences in rank order in a list based on differential gene expression. The advantage of GSEA is that it identifies pathways and cell processes more robustly by reducing the signal-to-noise ratio in a dataset, and there is higher resolution and ability to identify regulated gene groups. GSEA has been utilized in toxicogenomics, for example to study the neurotoxic effects of aquatic pollutants such as fluoxetine, venlafaxine, and carbamazepine (Thomas et al., 2012). In this study, GSEA identified central nervous system development, axonogenesis, brain development, and neurogenesis as the main biological pathways altered in fathead minnows exposed to these three neuroactive contaminants. Until now, GSEA has had limited use in fish transcriptomics studies but it promises to be an important bioinformatics methodology to characterize adverse outcome pathways (AOPs). Another enrichment analysis method called sub-network enrichment analysis (SNEA) can be used in fish transcriptomics studies (Trudeau et al., 2012), but it is not yet widely utilized in fish ecotoxicology. This approach identifies gene regulatory pathways underlying chemical perturbation and one can construct informative gene networks in a method similar to pathway analysis (outline below) but the networks are constructed in a *post hoc* fashion. GSEA and SNEA have shown high potential to characterize biological pathways affected by contaminants but their application remains limited for examining NGS data in fish toxicology as they require high quality gene annotation.

PATHWAY ANALYSIS

Biochemical pathways are important for characterizing AOPs in toxicogenomics. A number of bioinformatics tools are available to link transcriptomics data to pathway categories such as disease progression, drug effects, and biochemical processes among others. These tools include the Database for Annotation, Visualization and Integrated Discovery (DAVID[6]), Connectivity Map[7], and the Kyoto Encyclopedia of Genes and Genomes (KEGG[8]). Of interest to ecotoxicogenomics, KEGG MAPPER and Babelomics can be used to integrate metagenomic and transcriptomics with chemical and pathway information (Kawashima et al., 2008; Medina et al., 2010; Kanehisa et al., 2012). Other programs used for pathway analysis in fish ecotoxicology studies include Ingenuity Pathways Analysis (Ingenuity® Systems) and Pathway Studio® (Nikitin et al., 2003; Ariadne Genomics). Networks are built based upon relationships extracted from primary literature and algorithm searches for entity connections based on regulation, interaction, and binding between proteins or cell processes. In ecotoxicology studies, pathway analysis has been used with success to explore relationships among genes that are impacted by aquatic pollutants. Gene interaction pathways have been constructed after exposure to pollutants of concern such as ethinylestradiol, 17β-trenbolone, and fipronil in the hypothalamic–pituitary–gonadal axis of zebrafish (Wang et al., 2010), the pesticide methoxychlor in largemouth bass liver (Martyniuk et al., 2011) and environmental estrogens in fathead minnow ovary (Garcia-Reyero et al., 2009). NGS approaches in ecotoxicogenomics will benefit from these bioinformatics tools to integrate both DNA and transcriptomics data and better predict the adverse effects in non-target aquatic organisms.

The successes of building meaningful interaction pathways in fish toxicology reported in the literature are impressive as genomics information is limited for non-model fish species. Researchers using fish model in genomics studies have to consider the fact that many gene–gene interaction pathways are based on mammalian literature. Therefore, to extract significant functional gene information for pathway analysis, mammalian homologs for fish genes must be retrieved. Fish specific databases for model fishes such as the zebrafish, are currently under development and will include gene information not found in mammals.

REVERSE ENGINEERING

Reverse engineering offers a new way of characterizing AOPs in fish toxicology (Perkins et al., 2011). The theory behind reverse engineering and the potential applications in ecotoxicology are well described by Garcia-Reyero and Perkins (2011). Generally, the process of reverse engineering, borrowed from computing sciences and engineering, is to identify the working parts of a system in order to better understand how it functions. This methodology increases the potential to study this system in a different context. In toxicogenomics, multiple Omics datasets can be statistically evaluated to identify key nodes (genes or proteins) that regulate gene networks. A framework for reverse engineering of AOPs in ecotoxicology has been introduced by Perkins et al. (2011). This framework consists of building and integrating gene networks, interrogating the networks with chemical perturbations, defining the AOPs, and predicting phenotypic consequences to the perturbation. The authors provide an example using an impressive 868 microarray datasets from female fathead minnow ovary to investigate environmental contaminants able to disrupt the hypothalamic–pituitary–gonadal axis. The analysis permitted to identify gene networks affected by the anti-androgen flutamide, which were composed of several signaling and receptor genes (both estrogen and androgen responsive) and associated with cell regeneration, development, and antioxidant response. Some of the network nodes included activin A receptor (type 1), aryl hydrocarbon receptor (AHR) interacting protein, and Wnt1 inducible signaling pathway protein 1. This approach offers unique biological perspective on the regulatory pathways affected by flutamide.

Some challenges for reverse engineering, and other methods such as SNEA, have been addressed over a decade ago in the early stages of transcriptomics and network analysis (Szallasi, 1999). These include the stochastic nature of the transcriptome (or variation in the time sequences of gene activation/inhibition), the effective size of the network (i.e., how many interacting entities comprise a "network"), the compartmentalization of genetic networks (e.g., a highly compartmentalized gene network will have few regulators and may be more "buffered" from environmental perturbations), and information content of gene expression matrices (i.e., what information is present on a temporal scale about variation in gene–gene or gene–protein relationships). Despite these challenges, there have been great strides in adopting reverse engineering into

[6]http://david.abcc.ncifcrf.gov/

[7]www.broadinstitute.org/cmap/

[8]www.genome.jp/kegg/

aquatic toxicology. It should also be noted that many of the algorithms described above depend upon, by definition, the annotations and curated gene descriptions available. Nevertheless, fish ecotoxicogenomics studies using high-throughput transcript sequencing have benefited tremendously from these bioinformatics approaches and they have been extremely useful for characterizing genes and pathways altered by aquatic pollutants.

RESEARCH IN NON-MODEL FISH SPECIES USING NEXT-GENERATION SEQUENCING

Over the last few years, NGS has been used to examine DNA and RNA from over 20 fish species including Atlantic salmon (*Salmo salar*), bighead carp (*Hypophthalmichthys nobilis*), European seabass (*Dicentrarchus labrax*), lake sturgeon (*Acipenser fulvescens*), mangrove killifish (*Kryptolebias marmoratus*), pygmy perch (*Nannoperca* spp.), and spotted gar (*Lepisosteus oculatus*; **Table 1**). Most fish sequencing projects have employed the 454 pyrosequencer and have been successful using multiple tissue types (e.g., liver, gonad, kidney, brain) as well as different life stages (e.g., adult versus embryo). This demonstrates that NGS platforms are versatile and can be used to address a range of biological questions in fish. To date, only a few studies have used NGS technologies to research the impact of environmental contaminants in aquatic organisms. Below, we discuss the application and advantages of these methodologies in fish toxicogenomics.

NGS IN FISH TRANSCRIPTOMIC ANALYSES

Next-generation sequencing has already started to have a positive impact in the field of fish transcriptomics. Microarrays are frequently used in the field of fish ecotoxicology (Douglas, 2006; Falciani et al., 2008; Garcia-Reyero et al., 2009; Villeneuve et al., 2010; Gust et al., 2011; Sellin Jeffries et al., 2012). Typically microarray probes were designed based on cDNA contigs produced by suppressive subtractive hybridization (SHH) and cDNA libraries (Blum et al., 2004; Williams et al., 2006; Larkin et al., 2007; Cairns et al., 2008; Katsiadaki et al., 2010). However, a few studies have applied NGS technology to produce oligonucleotide microarrays. Because NGS platforms can generate high numbers of reads, the resulting sequences are often extended which increases the chance to find matching reads and correctly annotate them. Garcia-Reyero et al. (2008) used 454 pyrosequencing technologies to build a 44,000-oligonucleotide microarray for largemouth bass (*Micropterus salmoides*). This approach resulted in obtaining 31,391 unique sequences, which were compiled with sequences from SHH to produce nearly 16,000 gene sequences (half of them were annotated). The custom-designed microarray was then tested by assessing the impact of 17β-estradiol exposure on endocrine disruption and hormone signaling in adult largemouth bass. The combination of NGS and microarray analyses permitted characterizing several pathways perturbed by the estrogenic compound including gonad development, sex differentiation, signal transduction, and cell communication. In another study, Mirbahai et al. (2011) used NGS technology in combination with methylated DNA immunoprecipitation to design a 14,919-oligonucleotide microarray. This permitted to examine hepatic DNA methylation changes in common Dab (*Limanda limanda*) living in polluted environments and to correlate methylation levels with gene expression levels.

Recently, RNA-seq analyses are increasingly used (Xiang et al., 2010; Fraser et al., 2011) and the results suggest that this method could replace array-based technology in toxicogenomics research. Indeed, RNA-seq presents the advantage to quantify directly the expression level of mRNAs across the transcriptome from the number of reads for a particular cDNA contig in a sequencing run, allowing for the quantification of low-expressed transcripts. While currently cost-prohibitive, RNA-seq analysis can potentially provide a greater degree of resolution than microarrays and help to identify splice variants. Oleksiak et al. (2011) utilized this technique in supplement to a microarray experiment to determine the genomic differences between a polychlorinated biphenyl (PCB) sensitive and a PCB resistant population of Atlantic killifish (*Fundulus heteroclitus*). Using 454 pyrosequencing technology, they demonstrated that NGS data can be used to extend the length of array probes, which helped to find new matching sequences and to annotate previously unannotated probes. The RNA-seq study corroborated most of the microarray results and suggested that AHR regulatory pathway may be responsible for the PCB resistance of one of the killifish population. In another study, Whitehead et al. (2011) applied similar techniques to examine the effects of the *Deepwater Horizon* oil spill on Gulf killifish. RNA sequencing data was acquired using the Illumina platform and over 6000 unique EST sequences were obtained. Both microarray and RNA-seq analyses identified zona pellucida, choriogenin, and vitellogenin as PCB-responsive genes. These early studies provide strong evidence that RNA-seq methods are suitable to investigate the adverse effects of pollutants present in the aquatic environment.

The usefulness of this approach was further demonstrated by Pierron et al. (2011) who conducted RNA-seq to examine the effects of chronic metal exposure in four wild populations of yellow perch. NGS data generated from the yellow perch yielded over 9,000 gene sequences among which 6,000 were annotated. As mentioned previously, annotating EST sequences from non-model fish species is one of the main challenges for fish biologists and NGS technologies could facilitate this task. Pierron et al. (2011) were able to establish relationships between the hepatic expression levels of specific transcripts and the concentrations of copper and cadmium measured in the fish as well as to identify potential adverse effects. In general, these studies have successfully shown that NGS is a powerful technique to study the ecotoxicological responses of non-model fish species living in polluted environments.

NGS IN FISH EVOLUTION AND PHYSIOLOGY

Interestingly, most of the research published on the application of NGS in fish has focused on different aspects of fish evolution such as genome evolution (Hale et al., 2010; Amores et al., 2011), phenotypic evolution (Elmer et al., 2010; Goetz et al., 2010; Jeukens et al., 2010), and evolution of immune system (Star et al., 2011; Zhang et al., 2011a). This subject area is outside of the scope of this article, but a few of these studies are reviewed below.

Next-generation sequencing technologies have had a significant impact in the field of ecological divergence and have contributed

in elucidating the links between genetic and environmental factors leading to species evolution (Elmer et al., 2010). Most of the research in this field used the 454 GS FLX pyrosequencer for RNA-seq analyses to uncover the molecular basis for the phenotypic and ecological divergences between endemic species. For example, Jeukens et al. (2010) employed this methodology to investigate the genomic differences behind the phenotypic divergence of two populations of lake whitefish (*Coreons clupeaformis* spp.). The authors discovered that dwarf fish had an over-representation of genes linked to immunity, DNA replication and repair while normal fish over-expressed genes linked to protein synthesis. Elmer et al. (2010) used the same approach to correlate the genomic and phenotypic differences between crater lake cichlids: the benthic species *Amphilophus astorquii* and the limnetic species *Amphilophus zaliosus*. Their study revealed that a number of transcripts associated with development, biosynthesis, and metabolic processes were differentially expressed between the two species. Other studies have employed NGS technologies to characterize fish immune system and its evolution. All these studies have concluded that NGS technologies provide a greater scope of understanding of the genetic events that preceded natural selection and fish species evolution. The significant advancement made in these disciplines may provide valuable genetic insights to facilitate ecotoxicogenomic analyses. For example, studies on the Atlantic killifish (Oleksiak et al., 2011) combined principles in ecotoxicology and evolution to better understand adaptation of fish in polluted environments. Combining data on both genetic variation (SNPs) in fish genomes and transcriptomic responses will lead to the characterization of expression quantitative trait loci (eQTL) and genetic architecture that underlies adaptation.

CONCLUSION

There is great promise for toxicogenomics in non-model fish species. Fish offer unique challenges compared to mammals due to genome duplication events and the presence of multiple isoforms for many genes. Nevertheless, teleost fish are important model organisms for assessing the impact of anthropogenic pollutants in the environment as well as studying certain human diseases (Albertson et al., 2009; Zhang et al., 2010). As the costs for DNA and RNA sequencing decrease, the combination of several NGS platforms should facilitate whole genome sequencing projects and expand our knowledge of ecologically relevant species. Understanding the relationships between environmental chemical exposure and gene expression will provide valuable data for environmental risk assessments (ERA). In 2011, Piña and Barata reviewed the potential for ecotoxicogenomics studies to improve the tests necessary for ERA by discovering biological assays and biomarkers relevant to environmental conditions (Piña and Barata, 2011). Thus, the development of ecotoxicogenomics and bioinformatics tools will greatly benefit the assessment of the impacts of environmental pollutants. In the future, it will be necessary to integrate the extensive genomic data gathered from transcriptomics, gene regulation, and evolutionary biology into a working framework in order to propose new hypotheses in fish research.

ACKNOWLEDGMENTS

This review was supported by grant R01 ES015449 from the Superfund Basic Research Program from the National Institute of Environmental Health Sciences (Nancy D. Denslow) and by a Canada Research Chair and NSERC Discovery Grant to Christopher J. Martyniuk. Publication of this article was funded in part by the University of Florida Open-Access Publishing Fund.

REFERENCES

Albertson, R. C., Cresko, W., Detrich, H. W. III, and Postlethwait, J. H. (2009). Evolutionary mutant models for human disease. *Trends Genet.* 25, 74–81.

Amores, A., Catchen, J., Ferrara, A., Fontenot, Q., and Postlethwait, J. H. (2011). Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. *Genetics* 188, 799–808.

Benninghoff, A. D., and Williams, D. E. (2008). Identification of a transcriptional fingerprint of estrogen exposure in rainbow trout liver. *Toxicol. Sci.* 101, 65–80.

Blum, J. L., Knoebl, I., Larkin, P., Kroll, K. J., and Denslow, N. D. (2004). Use of suppressive subtractive hybridization and cDNA arrays to discover patterns of altered gene expression in the liver of dihydrotestosterone and 11-ketotestosterone exposed adult male largemouth bass (*Micropterus salmoides*). *Mar. Environ. Res.* 58, 565–569.

Boyce-Derricott, J., Nagler, J. J., and Cloud, J. G. (2010). The ontogeny of nuclear estrogen receptor isoform expression and the effect of 17beta-estradiol in embryonic rainbow trout (*Oncorhynchus mykiss*). *Mol. Cell. Endocrinol.* 315, 277–281.

Bräutigam, A., and Gowik, U. (2010). What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biol.* 12, 831–841.

Cairns, M. T., Johnson, M. C., Talbot, A. T., Pemmasani, J. K., McNeill, R. E., Houeix, B., Sangrador-Vegas, A., and Pottinger, T. G. (2008). A cDNA microarray assessment of gene expression in the liver of rainbow trout (*Oncorhynchus mykiss*) in response to a handling and confinement stressor. *Comp. Biochem. Physiol. D Genomics Proteomics* 3, 51–66.

Chi, W., Tong, C., Gan, X., and He, S. (2011). Characterization and comparative profiling of miRNA transcriptomes in bighead carp and silver carp. *PLoS ONE* 6, e23549. doi: 10.1371/journal.pone.0023549

Cumbie, J. S., Kimbrel, J. A., Di, Y., Schafer, D. W., Wilhelm, L. J., Fox, S. E., Sullivan, C. M., Curzon, A. D., Carrington, J. C., Mockler, T. C., and Chang, J. H. (2011). GENECOUNTER: a computational pipeline for the analysis of RNA-seq data for gene expression differences. *PLoS ONE* 6, e25279. doi: 10.1371/journal.pone.0025279

Denslow, N. D., Garcia-Reyero, N., and Barber, D. S. (2007). Fish 'n' chips: the use of microarrays for aquatic toxicology. *Mol. Biosyst.* 3, 172–177.

Douglas, S. E. (2006). Microarray studies of gene expression in fish. *OMICS* 10, 474–489.

Droege, M., and Hill, B. (2008). The Genome Sequencer FLX^{TM} System – longer reads, more applications, straight forward bioinformatics and more complete data sets. *J. Biotechnol.* 136, 3–10.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin. D., Zhao, P., Zhong, F., Korlach, J., and Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. *Science* 323, 133–138.

Elmer, K. R., Fan, S., Gunter, H. M., Jones, J. C., Boekhoff, S., Kuraku, S., and Meyer A. (2010). Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. *Mol. Ecol.* 19, 197–211.

Falciani, F., Diab, A. M., Sabine, V., Williams, T. D., Ortega, F., George, S. G., and Chipman, J. K. (2008). Hepatic transcriptomics profiles of European flounder (*Platichthys flesus*) from field sites and computational approaches to predict site from stress gene responses following exposure to model toxicants. *Aquat. Toxicol.* 90, 92–101.

Fraser, B. A., Weadick, C. J., Janowitz, I., Rodd, H., and Hughes, K. A. (2011).

Sequencing and characterization of the guppy (*Poecilia reticulata*) transcriptome. *BMC Genomics* 12, 202. doi: 10.1186/1471-2164-12-202

Garber, M., Grabherr, M. G., Guttman, M., and Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat. Methods* 8, 469–477.

Garcia-Reyero, N., Griffitt, R. J., Liu, L., Kroll, K. J., Farmerie, W. G., Barber, D. S., and Denslow, N. D. (2008). Construction of a robust microarray from a non-model species (largemouth bass) using pyrosequencing technology. *J. Fish Biol.* 72, 2354–2376.

Garcia-Reyero, N., and Perkins, E. J. (2011). Systems biology: leading the revolution in ecotoxicology. *Environ. Toxicol. Chem.* 30, 265–273.

Garcia-Reyero, N., Villeneuve, D. L., Kroll, K. J., Liu, L., Orlando, E. F., Watanabe, K. H., Sepúlveda, M. S., Ankley, G. T., and Denslow, N. D. (2009). Gene expression responses in male fathead minnows exposed to binary mixtures of an estrogen and antiestrogen. *BMC Genomics* 10, 308. doi: 10.1186/1471-2164-10-308

Goetz, F., Rosauer, D., Sitar, S., Goetz, G., Simchick, C., Roberts, S., Johnson, R., Murphy, C., Bronte, C. R., and Mackenzie, S. (2010). A genetic basis for the phenotypic differentiation between siscowet and lean lake trout (*Salvelinus namaycush*). *Mol. Ecol.* 19, 176–196.

Goncalves, A., Tikhonov, A., Brazma, A., and Kapushesky, M. (2011). A pipeline for RNA-seq data processing and quality assessment. *Bioinformatics* 27, 867–869.

Gust, K. A., Wilbanks, M. S., Guan, X., Pirooznia, M., Habib, T., Yoo, L., Wintz, H., Vulpe, C. D., and Perkins, E. J. (2011). Investigations of transcript expression in fathead minnow (*Pimephales promelas*) brain tissue reveal toxicological impacts of RDX exposure. *Aquat. Toxicol.* 101, 135–145.

Hale, M. C., Jackson, J. R., and Dewoody, J. A. (2010). Discovery and evaluation of candidate sex-determining genes and xenobiotics in the gonads of lake sturgeon (*Acipenser fulvescens*). *Genetica* 138, 745–756.

Hoffmann, J. L., Thomason, R. G., Lee, D. M., Brill, J. L., Price, B. B., Carr, G. J., and Versteeg, D. J. (2008). Hepatic gene expression profiling using GeneChips in zebrafish exposed to 17alpha-methyldihydrotestosterone. *Aquat. Toxicol.* 87, 69–80.

Holt, K. E., Parkhill, J., Mazzoni, C. J., Roumagnac, P., Weill, F. X., Goodhead, I., Rance, R., Baker, S., Maskell, D. J., Wain, J., Dolecek, C., Achtman, M., and Dougan, G. (2008). High-throughput sequencing provides insights into genome variation and evolution in *Salmonella typhi*. *Nat. Genet.* 40, 987–993.

Huang, Y., Huang, X., Yan, Y., Cai, J., Ouyang, Z., Cui, H., Wang, P., and Qin, Q. (2011). Transcriptome analysis of orange-spotted grouper (*Epinephelus coioides*) spleen in response to Singapore grouper iridovirus. *BMC Genomics* 12, 556. doi: 10.1186/1471-2164-12-556

Jeukens, J., Renaut, S., St-Cyr, J., Nolte, A. W., and Bernatchez, L. (2010). The transcriptomics of sympatric dwarf and normal lake whitefish (*Coregonus clupeaformis* spp., Salmonidae) divergence as revealed by next-generation sequencing. *Mol. Ecol.* 19, 5389–5403.

Jiang, Y., Lu, J., Peatman, E., Kucuktas, H., Liu, S., Wang, S., Sun, F., and Liu, Z. (2011). A pilot study for channel catfish whole genome sequencing and *de novo* assembly. *BMC Genomics* 12, 629. doi: 10.1186/1471-2164-12-629

Johansen, S. D., Karlsen, B. O., Furmanek, T., Andreassen, M., Jørgensen, T. E., Bizuayehu, T. T., Breines, R., Emblem, A., Kettunen, P., Luukko, K., Edvardsen, R. B., Nordeide, J. T., Coucheron, D. H., and Moum, T. (2011). RNA deep sequencing of the Atlantic cod transcriptome. *Comp. Biochem. Physiol. D Genomics Proteomics* 6, 18–22.

Ju, Z., Wells, M. C., and Walter, R. B. (2007). DNA microarray technology in toxicogenomics of aquatic models: methods and applications. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* 145, 5–14.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40, D109–D114.

Katsiadaki, I., Williams, T. D., Ball, J. S., Bean, T. P., Sanders, M. B., Wu, H., Santos, E. M., Brown, M. M., Baker, P., Ortega, F., Falciani, F., Craft, J. A., Tyler, C. R., Viant, M. R., and Chipman, J. K. (2010). Hepatic transcriptomic and metabolomic responses in the stickleback (*Gasterosteus aculeatus*) exposed to ethinyl-estradiol. *Aquat. Toxicol.* 97, 174–187.

Kawashima, S., Okuda, S., Tokimatsu, T., and Yamanishi, Y. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, D480–D484.

Keller, I., Taverna, A., and Seehausen, O. (2011). A pyrosequencing assay for the rapid discrimination of mitochondrial lineages in the *Salmo trutta* species complex. *Mol. Ecol. Resour.* 11, 196–199.

Kristiansson, E., Asker, N., Förlin, L., and Larsson, D. G. (2009). Characterization of the *Zoarces viviparus* liver transcriptome using massively parallel pyrosequencing. *BMC Genomics* 10, 345. doi: 10.1186/1471-2164-10-345

Kuhl, H., Tine, M., Beck, A., Timmermann, B., Kodira, C., and Reinhardt, R. (2011). Directed sequencing and annotation of three *Dicentrarchus labrax* L. chromosomes by applying Sanger- and pyrosequencing technologies on pooled DNA of comparatively mapped BAC clones. *Genomics* 98, 202–212.

Larkin, P., Villeneuve, D. L., Knoebl, I., Miracle, A. L., Carter, B. J., Liu, L., Denslow, N. D., and Ankley, G. T. (2007). Development and validation of a 2,000-gene microarray for the fathead minnow (*Pimephales promelas*). *Environ. Toxicol. Chem.* 26, 1497–1506.

Levene, M. J., Korlach, J., Turner, S. W., Foquet, M., Craighead, H. G., and Webb, W. W. (2003). Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* 299, 682–686.

Li, S. C., Chan, W. C., Ho, M. R., Tsai, K. W., Hu, L. Y., Hsu, C. N., Hwang, P. P., and Lin, W. C. (2010). Discovery and characterization of medaka miRNA genes by next generation sequencing platform. *BMC Genomics* 11(Suppl. 4), S8. doi: 10.1186/1471-2164-11-S4-S8

Liu, S., Zhou, Z., Lu, J., Sun, F., Wang, S., Liu, H., Jiang, Y., Kucuktas, H., Kaltenboeck, L., Peatman, E., and Liu, Z. (2011). Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. *BMC Genomics* 12, 53. doi: 10.1186/1471-2164-12-53

Mardis, E. R. (2008). Next generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.* 9, 387–402.

Martyniuk, C. J., Gerrie, E. R., Popesku, J. T., Ekker, M., and Trudeau, V. L. (2007). Microarray analysis in the zebrafish (*Danio rerio*) liver and telencephalon after exposure to low concentration of 17alpha-ethinylestradiol. *Aquat. Toxicol.* 84, 38–49.

Martyniuk, C. J., Spade, D. J., Blum, J. L., Kroll, K. J., and Denslow, N. D. (2011). Methoxychlor affects multiple hormone signaling pathways in the largemouth bass (*Micropterus salmoides*) liver. *Aquat. Toxicol.* 101, 483–492.

Medina, I., Carbonell, J., Pulido, L., Madeira, S. C., Conesa, A., Tárraga, J., Pascual-Montano, A., Nogales-Cadenas, R., Santoyo, J., García, F., Marbà, M., Montaner, D., and Dopazo, J. (2010). Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomics data with advanced functional profiling. *Nucleic Acids Res.* 38, W210–W213.

Mirbahai, L., Yin, G., Bignell, J. P., Li, N., Williams, T. D., and Chipman, J. K. (2011). DNA methylation in the liver tumorigenesis in fish from the environment. *Epigenetics* 6, 1319–1333.

Morozova, O., and Marra, M. A. (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics* 92, 255–264.

Nikitin, A., Egorov, S., Daraselia, N., and Mazo, I. (2003). Pathway studio – the analysis and navigation of molecular networks. *Bioinformatics* 19, 2155–2157.

Nowrousian, M. (2010). Next-generation sequencing techniques for eukaryotic microorganisms: sequencing-based solutions to biological problems. *Eukaryot. Cell* 9, 1300–1310.

Oleksiak, M. F., Karchner, S. I., Jenny, M. J., Franks, D. G., Welch, D. B., and Hahn, M. E. (2011). Transcriptomic assessment of resistance to effects of an aryl hydrocarbon receptor (AHR) agonist in embryos of Atlantic killifish (*Fundulus heteroclitus*) from a marine Superfund site. *BMC Genomics* 12, 263. doi: 10.1186/1471-2164-12-263

Pareek, C. S., Smoczynski, R., and Tretyn, A. (2011). Sequencing technologies and genome sequencing. *J. Appl. Genet.* 52, 413–435.

Perkins, E. J., Chipman, J. K., Edwards, S., Habib, T., Falciani, F., Taylor, R., Van Aggelen, G., Vulpe, C., Antczak, P., and Loguinov, A. (2011). Reverse engineering adverse outcome pathways. *Environ. Toxicol. Chem.* 30, 22–38.

Pierron, F., Normandeau, E., Defo, M. A., Campbell, P. G., Bernatchez, L., and Couture, P. (2011). Effects of chronic metal exposure on wild fish populations revealed by high-throughput cDNA sequencing. *Ecotoxicology* 20, 1388–1399.

Piña, B., and Barata, C. (2011). A genomic and ecotoxicological

perspective of DNA array studies in aquatic environmental risk assessment. *Aquat. Toxicol.* 105S, 40–49.

Pomraning, K. R., Smith, K. M., and Freitag, M. (2009). Genome-wide high throughput analysis of DNA methylation in eukaryotes. *Methods* 47, 142–150.

Prosdocimi, F., de Carvalho, D. C., de Almeida, R. N., and Beheregaray, L. B. (2011). The complete mitochondrial genome of two recently derived species of the fish genus *Nannoperca* (Perciformes, Percichthyidae). *Mol. Biol. Rep.* 39, 2767–2772.

Quinn, N. L., Levenkova, N., Chow, W., Bouffard, P., Boroevich, K. A., Knight, J. R., Jarvie, T. P., Lubieniecki, K. P., Desany, B. A., Koop, B. F., Harkins, T. T., and Davidson, W. S. (2008). Assessing the feasibility of GS FLX pyrosequencing for sequencing the Atlantic salmon genome. *BMC Genomics* 9, 404. doi: 10.1186/1471-2164-9-404

Renaut, S., Nolte, A. W., and Bernatchez, L. (2010). Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Mol. Ecol.* 19, 115–131.

Rhee, J. S., Kim, B. M., Kim, R. O., Choi, B. S., Choi, I. Y., Lee, Y. M., and Lee, J. S. (2011). Analysis of expressed sequence tags from the liver and ovary of the euryhaline hermaphroditic fish, *Kryptolebias marmoratus*. *Comp. Biochem. Physiol. D Genomics Proteomics* 6, 244–255.

Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., Leamon, J. H., Johnson, K., Milgrew, M. J., Edwards, M., Hoon, J., Simons, J. F., Marran, D., Myers J. W., Davidson, J. F., Branting, A., Nobile, J. R., Puc, B. P., Light, D., Clark, T. A., Huber, M., Branciforte, J. T., Stoner, I. B, Cawley, S. E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J. A., Namsaraev, E., McKernan, K. J., Williams, A., Roth, G. T., and Bustillo, J. (2011). An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475, 348–352.

Rothberg, J. M., and Leamon, J. H. (2008). The development and impact of 454 sequencing. *Nat. Biotechnol.* 26, 1117–1124.

Rougemont, J., Amzallag, A., Iseli, C., Farinelli, L., Xenarios, I., and Naef, F. (2008). Probabilistic base calling of Solexa sequencing data. *BMC Bioinformatics* 9, 431. doi: 10.1186/1471-2105-9-431

Rusk, N., and Kiermer, V. (2008). Primer: Sequencing – the next generation. *Nat. Methods* 5, 15.

Salem, M., Rexroad, C. E. III, Wang, J., Thorgaard, G. H., and Yao, J. (2010). Characterization of the rainbow trout transcriptome using Sanger and 454-pyrosequencing approaches. *BMC Genomics* 11, 564. doi: 10.1186/1471-2164-11-564

Seeb, J. E., Pascal, C. E, Grau, E. D., Seeb, L. W., Templin, W. D., Harkins, T., and Roberts, S. B. (2011). Transcriptome sequencing and high-resolution. melt analysis advance single nucleotide polymorphism discovery in duplicated salmonids. *Mol. Ecol. Resour.* 11, 335–348.

Sellin Jeffries, M. K., Mehinto, A. C., Carter, B., Denslow, N. D., and Kolok, A. S. (2012). Taking microarrays to the field: differential hepatic gene expression of caged fathead minnows from Nebraska watersheds. *Environ. Sci. Technol.* 46, 1877–1885.

Smith, A. D., Xuan, Z., and Zhang, M. Q. (2008). Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC Bioinformatics* 9, 128. doi: 10.1186/1471-2105-9-128

Soares, A. R., Pereira, P. M., Santos, B., Egas, C., Gomes, A. C., Arrais, J., Oliveira, J. L., Moura, G. R., and Santos, M. A. (2009). Parallel DNA pyrosequencing unveils new zebrafish microRNAs. *BMC Genomics* 10, 195. doi: 10.1186/1471-2164-10-195

Star, B., Nederbragt, A. J., Jentoft, S., Grimholt, U., Malmstrøm, M., Gregers, T. F., Rounge, T. B., Paulsen, J., Solbakken, M. H., Sharma, A., Wetten, O. F., Lanzén A., Winer, R., Knight, J., Vogel, J. H., Aken, B., Andersen, O., Lagesen, K., Tooming-Klunderud, A., Edvardsen, R. B., Tina, K. G., Espelund, M., Nepal, C., Previti, C., Karlsen, B. O., Moum, T., Skage, M., Berg, P. R., Gjøen, T., Kuhl, H., Thorsen, J., Malde, K., Reinhardt, R., Du, L., Johansen, S. D., Searle, S., Lien, S., Nilsen, F., Jonassen, I., Omholt, S. W., Stenseth, N. C., and Jakobsen, K. S. (2011). The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477, 207–210.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–1550.

Szallasi, Z. (1999). Genetic network analysis in light of massively parallel biological data acquisition. *Pac. Symp. Biocomput.* 4, 5–16.

Thomas, M. A., Joshi, P. P., and Klaper, R. D. (2012). Gene-class analysis of expression patterns induced by psychoactive pharmaceutical exposure in fathead minnow (*Pimephales promelas*) indicates induction of neuronal systems. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* 155, 109–120.

Trudeau, V. L., Martyniuk, C. J., Zhao, E., Hu, H., Volkoff, H., Decatur, W. A., and Basak, A. (2012). Is secretoneurin a new hormone? *Gen. Comp. Endocrinol.* 175, 10–18.

Villeneuve, D. L., Garcia-Reyero, N., Martinović, D., Cavallin, J. E., Mueller, N. D., Wehmas, L. C., Kahl, M. D., Linnum, A. L., Perkins, E. J., and Ankley, G. T. (2010). Influence of ovarian stage on transcript profiles in fathead minnow (*Pimephales promelas*) ovary tissue. *Aquat. Toxicol.* 98, 354–366.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476.

Wang, R. L., Bencic, D., Villeneuve, D. L., Ankley, G. T., Lazorchak, J., and Edwards, S. (2010). A transcriptomics-based biological framework for studying mechanisms of endocrine disruption in small fish species. *Aquat. Toxicol.* 98, 230–244.

Whitehead, A., Dubansky, B., Bodinier, C., Garcia, T. I., Miles, S., Pilley, C., Raghunathan, V., Roach, J. L., Walker, N., Walter, R. B., Rice, C. D., and Galvez, F. (2011). Genomic and physiological footprint of the *Deepwater Horizon* oil spill on resident marsh fishes. *Proc. Natl. Acad. Sci. U.S.A.* doi: 10.1073/pnas.1109545108 [Epub ahead of print].

Williams, T. D., Diab, A. M., George, S. G., Godfrey, R. E., Sabine, V., Conesa, A., Minchin, S. D., Watts, P.C., and Chipman, J. K. (2006). Development of the GENIPOL European flounder (*Platichthys flesus*) microarray and determination of temporal transcriptional responses to cadmium at low dose. *Environ. Sci. Technol.* 40, 6479–6488.

Xiang, L. X., He, D., Dong, W. R., Zhang, Y. W., and Shab, J. Z. (2010). Deep sequencing-based transcriptome profiling analysis of bacteria-challenged *Lateolabrax japonicus* reveals insight into the immune-relevant genes in marine fish. *BMC Genomics* 11, 472. doi: 10.1186/1471-2164-11-472

Zhang, G., Hoersch, S., Amsterdam, A., Whittaker, C. A., and Lees, J. A. (2010). Highly aneuploid zebrafish malignant peripheral nerve sheath tumors have genetic alterations similar to human cancers. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16940–16945.

Zhang, Y., Stupka, E., Henkel, C. V., Jansen, H. J., Spaink, H. P., and Verbeek, F. J. (2011a). Identification of common carp innate immune genes with whole-genome sequencing and RNA-seq data. *J. Integr. Bioinform.* 8, 169.

Zhang, Z., Wang, Y., Wang, S., Liu, J., Warren, W., Miltreva, M., and Walter, R. B. (2011b). Transcriptome analysis of female and male *Xiphophorus maculatus* Jp 163 A. *PLoS ONE* 6, e18379. doi: 10.1371/journal.pone.0018379

Conflict of Interest Statement: 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.

Received: 30 January 2012; paper pending published: 24 February 2012; accepted: 02 April 2012; published online: 25 April 2012.

*Citation: Mehinto AC, Martyniuk CJ, Spade DJ and Denslow ND (2012) Applications of next-generation sequencing in fish ecotoxicogenomics. Front. Gene. **3**:62. doi: 10.3389/fgene.2012.00062*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in
GENETICS

MINI REVIEW ARTICLE
published: 14 March 2012
doi: 10.3389/fgene.2012.00037

The metamorphosis of amphibian toxicogenomics

Caren C. Helbing*

Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC, Canada

Edited by:
Jonathan Freedman, National Institute of Health, USA

Reviewed by:
Tara Sabo-Attwood, University of Florida, USA
Raja S. Settivari, The Dow Chemical Company, USA

****Correspondence:***
Caren C. Helbing, Department of Biochemistry and Microbiology, University of Victoria, P.O. Box 3055 Stn CSC, Victoria, BC, Canada V8W 3P6.
e-mail: chelbing@uvic.ca

Amphibians are important vertebrates in toxicology often representing both aquatic and terrestrial forms within the life history of the same species. Of the thousands of species, only two have substantial genomics resources: the recently published genome of the Pipid, *Xenopus* (*Silurana*) *tropicalis*, and transcript information (and ongoing genome sequencing project) of *Xenopus laevis*. However, many more species representative of regional ecological niches and life strategies are used in toxicology worldwide. Since *Xenopus* species diverged from the most populous frog family, the Ranidae, ∼200 million years ago, there are notable differences between them and the even more distant Caudates (salamanders) and Caecilians. These differences include genome size, gene composition, and extent of polyploidization. Application of toxicogenomics to amphibians requires the mobilization of resources and expertise to develop *de novo* sequence assemblies and analysis strategies for a broader range of amphibian species. The present mini-review will present the advances in toxicogenomics as pertains to amphibians with particular emphasis upon the development and use of genomic techniques (inclusive of transcriptomics, proteomics, and metabolomics) and the challenges inherent therein.

Keywords: amphibian, frog, molecular techniques, polyploid, endocrine disruptor, estrogen, thyroid hormone, toxicogenomics

Amphibians have an undeniable, yet understated, role in toxicology. They diverged from other vertebrates 360 million years ago (Frost et al., 2006) and, currently, over 60% of the >6,900 known species are threatened or declining in numbers (AmphibiaWeb, 2012). Recent amphibians are comprised of three orders: Anura (frogs and toads), Caudata (salamanders), and Gymnophiona (caecilians; Dubois, 2004; Frost et al., 2006), of which ∼90% of species are Anura. Toxicological studies are primarily on anurans with some studies on caudates. Caecilian representation is lacking. In contrast to the limited ranges of caudate and caecilians, anurans are found on every continent except Antarctica (AmphibiaWeb, 2012).

Amphibians represent the only vertebrate group where a large majority of its members exhibit a life history that includes distinct independent aquatic larval and terrestrial juvenile/adult phases. The transition between the larval and juvenile phases requires substantial or complete remodeling of the organism (metamorphosis) in anticipation of a terrestrial lifestyle. Thus this places amphibians in a unique position for the assessment of toxicological effects in both aquatic and terrestrial environments. Over 10,000 study records on amphibians are currently available on the US Environmental Protection Agency's ECOTOX database (US EPA, 2012). Ninety percent of records represent aquatic exposures and these are biased toward a single species (*Xenopus laevis*). The remaining 10% of records comprise terrestrial exposures of frog, toad, and salamander species (US EPA, 2012). Less than 300 records include any mRNA expression data.

The exquisite sensitivity of frogs to hormones fostered the launching of several initiatives to develop standardized testing methods. Environment Canada is developing a frog tadpole exposure assay that uses native species (*Rana catesbeiana* and *pipiens*) and a combination of molecular and morphological criteria (Veldhoen et al., 2006c). The Organization for Economic Cooperation and Development (OECD, 2009) has recently established guidelines for a standardized assay for evaluating thyroid hormone (TH) active chemicals using a *X. laevis* metamorphosis assay (XEMA) and is in the process of evaluating a multigenerational reproductive assay in *X. tropicalis*. The XEMA assay has been adapted for *X. tropicalis* (Mitsui et al., 2006) and served as a template for the development of some native frog metamorphosis assays for *R. rugosa* (Oka et al., 2009), *Bombina orientalis* (Park et al., 2010), and *Pseudacris regilla* (Marlatt et al., submitted). However, the recommended XEMA assay relies upon morphological criteria and the input of toxicogenomic endpoints is not standard practice. Nevertheless, movement toward inclusion of molecular endpoints to reduce assay time and provide greater information regarding test chemical mode of action is evident in the literature (**Table 1**).

Amphibians are used in two general ways in the context of toxicology: in laboratory exposure settings where individual chemicals or complex mixtures are tested and in the field setting. Although availability of appropriate life stages is year round for some species (e.g., *Xenopus* tadpoles can be bred on demand, *R. catesbeiana* tadpoles can be collected from the wild and housed in aquatics facilities year round), many have limited availability throughout the year. Moreover, field sampling of threatened or endangered species necessitates the development of non-lethal sampling methods (fin biopsies) combined with molecular analyses (Veldhoen and Helbing, 2005). Efforts have also been made to combine transcript analysis with cultured tail fin biopsies

Table 1 | Representative studies using amphibian toxicogenomics.

Approach	Method	Species	Representative publications[a]
Transcriptomics	Microarray	*Rana catesbeiana*	Veldhoen et al. (2006b)
		Xenopus laevis	Helbing et al. (2007), Heimeier et al. (2009), Gohin et al. (2010), Searcy et al. (2012)
	QPCR	*Andrias japonicus*[b]	Katsu et al. (2006)
		Pleurodeles waltl	Ko et al. (2008)
		Pseudacris regilla	Veldhoen et al. (2006a), Marlatt et al. (submitted)
		Rana catesbeiana	Gunderson et al. (2011)
		Rana pipiens	Howe et al. (2004), Duarte-Guterman and Trudeau (2010), Langlois et al. (2010)
		Rana rugosa	Suda et al. (2011)
		Rana temporaria	Mortensen et al. (2006)
		Xenopus tropicalis	Langlois et al. (2011)
		Xenopus laevis	Zhang et al. (2006), Oka et al. (2008), Zimmermann et al. (2008), Baba et al. (2009), Rossi et al. (2009), Massari et al. (2010), Qin et al. (2010), Urbatzka et al. (2010), Zaya et al. (2011)
Proteomics	Two dimensional polyacrylamide gel electrophoresis followed by liquid chromatography and tandem mass spectrometry (2D-PAGE; LC-MS/MS)	*Rana catesbeiana* *Xenopus laevis*	Domanski and Helbing (2007), Serrano et al. (2010)
	Isobaric tags for relative and absolute quantitation (iTRAQ)	*Rana catesbeiana* *Xenopus laevis*	Domanski and Helbing (2007), Serrano et al. (2010)
Metabolomics	Ultra performance liquid chromatography Mass Spectrometry (UPLC-MS)	*Rana catesbeiana*	Helbing et al. (2010)
	Inductively coupled plasma mass spectrometry (ICP-MS)	*Xenopus laevis*	Tietge et al. (2010)

[a] *Due to space constraints, the author regrets that all studies could not be included. She has provided selected references representative of key laboratories active in the area and encourages the interested reader to consult the literature for additional published works by the highlighted laboratories.*

[b] *Due to the paucity of publications for salamanders, this work was included in the table even though this study used gel-based quantitation methods for PCR products.*

for rapid screening of chemicals and effluents (Hinther et al., 2010).

Toxicogenomics are best suited for identifying and evaluating factors categorized as sublethal deleterious effects that influence survival and recruitment; the primary factors contributing to amphibian population declines (Hayes et al., 2010). Such factors include: stress, susceptibility to disease, climate change, and environmental pollutants. Typically, molecular responses precede morphological endpoints giving early indications of response and modes of action. To date, application of toxicogenomics to amphibians has been extremely limited due to restrictions in resources/expertise and the difficulty in obtaining consensus on which toxicologically important species to develop large-scale genomics resources for. Approaches for transcriptomics, proteomics, and metabolomics pertaining to frogs has been previously reviewed elsewhere (Helbing et al., 2010). Research efforts have largely focused on the evaluation of endocrine disruption in frogs, primarily with respect to xenoestrogens and TH-active chemicals, including hormonal cross-talk and their complex interactions with environmental factors (**Table 1**). Sex reversal and/or intersex conditions in response to chemical exposures have been reported and some laboratories have begun to examine their molecular basis (**Table 1**). The absolute dependence of frog tadpoles upon TH during metamorphosis into a juvenile frog (Shi, 2000) provides for the most comprehensive and drastic response known to the hormone; although all vertebrates require THs for development, nervous system function, and metabolism (Oppenheimer, 1999). Indeed, the use of frog tadpoles as surrogate species for TH disruption in mammals has been explored at the molecular level (Searcy et al., 2012).

Even where more resources are available (for, e.g., commercially available *Xenopus* oligo microarrays), restrictions in cost and lack of utility across species platforms (Helbing et al., 2010) have greatly limited application of toxicogenomics tools beyond quantitative real time polymerase chain reaction (QPCR; **Table 1**). It is notable that very few relevant studies have been performed using salamanders and none with caecilians (**Table 1**). For labs having the necessary expertise, efforts have concentrated upon the production of transcript-based biomarkers (although some data on proteomics and metabolomics have been published; **Table 1**),

definition of baseline responses to model hormones, identification of appropriate sampling times for molecular biomarker use, and determination of natural variation (**Table 1**).

Most of the toxicogenomic studies that have been published so far are laboratory-based using *X. laevis* strains that are akin to the white mouse of the frog world. *Xenopus* are Pipidae; one of 38 families within the Anura (Frost et al., 2006). Pipidae are distantly related to most Anura and diverged from the largest family, the Ranidae (the true frogs), ∼200 million years ago (Sumida et al., 2004). Marked differences in life histories (for, e.g., *Xenopus* remain aquatic after metamorphosis, whereas Ranids become terrestrial), species sensitivities (Relyea and Jones, 2009), and genome compositions (discussed below) have prompted a call for developing toxicogenomics tools and approaches applicable to environmentally relevant species (Denslow et al., 2007).

Amphibian toxicogenomics has largely been driven through the adoption of tools made available through disciplines outside of toxicology, namely developmental, cell, and molecular biology. This provided, firstly, gene information from *X. laevis*, and, subsequently, a genome from *X. (Silurana) tropicalis*. However, this has not been without difficulty. *X. laevis* has been the most-used amphibian toxicological model due to the ease of husbandry in laboratory settings. However, genetically, *X. laevis* is pseudotetraploid, derived from an ancient tetraploid lineage with incomplete diploidization across a large portion of the genome (Mable et al., 2011). In fact, Pipidae have the highest number of polyploid species identified within amphibians (Mable et al., 2011). This created significant problems in initiating a genome sequencing project for this species due to the bioinformatic challenge of assembling a tetraploid genome *de novo*. Fortunately, a diploid *Xenopus* species did exist with similar husbandry benefits in this family. Thus the sequencing of the first frog genome was performed on *X. tropicalis* and completed in 2010 (Hellsten et al., 2010). The problem for toxicology is that the availability of the *X. tropicalis* genome information is now driving scientists to use this as a test species for use in toxicogenomics. Although an enormous opportunity, it is not clear how suitable this species will be as a representative of native frog species or amphibians in general.

Amphibians present a wide range of species diversity. They contain keystone species within a plethora of ecosystems throughout the world. They are important food sources for humans and wildlife, instrumental in pest control, and serve as sensitive indicators within a variety of ecosystems. With the exception of *R. catesbeiana* which is distributed worldwide, amphibians tend to have regional representation. Therefore, toxicological evaluations have often tended toward regionalism as well [for example, common frog ecotox species are *R. rugosa* (Japan), *R. temporaria* (Europe), and *R. pipiens* (North America)]. Moreover, amphibians have representatives of different sex determination systems (e.g., XX/XY, ZW/ZZ; Eggert, 2004) that could influence sensitivity to environmental contaminants. Genome organization is similar within amphibian subgroups, but varies substantially between subgroups. In addition to polyploidy in a few amphibian species, the genome sizes of amphibians span four orders of magnitude from one-quarter of the human genome (0.9 Gb, *Limnodynastes ornatus* = ornate burrowing frog) to among the largest known in animals (118 Gb, *Necturus lewisi* = gulf coast waterdog; Gregory, 2012). The estimated genome sizes and chromosome numbers of commonly used amphibian species in toxicology are presented in **Table 2**. Thus, coupling toxicological demands with genetics result in logistical and bioinformatic challenges. These have hampered building consensus and concerted effort to further genomics tools.

Despite this, current efforts are focused upon addressing the dearth in (1) application of available toxicogenomics resources to amphibians, and (2) genome sequence information representing amphibian species beyond *Xenopus*. Access to genome

Table 2 | Estimated genome sizes[a] of representative amphibians.

Species (common name)	Genome size (Gb)	Chromosome number
ANURA		
Bombina orientalis (oriental fire-bellied toad)	8.0	24
Bufo americanus (American toad)	5.1	22
Bufo bufo (common toad)	6.6	22
Bufo marinus (cane toad)	4.8	22
Hyla arborea (European tree frog)	4.7	24
Hyla versicolor (gray tree frog)	9.6	48
Pelobates fuscus (European spadefoot toad)	4.4	26
Pseudacris regilla (Pacific tree frog)	3.7	24
Rana aurora (red-legged frog)	9.0	26
Rana catesbeiana (North American bullfrog)	7.4	26
Rana clamitans (green frog)	6.7	26
Rana esculenta (edible frog)	6.8	26
Rana japonica (Japanese reddish frog)	5.7	26
Rana pipiens (northern leopard frog)	6.7	26
Rana rugosa (wrinkled frog)	8.0	26
Rana sylvatica (wood frog)	5.8	26
Rana temporaria (common European frog)	4.2	26
Spea hammondii (Western spadefoot toad)	1.6	26
Xenopus laevis (South African clawed frog)	3.2	36
Xenopus tropicalis (Western clawed frog)	1.7	20
CAUDATA		
Ambystoma maculatum (spotted salamander)	32.3	28
Ambystoma mexicanum (Mexican axolotl)	34.0	28
Ambystoma tigrinum (tiger salamander)	31.0	28
Andrias japonicus (Japanese giant salamander)	45.5	60
Dicamptodon ensatus (Pacific giant salamander)	55.6	28
Necturus maculosus (mudpuppy)	84.1	38
Notophthalmus viridescens (red spotted newt)	36.9	22
Pleurodeles waltl (Spanish ribbed newt)	20.0	24
Triturus vulgaris (common newt)	24.9	24
GYMNOPHIONA		
Geotrypetes seraphini (Gaboon caecilian)	4.6	38
Gymnopis multiplicata (Purple caecilian)	3.6	24–26
Siphonops annulatus (caecilian)	13.6	?

Adapted from Gregory (2012).

[a] *Genome sizes presented are the average of C-values from the Animal Genome Size database for a given species. C-values represent the haploid DNA amount in a gametic nucleus. The term is used interchangeably with genome size for diploids. However, when an organism is polyploid, the C-value may represent multiple genomes within the nucleus and may not represent the true haploid DNA amount.*

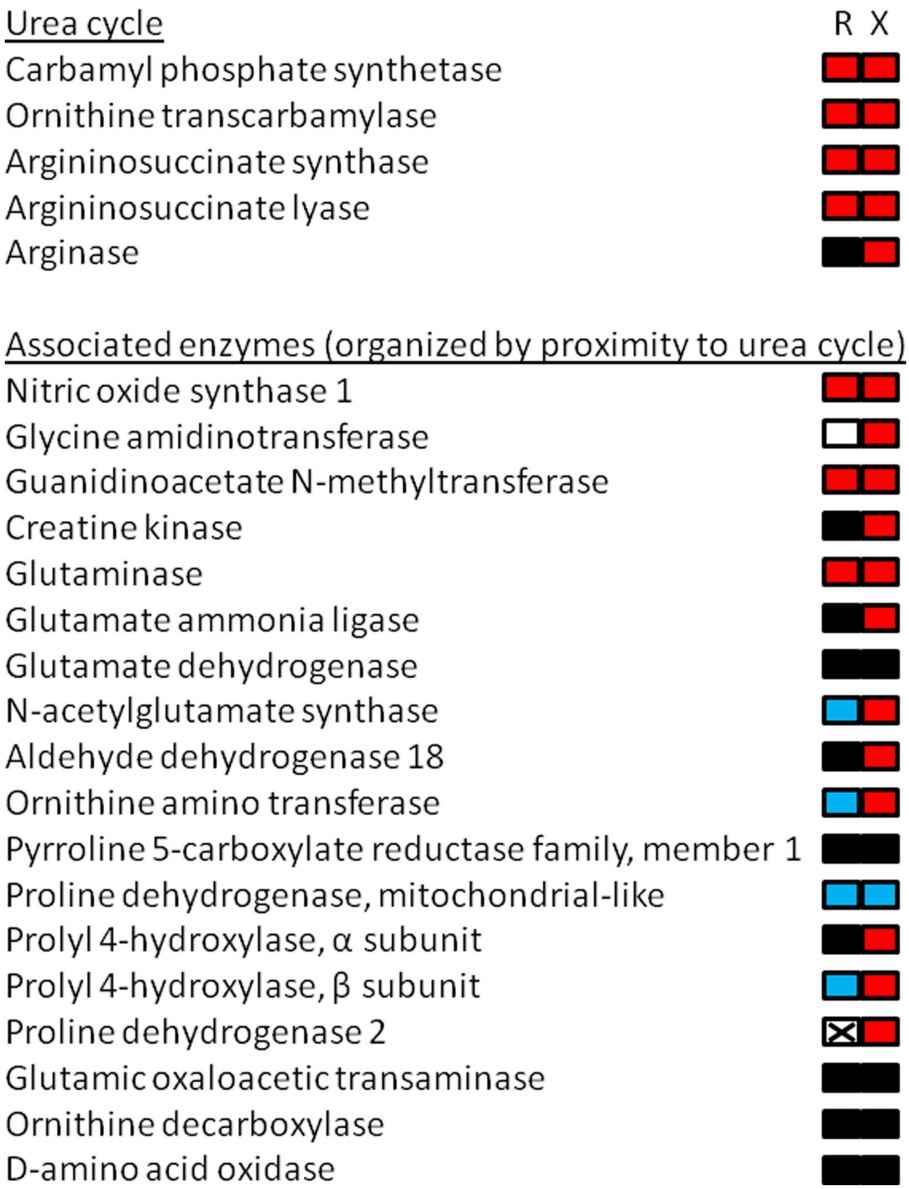


FIGURE 1 | Cartoon depiction of RNA-seq results from the liver of premetamorphic *Xenopus laevis* and *Rana catesbeiana* tadpoles focusing upon arginine and proline metabolism including the urea cycle. Tadpoles were exposed to 10 nM 3,5,3′-triiodothyronine (a thyroid hormone) or NaOH vehicle control for 48 h. The animals were treated and maintained in accordance with the guidelines of the Canadian Council on Animal Care. The liver transcriptomes were subjected to RNA-seq using 75 base HiSeq of paired end tagged (PET) libraries. The derived sequence information was assembled using the *X. tropicalis* genome as a scaffold and the contig identities were determined by a Blastx search against the *X. tropicalis* genome. The number of read counts (∼400 million) was normalized between samples and the relative count frequencies of the indicated pathway components were compared based upon the *X. tropicalis* arginine and proline metabolism KEGG pathway (xtr00330; www.genome.jp/kegg). The results are depicted as a bipartite rectangle beside the name of the enzymes corresponding with measured transcripts in the RNA-seq experiment that were identified in the KEGG pathway. The left side represents the relative change in transcript abundance levels of *Rana* (R) and the right side mRNA levels of *Xenopus* (X) where red is increased, black is no change, and blue is decreased relative to control animals. Non-detected transcript is depicted by a crossed-out white box. Use of the *X. tropicalis* genome as an assembly scaffold had limited utility since *X. laevis* and *R. catesbeiana* sequences aligned imperfectly to the *X. tropicalis* genome with *R. catesbeiana*, not surprisingly, having the least benefit of alignment. Nevertheless, some transcript identities linked to count frequencies were positively confirmed and the data obtained for the urea cycle enzymes, for example, matched well with previous observations (Helbing et al., 1992; Xu et al., 1993; Iwase et al., 1995). This validates the method for transcripts that are identifiable and quantifiable in this way.

and transcriptome sequence information is critical for the key toxicogenomics approaches today such as microarrays, QPCR, and proteomics techniques. The increasingly recognized role of epigenetic factors in toxicology necessitates the means for

analyzing genomes (Bilesimo et al., 2011). *De novo* high throughput sequencing of transcriptomes (RNA-seq) provides an unprecedented opportunity to obtain sequence and expression information of literally thousands of gene transcripts within a tissue (Martin and Wang, 2011). Access to resources such as Xenbase (www.xenbase.org) and deposition of amphibian expressed sequence tags (ESTs) and individually cloned sequences on publicly accessible databases have made it possible to garner limited information from RNA-seq experiments (**Figure 1**). However, accurate assembly and bioinformatics evaluation of RNA-seq data requires a genome sequence for the species of interest.

We will see more use of *X. tropicalis* for toxicogenomics; but we must be very careful not to let the genomics drive the toxicology and put resources and effort into developing appropriate tools for a wider range of toxicologically relevant species. Of the greatest urgency is obtaining genome sequence information from the Ranids, since they represent the largest group of amphibians worldwide, and the Caudata due to their unique genome structure. Now more than ever, we need the scaffolds and platforms for more amphibian species to move toxicogenomics forward.

ACKNOWLEDGMENTS

The author gratefully acknowledges grant support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Genome British Columbia. Thanks to N. Veldhoen for assistance in manuscript preparation, A. Carew and G. Taylor (Michael Smith Genome Sciences Centre, Vancouver, BC, Canada) for technical assistance, and N. Veldhoen, S. Maher, A. Carew, and P. Wojnarowicz for helpful discussions.

REFERENCES

AmphibiaWeb. (2012). *AmphibiaWeb: Information on Amphibian Biology and Conservation.* Available at: www.amphibiaweb.org [accessed January 12].

Baba, K., Okada, K., Kinoshita, T., and Imaoka, S. (2009). Bisphenol A disrupts Notch signaling by inhibiting gamma-secretase activity and causes eye dysplasia of *Xenopus laevis*. *Toxicol. Sci.* 108, 344–355.

Bilesimo, P., Jolivet, P., Alfama, G., Buisine, N., Le Mevel, S., Havis, E., Demeneix, B. A., and Sachs, L. M. (2011). Specific histone lysine 4 methylation patterns define TR-binding capacity and differentiate direct T3 responses. *Mol. Endocrinol.* 25, 225–237.

Denslow, N. D., Colbourne, J. K., Dix, D., Freedman, J. H., Helbing, C. C., Kennedy, S., and Williams, P. L. (2007). "Selection of surrogate animal species for comparative toxicogenomics," in *Genomic Approaches for Cross-Species Extrapolation in Toxicology*, eds W. Benson and R. Di Giulio (Pensacola, FL: SETAC and CRC Press), 33–75.

Domanski, D., and Helbing, C. C. (2007). Analysis of the *Rana catesbeiana* tadpole tail fin proteome and phosphoproteome during T3-induced apoptosis: identification of a novel type I keratin. *BMC Dev. Biol.* 7, 94. doi:10.1186/1471-213X-7-94

Duarte-Guterman, P., and Trudeau, V. L. (2010). Regulation of thyroid hormone-, oestrogen- and androgen-related genes by triiodothyronine in the brain of *Silurana tropicalis*. *J. Neuroendocrinol.* 22, 1023–1031.

Dubois, A. (2004). The higher nomenclature of recent amphibians. *Alytes* 22, 1–14.

Eggert, C. (2004). Sex determination: the amphibian models. *Reprod. Nutr. Dev.* 44, 539–549.

Frost, D. R., Grant, T., Faivovich, J., Bain, R. H., Haas, A., Haddad, C. F. B., De Sa, R. O., Channing, A., Wilkinson, M., Donnellan, S. C., Raxworthy, C. J., Campbell, J. A., Blotto, B. L., Moler, P., Drewes, R. C., Nussbaum, R. A., Lynch, J. D., Green, D. M., and Wheeler, W. C. (2006). The amphibian tree of life. *Bull. Am. Mus. Nat. Hist.* 297, 8–370.

Gohin, M., Bobe, J., and Chesnel, F. (2010). Comparative transcriptomic analysis of follicle-enclosed oocyte maturational and developmental competence acquisition in two non-mammalian vertebrates. *BMC Genomics* 11, 18. doi:10.1186/1471-2164-11-18

Gregory, T. (2012). *Animal Genome Size Database.* Available at: www.genomesize.com [accessed January 12].

Gunderson, M. P., Veldhoen, N., Skirrow, R. C., Macnab, M. K., Ding, W., van Aggelen, G., and Helbing, C. C. (2011). Effect of low dose exposure to the herbicide atrazine and its metabolite on cytochrome P450 aromatase and steroidogenic factor-1 mRNA levels in the brain of premetamorphic bullfrog tadpoles (*Rana catesbeiana*). *Aquat. Toxicol.* 102, 31–38.

Hayes, T. B., Falso, P., Gallipeau, S., and Stice, M. (2010). The cause of global amphibian declines: a developmental endocrinologist's perspective. *J. Exp. Biol.* 213, 921–933.

Heimeier, R. A., Das, B., Buchholz, D. R., and Shi, Y. B. (2009). The xenoestrogen bisphenol A inhibits postembryonic vertebrate development by antagonizing gene regulation by thyroid hormone. *Endocrinology* 150, 2964–2973.

Helbing, C., Bailey, C., Ji, L., Gunderson, M. P., Zhang, F., Veldhoen, N., Skirrow, R. C., Mu, R., Lesperance, M., Holcombe, G. W., Kosian, P. A., Tietge, J., Korte, J. J., and Degitz, S. J. (2007). Identification of gene expression indicators for thyroid axis disruption in a *Xenopus laevis* metamorphosis screening assay Part 1: effects on the brain. *Aquat. Toxicol.* 82, 227–241.

Helbing, C., Gergely, G., and Atkinson, B. G. (1992). Sequential up-regulation of thyroid hormone beta receptor, ornithine transcarbamylase, and carbamyl phosphate synthetase mRNAs in the liver of *Rana catesbeiana* tadpoles during spontaneous and thyroid hormone-induced metamorphosis. *Dev. Genet.* 13, 289–301.

Helbing, C. C., Maher, S. K., Han, J., Gunderson, M. P., and Borchers, C. (2010). Peering into molecular mechanisms of action with frogSCOPE. *Gen. Comp. Endocrinol.* 168, 190–198.

Hellsten, U., Harland, R. M., Gilchrist, M. J., Hendrix, D., Jurka, J., Kapitonov, V., Ovcharenko, I., Putnam, N. H., Shu, S., Taher, L., Blitz, I. L., Blumberg, B., Dichmann, D. S., Dubchak, I., Amaya, E., Detter, J. C., Fletcher, R., Gerhard, D. S., Goodstein, D., Graves, T., Grigoriev, I. V., Grimwood, J., Kawashima, T., Lindquist, E., Lucas, S. M., Mead, P. E., Mitros, T., Ogino, H., Ohta, Y., Poliakov, A. V., Pollet, N., Robert, J., Salamov, A., Sater, A. K., Schmutz, J., Terry, A., Vize, P. D., Warren, W. C., Wells, D., Wills, A., Wilson, R. K., Zimmerman, L. B., Zorn, A. M., Grainger, R., Grammer, T., Khokha, M. K., Richardson, P. M., and Rokhsar, D. S. (2010). The genome of the Western clawed frog *Xenopus tropicalis*. *Science* 328, 633–636.

Hinther, A., Domanski, D., Vawda, S., and Helbing, C. C. (2010). C-fin: a cultured frog tadpole tail fin biopsy approach for detection of thyroid hormone-disrupting chemicals. *Environ. Toxicol. Chem.* 29, 380–388.

Howe, C. M., Berrill, M., Pauli, B. D., Helbing, C. C., Werry, K., and Veldhoen, N. (2004). Toxicity of glyphosate-based pesticides to four North American frog species. *Environ. Toxicol. Chem.* 23, 1928–1938.

Iwase, K., Yamauchi, K., and Ishikawa, K. (1995). Cloning of cDNAs encoding argininosuccinate lyase and arginase from *Rana catesbeiana* liver and regulation of their mRNAs during spontaneous and thyroid hormone-induced metamorphosis. *Biochim. Biophys. Acta* 25, 139–146.

Katsu, Y., Kohno, S., Oka, T., Mitsui, N., Tooi, O., Santo, N., Urushitani, H., Fukumoto, Y., Kuwabara, K., Ashikaga, K., Minami, S., Kato, S., Ohta, Y., Guillette, L. J. Jr., and Iguchi, T. (2006). Molecular cloning of estrogen receptor alpha (ERalpha; ESR1) of the Japanese giant salamander, *Andrias japonicus*. *Mol. Cell. Endocrinol.* 257–258, 84–94.

Ko, C. I., Chesnel, A., Mazerbourg, S., Kuntz, S., Flament, S., and Chardard, D. (2008). Female-enriched expression of ERalpha during gonad differentiation of the urodele amphibian *Pleurodeles waltl*. *Gen. Comp. Endocrinol.* 156, 234–245.

Langlois, V. S., Carew, A. C., Pauli, B. D., Wade, M. G., Cooke, G. M., and Trudeau, V. L. (2010). Low levels of the herbicide atrazine alter sex ratios and reduce metamorphic success in *Rana pipiens* tadpoles raised in outdoor mesocosms. *Environ. Health Perspect.* 118, 552–557.

Langlois, V. S., Duarte-Guterman, P., and Trudeau, V. L. (2011). Expression profiles of reproduction- and thyroid hormone-related transcripts in the brains of chemically-induced intersex frogs. *Sex Dev.* 5, 26–32.

Mable, B. K., Alexandrou, M. A., and Taylor, M. I. (2011). Genome duplication in amphibians and fish: an extended synthesis. *J. Zool.* 284, 151–182.

Martin, J. A., and Wang, Z. (2011). Next-generation transcriptome assembly. *Nat. Rev. Genet.* 12, 671–682.

Massari, A., Urbatzka, R., Cevasco, A., Canesi, L., Lanza, C., Scarabelli, L., Kloas, W., and Mandich, A. (2010). Aromatase mRNA expression in the brain of adult *Xenopus laevis* exposed to Lambro river water and endocrine disrupting compounds. *Gen. Comp. Endocrinol.* 168, 262–268.

Mitsui, N., Fujii, T., Miyahara, M., Oka, T., Kashiwagi, A., Kashiwagi, K., Hanada, H., Urushitani, H., Santo, N., Tooi, O., and Iguchi, T. (2006). Development of metamorphosis assay using *Silurana tropicalis* for the detection of thyroid system-disrupting chemicals. *Ecotoxicol. Environ. Saf.* 64, 281–287.

Mortensen, A. S., Kortner, T. M., and Arukwe, A. (2006). Thyroid hormone-dependent gene expression as a biomarker of short-term 1,1-dichloro-2,2-bis(p-chlorophenyl)ethylene (DDE) exposure in European common frog (*Rana temporaria*) tadpoles. *Biomarkers* 11, 524–537.

Oka, T., Miyahara, M., Yamamoto, J., Mitsui, N., Fujii, T., Tooi, O., Kashiwagi, K., Takase, M., Kashiwagi, A., and Iguchi, T. (2009). Application of metamorphosis assay to a native Japanese amphibian species, *Rana rugosa*, for assessing effects of thyroid system affecting chemicals. *Ecotoxicol. Environ. Saf.* 72, 1400–1405.

Oka, T., Tooi, O., Mitsui, N., Miyahara, M., Ohnishi, Y., Takase, M., Kashiwagi, A., Shinkai, T., Santo, N., and Iguchi, T. (2008). Effect of atrazine on metamorphosis and sexual differentiation in *Xenopus laevis*. *Aquat. Toxicol.* 87, 215–226.

Oppenheimer, J. (1999). Evolving concepts of thyroid hormone action. *Biochimie* 81, 539–543.

Organization for Economic Cooperation and Development (OECD). (2009). *OECD Guideline for the (Testing) of Chemicals Test No. 231: The Amphibian Metamorphosis Assay* (Paris: OECD).

Park, C. J., Kang, H. S., and Gye, M. C. (2010). Effects of nonylphenol on early embryonic development, pigmentation and 3,5,3'-triiodothyronine-induced metamorphosis in *Bombina orientalis* (Amphibia: Anura). *Chemosphere* 81, 1292–1300.

Qin, X., Xia, X., Yang, Z., Yan, S., Zhao, Y., Wei, R., Li, Y., Tian, M., Zhao, X., Qin, Z., and Xu, X. (2010). Thyroid disruption by technical decabromodiphenyl ether (DE-83R) at low concentrations in *Xenopus laevis*. *J. Environ. Sci. (China)* 22, 744–751.

Relyea, R. A., and Jones, D. K. (2009). The toxicity of roundup original max (R) to 13 species of larval amphibians. *Environ. Toxicol. Chem.* 28, 2004–2008.

Rossi, F., Bernardini, G., Bonfanti, P., Colombo, A., Prati, M., and Gornati, R. (2009). Effects of TCDD on spermatogenesis related factor-2 (SRF-2): gene expression in *Xenopus*. *Toxicol. Lett.* 191, 189–194.

Searcy, B. T., Beckstrom-Sternberg, S. M., Beckstrom-Sternberg, J. S., Stafford, P., Schwendiman, A. L., Soto-Pena, J., Owen, M. C., Ramirez, C., Phillips, J., Veldhoen, N., Helbing, C. C., and Propper, C. R. (2012). Thyroid hormone-dependent development in *Xenopus laevis*: a sensitive screen of thyroid hormone signaling disruption by municipal wastewater treatment plant effluent. *Gen. Comp. Endocrinol.* doi:10.1016/j.ygcen.2011.12.036

Serrano, J., Higgins, L., Witthuhn, B. A., Anderson, L. B., Markowski, T., Holcombe, G. W., Kosian, P. A., Korte, J. J., Tietge, J. E., and Degitz, S. J. (2010). In vivo assessment and potential diagnosis of xenobiotics that perturb the thyroid pathway: proteomic analysis of *Xenopus laevis* brain tissue following exposure to model T4 inhibitors. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 5, 138–150.

Shi, Y.-B. (2000). *Amphibian Metamorphosis: From Morphology to Molecular Biology*. New York: Wiley-Liss.

Suda, M., Kodama, M., Oshima, Y., Yamamoto, K., Nakamura, Y., Tanaka, S., Kikuyama, S., and Nakamura, M. (2011). Up-regulation of FSHR expression during gonadal sex determination in the frog *Rana rugosa*. *Gen. Comp. Endocrinol.* 172, 475–486.

Sumida, M., Kato, Y., and Kurabayashi, A. (2004). Sequencing and analysis of the internal transcribed spacers (ITSs) and coding regions in the EcoR I fragment of the ribosomal DNA of the Japanese pond frog *Rana nigromaculata*. *Genes Genet. Syst.* 79, 105–118.

Tietge, J. E., Butterworth, B. C., Haselman, J. T., Holcombe, G. W., Hornung, M. W., Korte, J. J., Kosian, P. A., Wolfe, M., and Degitz, S. J. (2010). Early temporal effects of three thyroid hormone synthesis inhibitors in *Xenopus laevis*. *Aquat. Toxicol.* 98, 44–50.

United States Environmental Protection Agency (US EPA). (2012). *ECOTOX Database*. Available at: http://cfpub.epa.gov/ecotox/quick_query.htm [accessed January 12].

Urbatzka, R., Lorenz, C., Lutz, I., and Kloas, W. (2010). Expression profiles of LHbeta, FSHbeta and their gonadal receptor mRNAs during sexual differentiation of *Xenopus laevis* tadpoles. *Gen. Comp. Endocrinol.* 168, 239–244.

Veldhoen, N., Boggs, A., Walzak, K., and Helbing, C. C. (2006a). Exposure to tetrabromobisphenol-A alters TH-associated gene expression and tadpole metamorphosis in the Pacific Tree Frog *Pseudacris regilla*. *Aquat. Toxicol.* 78, 292–302.

Veldhoen, N., Skirrow, R., Ji, L., Domanski, D., Bonfield, E. R., Bailey, C. M., and Helbing, C. C. (2006b). Use of heterologous cDNA arrays and organ culture in the detection of thyroid hormone-dependent responses in a sentinel frog, *Rana catesbeiana*. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 1, 187–199.

Veldhoen, N., Skirrow, R., Osachoff, H., Wigmore, H., Clapson, D. J., Gunderson, M. P., Van Aggelen, G., and Helbing, C. C. (2006c). The bactericidal agent Triclosan modulates thyroid hormone-associated gene expression and disrupts postembryonic anuran development. *Aquat. Toxicol.* 80, 217–227.

Veldhoen, N., and Helbing, C. C. (2005). "Monitoring gene expression in *Rana catesbeiana* tadpoles using a tail fin biopsy technique and its application to the detection of environmental endocrine disruptor effects in wildlife species," in *Techniques in Aquatic Toxicology*, Vol. 2, ed. G. K. Ostrander (Boca Raton: CRC Press), 315–327.

Xu, Q., Baker, B. S., and Tata, J. R. (1993). Developmental and hormonal regulation of the Xenopus liver-type arginase gene. *Eur. J. Biochem.* 211, 891–898.

Zaya, R. M., Amini, Z., Whitaker, A. S., and Ide, C. F. (2011). Exposure to atrazine affects the expression of key genes in metabolic pathways integral to energy homeostasis in *Xenopus laevis* tadpoles. *Aquat. Toxicol.* 104, 254–262.

Zhang, F., Degitz, S. J., Holcombe, G. W., Kosian, P. A., Tietge, J., Veldhoen, N., and Helbing, C. C. (2006). Evaluation of gene expression endpoints in the context of a *Xenopus laevis* metamorphosis-based bioassay to detect thyroid hormone disruptors. *Aquat. Toxicol.* 76, 24–36.

Zimmermann, A. L., King, E. A., Dengler, E., Scogin, S. R., and Powell, W. H. (2008). An aryl hydrocarbon receptor repressor from *Xenopus laevis*: function, expression, and role in dioxin responsiveness during frog development. *Toxicol. Sci.* 104, 124–134.

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

Received: 31 January 2012; accepted: 27 February 2012; published online: 14 March 2012.

Citation: Helbing CC (2012) The metamorphosis of amphibian toxicogenomics. Front. Gene. ***3****:37. doi: 10.3389/fgene.2012.00037*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in
GENETICS

ORIGINAL RESEARCH ARTICLE
published: 05 April 2012
doi: 10.3389/fgene.2012.00048

Meta-analysis of global transcriptomics suggests that conserved genetic pathways are responsible for Quercetin and Tannic acid mediated longevity in *C. elegans*

Kerstin Pietsch[1,2]*, Nadine Saul[1,2], Suresh C. Swain[2†], Ralph Menzel[1,2], Christian E. W. Steinberg[1] and Stephen R. Stürzenbaum[2]*

[1] *Laboratory of Freshwater and Stress Ecology, Department of Biology, Humboldt-Universität zu Berlin, Berlin, Germany*
[2] *Analytical and Environmental Science Division, School of Biomedical Sciences, King's College London, London, UK*

Edited by:
Michael Aschner, Vanderbilt University Medical Center, USA

Reviewed by:
Kim A. Caldwell, The University of Alabama, USA
Chris Vulpe, University of California, USA
Pan Chen, Vanderbilt University, USA

****Correspondence:***
Kerstin Pietsch, Laboratory of Freshwater and Stress Ecology, Department of Biology, Humboldt-Universität zu Berlin, Spaethstr. 80/81, 12437 Berlin, Germany.
e-mail: kpietsch@gmx.de;
Stephen R. Stürzenbaum, Analytical and Environmental Science Division, School of Biomedical Sciences, King's College London, 150 Stamford Street, London SE1 9NH, UK.
e-mail: stephen.sturzenbaum@kcl.ac.uk

†Present address:
Suresh C. Swain, Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Penglais, Aberystwyth, Ceredigion SY23 3DA, UK.

The microarray data is MIAME compliant and has been deposited in the NCBI Gene Expression Omnibus, GEO (http://www.ncbi.nlm.nih.gov/geo) accession number GSE35354.

Recent research has highlighted that the polyphenols Quercetin and Tannic acid are capable of extending the lifespan of *Caenorhabditis elegans*. To gain a deep understanding of the underlying molecular genetics, we analyzed the global transcriptional patterns of nematodes exposed to three concentrations of Quercetin or Tannic acid, respectively. By means of an intricate meta-analysis it was possible to compare the transcriptomes of polyphenol exposure to recently published datasets derived from (i) longevity mutants or (ii) infection. This detailed comparative *in silico* analysis facilitated the identification of compound specific and overlapping transcriptional profiles and allowed the prediction of putative mechanistic models of Quercetin and Tannic acid mediated longevity. Lifespan extension due to Quercetin was predominantly driven by the metabolome, TGF-beta signaling, Insulin-like signaling, and the p38 MAPK pathway and Tannic acid's impact involved, in part, the amino acid metabolism and was modulated by the TGF-beta and the p38 MAPK pathways. DAF-12, which integrates TGF-beta and Insulin-like downstream signaling, and genetic players of the p38 MAPK pathway therefore seem to be crucial regulators for both polyphenols. Taken together, this study underlines how meta-analyses can provide an insight of molecular events that go beyond the traditional categorization into gene ontology-terms and Kyoto encyclopedia of genes and genomes-pathways. It also supports the call to expand the generation of comparative and integrative databases, an effort that is currently still in its infancy.

Keywords: Quercetin, Tannic acid, TGF-beta, ILS, DAF-12, p38 MAPK, *C. elegans*

INTRODUCTION

Although numerous molecules of herbal origin have been hailed as powerful decelerators of aging, few studies have focused on the precise physiological and genetic mechanisms that drive this process. The nematode *Caenorhabditis elegans* is ideally suited for biogerontological research, not only because of its short life cycle (Gami and Wolkow, 2006; Gill, 2006; Kaletta and Hengartner, 2006) but also due to the considerable conservation of basic cellular and molecular principles (The *C. elegans* Sequencing Consortium, 1998).

Of the many bioactive polyphenols (PPs) previously shown to extend the lifespan of *C. elegans* we selected two well characterized compounds, namely Quercetin (Q; work in *C. elegans*: Kampkötter et al., 2007a,b, 2008; Saul et al., 2008; Pietsch et al., 2009, 2011; Surco-Laos et al., 2011; Grünz et al., 2012; for a detailed review about general beneficial effects of Q, tested *in vivo* and *ex vivo*, be referred to Boots et al., 2008) and Tannic acid (TA; work in *C. elegans*: Saul et al., 2010, 2011; Lublin et al., 2011; for a detailed review be referred to Koleckar et al., 2008). Both PPs are characterized by inverted J-shaped concentration–response curves, typical for a hormetic effect (Calabrese and Baldwin, 2001). Moreover, the exposure to PPs enhances the oxidative and thermal stress resistance. Both PPs exert no major negative impact on the reproductive output but result in a reduction of body length and fat

content, two phenotypes that align well with the disposable soma patterns described by Kirkwood (1977, 1988). Utilizing various nematode mutant strains, preliminary genetic analyses identified *daf-2*, *age-1*, *sek-1*, and *unc-43* as mediators of Q induced longevity and stress resistance (Pietsch et al., 2009) and *sek-1* was shown to be important for TA mediated extension of lifespan (Saul et al., 2010). Notwithstanding these differences, some of the data obtained suggests that Q and TA induce longevity by similar genetic mechanisms, e.g., the involvement of *sek-1*, altering the fat metabolism, or hormesis-based dose–responses. An overview of results from previous studies can be found in Table S1 in Supplementary Material 1.

To pinpoint molecular genetic pathways, we performed genome-wide DNA microarray experiments spanning different concentrations of Q and TA, including at least two lifespan prolonging concentrations for each PP. In addition, we tested a low dose of Q (50 μM) because (i) the limitation of solubility precluded the use of doses above 200 μM Q and (ii) this concentration was shown to enhance thermal tolerance (unpublished data) possibly indicative of elevated stress resistance. For TA we included a post-effective, in single trials already toxic (Saul et al., 2010) concentration. Initially, this identified differentially expressed genes (DEGs), which were found to be significantly up- or down-regulated in response to Q and/or TA, many displayed a concentration dependent change in expression. A second layer of analysis included gene ontology (GO), Kyoto encyclopedia of genes and genomes (KEGG) pathways, as well as gene expression mountain algorithms (as introduced by Kim et al., 2001). By analyzing the overlapping transcripts identified in the respective lifespan prolonging concentrations of each PP (Q 100 and 200 μM = $Q_{longevity}$; TA 100 and 200 μM = $TA_{longevity}$), we were able to predict putative master regulators involved in PP mediated longevity. A final meta-analysis compared these core-genes with recently published transcriptional profile data linked to age-related gene expression, the genetic background of longevity mutants or infection. The alignment of common (condition-overlapping) DEGs facilitated the identification of genes and associated pathways that may act as master switches of longevity.

The primary goal of this study was to highlight that the meta-analysis of large datasets is not restricted to complement a simplistic database evaluation but has the potential to uncover compound overlapping molecular switches.

MATERIALS AND METHODS

STRAINS AND SAMPLE GENERATION

N2 wild type *C. elegans* were maintained on nematode growth medium (NGM) plates using *Escherichia coli* OP50 as food source (Brenner, 1974; Sulston and Hodgkin, 1988; Lewis and Fleming, 1995). Untreated nematodes (the P0 generation) were chunked onto control and treatment plates (50, 100, and 200 μM Q; 100, 200, and 300 μM TA) and incubated at 20°C for 4 days. A synchronous culture was generated through egg preparation (Strange et al., 2007) with sodium hypochlorite (Sigma, Germany). Eggs were rotated over night (20 rpm) and the resultant hatched L1s subsequently transferred to plates containing the respective doses of Q or TA. Worms were grown to the young (pre-reproductive) adult stage, harvested by rinsing off with M9 buffer, washed at least three times with M9, shock frozen in liquid nitrogen and stored at −80°C. For each condition, samples were cultivated in triplicate.

RNA PREPARATION

RNA was isolated following the standard procedures as defined by the Trizol protocol (Invitrogen, Germany) but modified to include a homogenization step with 0.5 mm glass beads to maximize cell breakage. The resultant RNA was further processed by means of a RNeasy purification followed by DNase digestion (Qiagen, Germany). All samples were stored at −80°C until further use.

RNA AMPLIFICATION, BIOTIN LABELING, AND DNA MICROARRAY ANALYSIS

RNA was processed with the MessageAMP™Premier RNA Amplification Kit (Ambion, Austin, TX, USA) which relies on the T7 *in vitro* transcription (IVT) amplification technology (Van Gelder et al., 1990). First- and second-strand cDNA synthesis, cRNA synthesis, labeling, fragmentation, GeneChip hybridization, and scanning were performed according to the manufacturer's specifications (Affymetrix, Santa Clara, CA, USA). We utilized the *C. elegans* array chip (Affymetrix) which covers the whole genome (22,548 transcripts). Triplicate chips were run for each condition (designated as Q0, Q50, Q100, Q200 and TA0, TA100, TA200, TA300). Whole RNA, cDNA, and cRNA qualities and quantities were assessed at each step using capillary electrophoresis (Bioanalyzer, Agilent Technologies, UK) and micro volume spectrophotometry (NanoDrop1000, Thermo Scientific, UK).

DATA INTERPRETATION AND STATISTICAL ANALYSIS

Processing of global transcription expression values (DNA microarray)

Pre-processing of raw microarray data included probe-specific background correction, summation of probe set values, and normalization using the GCRMA algorithm with CARMAweb 1.4, an R- and Bioconductor-based web service for microarray data analysis[1] (Rainer et al., 2006). The quality of normalization was assessed by box plot and MA-graph analyses. Differences between treatments were visualized by principal component analysis (PCA) plotting with multiexperiment viewer (MeV)[2] (Saeed et al., 2003). Data was initially filtered for missing values and then subjected to a CLEAR-test that combines differential expression and variability using the GEPAS web server[3] (Tárraga et al., 2008). An unpaired *t*-test was followed by a significance analysis of microarray (SAM) test including a calculation that estimates the false discovery rate (FDR). The FDR was set to a non-stringent level of <12.5%. DEGs showing a fold change of at least 1.25 were analyzed for their molecular functions, biological processes, and cellular components using the software packages GoMiner (Zeeberg et al., 2005) and DAVID[4]. Due to the sound technical and experimental quality of the data, which returned strong statistically significant signal intensities, the chosen fold-cut-off values were deemed to be biologically meaningful and aligned to previous data analyses (Grigoryev et al., 2004; McCarthy and Smyth, 2009).

[1] https://carmaweb.genome.tugraz.at/carma
[2] http://www.tm4.org/mev
[3] http://www.gepas.org
[4] http://david.abcc.ncifcrf.gov/

Comparison of datasets to screen for significant overlaps: representation factor

The overlap between differing conditions (e.g. PPs and gene expression mountains, gene classes, or datasets taken from the literature) was determined by computing the representation factors (RFs). The RFs define the fold enrichment between gene lists (Kim et al., 2001; Evans et al., 2008). The choice of N (genome) was based on the values recommended by the authors. Intersection p-values were calculated from the hypergeometric distribution. RFs were considered significant when RF >1 and $^{*}p < 0.05$, $^{**}p < 0.005$ or $^{***}p < 0.001$.

RESULTS

THE RELATIONSHIP BETWEEN EXPOSURE TO Q OR TA, LIFESPAN EXTENSION AND THE OUTPUT OF TRANSCRIPTIONAL RESPONSES

Previous reports have stated that, whilst 50 μM of Q evoked no significant change in lifespan, 100 and 200 μM Q prolonged the mean lifespan by 11 and 10%, respectively (Pietsch et al., 2009, 2011). However, at 250 μM of Q the lifespan was reduced by −7% (see Table S1 in Supplementary Material 1) and therefore can be considered to be a toxic concentration. In addition, because concentrations above 200 μM of Q did not dissolve completely in the agar and the bacteria, we refrained from using doses above 200 μM Q for the microarray experiment. Thus, for Q, we utilized two lifespan prolonging concentrations (100 and 200 μM), and a non-effective (50 μM) pre-longevity dose for global transcriptomics.

For TA, the most effective concentration was 100 μM, which resulted in an increase in mean lifespan of 18%, an effect that was still significant but notably less pronounced at 200 μM (8%) and was absent at 300 μM (Saul et al., 2010; see Table S1 in Supplementary Material 1). Hence, we used two life-extending concentrations (100 and 200 μM of TA), and in addition one post-effective (borderline toxic) concentration (300 μM TA).

We chose this test design to cover a large dose-range: from pre-effective to effective in the case of Q (Q0, Q50, Q100, Q200), and from effective to post-effective/toxic for TA (TA0, TA100, TA200, TA300; Note: the numerical identifiers represent the concentration in micromolar). Pre-processing of raw microarray data included probe-specific background correction, summation of probe set values, and normalization using the GCRMA algorithm with 93 CARMAweb 1.4, an R- and Bioconductor-based web service for microarray data analysis (see text footnote 1) (Rainer et al., 2006). The CARMAweb reports can be viewed in Supplementary Material 2 and 3.

Venn diagrams (**Figure 1**) were compiled to summarize the level of overlap between doses (statistics for significant DEGs can be found in Supplementary Material 4: sheet "Statistics"; full genome fold change (FC-) values are listed in Supplementary Material 5). Investigating the transcriptional response of Q treatment (**Figure 1A**), revealed that the number of DEGs increased markedly at the lifespan modulating doses of Q100 and Q200 (3.3- and 3.8-fold compared to the non-longevity dose Q50). Interestingly, more genes were up-regulated than down-regulated at Q50, but the opposite was observed at Q100 and Q200. It is worthy of note that the majority of DEGs overlapped in Q100 and Q200 exposures, thus suggesting the presence of common response pathways.

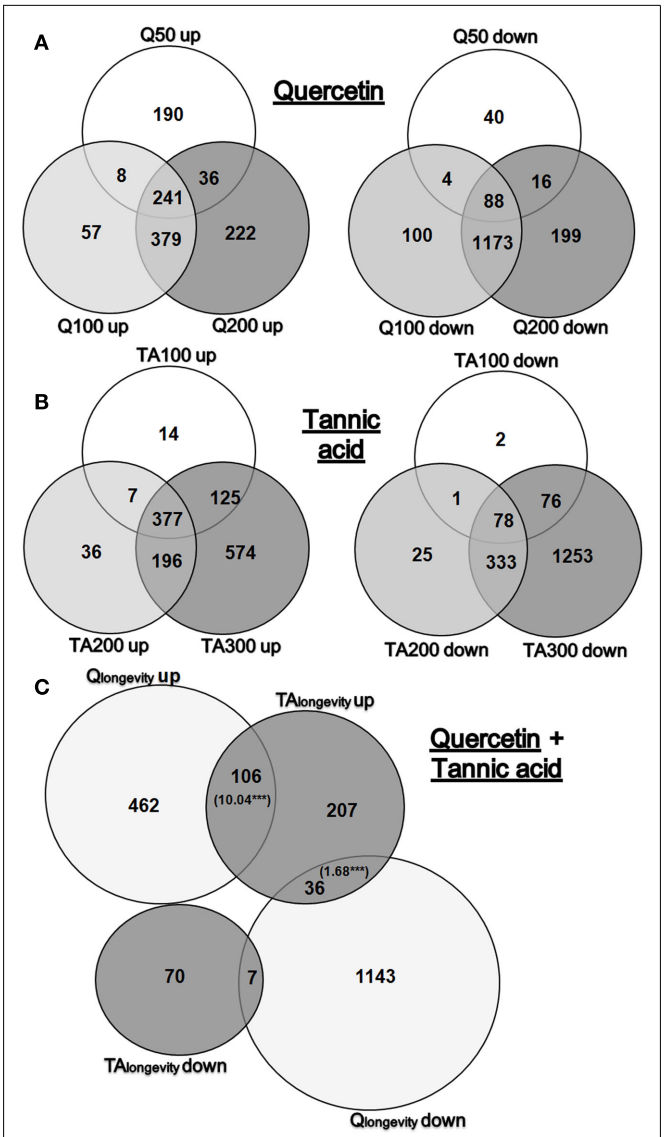


FIGURE 1 | Venn diagrams of DEGs. Overview of significant DEGs (fold change >1.25 or <0.8) in response to Q50, Q100, Q200 or TA100, TA200, TA300. Shown is the overlap of either up- or down-regulated genes in Q **(A)** or TA **(B)** treated nematodes. **(C)** Displays the overlap of significantly up- and down-regulated genes in $Q_{longevity}$ and $TA_{longevity}$. The bracketed numbers in the intersections display significant overrepresented RFs, $^{***}p < 0.001$.

As with Q, exposure to TA resulted in a dose dependent increase in DEGs which was most pronounced at the highest dose tested. More genes were up-regulated than down-regulated in the lifespan prolonging conditions TA100 and TA200, a ratio that was reversed at TA300 where more genes were down-regulated (**Figure 1B**). Striking was the proportion of DEGs that responded exclusively to TA300 (61%), a substantial proportion compared to TA100 (2%) or TA200 (6%).

By taking the intersection of DEGs shared by the lifespan prolonging concentrations of each PP, it was possible to extract gene lists for $Q_{longevity}$ (Q100 and Q200 overlap), $TA_{longevity}$ (T100 and T200 overlap) and $Q\&TA_{longevity}$ (overlap of all four groups). A Venn diagram that incorporates $Q_{longevity}$ and $TA_{longevity}$ identified a significant overlap (**Figure 1C**). As our main focus was the identification of mechanisms involved in PP triggered lifespan extension, downstream data processing focused primarily on $Q_{longevity}$ and $TA_{longevity}$ gene lists, however, individual GO-, KEGG-, gene expression mountains-, and gene class-analyses are provided for all concentrations in the Supplementary Material 4.

OVERREPRESENTED GO-TERMS

The summary results of the GO-cluster analysis with DEGs from the $Q_{longevity}$, $TA_{longevity}$, and $Q\&TA_{longevity}$ gene lists are presented in Table S2 in Supplementary Material 1. The complete analyses for all single concentrations are provided in the Supplementary Material 4: sheet "GO-Analysis." The $Q_{longevity}$ analysis returned, due to the large input list of DEGs, more overrepresented GO-terms than the equivalent analysis with $TA_{longevity}$. Overrepresented GO-terms derived from the up-regulated transcripts in $Q_{longevity}$ included *chromatin assembly*, *lipid metabolic process*, *monooxygenase activity*, and *nucleosome*. GO-terms from down-regulated genes included, besides others, *nervous system development*, *regulation of multicellular organism growth*, *Dauer entry*, *regulation -of transcription*, *-of response to stimulus*, *-of cell communication*, *-of biological quality*, *-of locomotion*, and *-of programmed cell death*. In $TA_{longevity}$, significant GO-terms were linked to *muscle contraction*, *neurotransmitter transporter activity and cytoskeleton*, *embryonic development ending in birth or egg hatching*, *positive regulation of biological process*, and *chromatin*. GO-terms in $Q\&TA_{longevity}$ were restricted to Cellular Components (*pseudopodium*).

Scrutinizing the GO-analysis of single concentrations (Supplementary Material 4: sheet "GO-Analysis") revealed, for example, that the GO-terms in the pre-lifespan extending concentration Q50 (*lysozyme activity and oxidoreductase activity acting on the CH–CH group of donors*) may reflect an early onset of induced immunity and stress resistance. In contrast, a striking accumulation of overrepresented GO-terms were observed in TA300, which included numerous categories indicative of an unfavorable condition (*DNA damage response*, *signal transduction*, *mismatch DNA binding*, *cell death*, and others).

OVERREPRESENTED KEGGs

Whilst the assignment of GO-terms is defined by automatic/electronic annotation that is based on sequence homology, KEGG-pathways are manually curated from the literature. As with the GO-ontology, KEGG-analysis on the Q treatment lists returned a multitude of pathways (**Table 1**). For further details about regulated genes in the respective KEGG-terms see Table S3 in Supplementary Material 1 for $Q_{longevity}$ and $TA_{longevity}$, and Supplementary Material 4: sheet "KEGG-Analysis" for single concentrations. Overrepresented KEGGs in Q50 included pathways involved in the metabolism of amino acids, glutathione, and xenobiotics as well as fatty acid elongation. At higher exposures seven (Q100) and eight (Q200) KEGG-pathways were found to be overrepresented, notably most (six) were present at both concentrations (see $Q_{longevity}$, **Table 1**) with an analogous mode/direction of regulation, i.e., either repressed or induced (Table S3 in Supplementary Material 1). Pathways include metabolism of amino acids, xenobiotics or drugs, transport processes (*Lysosome*), and signal transduction processing (*Wnt* and *TGF-beta signaling*). Whilst DEGs from amino acid metabolism and the *lysosome* displayed heterogeneous expression levels, transcripts belonging to the signaling pathways were predominantly repressed, whereas almost all genes associated with *drug/xenobiotic metabolism* were found to be induced.

The dynamic response patterns of Q exposure were not mirrored by the TA treatments. No KEGG-pathways could be linked to the most potent lifespan extending concentration, TA100; hence, none were assigned to $TA_{longevity}$. For TA200 only two KEGG-pathways were found to be overrepresented, both derived from amino acid metabolism. Analysis of the TA300 gene list returned six KEGG-pathways, notably *DNA replication*, *mismatch repair*, and *Ubiquitin-mediated proteolysis*.

META-ANALYSIS: COMPARISON OF GLOBAL TRANSCRIPTIONAL PATTERNS IN $Q_{LONGEVITY}$ AND $TA_{LONGEVITY}$ TO SELECTED DATASETS IN THE LITERATURE

To date, several microarray studies have identified age-related transcriptional changes in the nematode *C. elegans* (summarized in Golden and Melov, 2007). To investigate if a systematic comparison would reveal common age-related pathways, we correlated our datasets with the published expression profiles obtained from long-lived *daf-12(rh273)* (Fisher and Lithgow, 2006), *daf-2* mutants (Evans et al., 2008), TGF-beta mutants (Shaw et al., 2007), and immune-responsive *Pseudomonas aeruginosa* infected nematodes (Evans et al., 2008). The dataset from Evans et al. (2008) includes a meta-analysis of different *daf-2* alleles (Murphy et al., 2003; McElwee et al., 2004) and immune-challenged nematodes (Shapira et al., 2006; Troemel et al., 2006).

The meta-analysis aimed to indentify similar molecular mechanisms of longevity in $Q_{longevity}$, as well as $TA_{longevity}$ and define (if possible) the correlation to long-living mutants. Datasets were analyzed by focusing on (i) overrepresented gene expression mountains (according to Kim et al., 2001), which characterize the global patterns of a transcriptional response and (ii) DEGs, to identify new target genes. The significance of overlap between two gene lists were assigned by calculating the RF, and significances were determined by hypergeometric probability scoring. Supplementary Material 5 provides Gene Expression Mountains, Gene Classes and Groups, and further studies in a table. By using the filter function in Excel, all results that are described in the following sections can be reconstructed.

Overrepresented expression mountains

Kim et al. (2001) assembled data from several hundred *C. elegans* DNA microarray experiments. This allowed a three-dimensional gene expression map to be computed consisting of 44 co-regulated gene-groups (mountains), 30 of which could be assigned to specific gene classes and therefore to a potential physiological importance.

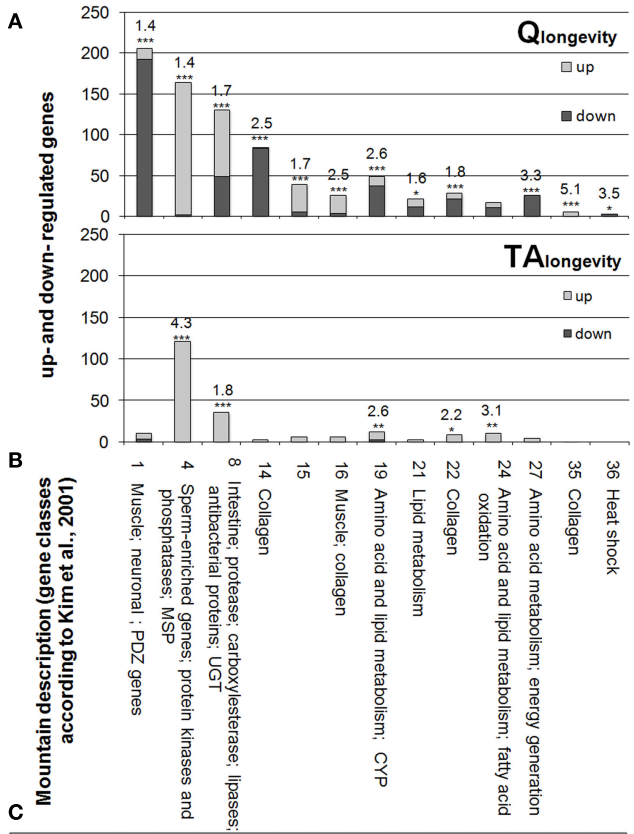
Figure 2 displays all gene expression mountains, in which DEGs of $Q_{longevity}$ and/or $TA_{longevity}$ (**Figures 2A,B**) are overrepresented

Table 1 | Overview of significant KEGG-pathways (see Table S3 in Supplementary Material 1 for pooled groups and Supplementary Material 4: sheet "KEGG-Analysis" for genes with FC-values).

	KEGG-pathway	No. DEGs/listed genes	*p*-Value
Q50	cel00330: arginine and proline metabolism	6/30	0.002
	cel00062: fatty acid elongation in mitochondria	4/13	0.007
	cel00480: glutathione metabolism	5/35	0.022
	cel00270: cysteine and methionine metabolism	4/25	0.041
	cel00980: metabolism of xenobiotics by cytochrome P450	4/26	0.045
Q100	cel04142: lysosome	15/68	0.000
	cel04350: TGF-beta signaling pathway	11/36	0.001
	cel00600: sphingolipid metabolism	8/20	0.001
	cel04310: Wnt signaling pathway	14/60	0.001
	cel00980: metabolism of xenobiotics by cytochrome P450	8/26	0.005
	cel00982: drug metabolism	8/29	0.009
	cel00330: arginine and proline metabolism	8/30	0.011
Q200	cel04350: TGF-beta signaling pathway	13/36	0.000
	cel04310: Wnt signaling pathway	15/60	0.001
	cel00982: drug metabolism	9/29	0.005
	cel00480: glutathione metabolism	9/35	0.015
	cel00330: arginine and proline metabolism	8/30	0.020
	cel04142: lysosome	14/68	0.026
	cel00980: metabolism of xenobiotics by cytochrome P450	7/26	0.033
	cel00340: histidine metabolism	3/9	0.048
$Q_{longevity}$	cel00330: arginine and proline metabolism	9/30	0.000
	cel04350: TGF-beta signaling pathway	9/36	0.001
	cel04310: Wnt signaling pathway	11/60	0.004
	cel00982: drug metabolism	7/29	0.008
	cel04142: lysosome	11/68	0.009
	cel00980: metabolism of xenobiotics by cytochrome P450	6/26	0.022
TA100	x	x	x
TA200	cel00450: selenoamino acid metabolism	6/20	0.006
	cel00270: cysteine and methionine metabolism	6/25	0.015
TA300	cel04914: progesterone-mediated oocyte maturation	15/38	0.004
	cel04120: ubiquitin-mediated proteolysis	26/86	0.006
	cel00534: heparan sulfate biosynthesis	7/8	0.006
	cel03430: mismatch repair	9/18	0.008
	cel03030: DNA replication	13/33	0.008
	cel00500: starch and sucrose metabolism	9/24	0.050
$TA_{longevity}$	x	x	x
$Q\&TA_{longevity}$	x	x	x

(additional RF values for single concentrations in all 44 gene expression mountains can be found in Supplementary Material 4: sheet "Gene Expression Mountains" and "Gene Classes and Groups"). DEGs from $Q_{longevity}$ and $TA_{longevity}$ could both be assigned to mounts 4 (sperm-enriched genes, protein kinases and phosphatases, MSPs), 8 (enriched for intestine-, protease-, carboxylesterase-, lipase-genes, antibacterial proteins and UGTs), 19 (enriched for genes from amino acid and lipid metabolism and CYPs), 22 (collagen enriched), and 24 (also enriched for amino acid and lipid metabolism genes, as well as fatty acid oxidation). In addition, $Q_{longevity}$ was significantly linked to mounts 1 (enriched for muscle, neuronal, and PDZ genes), 14 and 35 (enriched for collagen), 15 (no specific gene-groups), 21 (enriched for lipid metabolism genes), and 32 (enriched for nucleosomal histones). Comparing the direction of regulation, it was noticeable that in $TA_{longevity}$ the majority of DEGs in the respective mounts were up-regulated, whereas no trend was observed in the $Q_{longevity}$ gene list (with the exception of mountain 4).

The comparison of PP treatments to overrepresented mounts from experiments with long-lived mutants and *P. aeruginosa* infected nematodes (**Figure 2C**; Supplementary Material 4: sheet "Gene Expression Mountains") revealed large overlaps. For example, Fisher and Lithgow (2006) reported for *daf-12(rh273)* an enrichment of induced DEGs in mount 4, a finding that matches well with both PPs tested. Further differentially regulated mounts in *daf-12(rh273)* are also overrepresented in $Q_{longevity}$ and $TA_{longevity,}$ suggesting parallels in molecular mechanisms to



mountain	1	4	8	14	15	16	19	21	22	24	27	35	36
daf-12(rh273) **(Fisher and Lithgow, 2006)**													
up		x											
down			x				x		x	x	x		
daf-2 **(metaanalysis Evans et al., 2008)**													
up			x		x	x		x	x				
down			x				x	x	x	x	x		
TGFbeta adults (Shaw et al., 2007)													
up			x	x	x	x	x	x	x	x	x		
down													
P. aeruginosa **infection (metaanalysis Evans et al., 2008)**													
up			x		x		x	x		x	x		x
down			x	x		x	x	x	x	x	x		

FIGURE 2 | Overrepresented gene expression mountains. Identification of overrepresented gene expression mountains (based on Kim et al., 2001) from $Q_{longevity}$ **(A)** and $TA_{longevity}$ **(B)**, as well as from datasets of longevity mutants and *P. aeruginosa* infected *C. elegans* (Fisher and Lithgow, 2006; Shaw et al., 2007; Evans et al., 2008) **(C)**. Depicted are only RFs from mounts which were significantly affected by at least either $Q_{longevity}$ or $TA_{longevity}$. $^{*}p < 0.05$, $^{**}p < 0.005$, $^{***}p < 0.001$. (The complete dataset and RFs for all 44 gene expression mountains in each and pooled PP concentrations, as well as RFs for previously published datasets can be found in Supplementary Material 4: sheet "Gene Expression Mountains"; results can be reconstructed using the filter function in Supplementary Material 5.)

daf-12(rh273) and possibly DAF-12's involvement in Q and TA mediated longevity.

Similarly, distinct mounts of $Q_{longevity}$ are highly similar to *daf-2* (Evans et al., 2008) and TGF-beta mutants (Shaw et al., 2007), as well as *P. aeruginosa* infected nematodes (Evans et al., 2008). The pattern of overrepresented mounts in $TA_{longevity}$ best resembles the results from adult TGF-beta mutants (Shaw et al., 2007), but also overlaps to some extent with the other conditions. These marked similarities will be probed in more detail in the next section.

Overlapping DEGs from selected microarray studies in C. elegans

Figure 3 demonstrates the relationship between $Q_{longevity}$ (left section), $TA_{longevity}$ (right section), and $Q\&TA_{longevity}$ (middle section) to *daf-12(rh273)* mutants (**Figure 3A**), *daf-2* mutants (**Figure 3B**), adult TGF-beta mutants (**Figure 3C**), and *P. aeruginosa* infected nematodes, respectively. [A comparison to further microarray studies (Hill et al., 2000; Wang and Kim, 2003; Viswanathan et al., 2005) and single concentration results can be found in Supplementary Material 4: sheet "Meta-Analysis."] To gain a deep functional insight into aging-related genes, significantly overlapping DEGs were examined further with a GO- and InterPro-cluster analysis (Table S4 in Supplementary Material 1).

A comparison to daf-12(rh273). The RFs for overlapping genes (both up- and down-regulated) of $Q_{longevity}$ and *daf-12(rh273)* were significantly enriched (**Figure 3A**), suggesting that they share similar genetic mechanisms. Interestingly, eight DEGs (*bre-1*, *clec-4*, *clec-66*, *dod-22*, F35E12.5, F55G11.5, M02F4.7, *pho-1*) were down-regulated in both conditions, transcripts which were also repressed in *daf-2* and previously thought to be down-regulated in several longevity backgrounds (Fisher and Lithgow, 2006). The downstream analysis revealed one prominent cluster in each intersection (*major sperm protein* for commonly up-regulated genes and *C-type lectin-like* for commonly down-regulated genes). It is however noteworthy that a group of eight genes of the *CUB-like region* proteins were up-regulated in $Q_{longevity}$ but down-regulated in *daf-12(rh273)* mutants.

In $TA_{longevity}$, the overlap with *daf-12(rh273)* was significant for common up-regulated DEGs and also the category "up in $TA_{longevity}$/down in *daf-12*" (**Figure 3A**). Cluster analyses for these intersections returned transcripts involved in *major sperm protein* and *hydrolase activity*, respectively. Comparing DEGs of $Q\&TA_{longevity}$ and *daf-12(rh273)* revealed a strong overrepresentation of common up- and down-regulated transcripts. All up-regulated genes, 18 in total, assign to the gene-groups *msp*, cell structure, and/or sperm- and male enriched (notably 17 are members of gene expression mountain 4).

A comparison to daf-2. $Q_{longevity}$ and *daf-2* mutants share a significant proportion of DEGs (**Figure 3B**). The GO- and InterPro-cluster analyses of the intersection of up-regulated genes revealed five statistically significant clusters, including *determination of adult lifespan*, *lipid metabolic process*, and *oxidoreductase activity* (Table S4 in Supplementary Material 1). The intersection of down-regulated transcripts returned four clusters, again *determination of adult lifespan* and *UDP-glycosyltransferase activity*. In contrast, $TA_{longevity}$ differs from $Q_{longevity}$ when compared to *daf-2* mutants; only the section "up in $TA_{longevity}$/down in *daf-2*" revealed an overrepresentation (**Figure 3B**).

Since a substantial part of the category "up in $TA_{longevity}$/down in *daf-2*" are also up-regulated in TGF-beta mutants (17 genes)

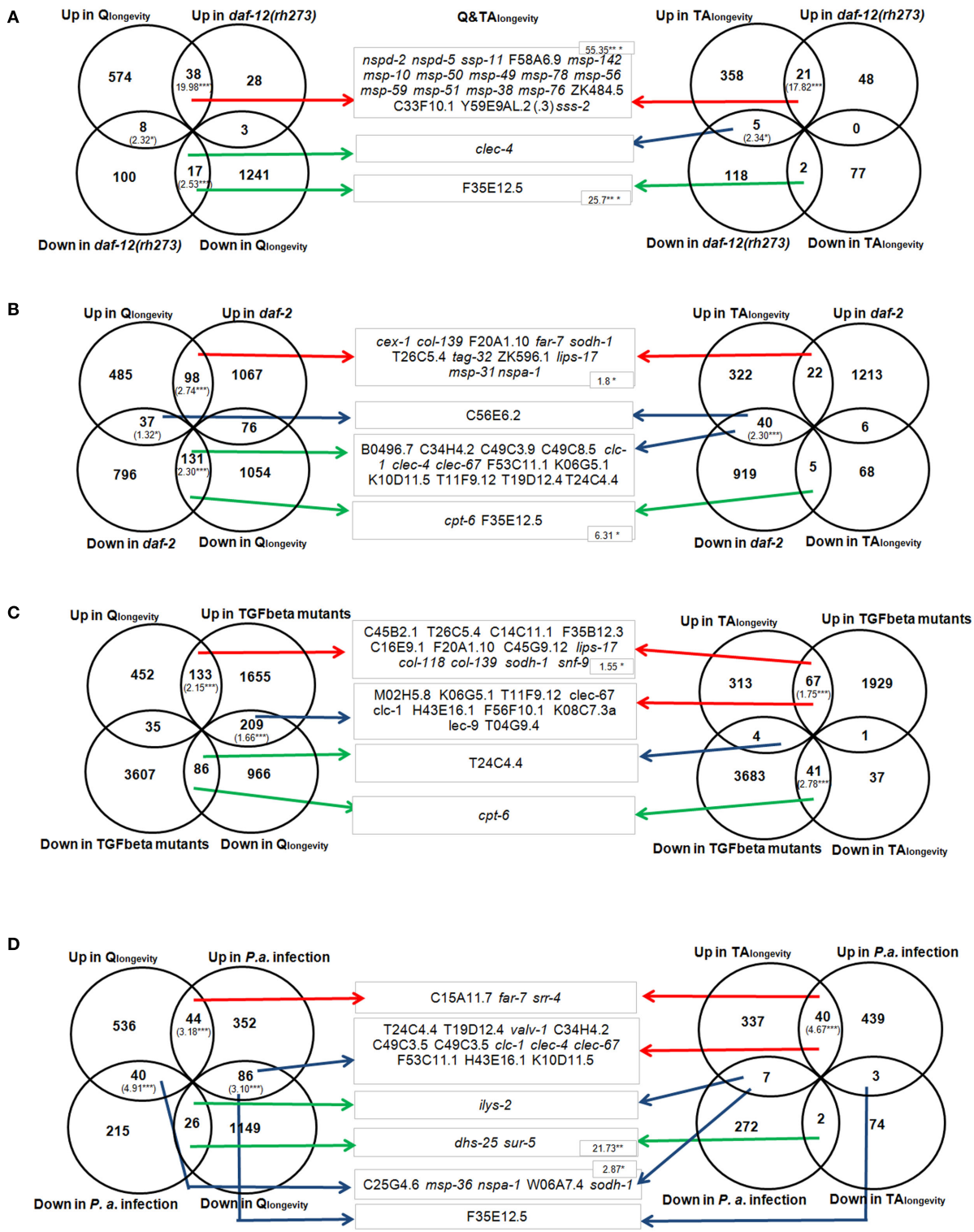


FIGURE 3 | Comparison of DEGs derived from $Q_{longevity}$, $TA_{longevity}$, and *daf-12(rh273)* mutants (Fisher and Lithgow, 2006) (A); *daf-2* mutants (Evans et al., 2008) (B); TGF-beta adults (Shaw et al., 2007) (C); and *P. aeruginosa* challenged *C. elegans* (Evans et al., 2008) (D). Significant overlap is indicated by RF values in brackets. The middle section represents the overlap of pooled DEGs in $Q\&TA_{longevity}$ and respective conditions. Arrow colors define the mode of regulation, where red defines intersections containing jointly up-regulated DEGs, green down-regulated DEGs and blue DEGs that are regulated in opposite directions. The overlap is significant where RF values (numbers in brackets) are shown. $^{*}p < 0.05$, $^{**}p < 0.005$, $^{***}p < 0.001$. (Additional information about all RF values in single PP concentrations and published data can be found in Supplementary Material 4: sheet "Meta-Analysis." A GO- and InterPro-cluster analysis of overlapping DEGs in respective intersections can be found in Table S4 in Supplementary Material 1. All results can be reconstructed using the filter function in Supplementary Material 5.)

and in *P. aeruginosa* infected nematodes (18 genes), suggests that a network of DEGs are present in all three conditions. Four clusters emerge from this intersection, notably as in $Q_{longevity}$, the term *lipid metabolic process like* stands out. The fact, that no gene of this GO-term was present in the intersection $Q_{longevity}$/*daf-2* confirms that both PPs modulate lipid metabolism, but via differing routes. Overall, $Q\&TA_{longevity}$ and *daf-2* share 11 up-regulated genes and two down-regulated genes (**Figure 3B**, middle) but 13 genes are expressed in opposite direction (two CUB-like genes (F53C11.1, K10D11.5), two belonging to the defense response category (C49C3.9, T19D12.4) and two coding for proteins with a von Willebrand factor type A (vWA) domain).

A comparison to adult TGF-beta mutants. The overlap between the gene lists for $Q_{longevity}$ and TGF-beta mutants is significant both in the up-regulated section and the section "$Q_{longevity}$ down/TGF-beta mutants up" (**Figure 3C**). The cluster analysis revealed, e.g., the terms *catalytic activity*, *active transmembrane transporter activity*, and *peptidase 28*, *carbohydrate metabolic process*, respectively (Table S4 in Supplementary Material 1). Indeed, 45 genes of the "$Q_{longevity}$ down/TGF-beta mutants up" section are also down-regulated in *daf-2*, suggesting that some overlap of TGF-beta and ILS signaling prevails, albeit with contrasting regulation patterns.

The comparison of $TA_{longevity}$ and TGF-beta mutants (**Figure 3C**) revealed that approximately 25% of the DEGs overlapped. The cluster analysis highlighted two clusters for each intersection: *peptidase 28* and *lipid metabolic process* in the section comprising induced DEGs and *chromatin organization* and *embryonic development ending in birth or egg hatching* in the group of repressed DEGs. Comparing DEGs from $Q\&TA_{longevity}$ and TGF-beta mutants (**Figure 3C**, middle) returned only 12 genes that were marginally overrepresented in the up-regulated intersection which were also up-regulated in *daf-2* mutants.

A comparison to P. aeruginosa infected nematodes. $Q_{longevity}$ or $TA_{longevity}$ were compared to the transcriptional response induced by the infection with the pathogenic bacteria *P. aeruginosa* (Shapira et al., 2006; Troemel et al., 2006; summarized in Evans et al., 2008). Venn diagrams (**Figure 3D**) illustrate the presence of significant intersections of DEGs in $Q_{longevity}$ and immune-challenged nematodes (both up; up in $Q_{longevity}$/down in infection; both down; up in infection/down in $Q_{longevity}$). The cluster analysis revealed a common up-regulation in, for example, *lipid metabolic processes*, *monooxygenase activity*, and genes coding for *UDP-glucuronosyl/UDP-glucosyltransferase* as well as a concurrent down-regulation in particular parts of *catalytic activity*. The intersections of oppositely responding DEGs contain genes coding for *oxidoreductases* and *serine-type peptidase activity* ($Q_{longevity}$ up/infection down) and for structural constitutes of cuticle ($Q_{longevity}$ down/infection up), respectively. A large proportion of DEGs from the intersection "up in infection/down in $Q_{longevity}$" is down-regulated in TGF-beta, *daf-2* and *daf-12(rh273)* (for exact numbers and genes see Supplementary Material 5 in respective filter condition), and 21 genes of "up in $Q_{longevity}$/down in infection" are up-regulated in TGF-beta mutants.

$TA_{longevity}$ shares highly significant patterns with the list of up-regulated DEGs from infected nematodes. The cluster analysis returned only one term, genes with a *CUB-like region*, however, interestingly, CUB-like genes were previously found to be activated in the induced immune response pathway modulated by PMK-1 (Troemel et al., 2006).

Only five genes are regulated in $Q\&TA_{longevity}$ as well as in nematodes infected with *P. aeruginosa* (up-regulated: C15A11.7, *far-7* and *srr-4*; down-regulated: *dhs-25*, *sur-5*). Of the 15 common up-regulated DEGs in "$TA_{longevity}$ and *P. aeruginosa* infection and down-regulated in $Q_{longevity}$," 11 are also down-regulated in *daf-2* mutants, including two C-type lectins (*clec-4*, *clec-67*), two genes coding for proteins with a CUB-like region (F53C11.1, K10D11.5) and two defense response genes (C49C3.9, T19D12.4).

Finally, we evaluated the overlap of DEGs in Q and TA treated nematodes with the DEGs regulated by PMK-1 (Troemel et al., 2006; Supplementary Material 4: sheet "Meta-Analysis") and found a significant overlap for $Q_{longevity}$ and $TA_{longevity}$ in the up-regulated intersection and the "down-regulated by PP/up due to PMK-1." These results underline the involvement of the p38 MAP kinase pathway in both PP actions.

DISCUSSION

Polyphenols transform the transcriptional output and thereby alter the physiological status of an animal. By performing DNA microarray experiments and comparing the resulting data with previously published gene expression screens, we revealed an interlaced interplay of genetic pathways affected by Q and TA. Moreover, our findings support results from previous studies (Pietsch et al., 2009, 2011; Saul et al., 2010).

CONCENTRATION DEPENDENT VARIATIONS IN GLOBAL TRANSCRIPTIONAL RESPONSES

Low concentrations of Quercetin (Q50) modulate processes which may contribute to the wellbeing of the nematode (e.g. the gene classes *gst*s, peroxidases, lysozymes; GO-terms: *oxidoreductase activity*; KEGG-analysis: *glutathione metabolism* and *metabolism of xenobiotics by cytochrome P450*), but not sufficiently to significantly extend lifespan (Pietsch et al., 2011). In contrast, Q100 and Q200 impact significantly on the global transcriptome which manifests itself in a strong longevity phenotype. This effect is, at least in part, driven by transcriptional repression mechanisms as indicated by the high number of down-regulated genes. As shown recently, increasing the concentration to Q250 reverts the beneficial effects on longevity (Pietsch et al., 2011). High dose toxicity is particularly apparent at, for example, TA300 (GO-terms: *DNA repair*, *DNA damage response*, *response to stress*, and *cell death* and KEGG-analysis: *Ubiquitin-mediated proteolysis*, *mismatch repair*, *DNA replication)*.

GENETIC BASIS OF $Q_{LONGEVITY}$ AND $TA_{LONGEVITY}$

$Q_{longevity}$ shares transcriptional patterns with long-lived *daf-12(rh273)*, *daf-2* mutants, TGF-beta mutants, as well as *P. aeruginosa* challenged nematodes (as identified by the meta-analysis and overlapping gene expression mountains). These results suggest that Q operates through a complex network of interlinked pathways. Combining the microarray analysis with lifespan data from Q exposed nematode mutants (Pietsch et al., 2009) allowed

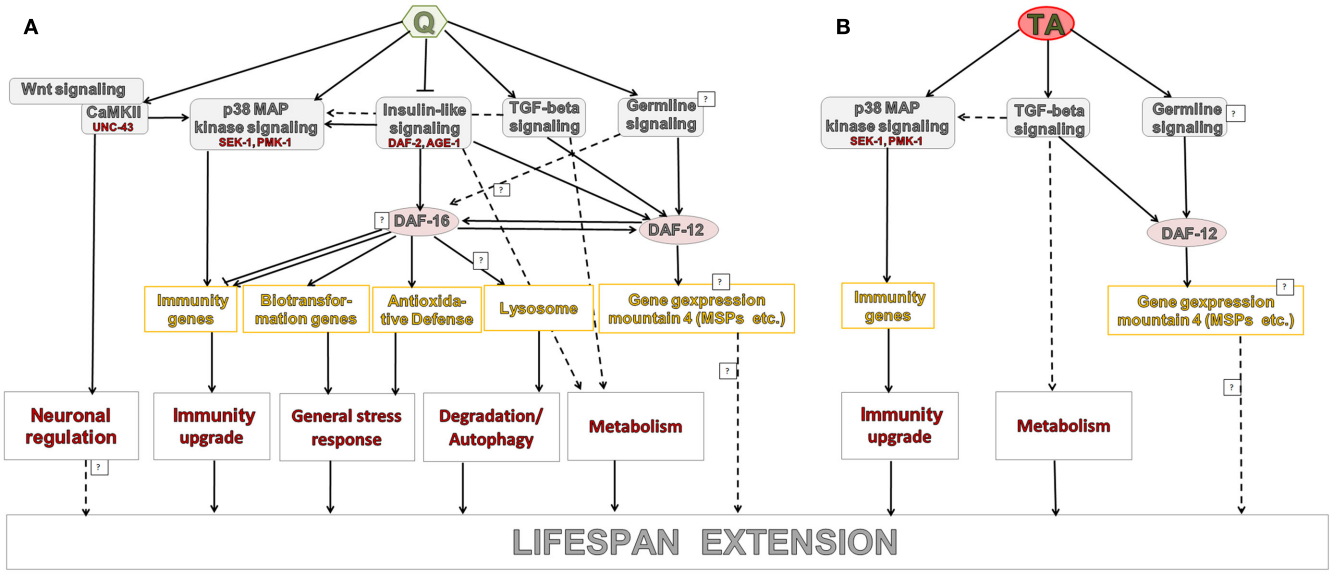


FIGURE 4 | Hypothetical model of Q's (A) and TA's (B) mode of action.

the construction of a hypothetical model that describes the mode of action of Q induced longevity (**Figure 4A**).

The overlap between $TA_{longevity}$ and the conditions assessed as part of the meta-analysis revealed a less pronounced, yet still significant correlation. Beside activation of amino acid metabolism pathways, the strongest overlap was with data from the TGF-beta mutants. This suggests that the TGF-beta pathway plays a prominent role in the TA mediated life extension which is summarized in **Figure 4B**.

It is intriguing to note that TA treatment has, compared to Q treatment, a lesser effect on the transcriptome (as defined by the number of significant DEGs) but a more marked life-extending property. It is currently not known if this is an indirect effect or whether transcriptional changes diminish/suppress the positive output of Q action. Notwithstanding the observed differences in genetic action modes, TGF-beta and p38 MAPK pathways, as well as the nuclear hormone receptor DAF-12 seem to be involved in both PPs. DAF-12, which is downstream of the TGF-beta- and ILS pathways, is a member of the steroid hormone receptor superfamily. It is linked to Dauer formation and, together with DAF-16, also influences gonad-dependent adult longevity. Given that neither Blueberry polyphenols (Wilson et al., 2006), Caffeic acid nor Rosmarinic acid (Pietsch et al., 2011) could extend the lifespan of *sek-1* mutants (*sek-1* is a genetic player in the p38 MAPK pathway) suggests that the innate immunity may act as a prominent downstream effector of PPs.

Based on the meta-analysis with $Q\&TA_{longevity}$ and relevant microarray studies (Hill et al., 2000; Wang and Kim, 2003; Viswanathan et al., 2005; Fisher and Lithgow, 2006; Shaw et al., 2007; Evans et al., 2008), we were able to identify a subset of transcripts that are possibly relevant for aging (see Table S5 in Supplementary Material 1). Clearly further investigations into their aging-modulating activities are warranted. Likewise, we call for further experimentation to establish the regulatory interlink between TGF-beta signaling and DAF-12, as well as p38 MAPK.

CONCLUSION

The meta-analysis displayed an extensive overlap between PP treatment and numerous mutants as well as immunity challenging conditions, however, frequently the mode of regulation was in opposite direction. This strongly suggests the presence of a complex regulatory interplay between the input and multiple downstream targets. Clearly, PPs action cannot be reduced to the activation or inhibition of single genes and pathways; nevertheless it is apparent that TGF-beta, ILS, and p38 MAPK play a prominent role in PPs' mode of action. Furthermore, we were able to demonstrate that an extensive comparison with data from the literature can provide a deep insight into the transcriptome to a level that goes beyond a simple GO- and KEGG-analysis. Given the convolution observed with single gene knockout alleles and exposures to pure compounds, one can envisage the complexity that will arise with multidimensional mixture toxicity experiments. The development of comparative databases and most importantly powerful, yet intuitive, bioinformatic tools will undoubtedly aid in the streamlining of large datasets. Overall, our results strengthen the notion that both PPs act in conserved genetic pathways that overlap, or at least correlate, with the longevity phenotypes and transcriptional fingerprints of certain mutant strains. Clearly, further future tests are needed to confirm single genetic players and specify the interplay of conserved pathways.

ACKNOWLEDGMENTS

We would like to thank members of the King's College Genomics Center (Dr. Matthew Arno, Dr. Estibaliz Aldecoa-Otalora Astarloa) for assistance in the microarray experimentation. Furthermore, we thank the Caenorhabditis Genetics Centre, which is funded by the National Institutes of Health National Centre for Research Resources, for the supply of the *Caenorhabditis elegans* strains. This work was supported, in part, by a grant awarded by

the Biotechnology and Biological Sciences Research Council (BBSRC grant BB/E025099 and a BBSRC Underwood Fellowship). Ralph Menzel was supported by the DFG grant STE673/16-1.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Toxicogenomics_/10.3389/fgene.2012.00048/abstract

REFERENCES

Boots, A. W., Haenen, G. R. M. M., and Bast, A. (2008). Health effects of quercetin: from antioxidant to nutraceutical. *Eur. J. Pharmacol.* 585, 325–337.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71–94.

Calabrese, E. J., and Baldwin, L. A. (2001). Hormesis: U-shaped dose responses and their centrality in toxicology. *Trends Pharmacol. Sci.* 22, 285–291.

Evans, E. A., Kawli, T., and Tan, M. W. (2008). *Pseudomonas aeruginosa* suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in *Caenorhabditis elegans*. *PLoS Pathog.* 4, e1000175. doi:10.1371/journal.ppat.1000175

Fisher, A. L., and Lithgow, G. J. (2006). The nuclear hormone receptor DAF-12 has opposing effects on *Caenorhabditis elegans* lifespan and regulates genes repressed in multiple long-lived worms. *Aging Cell* 5, 127–138.

Gami, M. S., and Wolkow, C. A. (2006). Studies of *Caenorhabditis elegans* DAF-2/insulin signaling reveal targets for pharmacological manipulation of lifespan. *Aging Cell* 5, 31–37.

Gill, M. S. (2006). Endocrine targets for pharmacological intervention in ageing in *Caenorhabditis elegans*. *Aging Cell* 5, 23–30.

Golden, T. R., and Melov, S. (2007). "Gene expression changes associated with aging in *C. elegans*," in *WormBook*, ed. The *C. elegans* Research Community (WormBook). Available at: http://www.wormbook.org

Grigoryev, D. N., Ma, S.-F., Irizarry, R. A., Ye, S. Q., Quackenbush, J., and Garcia, J. G. N. (2004). Orthologous gene-expression profiling in multi-species models: search for candidate genes. *Genome Biol.* 5, R34.

Grünz, G., Haas, K., Soukup, S., Klingenspor, M., Kulling, S. E., Daniel, H., and Spanier, B. (2012). Structural features and bioavailability of four flavonoids and their implications for lifespan-extending and antioxidant actions in *C. elegans*. *Mech. Ageing Dev.* 133, 1–10.

Hill, A. A., Hunter, C. P., Tsung, B. T., Tucker-Kellogg, G., and Brown, E. L. (2000). Genomic analysis of gene expression in *C. elegans*. *Science* 290, 809–812.

Kaletta, T., and Hengartner, M. O. (2006). Finding function in novel targets: *C. elegans* as a model organism. *Nat. Rev. Drug Discov.* 5, 387–399.

Kampkötter, A., Nkwonkam, C. G., Zurawski, R. F., Timpel, C., Chovolou, Y., Wätjen, W., and Kahl, R. (2007a). Effects of the flavonoids kaempferol and fisetin on thermotolerance, oxidative stress and FoxO transcription factor DAF-16 in the model organism *Caenorhabditis elegans*. *Arch. Toxicol.* 81, 849–858.

Kampkötter, A., Nkwonkam, C. G., Zurawski, R. F., Timpel, C., Chovolou, Y., Wätjen, W., and Kahl, R. (2007b). Investigations of protective effects of the flavonoids quercetin and rutin on stress resistance in the model organism *Caenorhabditis elegans*. *Toxicology* 234, 113–123.

Kampkötter, A., Timpel, C., Zurawski, R. F., Ruhl, S., Chovolou, Y., Proksch, P., and Wätjen, W. (2008). Increase of stress resistance and lifespan of *Caenorhabditis elegans* by quercetin. *Comp. Biochem. Physiol. B* 149, 314–323.

Kim, S. K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J. M., Eizinger, A., Wylie, B. N., and Davidson, G. S. (2001). A gene expression map for *Caenorhabditis elegans*. *Science* 293, 2087–2092.

Kirkwood, T. B. (1977). Evolution of ageing. *Nature* 270, 301–304.

Kirkwood, T. B. (1988). The nature and causes of ageing. *Ciba Found. Symp.* 134, 193–207.

Koleckar, V., Kubikova, K., Rehakova, Z., Kuca, K., Jun, D., Jahodar, L., and Opletal, L. (2008). Condensed and hydrolysable tannins as antioxidants influencing the health. *Mini Rev. Med. Chem.* 8, 436–447.

Lewis, J. A., and Fleming, J. T. (1995). "Basic culture methods," in *Caenorhabditis elegans, Modern Biological Analysis of an Organism*, eds H. F. Epstein and D. C. Shakes (San Diego, CA: Academic Press), 3–29.

Lublin, A., Isoda, F., Patel, H., Yen, K., Nguyen, L., Hajje, D., Schwartz, M., and Mobbs, C. (2011). FDA-approved drugs that protect mammalian neurons from glucose toxicity slow aging dependent on cbp and protect against proteotoxicity. *PLoS ONE* 6, e27762. doi:10.1371/journal.pone.0027762

McCarthy, D. J., and Smyth, G. (2009). Testing significance relative to a fold-change threshold. *Bioinformatics* 25, 765–771.

McElwee, J. J., Schuster, E., Blanc, E., Thomas, J. H., and Gems, D. (2004). Shared transcriptional signature in *Caenorhabditis elegans* Dauer larvae and long-lived daf-2 mutants implicates detoxification system in longevity assurance. *J. Biol. Chem.* 279, 44533–44543.

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., and Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* 424, 277–283.

Pietsch, K., Saul, N., Chakrabarti, S., Stürzenbaum, S. R., Menzel, R., and Steinberg, C. E. W. (2011). Hormetins, antioxidants and prooxidants: defining quercetin-, caffeic acid- and rosmarinic acid- mediated life extension in *C. elegans*. *Biogerontology* 12, 329–347.

Pietsch, K., Saul, N., Menzel, R., Stürzenbaum, S. R., and Steinberg, C. E. W. (2009). Quercetin mediated lifespan extension in *Caenorhabditis elegans* is modulated by age-1, daf-2, sek-1, and unc-43. *Biogerontology* 10, 565–578.

Rainer, J., Sanchez-Cabo, F., Stocker, G., Sturn, A., and Trajanoski, Z. (2006). CARMAweb: comprehensive R- and bioconductor-based web service for microarray data analysis. *Nucleic Acids Res.* 34, W498–W503.

Saeed, A. I, Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003). MeV/TM4:a free, open-source system for microarray data management and analysis. *BioTechniques* 34, 374–378.

Saul, N., Pietsch, K., Menzel, R., and Steinberg, C. E. W. (2008). Quercetin-mediated longevity in *Caenorhabditis elegans*: is DAF-16 involved? *Mech. Ageing Dev.* 129, 611–613.

Saul, N., Pietsch, K., Menzel, R., Stürzenbaum, S. R., and Steinberg, C. E. W. (2010). The longevity effect of tannic acid in *Caenorhabditis elegans*: disposable soma meets hormesis. *J. Gerontol. A Biol. Sci. Med. Sci.* 65, 626–635.

Saul, N., Pietsch, K., Stürzenbaum, S. R., Menzel, R., and Steinberg, C. E. W. (2011). Diversity of polyphenol action in *Caenorhabditis elegans*: between toxicity and longevity. *J. Nat. Prod.* 74, 1713–1720.

Shapira, M., Hamlin, B. J., Rong, J., Chen, K., Ronen, M., and Tan, M. W. (2006). A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14086–14091.

Shaw, W. M., Luo, S., Landis, J., Ashraf, J., and Murphy, C. T. (2007). The *C. elegans* TGF-beta Dauer pathway regulates longevity via insulin signaling. *Curr. Biol.* 17, 1635–1645.

Strange, K., Christensen, M., and Morrison, R. (2007). Primary culture of *Caenorhabditis elegans* developing embryo cells for electrophysiological, cell biological and molecular studies. *Nat. Protoc.* 2, 1003–1012.

Sulston, J. E., and Hodgkin, J. (1988). "Methods," in *The Nematode Caenorhabditis elegans*, ed. W. B. Wood (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), 587–606.

Surco-Laos, F., Cabello, J., Gómez-Orte, E., González-Manzano, S., González-Paramás, A. M., Santos-Buelga, C., and Dueñas, M. (2011). Effects of O-methylated metabolites of quercetin on oxidative stress, thermotolerance, lifespan and bioavailability on *Caenorhabditis elegans*. *Food Funct.* 2, 445–456.

Tárraga, J., Medina, I., Carbonell, J., Huerta-Cepas, J., Minguez, P., Alloza, E., Al-Shahrour, F., Vegas-Azcárate, S., Goetz, S., Escobar, P., Garcia-Garcia, F., Conesa, A., Montaner, D., and Dopazo, J. (2008). GEPAS, a web-based tool for microarray data analysis and interpretation. *Nucleic Acids Res.* 36, W308–W314.

The *C. elegans* Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018.

Troemel, E. R., Chu, S. W., Reinke, V., Lee, S. S., Ausubel, F. M., and Kim, D. H. (2006). p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genet.* 2, e183. doi:10.1371/journal.pgen.0020183

Van Gelder, R. N., von Zastrow, M. E., Yool, A., Dement, W. C., Barchas, J. D., and Eberwine, J. H. (1990). Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 87, 1663–1667.

Viswanathan, M., Kim, S. K., Berdichevsky, A., and Guarente, L. (2005). A role for SIR-2.1 regulation of ER stress response genes in determining *C. elegans* lifespan. *Dev. Cell* 5, 605–615.

Wang, J., and Kim, S. K. (2003). Global analysis of Dauer gene expression in *Caenorhabditis elegans*. *Development* 130, 1621–1634.

Wilson, M. A., Shukitt-Hale, B., Kalt, W., Ingram, D. K., Joseph, J. A., and Wolkow, C. A. (2006). Blueberry polyphenols increase lifespan and thermotolerance in *Caenorhabditis elegans*. *Aging Cell* 5, 59–68.

Zeeberg, B. R., Qin, H., Narasimhan, S., Sunshine, M., Cao, H., Kane, D. W., Reimers, M., Stephens, R. M., Bryant, D., Burt, S. K., Elnekave, E., Hari, D. M., Wynn, T. A., Cunningham-Rundles, C., Stewart, D. M., Nelson, D., and Weinstein, J. N. (2005). High-throughput gominer, an "industrial-strength" integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of common variable immune deficiency (CVID). *BMC Bioinformatics* 6, 168. doi:10.1186/1471-2105-6-168

Conflict of Interest Statement: 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.

Received: 27 January 2012; paper pending published: 09 February 2012; accepted: 16 March 2012; published online: 05 April 2012.

Citation: Pietsch K, Saul N, Swain SC, Menzel R, Steinberg CEW and Stürzenbaum SR (2012) Meta-analysis of global transcriptomics suggests that conserved genetic pathways are responsible for Quercetin and Tannic acid mediated longevity in C. elegans. Front. Gene. ***3****:48. doi: 10.3389/fgene.2012.00048*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in
GENETICS

ORIGINAL RESEARCH ARTICLE
published: 12 April 2012
doi: 10.3389/fgene.2012.00050

The nematode *Caenorhabditis elegans*, stress and aging: Identifying the complex interplay of genetic pathways following the treatment with humic substances

Ralph Menzel[1]*, Stefanie Menzel[1,2], Suresh C. Swain[3†], Kerstin Pietsch[1], Sophie Tiedt[1], Jördis Witczak[1], Stephen R. Stürzenbaum[3] and Christian E. W. Steinberg[1]

[1] Laboratory of Freshwater and Stress Ecology, Department of Biology, Humboldt-Universität zu Berlin, Berlin, Germany
[2] Institute of Biochemistry, Freie Universität Berlin, Berlin, Germany
[3] Analytical and Environmental Science Division, School of Biomedical Sciences, King's College London, London, UK

Edited by:
Jonathan Freedman, National Institute of Environmental Health Science, National Institute of Health, USA

Reviewed by:
Tara Sabo-Attwood, University of Florida, USA
Wyatt McMahon, Virginia Bioinformatics Institute, USA

***Correspondence:**
Ralph Menzel, Laboratory of Freshwater and Stress Ecology, Department of Biology, Humboldt-Universität zu Berlin, Späthstr. 80/81, 12437 Berlin, Germany.
e-mail: ralph.menzel@biologie.hu-berlin.de

†**Present address:**
Suresh C. Swain, Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Penglais, Aberystwyth, Ceredigion SY23 3DA, UK.

Low concentrations of the dissolved leonardite humic acid HuminFeed® (HF) prolonged the lifespan and enhanced the thermal stress resistance of the model organism *Caenorhabditis elegans*. However, growth was impaired and reproduction delayed, effects which have also been identified in response to other polyphenolic monomers, including Tannic acid, Rosmarinic acid, and Caffeic acid. Moreover, a chemical modification of HF, which increases its phenolic/quinonoid moieties, magnified the biological impact on *C. elegans*. To gain a deep insight into the molecular basis of these effects, we performed global transcriptomics on young adult (3 days) and old adult (11 days) nematodes exposed to two different concentrations of HF. We also studied several *C. elegans* mutant strains in respect to HF derived longevity and compared all results with data obtained for the chemically modified HF. The gene expression pattern of young HF-treated nematodes displayed a significant overlap to other conditions known to provoke longevity, including various plant polyphenol monomers. Besides the regulation of parts of the metabolism, transforming growth factor-beta signaling, and Insulin-like signaling, lysosomal activities seem to contribute most to HF's and modified HF's lifespan prolonging action. These results support the notion that the phenolic/quinonoid moieties of humic substances are major building blocks that drive the physiological effects observed in *C. elegans*.

Keywords: humic substances, hydroxybenzene, gene expression, aging, longevity, stress, TGF-beta, *Caenorhabditis elegans*

INTRODUCTION

Several studies with different model organisms have demonstrated that mild chemical stress trains cellular stress response pathways, e.g., biotransformation and antioxidant systems, which can ultimately result in lifespan extension; for a review see Kourtis and Tavernarakis (2011). Recently, we were able to show that exposure to a specific humic substance preparation, HuminFeed® (HF), significantly extends the lifespan of the nematode *Caenorhabditis elegans* (Steinberg et al., 2007). HF is weathered leonardite humic material characterized by high functional group content (Meinelt et al., 2007). By analogy it has been concluded that the effective building blocks may be hydroxybenzene groups. To confirm the biological impact of these structures, HF was chemically modified by increasing the concentrations of phenolic and quinonoid functional groups (Menzel et al., 2011). This chemical modification boosted the antioxidant properties of HF both *in vitro* and *in vivo*. Moreover, modified HF caused a significantly increased tolerance toward thermal stress in *C. elegans* and extended its lifespan (Menzel et al., 2011). In contrast, HF and the modified substances delayed the onset of reproduction and caused a reduction in overall body length. The underlying molecular basis of these HF mediated effects is, to date, unknown.

To define the transcriptional responses of HF exposure, we conducted global gene expression analyses using the Affymetrix® whole genome DNA microarray platform. Nematodes were exposed to two different concentrations of HF over a 3- or 11-days incubation period. We also assessed the effect of Huminfeed–Hydroquinone (HF-HQ), a HF derivate chemically enriched with hydroquinone.

Initially, we defined the differently expressed genes (DEGs), many of which displayed concentration dependent changes in expression. Selected results were confirmed by quantitative real-time RT-PCR. Subsequent investigations included gene ontology (GO; Ashburner et al., 2000) and Kyoto encyclopedia of genes and genomes (KEGG; Kanehisa, 2002) pathway analyses. Moreover, over-represented gene expression mountains and gene classes were evaluated according to Kim et al. (2001). In doing so, we were able to identify gene classes and pathways that returned a significant over-representation in HF or HF-HQ treated nematodes. A meta-analysis compared our findings with recently published data specific to either age-related gene expression, the genetic background of longevity mutants, the transcriptional profile of polyphenol treated nematodes or infection/immunity-related gene expression. This allowed us to pinpoint genes and

associated pathways predicted to be key players in HF mediated longevity.

To substantiate the importance of these genes and pathways, loss of function mutants were tested for their ability to extend the lifespan in response to HF or HF-HQ exposure. In summary, this study provides new evidence that specific humic substances induce a complex mode of action. Moreover, humic substances are not limited (as previously thought) to act indirectly, e.g., via the unspecific binding to organic and inorganic compounds or the shuttling of electrons in microbial redox reactions, but rather extends the lifespan of *C. elegans* by means of regulatory and stress response pathways.

MATERIALS AND METHODS

NEMATODES

Maintenance of large synchronous cultures of old nematodes is challenging due to the offspring generated during the onset of reproductive output. Rather than using fluorodeoxyuridine to inhibit embryonic development, which has recently been shown to affect the worm (Aitlhadj and Stürzenbaum, 2010; Davies et al., 2012), this study utilized the *C. elegans* mutant strain GE24, *pha-1(e2123)*, a putative transcriptional regulator of the pharyngeal precursor cells (Granato et al., 1994a). The mutant allele *pha-1(e2123)* is temperature sensitive; reproduction resembles wild type at 15°C, but is 100% embryonic lethal at 25°C. At the restrictive temperature, pharyngeal tissues of mutant embryos fail to undergo terminal differentiation and morphogenesis. After passing embryogenesis at the permissive temperature, however, a temperature shift does not affect pharyngeal functionality. Previously introduced as a selectable genetic marker (Granato et al., 1994b), we used *pha-1(e2123)* to maintain and follow a bulk preparation of synchronized nematodes. For reasons of comparison, we used the wild type strain Bristol N2 for all qRT-PCR experiments.

The lifespan assay included, besides N2 and *pha-1(e2123)*, the following mutant strains: *asah-1(tm495)*; RB1855, *cyp-34A9(ok2401)*; DA465, *eat-2(ad465)*; TK22, *mev-1(kn1)*; AM1, *osr-1(rm1)*; AU1, *sek-1(ag1)*; VC199, *sir-2.1(ok434)*; and MT2605, *unc-43(n498n1186)*. All nematode strains were maintained on nematode growth medium (NGM) plates using *Escherichia coli* OP50 as food source according to standard procedures (Brenner, 1974; Sulston and Hodgkin, 1988).

HUMIC MATERIALS

HuminFeed® (HF; Humintech GmbH, Düsseldorf, Germany) was made by an alkaline extraction process of highly oxidized lignite (for a detailed and comparative physicochemical analysis see Meinelt et al., 2007). Our experiments used the same HF batch as previously chemically characterized. Moreover, we utilized a formaldehyde polycondensation product between HF and hydroquinone, namely HF-HQ, as described in Menzel et al. (2011). HF was used solely for practical reasons; it does not constitute an advertisement for this product.

CULTIVATION FOR THE GENE EXPRESSION SCREEN

Untreated nematodes were chunked onto control plates (no HF) and treatment plates [0.2 and 2.0 mM dissolved organic carbon (DOC) of HF and HF-HQ, respectively] and incubated at 15°C for 5 days. Then, a synchronous culture was generated by filtering worms through a 10-μm membrane (SM 16510/11, Sartorius, Germany), a pore size that retains all but first stage juveniles (L1). For each individual experiment, 15,000 larvae were distributed to three freshly prepared plates (∅ = 94 mm) and cultivated at 25°C to the young adult stage (3 days) or an older adult stage (11 days). Nematodes were fed every third day by replenishing the bacterial suspension and supplemented with a fresh preparation of humic substances. Following the respective exposures, nematodes were harvested by rinsing off with M9 buffer, rewashed twice, frozen in liquid nitrogen, and stored at −80°C until use. Each condition was cultivated in triplicate.

RNA PREPARATION

Total RNA of each individual condition ($n = 3$) was isolated using an innuSPEED Tissue RNA Kit (AnalytikJena, Jena, Germany), which included an improved homogenization step with a SpeedMill (AnalytikJena, Jena, Germany) and the removal of genomic DNA through an initial spin filter column step. The RNA quality and quantity was analyzed both spectroscopically (NanoDrop 1000, ThermoScientific, UK) and by means of Agilent's Bioanalyzer 2100 equipped with a RNA 6000 Nano kit (Agilent Technologies, Santa Clara, CA, USA). All RNA-samples showed no signs of degradation as indicated by Agilent's RNA integrity numbers of 9 or 10.

DNA MICROARRAYS

Procedure

The processing of each RNA sample, the first- and second-strand cDNA synthesis as well as cRNA synthesis, labeling, and fragmentation was performed with a MessageAmp™Premier RNA Amplification Kit (Ambion, Austin, TX, USA). We used the GeneChip® *C. elegans* genome platform (Affymetrix, Santa Clara, CA, USA), representing 22,548 different transcripts. To conduct the microarray hybridization experiments, we followed the specifications from Affymetrix's GeneChip® hybridization, wash, and stain kit. The automated washing steps were performed in a GeneChip® fluidics station 450 (Affymetrix), and scans conducted by means of a GeneChip® scanner 3000 7G (Affymetrix). Triplicate GeneChips® were run for each condition.

Data interpretation and statistical analysis

Pre-processing of DNA microarray raw data included probe-specific background correction, summarization of probe set values, and normalization using the GCRMA algorithm with CARMAweb 1.4, an R- and Bioconductor-based web service for microarray data analysis (Rainer et al., 2006)[1]. Then, the data were initially filtered for missing values and subjected to a CLEAR-test that combines differential expression and variability using the GEPAS software (Herrero et al., 2003)[2]. For selection of DEGs, an unpaired *t*-test was performed followed by a significance analysis of microarray (SAM) test including a calculation that estimates the false discovery rate (FDR). FDR, reducing on

[1] https://carmaweb.genome.tugraz.at/carma/
[2] http://www.gepas.org

the one hand type I errors for null associations, was set to a non-stringent level of <12.5%, mainly to guard from an increase of type II error (Swain et al., 2010) and also based on findings by Levine et al. (2011), which described 12.5% as most acceptable optimum level of FDR, representing the 90th percentile of the normal distribution curve. DEGs exceeding a fold change of 1.25 were further analyzed with respect to their functional clustering. We chose this fold-cut-off to allow an interpretation that is biologically meaningful, akin to the notion that data of sound technical and experimental quality which returns strong, statistically significant, absolute signal intensities is sufficiently robust to justify a fold-cut-off of >1.2 (Grigoryev et al., 2004; McCarthy and Smyth, 2009). This analysis was conducted using the functional annotation clustering tool of the Database for Annotation, Visualization, and Integrated Discovery (DAVID; Huang et al., 2007)[3]. This tool identified annotation categories including, e.g., GO terms and bio-pathways that are significantly enriched within the gene list, followed by a multiple sample correction (Benjamini and Hochberg, 1995). The resultant annotation clusters were ranked according to the statistical significance of cluster enrichment.

Representation factor

To assess the level of overlap between different conditions we calculated the representation factor (RF) in order to explore the fold enrichment. The RF identifies the level of enrichment (of individual transcripts) between gene lists (Kim et al., 2001; Evans et al., 2008). The choice of N(genome) was based on the values recommended by the authors. Intersection *P*-values were calculated from the hypergeometric distribution. RF were considered significant when RF > 1.

VALIDATION OF DNA MICROARRAY DATA BY qRT-PCR

qRT-PCR analyses were conducted with samples from control and 2.0 mM DOC HF as well as HF-HQ exposed N2 wild type nematodes. The cultivation conditions of N2 wild type and *pha-1(e2123)* were identical. β-Actin (*act-1*) was used as reference gene, which did not change significantly in the DNA microarray data. A total of 1 μg RNA was reverse transcribed into cDNA (Menzel et al., 2005). Quantitative real-time amplification was performed in a MyiQ single color qPCR detection system (BIO-RAD, Germany) using the double-stranded DNA intercalating fluorescent agent EvaGreen for amplicon detection. Each reaction consisted of the qPCR Green Core Kit (Jena Bioscience, Germany), 200 nM of each primer pair, and cDNA template equivalent to 5 ng RNA starting material. The relative expression of the target genes was calculated by means of the comparative $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). All experiments for each selected gene were performed in duplicate; RT-negatives were also run for each sample and gene to confirm the absence of DNA contamination. The list of primers with their corresponding PCR-efficiencies (91–100%) is given in Table S1 in Supplementary Material; at least one primer of each pair spanned an intron to avoid amplification of genomic DNA.

LIFESPAN ASSAY AND STATISTICAL EVALUATION

The lifespan of *C. elegans* was investigated as previously described (Pietsch et al., 2009) using synchronized L4 larvae and a growth temperature of 20°C. However, *pha-1(e2123)* was pre-cultured at 15°C until the L4 state, and then maintained at 20°C. The concentrations of HF and HF-HQ were 0 and 0.4 mM DOC, respectively, mixed both into the agar and to the bacterial lawn. The first day of adulthood was defined as day 1. We performed three independent trials, each comprising 10 small agar plates (∅ = 35 mm) and 150 nematodes per trial. Animals were scored daily for survival until all worms had died. Median and mean lifespan and percentage changes (compared to controls) were determined. The statistical significance of alterations in the mean lifespan was calculated using the log-rank test (Azen et al., 1977), available online from the Bioinformatics group at the Walter and Eliza Hall Institute of Medical Research (Melbourne, Australia)[4]. Blinding of studies was not possible due to the color of humic material, which also stains the NGM agar.

RESULTS

TRANSCRIPT PROFILING BY WHOLE GENOME MICROARRAY FOLLOWING HF AND HF-HQ TREATMENT

The Venn diagrams in **Figure 1** present an overview of the number of genes that were significantly up- or down-regulated in response to the humic substances and the respective overlap between the two concentrations, 0.2 and 2.0 mM DOC. Due to the low threshold (a minimum fold change in gene expression of 1.25), many genes were classed as DEGs, especially in young adults exposed to HF. However, the extended incubation time of 11 days was characterized by a substantial decline in the number of DEGs, in particular down-regulated genes. The HF-HQ derived data resemble the results from young adults exposed to HF, albeit overall less DEGs were identified. The intersection between HF-HQ and HF at 0.2 and 2.0 mM DOC comprised of 174 and 532 DEGs, respectively. An extensive overlap was observed between the two concentrations per HF condition (**Figure 1**). The complete data can be viewed in the National Center for Biotechnology Information's (NCBI) Gene Expression Omnibus (GEO) database (accession number GSE35360) and are also given in Table S2 in Supplementary Material, including expression values, statistics, and gene annotations.

VALIDATION OF TRANSCRIPT PROFILES OF SELECTED GENES AND CONDITIONS BY qRT-PCR

A validation of the microarray experiment was deemed to be important to (i) allow a comparison between the *pha-1(e2123)* strain (used for microarray experiments) and the N2 wild type (used for qRT-PCR), and (ii) evaluate the expression levels at reduced exposure times (24 and 48 h as well as 72 h). Samples generated for qRT-PCR originated from worms exposed to 2.0 mM DOC of HF or HF-HQ, respectively. Overall, both methods (microarray and qPCR) and genotypes [*pha-1(e2123)* and wild type] returned analogous expression profiles at 72 h exposure in 9 of 10 genes tested; only F15E11.13 could not be confirmed (**Table 1**). This suggests that *pha-1(e2123)* and wild type are

[3]http://david.abcc.ncifcrf.gov/

[4]http://bioinf.wehi.edu.au/software/russell/logrank

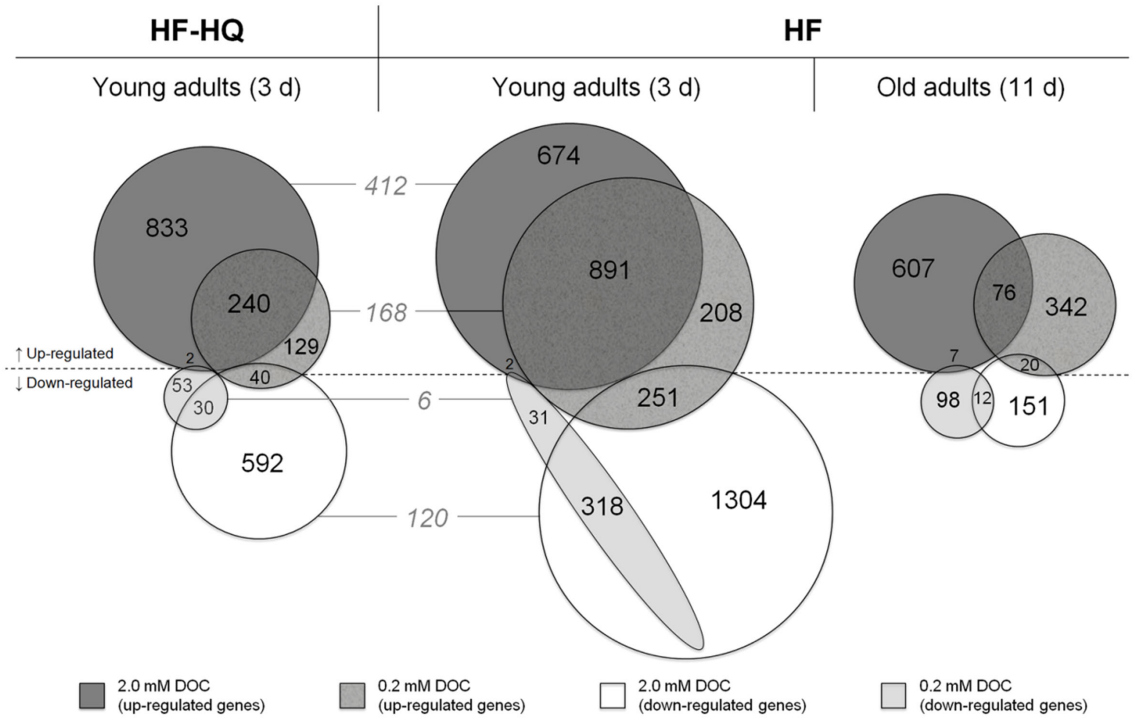


FIGURE 1 | Over-represented genes. The Venn diagrams show the overlap of significantly up- or down-regulated genes (fold change >1.25 or <0.8) per concentration in nematodes exposed for 3 days to HF (center), HF-HQ (left), and for 11 days to HF (right). The gray colored numbers given in italics represent the overlap between HF and HF-HQ derived DEG lists.

Table 1 | Quantitative PCR of 10 HF-responsive genes identified by DNA microarray.

Gene	Expression level							
	qRT-PCR N2 wild type						DNA microarray *pha-1(e2123)*	
	24 h		48 h		72 h		72 h	
2.0 mM DOC→	HF	HF-HQ	HF	HF-HQ	HF	HF-HQ	HF	HF-HQ
asah-1	○○	○○	○○	○●	○○	●●	●●●	●●●
ctl-3	○●	○●	●●	○●	○○	●○	○○○	●●●
cyp-34A9	○●	●○	○○	○○	●●	●○	●●●	●●●
F15E11.13	○●	●●	○○	●●	●●	●●	●●●	○○○
gst-24	○●	○○	●●	○○	○○	○●	●●●	●●●
lys-7	●○	○●	○●	●○	○●	●○	○●○	●●●
sek-1	●○	○○	○○	○○	○○	●○	○○○	○○○
skn-1	●●	●●	○●	○○	●○	○○	●●○	○○○
sod-4	○●	○○	●○	○○	○○	○●	○○○	●●●
sodh-1	●○	○●	○○	○○	○●	●●	●●●	●●●

Fold change: ●*<0.3,* ●*0.3–0.7,* ○ *0.7–1.5,* ●*1.5–3.0,* ●*>3.0.*

essentially interchangeable. The inclusion of further time points revealed that the majority of the selected genes did not respond rapidly to the exposure, exceptions to this were F15E11.13 and, to some extent, *cyp-34A9* and *skn*-1.

EVALUATION OF TRANSCRIPT PROFILES BASED ON ANNOTATION ENRICHMENT ANALYSES

The transcriptional profiles were subjected to detailed analyses to identify pathways linked to the mode of actions of HF and HF-HQ. The principal approach applied an annotation enrichment analysis within the different DEG sets, sub-divided in up- and down-regulated genes.

First, DEGs were assigned to KEGG pathways and mapped to known molecular interaction networks, such as metabolic pathways or environmental information processes (Table S3 in Supplementary Material). The analysis of the two HF derived lists for young adults (3 days) identified 10 KEGG pathways for 0.2 and 16 KEGG pathways for 2.0 mM DOC (**Table 2**, left part). DEG lists of 3-days-old HF-HQ treated nematodes returned 3 and 12 pathways for the low and the high concentration, respectively (**Table 2**). Despite some individual differences between both conditions (e.g., the down-regulation of spliceosome specific genes by HF), the overall overlap between concentrations and conditions was significant. Particularly noticeable was the induction of fatty acid metabolism, in particular arachidonic acid (AA) and sphingolipid metabolism, as well as the up-regulation of lysosome related genes. In the 11-days-old HF exposed worms, only four significantly enriched KEGG pathways were modulated, which were restricted to the lower concentration of HF (0.2 mM DOC). Besides the persistent induction of the AA metabolism, HF (0.2 mM DOC) was marked by a distinct up-regulation of the biotransformation machinery, which includes glutathione and cytochrome P450 (CYP) pathways. It should be noted that HF (2 mM DOC) also induced biotransformation associated genes, however, because of the large number of DEGs, they were not found to be significantly enriched.

Second, DEGs were classified using GO terms to obtain further functional insights into gene expression responses. The GO analysis produced a multitude of significantly enriched terms, many represented by the same genes across and within the three GO domains (biological process, cellular component, and molecular function). Redundancy was removed by applying the "GOTERM_XX_ALL" option in DAVID 3.0, a functional annotation clustering tool. Table S4 in Supplementary Material lists GO terms represented by the largest number of genes within individual functional clusters. A further selection of the 25 most striking terms are given in **Table 2** (right part). Again, results from young adult nematodes, exposed either to HF or HF-HQ, were more consistent compared to their older counterparts. As before, lysosomal processes, defense response as well as lipid and fatty acid metabolism were found to be enriched in the group of up-regulated genes. Moreover, humic substances induced the expression of genes coding for constituents of the cuticle and cytoskeleton. The persistent strong induction of the cellular components pseudopodium and extracellular region is caused by a comprehensive up-regulation of various major sperm proteins (MSP). Oxidative/reductive processes, determinants of adult life span, and neuropeptide signaling were found to be more enriched in response to the chemically modified HF-HQ. Both HF preparations seem to slow down the reproductive development of *C. elegans* (also cell cycle and gamete production) following a short term exposure of 3 days, a process which was seen to be reverted in older nematodes exposed for 11 days.

In a third step, we compared the DEG lists to a library of 59 different gene classes or functionally related groups of genes (for the complete comparison see Table S5 in Supplementary Material), which were assembled into a gene expression map (Kim et al., 2001). **Table 3** (left part) shows a selection of the 15 most relevant gene groups in which at least one dataset displays an over-representation. As before, genes coding for determinants of cell structure, lipid metabolism, glutathione transferases, and MSPs were significantly enriched. Lysozyme and protease encoding genes were found to be predominantly up-regulated after 3 days of exposure but down-regulated in 11-days-old HF-treated nematodes.

META-ANALYSIS: COMPARISON OF TRANSCRIPT PROFILES WITH SELECTED DATASETS TAKEN FROM THE LITERATURE

Datasets from HF and HF-HQ treated nematodes were compared to expression profiles obtained from long-lived mutants, dauer larvae, worms treated with lifespan-extending polyphenols and immunity challenged nematodes. As a control, the analysis included studies addressing the gene expression changes during the *C. elegans* life-cycle. Table S6 in Supplementary Material summarizes the results of the complete analysis comprising 40 individual data sets. **Table 3** (right part) reduces the meta-analysis to the 15 most overlapping data sets. The transcriptional profiles of HF and HF-HQ are closest to the results from Tannic acid treated nematodes and mutants of the transforming growth factor-beta (TGF-β) pathway. In contrast, significant overlaps were limited to up-regulated DEGs in long-lived *daf-2(e1370)* and *daf-12(rh273)*, as well as nematodes infected with *Pseudomonas aeruginosa* or exposed to another polyphenol, the flavonoid Quercetin. A significant proportion of genes up-regulated after 3 days were shown to be down-regulated after 11 days. This result demonstrates the level of dynamic transcriptional changes during the HF mediated aging process.

HF and HF-HQ derived transcriptional profiles were also analyzed by assigning DEGs to gene expression mountains originally assembled from 553 different *C. elegans* DNA microarray experiments (Kim et al. (2001). Table S7 in Supplementary Material displays the overlap of all gene expression mountains with the HF and HF-HQ derived datasets, respectively, as calculated by RFs. A summary is presented in **Figure 2** and includes a graphical overview of the data obtained for 14 mounts, distinguishing between up- and down-regulated genes as well as a summary from selected published data sets. Young adult nematodes exposed to the lower concentration of HF-HQ (**Figure 2A**, left side), and both HF concentrations (**Figure 2B**) resembled the gene expression mount map characteristic for Tannic acid exposed nematodes or mutants of the TGF-β pathway. The higher concentration of HF-HQ on the other hand seems to overlap, in part, with the gene expression patterns following *P. aeruginosa* infection (**Figure 2A**, right side). No clear categorization was possible for old adults exposed to HF.

A stringent GO term analysis on genes belonging to each of the significantly overlapping sections made it possible to focus on similarities within and between HF, HF-HQ, and long-lived mutant strains. Only gene lists derived from TGF-β mutants as well as *daf-2(e1370)* mutants produced significant results. **Figure 3**

Table 2 | Overview of over-represented KEGG pathways and gene ontology terms in DEG lists.

KEGG-pathway	Representation factor (RF)						GO-term	Representation factor (RF)						
	HF-HQ - 3d		HF - 3d		HF - 11d			HF-HQ - 3d		HF - 3d		HF - 11d		
DOC	0.2 mM	2.0 mM	0.2 mM	2.0 mM	0.2 mM	2.0 mM	DOC	0.2 mM	2.0 mM	0.2 mM	2.0 mM	0.2 mM	2.0 mM	
Alpha-Linolenic acid metabolism			5.2	4.4			**Cell Cycle**	4.2		2.8	2.9		2.4	Biological process
Aminoacyl-tRNA biosynthesis				2.3			**Defense response**	4.3	4.0	3.2	2.0	3.9		Biological process
Arachidonic acid metabolism	13.7	4.8		3.3	14.8		**Determination of adult life span**	2.2	2.4					Biological process
Arginine and proline metabolism				2.4					2.3	2.0		3.9		Biological process
Basal transcription factors				2.3			Fatty acid metabolic process		5.1	3.4	3.7			Biological process
Biosynthesis of unsaturated fatty acids			4.0	3.3			**Gamete generation**	3.6	2.6	3.0	2.1		2.6	Biological process
Cysteine and methionine metabolism		3.7					Lipid modification		3.0	3.2	3.4			Biological process
Drug metabolism		3.0	2.6		10.2		Lysosome organization		7.8		6.5			Biological process
Ether lipid metabolism				4.7			Neuropeptide signaling pathway	11.5	7.0		3.3			Biological process
Fatty acid elongation in mitochondria		5.5					**Oxidation reduction**	2.7	2.0			2.0		Biological process
Fatty acid metabolism		2.6										4.6		Biological process
		2.7					Reproductive develop. process	2.0		2.4	2.0			Biological process
Glutathione metabolism					6.9		**Cytoskeleton**			2.7		2.3	3.2	Cellular component
Glycerophospholipid metabolism				2.9						4.1	3.1			Cellular component
Homologous recombination				2.8			**Extracellular region**	3.2	2.8		2.0		4.7	Cellular component
Jak-STAT signaling pathway				2.9			Gap junction		2.6	8,9	3.6			Cellular component
Lysosome		4.6	2.5	3.3			Intracellular non-m.-b.* organelle			3.5	3.1		2.0	Cellular component
MAPK signaling pathway	3.5						**Lysosome**		10.1	5.8	7.7			Cellular component
Metabolism of xenobiotics by CYP					9.1		Peroxisome		5.6	7.6	6.9			Cellular component
Other glycan degradation		7.2		5.5			**Pseudopodium**	46.2		13.9		46.0	30.0	Cellular component
Progesterone-med. oocyte maturation			8.0	2.2			ATPase activity		7.2		2.8		2.4	Molecular function
Propanoate metabolism			2.7				ATP-dependent helicase activity			3.8	3.6		3.0	Molecular function
Proteasome				2.3			Ceramidase activity		7.5		5.1			Molecular function
Sphingolipid metabolism		5.7	3.4	2.8			Hydrolase activity (C-N bonds)		5.2		3.5			Molecular function
Spliceosome			3.1	2.5			Lipase activity		3.6		3.2			Molecular function
Starch and sucrose metabolism	20.8						**Lysozyme activity**	26.1		10.6		68.2		Molecular function
TGF-beta signaling pathway		2.4	2.8				**Oxidoreductase activity**	17.4	3.9	8.0	5.8			Molecular function
Tryptophan metabolism		2.7					**Structural constituent of cuticle**	4.4		2.0		2.8	3.0	Molecular function
Wnt signaling pathway		2.2	2.1					5.9						Molecular function

*Shown is the representation factor for DEGs in the respective list in relation to the overall count in significantly enriched KEGG and GO terms, respectively; red label stands for up-, green for down-regulated genes; respective pale coloring corresponds to P < 0.05, deep colors to P < 0.001; apparently most relevant terms are given in bolt; please find more details in Tables S3 and S4 in Supplementary Material. *Membrane-bound.*

Table 3 | Overview of over-represented gene classes in DEG lists and comparison to literature data.

Gene class	Representation factor (RF)*						Literature data	Representation factor (RF)*					
	HF-HQ - 3d		HF - 3d		HF - 11d			HF-HQ - 3d		HF - 3d		HF - 11d	
DOC	0.2 mM	2.0 mM	0.2 mM	2.0 mM	0.2 mM	2.0 mM	DOC	0.2 mM	2.0 mM	0.2 mM	2.0 mM	0.2 mM	2.0 mM
Cell structure	8.3		3.7		8.7	6.1	**TGF-β adults**[1] ↑	2.9	2.4	2.3	2.6	2.3	
		2.8		2.1					2.1				
Collagens	3.2					3.6	**TGF-β adults**[1] ↓						
		5.7						2.0		2.8	2.4		
CYPs					3.7		***P. aeruginosa* infection**[2] ↑	4.4	6.2	3.6	2.5	5.9	
					11.6							4.9	6.2
Dehydrogenases	2.8						***P. aeruginosa* infection**[2] ↓	6.4	2.1	2.7	3.4	4.1	3.7
					6.2				8.2			13.7	
Desaturases		5.0	5.4	5.9			***daf-2(e1370)***[2] ↑	2.2	2.9			2.0	
												3.4	2.0
Esterases	2.4	2.3					***daf-2(e1370)***[2] ↓	2.5	2.7	2.7	2.8	3.3	
						4.7			2.6			2.4	2.0
GSTs	3.3	2.0	2.8	2.8	5.9		*daf-12(rh273)*[3] ↑	13.1		8.1		12.4	7.1
		2.9			5.4						6.3		
Hydrolases			2.5	2.2			*daf-12(rh273)*[3] ↓	4.6	2.3	2.8		3.8	
		2.3				3.3			3.5			7.5	
Lipid metabolism		2.4	2.1	2.8	2.2		*pmk-1(km25)*[4] ↑		3.8	4.6	3.3	4.4	
		2.6			5.4				2.0			11.3	8.8
Lysozymes	17.4	9.0	4.9				*pmk-1(km25)*[4] ↓		2.5				
				4.1	42.6	29.0							
Msps	18.5		7.5		20.5	8.9	*daf-16* + Resveratrol[5] ↑	2.1		2.2		2.0	
				6.6									
Oocyte enriched						2.3	Quercetin[6] ↑	6.0	2.5	3.9		4.6	2.1
			3.5	2.6					2.5		2.7	7.3	
Proteases		3.1	2.4	2.1			Quercetin[6] ↓		2.8		2.2	4.2	
						4.7						2.7	3.6
Sperm-enriched			3.2				**Tannic acid**[6] ↑	7.1	3.9	5.1	2.5	5.5	
				2.4							2.5	2.4	2.6
Ubiquitin genes							**Tannic acid**[6] ↓						2.9
						2.4		6.0			2.6		

**Shown is the representation factor for DEGs in the respective list in relation to the overall count in significantly enriched gene classes or in comparison to literature data; red label and ↑ refer to up-, green label and ↓ to down-regulated genes; respective pale coloring corresponds to $P < 0.05$, deep colors to $P < 0.001$; gene glasses and datasets deemed to be most relevant are given in bold; for more details see Tables S5 and S6 in Supplementary Material. [1]Shaw et al. (2007), [2]Evans et al. (2008), [3]Fisher and Lithgow (2006), [4]Troemel et al. (2006), [5]Viswanathan et al. (2005), [6]Pietsch et al. (2012).*

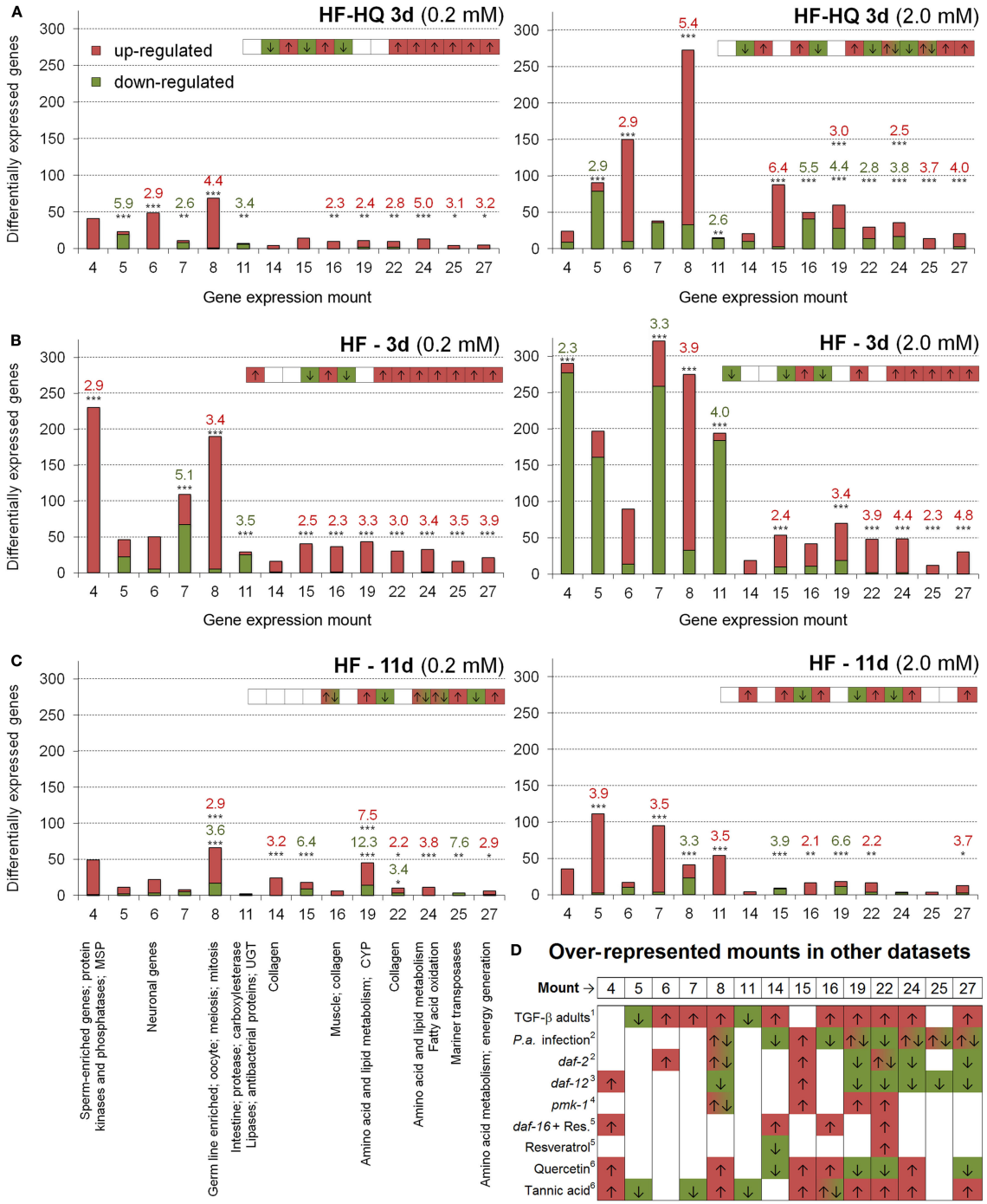


FIGURE 2 | Over-represented gene expression mountains. Identification of over-represented gene expression mountains of **(A)** HF-HQ (3 days), **(B)** HF (3 days), and **(C)** HF (11 days) treated nematodes; 0.2 mM DOC data are shown on the left, 2.0 mM data on the right. Only mounts which are significantly affected by at least one HS treatment are presented, red labels represent up-, green labels down-regulated genes. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. **(D)** Presents the associated term names on the left and a comparison to selected published datasets on the right (only mounts relevant to HS are shown). Note the short graphical overview in the right upper corner of each diagram. The complete dataset (gene numbers and RFs for all 44 gene expression mountains of all six HS conditions as well as RFs for previously published datasets) can be found in Table S7 in Supplementary Material. [1]Shaw et al. (2007), [2]Evans et al. (2008), [3]Fisher and Lithgow (2006), [4]Troemel et al. (2006), [5]Viswanathan et al. (2005), [6]Pietsch et al. (2012).

shows the Venn diagrams describing the overlap between TGF-β mutants and 0.2 mM (**Figure 3A**) as well as 2.0 mM (**Figure 3B**) HF or HF-HQ treated nematodes. The overlap between the profiles of *daf-2(e1370)*, and 2.0 mM HF or HF-HQ treated nematodes

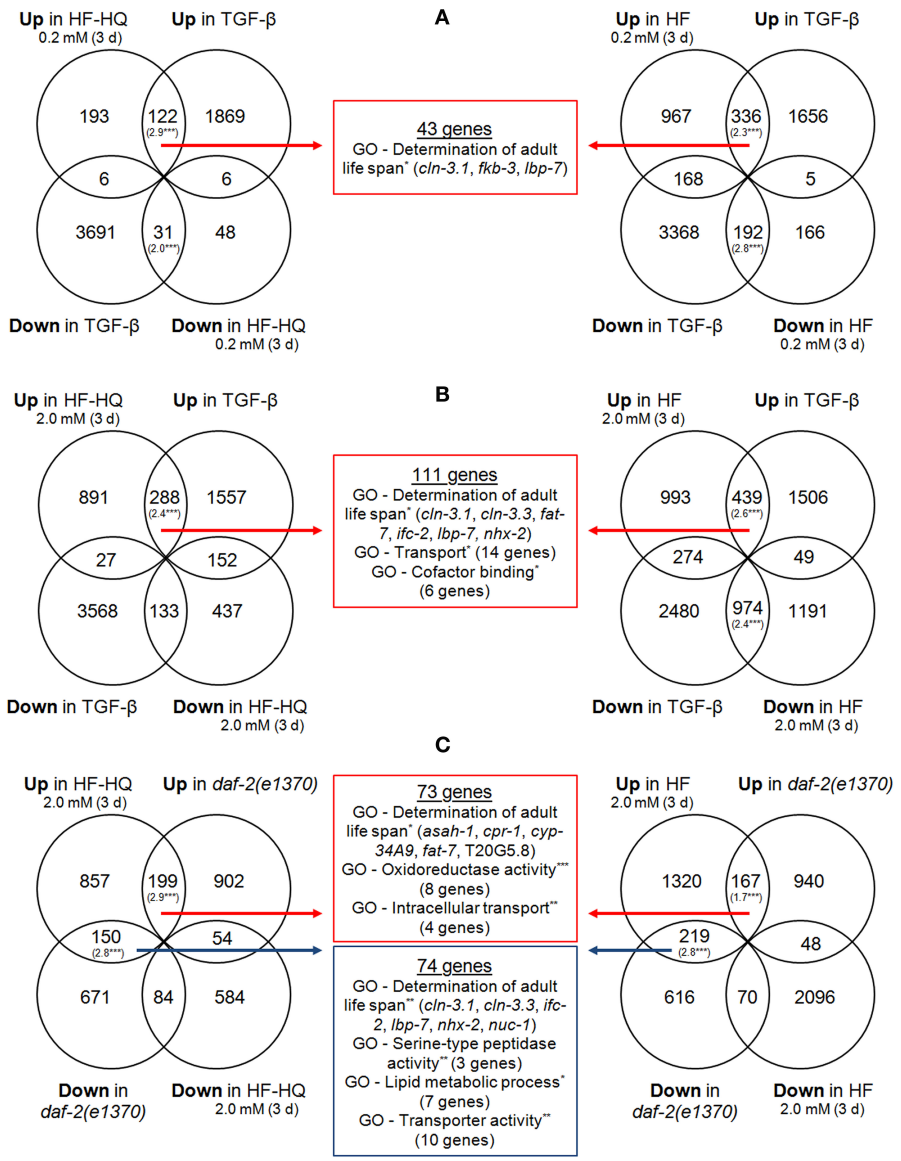


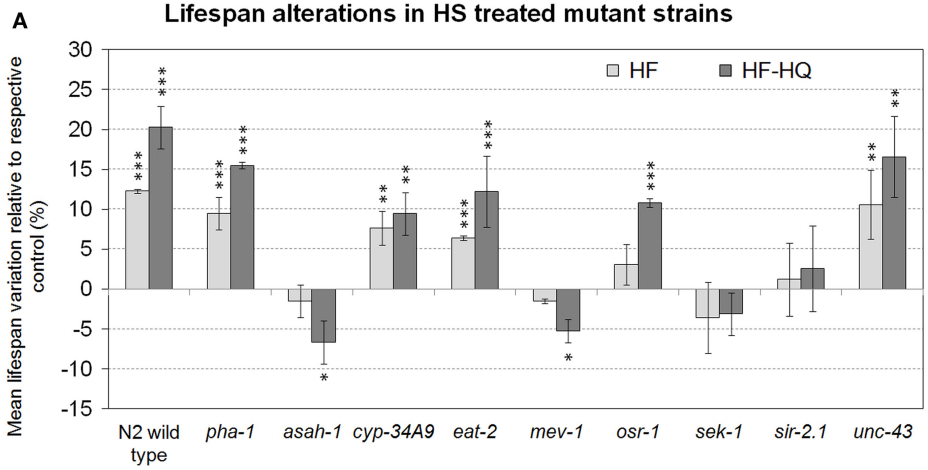
FIGURE 3 | Overlap of similarly regulated genes. Shown are comparisons of DNA microarray data sets of HF (right) and HF-HQ (left) treatments with the data set derived from TGF-β adults (Shaw et al., 2007) – 0.2 mM DOC HS **(A)** and 2.0 mM DOC HS **(B)** – as well as from *daf-2(e1370)* mutants (Evans et al., 2008) **(C)**. The RF values in brackets indicate a significant overlap between the data sets. The middle section represents the overlap of DEGs in both HF, HF-HQ, and the comparative condition, shown are significantly over-represented GO terms (Biological process), the associated number of genes, and the individual names of all genes which are part of the GO term "Determination of adult life span." Red arrows/boxes are intersections of commonly up-regulated transcripts; blue arrows/boxes are intersections derived from genes with opposite transcriptional responses. $^{*}P < 0.05$. $^{**}P < 0.005$. $^{***}P < 0.001$

is given in **Figure 3C**. All three analyses returned the term "Determination of adult life span," moreover, common transport processes were identified. Some of these genes act downstream of DAF-16 (Murphy et al., 2003) and/or thought to be involved in lysosomal metabolism (e.g., *asah-1*, encoding a putative *n*-acylsphingosine amidohydrolase) or a potential drug metabolizer (e.g., *cyp-34A9*, a cytochrome P450 monooxygenase). Both mutant strains were selected for lifespan assays as described below.

IDENTIFICATION OF GENES REQUIRED FOR HF/HF-HQ MEDIATED LONGEVITY

To substantiate the importance of stress response genes and genes relevant to aging, nine *C. elegans* mutants were tested for their ability to trigger longevity during humic substance exposure (**Figure 4A**). HF or HF-HQ exposure led to a significant increase in mean lifespan of N2 wild type and the *pha-1(e2123)* mutant strain, confirming that *pha-1* was a suitable test strain for this analysis. In addition, *cyp-34A9(ok2401)*, *eat-2(ad465)*,

A



B

	asah-1 (tm495)	*cyp-34A9* (ok2401)	*eat-2* (ad465)	*mev-1* (kn1)	*osr-1* (rm1)	*sek-1* (ag1)	*sir-2.1* (ok434)	*unc-43* (n498n1186)
HF	×			×	×	×	×	
HF-HQ	×			×		×	×	
Blueberry polyphenols[1]				×	×	×		×
Caffeic acid[2]					×	×	×	×
Catechin[3]			×	×				
Quercetin[4]						×		×
Rosmarinic acid[2]					×	×	×	×
Tannic acid[5]			×			×		

× – represents combinations where no significant extension of lifespan has been found

FIGURE 4 | Genetic modulators. (A) Lifespan alterations in nine mutant strains following the treatment with either 0.4 mM DOC HF (light gray) or 0.4 mM DOC HF-HQ (dark gray). Shown are the mean lifespan variations compared to the respective untreated control. Error bars represent SEM, $^{**}P < 0.01$, $^{***}P < 0.001$. **(B)** provides a graphical comparison of previously published datasets on polyphenols. Crosses mark combinations where no significant extension of lifespan was observed; gray shaded boxes are untested combinations. [1]Wilson et al. (2006), [2]Pietsch et al. (2011), [3]Saul et al. (2009), [4]Pietsch et al. (2009), [5]Saul et al. (2010).

and *unc-43(n498n1186)* responded to the exposure to HF or HF-HQ with an increase in lifespan. HF-HQ was seemingly a more effective trigger of longevity than the unmodified HF, a fact which may explain why *osr-1(rm1)* responded only to HF-HQ, but not to HF. No lifespan extension was observed in *sek-1(ag1)* and *sir-2.1(ok434)* and was even reduced in *asah-1(tm495)* and *mev-1(kn1)* (**Figure 4A**). An overview of comparative data from similar experiments testing the impact of individual polyphenols and one fruit extract suggest that the HF/HF-HQ mediated longevity align well with the key mechanisms identified for Caffeic acid and Rosmarinic acid as well as blueberry polyphenols. The overlap to Tannic acid treated nematodes was limited to *sek-1*, a frequently encountered effector in aging pathways (**Figure 4B**).

DISCUSSION

Caenorhabditis elegans exposed to the humic substance HF were shown to live longer and to be more stress resistant, but were impaired in their reproductive performance and growth (Menzel et al., 2011). Therefore, HF's impact resembles other polyphenol monomers, such as Quercetin (Kampkötter et al., 2008; Pietsch et al., 2009) and Tannic acid (Saul et al., 2010). Indeed, HF was found to be rich in functional group content, with possibly hydroxybenzenes as effective building blocks (Meinelt et al., 2007). First experimental evidence was offered by Menzel et al. (2011) who were able to demonstrate that a chemical modification of HF (an enrichment in phenolic and quinonoid functional groups) resulted in the amplification of the biological effects. The current study aimed to extend this knowledgebase by applying global

transcriptomics to compare the impact of HF and a modified HF (HF-HQ).

QUANTITATIVE ANALYSIS OF TRANSCRIPT PROFILES AND VALIDATION BY qRT-PCR

Exposure to HF or HF-HQ had a profound impact on global gene expression patterns. Clearly, the return of over 1000 DEGs can be attributed, in part, to the non-stringent cut-off value of 1.25-fold, which was purposefully chosen to facilitate a comprehensive secondary analysis. The overlap between the HF and the HF-HQ DEGs-lists is considerable, even though HF returned more DEGs than HF-HQ. A 10-fold increase in concentration (from 0.2 to 2.0 mM DOC) amplified the number of DEGs approximately twofold for HF and threefold for HF-HQ. Moreover, at least for HF exposure, it was apparent that the transcriptome was more responsive in young adults than old adults. This finding together with the observed delay in reproduction and the retarded growth (Menzel et al., 2011) suggests that the reproductive development is a target following HF exposure. The high level of overlap between the two concentrations in young adults was not observed in old adults, indicating a shift in gene expression dynamics, which in accordance with Van Straalen and Feder (2012) is separated by effects (concentration/dose) rather than different exposures (compounds). Given that the primary objective of this work was to distinguish between the two preparations of HF and compare the results to the genomics literature, the scenario "exposure" is more relevant. Therefore we focused on the 3-days derived data.

The Affymetrix DNA microarray system is a robust, reliable, and well established system (Dalma-Weiszhausz et al., 2006). Nonetheless, we analyzed the expression level of 10 genes by qRT-PCR, however in the N2 wild type rather than *pha-1(e2123)* which was used for the microarray experiments. By doing so, we were able to confirm the microarray data but also the validity of the mutant strain, which was required for the production of a large population of age-synchronized old adults. The addition of two additional time points, namely shorter exposure times of 24 and 48 h, revealed complex time resolved differences in transcription. Many genes were initially repressed or transcriptionally inactive but induced at 72 h. In contrast, *skn-1* was induced at 24 h but returned to base line levels thereafter. Interestingly, *skn-1* encodes a longevity-promoting transcription factor and is positioned in the p38 MAPK pathway. Its expression has been shown to be enhanced under conditions of stress (An and Blackwell, 2003) or reduced DAF-2 signaling.

QUALITATIVE ANALYSIS OF TRANSCRIPT PROFILES BY ABUNDANCE SCREENS

The generation of a dataset that describes quantitative changes in gene expression upon a certain condition, e.g., exposure to a chemical compound, is less challenging than the interpretation of its relevance, significance, and contribution to the physiology of the organism. Statistical methods aid in the identification of over- or under-represented transcripts which can be aligned to biological processes or functions via KEGG- and GO-term screens.

The over-representation of DEGs related to lipid metabolism and biotransformation in HF and HF-HQ derived transcript profiles suggests the presence of enhanced catabolism, possibly of toxic intermediates. This may contribute to the lifespan extension which, according to the green theory of aging, is due to the investment in cellular waste disposal and protein conservation (Gems and McElwee, 2005). The induction of several signaling pathways (MAPK, Wnt, TGF-β, neuropeptide) might reflect transcriptional changes of downstream targets. However, changes in heat shock protein (HSP) gene expression were not observed. Clearly, other transcripts responded to the HF/HF-HQ challenge, some are possibly involved in the observed longevity phenotype. For example the pronounced over-representation of lysosome specific genes may be linked to the process of autophagy, which mediates the degradation of cellular components, including whole organelles and protein aggregates. The importance of an efficient lysosomal activity is indicated by the finding that long-lived *C. elegans* mutants frequently display increased autophagy (Melendez et al., 2003; Hars et al., 2007).

The down-regulation of the GO terms "reproductive developmental process," "gamete generation," and "cell cycle" corroborate the notion that the reproductive development slows down in HF/HF-HQ exposed nematodes. Interestingly, this process seems to be dynamic in nature as "gamete generation" and "cell cycle" appear to be up-regulated (at least in 2.0 mM DOC HF) in old adults. In contrast, transcripts involved in the constitution of the cuticle (also the gene class "cell structure") or active in the extracellular region and the pseudopodium (here in particular MSPs) are consistently up-regulated. The cuticle of *C. elegans* can differ in layer numbers, relative thickness, and composition during development (particularly in larvae) and changing environmental conditions. Indeed, genes encoding for cuticle collagens were found to be induced in response to several bacterial species (Coolon et al., 2009) and under oxidative stress (Shin et al., 2011). In aging research, studies identified a large number of collagens as age regulated genes (Halaschek-Wiener et al., 2005; Budovskaya et al., 2008). These data suggest that cuticle collagens may be differentially regulated indirectly in defense against environmental perturbations and potentially in longevity. A comparable up-regulation of *msp*-genes were observed in the long-lived *daf-12(rh273)* mutant (Fisher and Lithgow, 2006), but also in Quercetin or Tannic acid exposed wild type worms (Pietsch et al., 2012).

The HF and HF-HQ derived data were remarkably similar. However, only HF-HQ returned an up-regulation of the Biological processes "oxidation/reduction activity" possibly due to the previously observed increase in oxidoreductive activity (Menzel et al., 2011).

QUALITATIVE ANALYSIS OF TRANSCRIPT PROFILES BY LITERATURE COMPARISON

The main problem of abundance screens (e.g., GO-term profiling) is the incomplete gene annotation of genomes and the risk of over-interpretation, as enrichment values can occur by chance (Rhee et al., 2008). The application of appropriate statistical tools minimizes, but cannot exclude, the frequency of false-positives. To offer a more independent verification, we conducted a meta-analysis to include published data and searched for overlapping gene clusters via the gene expression mount map created by Kim et al. (2001). This revealed that Tannic acid exposed wild type worms (Pietsch

et al., 2012), TGF-β mutants (Shaw et al., 2007), and worms subjected to humic substances produce similar expression pattern mountains. The comparison of DEGs-lists of either up- or down-regulated genes confirmed this result. The overlap to worms with a challenged immunity in response to an infection with *P. aeruginosa*, Quercetin treated nematodes as well as long-lived and more stress-resistant mutants [*daf-2(e1370)* and *daf-12(rh273)*] was, at large, restricted to the section of up-regulated DEGs. Both HF preparations and TGF-β mutants (and *daf-2* for 2.0 mM DOC) shared GO terms, suggesting that the negative regulation of the TGF-β pathway and, less pronounced, the insulin-like signaling (ILS) pathway, play prominent roles in the lifespan extension due to HF/HF-HQ. Both signaling cascades control, by responding to environmental conditions, whether *C. elegans* larvae grow to adults or to long-lived and stress-resistant dauer larvae. Based on our results, it seems that HF preparations are able to modulate these pathways, thereby facilitating the observed increase in stress resistance and longevity.

GENETIC PLAYERS THAT PROMOTE HF-DEPENDENT LONGEVITY

HF or HF-HQ treatment was not able to prolong the lifespan of the *C. elegans* mutants *sek-1*, *sir-2.1*, *mev-1*, *asah-1* and (in the case of HF only) *osr-1*. SEK-1, a MAP2K, is part of the p38 MAP kinase pathway and acts downstream of TIR-1 (toll and interleukin receptor) and NSY-1 (MAP3K). It phosphorylates the MAP kinases JNK-1 and PMK-1 (Tanaka-Hino et al., 2002), the latter results in an elevated immune response to pathogen infection (Kim et al., 2002) and also functions via SKN-1 to control resistance against metals, such as arsenic (An and Blackwell, 2003). OSR-1 is coupled to SEK-1 (through UNC-43) and regulates the osmotic stress response and survival in hyper-osmotic environments, where viability depends on activity of the CaMKII pathway (Solomon et al., 2004). Both *sek-1* and *osr-1* were shown to be essential for the blueberry polyphenol induced longevity (Wilson et al., 2006) as well as the Caffeic acid and Rosmarinic acid (Pietsch et al., 2011) mediated longevity. Surprisingly, *unc-43* was not found to be essential for the longevity effect by HF or HF-HQ.

A genetic analysis suggested that *sir-2.1* (which encodes the NAD^+-dependent deacetylase) extends lifespan via the ILS pathway and requires *daf-16* (Tissenbaum and Guarente, 2001). Berdichevsky et al. (2006) proposed the existence of a stress-dependent pathway in which SIR-2.1 acts in parallel to the ILS pathway, but still via an activation of DAF-16. Likewise, Caffeic acid, Rosmarinic acid, and Resveratrol mediated lifespan extension were all shown to be dependent on SIR-2.1. Viswanathan et al. (2005) described the involvement of *sir-2.1* in the up-regulation of stress response genes (especially *abu-11*), a mechanism which is thought to aid protein folding in the endoplasmic reticulum. SIR-2.1's involvement in stress response pathways during aging and HF/HF-HQ challenge may thus promote longevity and is possibly linked to the ILS-signaling cascades.

To further examine whether HF preparations could protect against acute oxidative stress, we examined *mev-1(kn1)*, a nematode which harbors a mutation in the cytochrome *b* large subunit of mitochondrial complex II (Ishii et al., 1998). The mutation causes an overproduction of superoxide and increased oxidative stress, resulting in accelerated aging and a reduced lifespan (Hosokawa et al., 1994; Senoo-Matsuda et al., 2001). Neither HF/HF-HQ, blueberry polyphenol (Wilson et al., 2006) nor Catechin (Saul et al., 2009) treatment were able to revert or protect against the oxidative stress encountered by the *mev-1(kn1)* mutants. The pro-oxidant properties, as described for different polyphenols by Akagawa et al. (2003) and Wiegant et al. (2009), may explain this result. Indeed, HF-HQ induced antioxidant enzymes, such as catalases (*ctl-2*, *ctl-3*) and a superoxide dismutase (*sod-3*).

The *asah-1* gene encodes a putative acid ceramidase, a lysosomal enzyme which catalyses the hydrolysis of ceramide to sphingosine and free fatty acid. Although the annotation data for *asah-1* is patchy, it is thought to act downstream from DAF-16 (Murphy et al., 2003), has been assigned to the GO-term "Determination of adult life span," and leads to Farber lipogranulomatosis, when mutated in humans. In affected individuals, harmful amounts of lipids accumulate in cells and tissues throughout the body (Mao and Obeid, 2008). Because several gene clusters coding for lysosome components, ceramidases, and sphingolipid metabolizing enzymes were found to be up-regulated by HF/HF-HQ, these processes may be essential for the effectiveness of humic substances, a hypothesis that requires further investigation.

CONCLUSION

The humic substance preparation HF significantly extends the lifespan of the nematode *C. elegans*. Here we were able to show that HF's mode of action resembles other polyphenol monomers: (i) the transcript profiles of HF are very similar to Tannic acid and, but less pronounced, to Quercetin; (ii) the enrichment of HF with hydroquinones (HF-HQ) enhances its phenotypic effects (Menzel et al., 2011) and returns a more streamlined transcript profile (i.e., HF-HQ and HF affected common gene clusters, even though the total number of DEGs was lower in the HF-HQ sample); and (iii) HF/HF-HQ and polyphenols induce similar effects on key mutant nematodes.

Despite these similarities, the mode of action of HF is complex (**Figure 5**) and is heavily influenced by the negative regulation of TGF-β- and ILS signaling as well as increased lysosomal activity. Longevity, according to Kirkwood and Austad (2000), is driven by an organism's ability to cope with extrinsic or intrinsic stressors. Clearly stress response pathways do not function in isolation but act, in concert, within a stress network where multiple hubs serve as coordinators of various modules (Kourtis and Tavernarakis, 2011). The process of aging both influences and is influenced by this stress network. Mild environmental stress, as triggered by low concentrations of polyphenols or polyphenol-containing humic substances, primes response pathways which in turn increase stress resistance and longevity. These mechanisms are multidimensional but one of the prime candidates involved in the impact of polyphenols and humic substances is, at least in the nematode *C. elegans*, *sir-2.1*.

ACKNOWLEDGMENTS

We would like to thank members of the King's College Genomics Center (Dr. Matthew Arno and Dr. Estibaliz Aldecoa-otalora Astarloa) for assistance in our DNA microarray experiments. Furthermore, we thank the *Caenorhabditis* Genetics Center, which

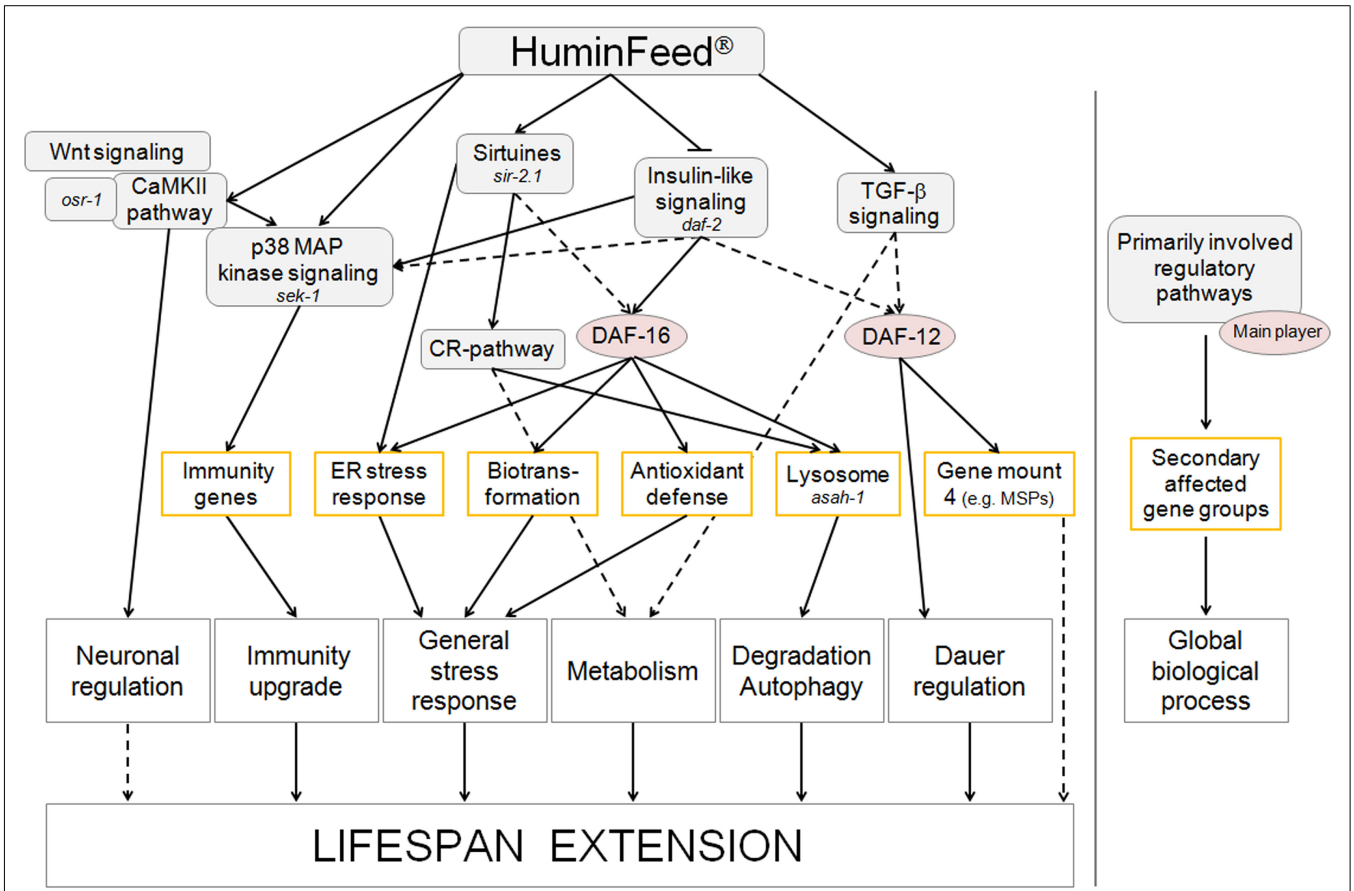


FIGURE 5 | Flow chart summarizing the multidimensional modes of action of HF/HF-HQ. Solid arrows denote direct, well-known connections; dotted lines highlight suggested relations. Genes identified in lifespan assays or selected by the literature meta-analysis are annotated in italics.

is funded by the National Institutes of Health National Centre for Research Resources (USA), for the supply of all *Caenorhabditis elegans* strains, except for *asah-1(tm495)*, which was kindly provided by the Mitani Lab (Department of Physiology, Tokyo Women's Medical University, School of Medicine, Japan) an its National BioResource Project for the nematode. Suresh C. Swain and Stephen R. Stürzenbaum were supported by the BBSRC (BB/E025099), Ralph Menzel, Sophie Tiedt, and Christian E. W. Steinberg by a DFG grant (STE673/16-1).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Toxicogenomics_/10.3389/fgene.2012.00050/abstract

REFERENCES

Aitlhadj, L., and Stürzenbaum, S. R. (2010). The use of FUdR can cause prolonged longevity in mutant nematodes. *Mech. Ageing Dev.* 131, 364–365.

Akagawa, M., Shigemitsu, T., and Suyama, K. (2003). Production of hydrogen peroxide by polyphenols and polyphenol-rich beverages under quasi-physiological conditions. *Biosci. Biotechnol. Biochem.* 67, 2632–2640.

An, J. H., and Blackwell, T. K. (2003). SKN-1 links *C.* elegans mesendodermal specification to a conserved oxidative stress response. *Genes Dev.* 17, 1882–1893.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29.

Azen, S. P., Roy, S., Pike, M. C., Casagrande, J., and Mishell, D. R. Jr. (1977). A new procedure for the statistical evaluation of intrauterine contraception. *Am. J. Obstet. Gynecol.* 128, 329–335.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300.

Berdichevsky, A., Viswanathan, M., Horvitz, H. R., and Guarente, L. (2006). *C. elegans SIR*-2.1 interacts with 14-3-3 proteins to activate DAF-16 and extend life span. *Cell* 125, 1165–1177.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71–94.

Budovskaya, Y. V., Wu, K., Southworth, L. K., Jiang, M., Tedesco, P., Johnson, T. E., and Kim, S. K. (2008). An elt-3/elt-5/elt-6 GATA transcription circuit guides aging in *C. elegans*. *Cell* 134, 291–303.

Coolon, J. D., Jones, K. L., Todd, T. C., Carr, B. C., and Herman, M. A. (2009). *Caenorhabditis elegans* genomic response to soil bacteria predicts environment-specific genetic effects on life history traits. *PLoS Genet.* 5, e1000503. doi:10.1371/journal.pgen.1000503

Dalma-Weiszhausz, D. D., Warrington, J., Tanimoto, E. Y., and Miyada, C. G. (2006). The Affymetrix GeneChip platform: an overview. *Meth. Enzymol.* 410, 3–28.

Davies, S. K., Leroi, A. M., and Bundy, J. G. (2012). Fluorodeoxyuridine affects the identification of

metabolic responses to daf-2 status in *Caenorhabditis elegans*. *Mech. Ageing Dev.* 133, 46–49.

Evans, E. A., Kawli, T., and Tan, M. W. (2008). *Pseudomonas aeruginosa* suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in *Caenorhabditis elegans*. *PLoS Pathog.* 4, e1000175. doi:10.1371/journal.ppat.1000175

Fisher, A. L., and Lithgow, G. J. (2006). The nuclear hormone receptor DAF-12 has opposing effects on *Caenorhabditis elegans* lifespan and regulates genes repressed in multiple long-lived worms. *Aging Cell* 5, 127–138.

Gems, D., and McElwee, J. J. (2005). Broad spectrum detoxification: the major longevity assurance process regulated by insulin/IGF-1 signaling? *Mech. Ageing Dev.* 126, 381–387.

Granato, M., Schnabel, H., and Schnabel, R. (1994a). Genesis of an organ: molecular analysis of the pha-1 gene. *Development* 120, 3005–3017.

Granato, M., Schnabel, H., and Schnabel, R. (1994b). pha-1, a selectable marker for gene transfer in *C. elegans*. *Nucleic Acids Res.* 22, 1762–1763.

Grigoryev, D. N., Ma, S. F., Irizarry, R. A., Ye, S. Q., Quackenbush, J., and Garcia, J. G. (2004). Orthologous gene-expression profiling in multi-species models: search for candidate genes. *Genome Biol.* 5, R34.

Halaschek-Wiener, J., Khattra, J. S., Mckay, S., Pouzyrev, A., Stott, J. M., Yang, G. S., Holt, R. A., Jones, S. J. M., Marra, M. A., Brooks-Wilson, A. R., and Riddle, D. L. (2005). Analysis of long-lived *C. elegans daf*-2 mutants using serial analysis of gene expression. *Genome Res.* 15, 603–615.

Hars, E. S., Qi, H., Ryazanov, A. G., Jin, S., Cai, L., Hu, C., and Liu, L. F. (2007). Autophagy regulates ageing in *C. elegans*. *Autophagy* 3, 93–95.

Herrero, J., Al-Shahrour, F., Diaz-Uriarte, R., Mateos, A., Vaquerizas, J. M., Santoyo, J., and Dopazo, J. (2003). GEPAS: a web-based resource for microarray gene expression data analysis. *Nucleic Acids Res.* 31, 3461–3467.

Hosokawa, H., Ishii, N., Ishida, H., Ichimori, K., Nakazawa, H., and Suzuki, K. (1994). Rapid accumulation of fluorescent material with aging in an oxygen-sensitive mutant mev-1 of *Caenorhabditis elegans*. *Mech. Ageing Dev.* 74, 161–170.

Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W., Lane, H. C., and Lempicki, R. A. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* 35, W169–W175.

Ishii, N., Fujii, M., Hartman, P. S., Tsuda, M., Yasuda, K., Senoo-Matsuda, N., Yanase, S., Ayusawa, D., and Suzuki, K. (1998). A mutation in succinate dehydrogenase cytochrome b causes oxidative stress and ageing in nematodes. *Nature* 394, 694–697.

Kampkötter, A., Timpel, C., Zurawski, R. F., Ruhl, S., Chovolou, Y., Proksch, P., and Watjen, W. (2008). Increase of stress resistance and lifespan of *Caenorhabditis elegans* by quercetin. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 149, 314–323.

Kanehisa, M. (2002). The KEGG database. *Novartis Found. Symp.* 247, 91–101. [Discussion 101–103, 119–128, 244–152].

Kim, D. H., Feinbaum, R., Alloing, G., Emerson, F. E., Garsin, D. A., Inoue, H., Tanaka-Hino, M., Hisamoto, N., Matsumoto, K., Tan, M. W., and Ausubel, F. M. (2002). A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* 297, 623–626.

Kim, S. K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J. M., Eizinger, A., Wylie, B. N., and Davidson, G. S. (2001). A gene expression map for *Caenorhabditis elegans*. *Science* 293, 2087–2092.

Kirkwood, T. B. L., and Austad, S. N. (2000). Why do we age? *Nature* 408, 233–238.

Kourtis, N., and Tavernarakis, N. (2011). Cellular stress response pathways and ageing: intricate molecular relationships. *EMBO J.* 30, 2520–2531.

Levine, M. T., Eckert, M. L., and Begun, D. J. (2011). Whole-genome expression plasticity across tropical and temperate *Drosophila melanogaster* populations from Eastern Australia. *Mol. Biol. Evol.* 28, 249–256.

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) Method. *Methods* 25, 402–408.

Mao, C., and Obeid, L. M. (2008). Ceramidases: regulators of cellular responses mediated by ceramide, sphingosine, and sphingosine-1-phosphate. *Biochim. Biophys. Acta* 1781, 424–434.

McCarthy, D. J., and Smyth, G. K. (2009). Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics* 25, 765–771.

Meinelt, T., Paul, A., Phan, T. M., Zwirnmann, E., Krüger, A., Wienke, A., and Steinberg, C. E. W. (2007). Reduction in vegetative growth of the water mold *Saprolegnia parasitica* (Coker) by humic substance of different qualities. *Aquat. Toxicol.* 83, 93–103.

Melendez, A., Talloczy, Z., Seaman, M., Eskelinen, E. L., Hall, D. H., and Levine, B. (2003). Autophagy genes are essential for dauer development and life-span extension in C. *elegans*. *Science* 301, 1387–1391.

Menzel, R., Menzel, S., Tiedt, S., Kubsch, G., Stößer, R., Bährs, H., Putschew, A., Saul, N., and Steinberg, C. E. W. (2011). Enrichment of humic material with hydroxybenzene moieties intensifies its physiological effects on the nematode *Caenorhabditis elegans*. *Environ. Sci. Technol.* 45, 8707–8715.

Menzel, R., Rödel, M., Kulas, J., and Steinberg, C. E. W. (2005). CYP35: xenobiotically induced gene expression in the nematode *Caenorhabditis elegans*. *Arch. Biochem. Biophys.* 438, 93–102.

Murphy, C. T., Mccarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., and Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* 424, 277–283.

Pietsch, K., Saul, N., Chakrabarti, S., Stürzenbaum, S. R., Menzel, R., and Steinberg, C. E. W. (2011). Hormetins, antioxidants and prooxidants: defining quercetin-, caffeic acid- and rosmarinic acid-mediated life extension in *C. elegans*. *Biogerontology* 12, 329–347.

Pietsch, K., Saul, N., Menzel, R., Stürzenbaum, S. R., and Steinberg, C. E. W. (2009). Quercetin mediated lifespan extension in *Caenorhabditis elegans* is modulated by age-1, daf-2, sek-1 and unc-43. *Biogerontology* 10, 565–578.

Pietsch, K., Saul, N., Swain, S., Menzel, R., Steinberg, C. E. W., and Stürzenbaum, S. R. (2012). Ruling aging with polyphenols: meta-analysis of global transcriptional responses reveals conserved genetic pathways as mediators for quercetin and tannic acid effected longevity in *C. elegans*. *Front. Genet.* 3, 48.

Rainer, J., Sanchez-Cabo, F., Stocker, G., Sturn, A., and Trajanoski, Z. (2006). CARMAweb: comprehensive R- and bioconductor-based web service for microarray data analysis. *Nucleic Acids Res.* 34, W498–W503.

Rhee, S. Y., Wood, V., Dolinski, K., and Draghici, S. (2008). Use and misuse of the gene ontology annotations. *Nat. Rev. Genet.* 9, 509–515.

Saul, N., Pietsch, K., Menzel, R., Stürzenbaum, S. R., and Steinberg, C. E. W. (2009). Catechin induced longevity in *C. elegans*: from key regulator genes to disposable soma. *Mech. Ageing Dev.* 130, 477–486.

Saul, N., Pietsch, K., Menzel, R., Stürzenbaum, S. R., and Steinberg, C. E. W. (2010). The longevity effect of tannic acid in *Caenorhabditis elegans*: disposable soma meets hormesis. *J. Gerontol. A Biol. Sci. Med. Sci.* 65, 626–635.

Senoo-Matsuda, N., Yasuda, K., Tsuda, M., Ohkubo, T., Yoshimura, S., Nakazawa, H., Hartman, P. S., and Ishii, N. (2001). A defect in the cytochrome b large subunit in complex II causes both superoxide anion overproduction and abnormal energy metabolism in *Caenorhabditis elegans*. *J. Biol. Chem.* 276, 41553–41558.

Shaw, W. M., Luo, S., Landis, J., Ashraf, J., and Murphy, C. T. (2007). The *C. elegans* TGF-beta dauer pathway regulates longevity via insulin signaling. *Curr. Biol.* 17, 1635–1645.

Shin, H., Lee, H., Fejes, A. P., Baillie, D. L., Koo, H. S., and Jones, S. J. (2011). Gene expression profiling of oxidative stress response of *C. elegans* aging defective AMPK mutants using massively parallel transcriptome sequencing. *BMC Res. Notes* 4, 34. doi:10.1186/1756-0500-4-34

Solomon, A., Bandhakavi, S., Jabbar, S., Shah, R., Beitel, G. J., and Morimoto, R. I. (2004). *Caenorhabditis elegans* OSR-1 regulates behavioral and physiological responses to hyperosmotic environments. *Genetics* 167, 161–170.

Steinberg, C. E. W., Saul, N., Pietsch, K., Meinelt, T., Rienau, S., and Menzel, R. (2007). Dissolved humic substances facilitate fish life in extreme aquatic environments and have the potential to extend lifespan of *Caenorhabditis elegans*. *Ann. Environ. Sci.* 1, 81–90.

Sulston, J. E., and Hodgkin, J. (1988). "Methods," in *The nematode Caenorhabditis elegans*, ed. W. B. Wood (Cold Spring Harbor: CSHL Press), 587–606.

Swain, S., Wren, J. F., Stürzenbaum, S. R., Kille, P., Morgan, A. J., Jager, T., Jonker, M. J., Hankard, P. K., Svendsen, C., Owen, J., Hedley, B. A., Blaxter, M., and Spurgeon, D. J. (2010). Linking toxicant physiological mode of action with induced gene expression changes in *Caenorhabditis elegans*. *BMC Syst. Biol.* 4, 32. doi:10.1186/1752-0509-4-32

Tanaka-Hino, M., Sagasti, A., Hisamoto, N., Kawasaki, M., Nakano, S.,

Ninomiya-Tsuji, J., Bargmann, C. I., and Matsumoto, K. (2002). SEK-1 MAPKK mediates Ca^{2+} signaling to determine neuronal asymmetric development in *Caenorhabditis elegans. EMBO Rep.* 3, 56–62.

Tissenbaum, H. A., and Guarente, L. (2001). Increased dosage of a sir-2 gene extends lifespan in *Caenorhabditis elegans. Nature* 410, 227–230.

Troemel, E. R., Chu, S. W., Reinke, V., Lee, S. S., Ausubel, F. M., and Kim, D. H. (2006). p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans. PLoS Genet.* 2, e183. doi:10.1371/journal.pgen.0020183

Van Straalen, N. M., and Feder, M. E. (2012). Ecological and evolutionary functional genomics-how can it contribute to the risk assessment of chemicals? *Environ. Sci. Technol.* 46, 3–9.

Viswanathan, M., Kim, S. K., Berdichevsky, A., and Guarente, L. (2005). A role for SIR-2.1 regulation of ER stress response genes in determining *C. elegans* life span. *Dev. Cell* 9, 605–615.

Wiegant, F. A. C., Surinova, S., Ytsma, E., Langelaar-Makkinje, M., Wikman, G., and Post, J. A. (2009). Plant adaptogens increase lifespan and stress resistance in *C. elegans. Biogerontology* 10, 27–42.

Wilson, M. A., Shukitt-Hale, B., Kalt, W., Ingram, D. K., Joseph, J. A., and Wolkow, C. A. (2006). Blueberry polyphenols increase lifespan and thermotolerance in *Caenorhabditis elegans. Aging Cell* 5, 59–68.

Conflict of Interest Statement: 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.

Received: 27 January 2012; accepted: 20 March 2012; published online: 12 April 2012.

Citation: Menzel R, Menzel S, Swain SC, Pietsch K, Tiedt S, Witczak J, Stürzenbaum SR and Steinberg CEW (2012) The nematode Caenorhabditis elegans, stress and aging: Identifying the complex interplay of genetic pathways following the treatment with humic substances. Front. Gene. ***3****:50. doi: 10.3389/fgene.2012.00050*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in
GENETICS

ORIGINAL RESEARCH ARTICLE
published: 16 May 2012
doi: 10.3389/fgene.2012.00085

Functional environmental genomics of a municipal landfill soil

Dick Roelofs[1]*, Muriel de Boer[1], Valeria Agamennone[1], Pascal Bouchier[1], Juliette Legler[2] and Nico van Straalen[1]

[1] *Department of Ecological Science, VU University, Amsterdam, Netherlands*
[2] *Institute for Environmental Studies, VU University, Amsterdam, Netherlands*

Edited by:
Stephen Sturzenbaum, King's College London, UK

Reviewed by:
Alok K. Sharma, Covance Laboratories Inc., USA
Olga Tsyusko, University of Kentucky, USA
David Spurgeon, Centre for Ecology and Hydrology, UK

***Correspondence:**
Dick Roelofs, Department of Ecological Science, VU University, de Boelelaan 1085, 1081HV, Amsterdam, Netherlands.
e-mail: dick.roelofs@vu.nl

We investigated the toxicity of soil samples derived from a former municipal landfill site in the South of the Netherlands, where a bioremediation project is running aiming at reusing the site for recreation. Both an organic soil extract and the original soil sample was investigated using the ISO standardized *Folsomia* soil ecotoxicological testing and gene expression analysis. The 28 day survival/reproduction test revealed that the ecologically more relevant original soil sample was more toxic than the organic soil extract. Microarray analysis showed that the more toxic soil samples induced gene regulatory changes in twice as less genes compared to the soil extract. Consequently gene regulatory changes were highly dependent on sample type, and were to a lesser extent caused by exposure level. An important biological process shared among the two sample types was the detoxification pathway for xenobiotics (biotransformation I, II, and III) suggesting a link between compound type and observed adverse effects. Finally, we were able to retrieve a selected group of genes that show highly significant dose-dependent gene expression and thus were tightly linked with adverse effects on reproduction. Expression of four cytochrome P450 genes showed highest correlation values with reproduction, and maybe promising genetic markers for soil quality. However, a more elaborate set of environmental soil samples is needed to validate the correlation between gene expression induction and adverse phenotypic effects.

Keywords: microarray, *Folsomia candida*, biotransformation, cytochrome P450

INTRODUCTION

The focus of ecotoxicological research is aimed at understanding toxicological phenomena in a variety of biota (Fent, 2004). Much emphasis is placed on lab-controlled testing of single compounds to address regulatory issues of chemical registration. However, testing the toxicity of complex environmental samples such as fresh water, river sediments, and natural soils remains challenging for several reasons. One of the major problems is that chemical analysis of pollutants in ecosystems often reveals an extensive list of toxicants that are potentially hazardous. Although legislation is based on the concept of concentration thresholds that must not be exceeded to ensure that the site is safe, such analysis cannot provide evidence for the real toxicological consequences of complicated mixtures. One of the valuable tools to assess ecotoxicological consequences of complex environmental mixtures is to apply bioassays. In such assays, survival and reproduction is studied of model organisms (validated in international standard tests), exposed to samples from the environment under controlled conditions. However, identification of the compound(s) causing the adverse effects among the potential list of compounds in a mixture is challenging, due to the fact that the endpoints survival, growth, and reproduction are not specific to the type of stress exerted on the test animals. Traditional bioassays do not allow conclusions on the nature of the chemicals causing the effects.

Transcriptional profiling seems to have important advantages over traditional bioassays. Several recent studies provided evidence that transcriptome profiles bear a signature of the type of pollution (Owen et al., 2008; Nota et al., 2010). If combined with traditional endpoints, genomics analysis of exposed animals can link adverse effects at the organismal level to mechanistic explanation. However, up to now this is only exemplified for single-compound exposures (van Straalen and Roelofs, 2008).

To simulate more realistic ecotoxicological scenarios, some recent studies have investigated toxic effects at the gene regulatory level of compounds presented in binary mixtures. For instance, trinitrotoluene (TNT) mixed with an additional explosive trinitrotriazacyclohexane (RDX) radically altered the gene expression profile of the ecotoxicological model organism *Eisenia fetida* when compared to single TNT exposure (Gong et al., 2007). While TNT alone regulated 321 genes, the mixture decreased the count to only three genes. These results implied a strong antagonistic effect of RDX on gene expression induced by TNT. In contrast, mixture toxicity studies with compounds proposed to have comparable modes of action should generate comparable transcriptional responses. Indeed, when *Daphnia magna* was exposed to two polycyclic aromatic hydrocarbons (fluoranthene and pyrene) no clear distinction could be made between the compounds, suggesting similar molecular modes of action (Vandenbrouck et al., 2010).

Furthermore, cluster analysis with both the single compounds and the binary mixture treatments resulted in a separation of treatments based on differences in toxic ratios rather than component differences. However, the results were highly dependent on the composition of the binary mixture. In any case, these lab-controlled experiments suggest that transcriptomics may prove valuable in determining the most toxic substances among complex environmental samples.

Only few studies (restricted to aquatic samples and river sediments) have addressed gene regulatory consequences of exposure to complex environmental samples. Menzel et al. (2009) studied exposure of nematodes to polluted and clean river sediments and showed that several biological processes, such as oxidative phosphorylation, xenobiotics, and development in response to exposure to the most polluted samples. This demonstrates that ecotoxicogenomics can be used to distinguish pollution levels in river sediments. To our knowledge, such an approach has not yet been applied to assess soil quality.

In the present study we investigated the toxicity of soil samples derived from a former municipal landfill site in the South of the Netherlands, where a bioremediation project is running aiming at reusing the site for recreation. Very recently, Legler et al. (2011) investigated this complex environmental sample to study the substances that cause toxicity using effect-directed analysis. They identified the presence of compounds (11H-benzo[*b*]fluorene, 9-methylacridine, 4-azapyrene, and 2-phenylquinoline) with previously unknown teratogenic toxicity in zebrafish. They concluded that these compounds may have been missed by current soil chemical quality assessment.

Here we present toxicogenomic data using the soil ecotoxicological model organism *Folsomia candida*. We asked the question whether an original soil sample exerts comparable toxic responses in our soil ecotoxicological model when compared to the toxic responses exerted upon exposure to the organic extract from that soil (Legler et al., 2011). If this is the case, analysis of extracts will have predictive power to estimate adverse effects in the field (Fent, 2004). To that end, the arthropods were exposed to an organic soil extract and the original Vlagheide soil sample. Results from a 28 day survival/reproduction test revealed differences in toxicity between the organic extract and the ecologically more relevant original soil sample. The more toxic soil samples induced gene regulatory changes in twice as less genes compared to the soil extract. Despite these differences several gene categories (biological processes) were shared among the two samples. In addition, a substantial number of genes were dependent on sample type (soil or extract), potentially explaining the difference in toxicity. Our results show that bioassays deploying functional genomics can reveal crucial information on the nature of the toxicants. Furthermore, we argue that it is essential to include ecologically relevant test organisms in order to properly assess the risk of environmental samples.

MATERIALS AND METHODS

STUDY SITE

The Vlagheide municipal waste landfill site is located about 10 miles South-East from 's-Hertogenbosch, the Netherlands. Haskoning B.V. sampled the site in October 2005 at depths varying from 3 to 18 m. In total seven soil samples were taken and pooled, sieved (mesh size 250 μm), homogenized, and freeze-dried to end up as a composite sample of approximately 1 kg dry weight. In order to retrieve an extract the sample was subjected to pressurized liquid extraction an accelerated solvent extraction (ASE) apparatus (Dionex, ASE200, Sunnyvale, CA, USA) with a mixture of acetone and dichloromethane in a 1:3 ratio. The extract was then subjected to gel permeation chromatography (GPC) clean-up with dichloromethane (Legler et al., 2011). An overview of concentrations of persistent organic pollutants such as polychlorinated biphenyls, polycyclic aromatic hydrocarbons, brominated flame retardants, and organochlorine pesticides in the composite sample have been published elsewhere (Legler et al., 2011). Several metals were measured in three samples and were present at variable levels depending on the depth of sampling. Mean concentrations of lead, cadmium, zinc, and copper were 1751, 29, 3060, and 761 mg/kg soil respectively. A full overview of metal measurements and soil parameters is given in datasheet Table S1 in Supplementary Material.

ECOTOXICITY TEST

Treatments consisted of two separate dilution series, the first being a series of diluted whole environmental Vlagheide soil, and the second being a series of LUFA 2.2 reference soil spiked with the extract derived from composite Vlagheide soil sample according to Droge et al. (2006) using acetone as solvent.

For the 100% extract treatment, an extract derived from 100 g (d.w.) Vlagheide soil was spiked-in 100 g (d.w.) LUFA 2.2 soil. For the 100% soil treatment, we used sieved and freeze-dried Vlagheide soil. Using LUFA 2.2 soil for dilution, we employed a dilution factor of 2.5 to prepare five additional treatments within each series, resulting in a 40, 16, 6.4, 2.56, and 0% dilution of both undiluted extract and soil treatments. The control (0%) sample was 100% LUFA 2.2 soil in case of the Soil sample dilution series, whereas the control for the Extract dilutions was a solvent control consisting of LUFA 2.2 soil including acetone (the solvent) in an identical amount as was used for the extract dilution series. In this way we were able to normalize for the effect of the solvent during the spike-in procedure of the extract samples (Droge et al., 2006).

Preparation of the test soils and experimental set-up was done following the standard ISO protocol 11267 (ISO, 1999), with four biologically replicated test jars per treatment. The standard ISO test procedure for inhibition of reproduction after 28 days was followed. In parallel, we exposed 10 animals in each of four biological replicate jars per treatment, for 4 days on top of a compressed layer of test soil. These arthropods were snap frozen in liquid nitrogen immediately after exposure so that total RNA could be isolated (see below) for subsequent Microarray and QPCR analysis.

A logistic model was fitted to estimate toxicity end points no-effect concentration (NOEC) and a significant sublethal decrease in reproduction (DiR) for extract samples and soil samples of 50% (further deduced as DiR).

RNA PREPARATION, AMPLIFICATION, LABELING, AND HYBRIDIZATION

Samples (Lufa control, NOEC; sublethal effects DiR) were subjected to RNA extraction using the SV Total RNA Isolation System (Promega) according to the manufacturer's instructions. Agilent

RNA Spike-In Kit (Agilent Technologies) was used to prepare Spike A Mix and Spike B Mix in order to normalize the hybridization measurements. Approximately, 50 ng input of total RNA was used for amplification and labeling with Agilent's Low Input Quick Amp Labeling Kit. The RNA was reverse transcribed into cDNA and treated with a T7 RNA polymerase to incorporate cyanine 3- or cyanine 5-labeled CTPs in the synthesized cRNA, which was purified with RNeasy (Qiagen) and quality controlled using spectrophotometric measurements on a NanoDrop 2000 (Thermo scientific). Hybridization of 8 × 15 K format microarray slides was performed with 300 ng cyanine 3-labeled cRNA and 300 ng cyanine 5-labeled cRNA according to manufacturer's protocol (Agilent). The design of the microarray is described by Nota et al. (2009) and details can be found under Gene Expression Omnibus (GEO) platform number GPL7150. A replicate reference design was used (**Figure 1**) so that each treatment sample was competitively hybridized against a control sample (Lufa 2.2). The design included dye-swapped biological replicates. Two slides of the 8 × 15 K Agilent microarray platform, containing 5069 unique gene probes in triplicate, was used for soil samples and extract samples (**Figure 1**). Hybridization was performed at 65°C for 17 h rotating at 10 rpm in an incubator. Following the hybridization, the slides were washed using Gene Expression wash buffers and scanned on an Agilent DNA Microarray Scanner. The microarray scan images were preprocessed with Feature Extraction software (version 10.5.1.1.) and the obtained Fold changes were subjected to further statistical analyses. The data was submitted to NCBI's GEO and can be retrieved under accession number GSE37154.

MICROARRAY DATA ANALYSIS

Statistical analysis of microarray data was performed using the Limma package in R environment (version 2.13.0, Wettenhall and Smyth, 2004). The data were normalized to account for dye bias with the global loess method and the significance of gene expression was verified for each of the soil and extract dataset by a modified *t*-test using Bayesian statistics two-way analysis of variance (ANOVA) with factors sample type and treatment as main factors. All calculated probabilities were corrected for multiple testing using Benjamini and Hochberg's false discovery rate procedure at the level of $p < 0.05$ (Benjamini and Hochberg, 1995). Each probe was assigned a mean $\log_2$ expression ratio and an adjusted *p*-value. Gene annotation was performed in R using a Blast2go script. Subsequently, a GO term Enrichment Analysis was performed by applying the TopGO algorithm on significant gene lists (Alexa et al., 2006; de Boer et al., 2011a) to assess which biological processes, molecular functions, and cellular components were mostly affected. The TIGR MultiExperiment Viewer (TIGR Mev version 4.6.2; Saeed et al., 2006) was used to perform cluster analysis in order to define groups of genes that share common patterns of expression. Hierarchical clustering was done using Euclidean distance and average linkage method. Heat maps were used to represent the data. A general linear model was used to investigate the interaction between factors affecting the variability in the data with the factors Sample type (Extract or Soil), treatment (NOEC or DiR), and the Sample X Treatment interaction according to de Boer et al. (2011a). Finally, a Principal Component Analysis (PCA) in TIGR Mev allowed the identification of factors that most contribute to the variability in the data.

Extract

Array 1	Array 2	Array 3	Array 4
control	control	NOEC	NOEC
NOEC	NOEC	control	Control
control	control	DiR	DiR
DiR	DiR	control	control
Array 5	Array 6	Array 7	Array 8

Soil

Array 1	Array 2	Array 3	Array 4
control	control	DiR	DiR
DiR	DiR	control	Control
control	control	NOEC	NOEC
NOEC	NOEC	control	control
Array 5	Array 6	Array 7	Array 8

FIGURE 1 | Hybridization scheme for gene expression analysis. On each of the arrays, a test sample (NOEC or DiR) of either Soil or Extract is hybridized against a control sample (Lufa 2.2 for Soils; acetone spiked-in Lufa 2.2 for extracts). Green Cy-3, Red Cy-5.

QUANTITATIVE PCR ANALYSIS

Quantitative PCR (QPCR) was performed using a selected group of genes according to de Boer et al. (2011b) on a Biomark HD system (Fluidigm). Information on gene description and PCR primer sequence can be found in **Table 2** of the results section. Quantitative analyses of cycle threshold (C_t) values were performed with the software package Genex Light 4.3.5 (Multi ID analysis) according to de Boer et al. (2011b). First, the three technical replicates were averaged over each sample. Efficiency corrections were applied on the mean C_t values using PCR efficiency values previously established and published by de Boer et al. (2011b). Subsequently, gene expression values were assessed relative to an internal reference by normalization with the geometric mean of the housekeeping genes SDHA and YWAZ (de Boer et al., 2009). Finally, the Log2 transformed normalized gene expression values were subjected to statistical analysis in SPSS version 17 (IBM). Linear regression was applied to assess dose-dependency. The residuals were tested for normality using a Kolmogorov–Smirnov test. A one-way ANOVA was applied to test whether expression values were significantly different between sample type (Extract or Soil).

RESULTS

ECOTOXICITY TEST

The effects of the original soil from the landfill and its extract on reproduction of *Folsomia candida* were assessed in an ISO standardized 28 days toxicity test. **Figure 2** shows the dose-response curves resulting from exposure to Soil (blue) and Extract (red) samples. The different shapes of the lines clearly indicate a difference in toxicity between the two kinds of samples. The original soil samples show a clearer dose-dependence and appear to be more toxic compared to the organic extracts.

The NOEC and the 50% DiR were deduced for each sample type and a spike-in concentration of 6.4% was taken as NOEC for both Extract and Soil exposures. Furthermore, DiR was observed at 100% spike-in Extract, whereas 16% of spiked-in Soil concentration did not significantly deviate from the 50% DiR estimated by the logistic model for Soil toxicity and was thus taken for further investigations at the molecular level.

MICROARRAY ANALYSIS

Figure 3 shows the number of genes that were differentially expressed in response to toxic exposure in Soil and Extract samples. Regarding the Soil exposure 109 genes were significantly regulated in response to both NOEC and DiR concentration, of which 76 were up-regulated and 33 were down-regulated. Moreover, 354 genes were only regulated at the NOEC level and 521 genes were only regulated at the DiR level. Intriguingly, exposure to Extracts caused differential expression of an increased amount of genes as compared to Soil exposure. As much as 1581 genes were significantly regulated in response to both NOEC and DiR, of which 747 were up-regulated and 834 down-regulated. Moreover, 613 genes were only regulated at NOEC level and 536 exclusively regulated at DiR level in the Extracts.

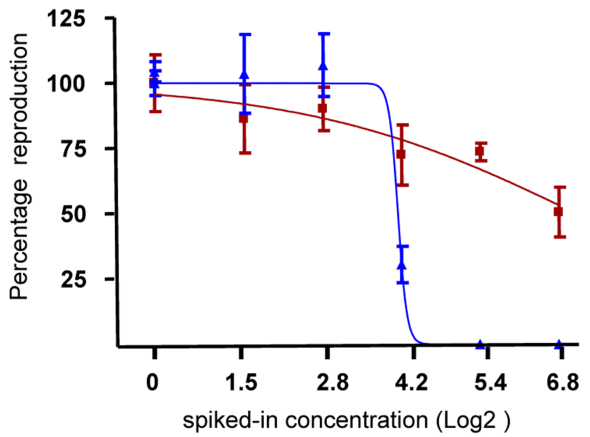


FIGURE 2 | Effect of soil and extract on reproduction of *F. candida*. Red dots indicate the number of *F. candida* juveniles in the jars after 28 days exposure to six dilutions of soil samples. Blue dots indicate the number of juveniles retrieved after 28 days exposure to dilutions of extract samples. The lines indicate the dose–response curves derived from a logistic model. *x*-Axis, Log2 transformed spiked-in concentrations; *y*-axis, percentage reproduction scaled to the control samples (set at 100%).

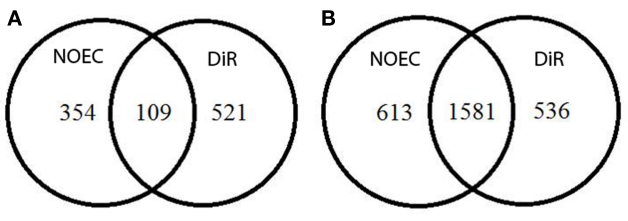


FIGURE 3 | Venn diagrams showing the number of genes responding to NOEC and DiR levels of toxicants in the Soil exposure (A) and in Extract exposure (B).

Subsequently, the two-way ANOVA with factors Sample type (Extract or Soil) and Treatment (NOEC or DiR) was applied to identify genes only affected by Sample type or Treatment, and to assess whether genes exerted a Sample X Treatment interaction. In total 1929 were affected onle by Sample type (Soil, Extract), while 396 genes were exclusively regulated by Treatment (NOEC, DiR). In addition, 400 genes showed a significant Sample x Treatment interaction. Heat maps of significantly regulated genes for these three factors are represented in Supplementary data.

Table 1 shows the Gene Ontology terms for biological processes enriched in the significant gene lists for treatment, sample type, and the interaction.

Some of the biological processes that were differentially affected by different treatments (low and high level of exposure) are lipid metabolism, response to nutrient level, response to chemical stimulus. The exposure to contaminated samples had a strong impact on 368 genes in both kinds of samples (soil and extract). Glucosyl glucuronosyl transferases (Fcc00734, biological process "lipid metabolism") and superoxide dismutase (Fcc01344, biological process "response to nutrient level"), for example, were both affected by the exposure, being differentially expressed in response to different toxic levels.

On the other side, other biological processes were found to be affected mainly by the difference in sample type: 1929 genes showed different responses between soil and extract samples. These genes are referable to biological processes such as RNA processing, biosynthetic processes, translation, and translational elongation. For instance, isopenicillin-*N*-synthetase (Fcc00057), translation initiation factor (Fcc00062), and ribosomal proteins (Fcc00498, Fcc00410) were differentially affected by the exposure in soil and extract samples.

Figure 4 shows an acyclic graphs resulting from the enrichment analysis referring to biological processes affected in the context of the interaction between sample type and level of exposure. The most significant ones are lipid metabolic processes (such as fatty acid metabolism and fatty acid oxidation), cellular biosynthetic processes, organic acid metabolic processes, regulation of body fluid, vascular development, and response to wounding.

Some significant genes were found to respond to the toxic exposure in either kind of sample type, at both high and low levels of exposure. Up-regulation of heat-shock proteins (Fcc05793), ubiquitins (Fcc02887), biotransformation enzymes (Fcc01651, Fcc04073, Fcc5260), and components of the antibiotic biosynthetic pathway (Fcc00057, Fcc00170, Fcc05968) was observed. On the contrary, hedgehog, antimicrobial genes, and molecular chaperones were down-regulated.

A PCA was then performed in order to explore how the two factors affect each other and to identify prevalent expression profiles among samples. **Figure 5** is the output of the PCA and shows that the 58% of the variation in the data can be explained by the two factors sample type and exposure level.

QUANTITATIVE PCR ANALYSIS

From the microarray analysis it became apparent that some genes show a dose-dependent regulation in response to both soil

Table 1 | Gene Ontology (GO) terms for biological processes that are over represented in the lists of significant transcripts and their *P*-values as obtained using the R package topGO (Alexa et al., 2006).

GO ID	GO term	*p* Value	# in GO term	# Significant
TREATMENT				
GO:0006629	Lipid metabolic process	1.78e-06	173	30
GO:0009605	Response to external stimulus	3.80e-07	103	23
GO:0009991	Response to extracellular stimulus	5.02e-07	36	13
GO:0031667	Response to nutrient level	9.71e-08	32	13
GO:0042221	Response to chemical stimulus	0.00052	207	28
GO:0007584	Response to nutrients	1.57e-06	23	10
SAMPLE TYPE				
GO:0006996	Organelle organization	0.150994	292	126
GO:0042254	Ribosome biogenesis	0.000193	56	36
GO:0006364	RNA processing	0.000367	38	26
GO:0009058	Biosynthetic process	0.002262	363	170
GO:0009059	Macromolecule biosynthetic process	0.003222	180	90
GO:0006412	Translation	4.86e-05	128	73
GO:0006414	Translational elongation	6.04e-06	28	23
INTERACTION				
GO:0044249	Cellular biosynthetic process	0.001781	210	24
GO:0019395	Fatty acid oxidation	0.001658	15	5
GO:0006629	Lipid metabolic process	2.77e-05	173	25
GO:0001944	Vasculature development	0.000117	26	8

Treatment, No-Effect Concentration (NOEC) versus Decrease in Reproduction (DiR); Sample type, Extract versus Soil; Interaction, Treatment X Sample Type.

and extract. For instance, CYP6N4v1 (Fcc01651) transcription, significantly increased two fold with increasing exposure level. We therefore decided to assay this gene using a QPCR assay in all extract- and soil sample concentrations (expression data are provided in datasheet Table S2 in Supplementary Material). **Figure 6** shows CYP6N4v1 expression as all exposure levels in extract (**Figure 6A**) and soil (**Figure 6B**) samples. The QPCR profiles significantly correlated with the dose-dependent induction as observed in the microarray data (Spearman's Rho 0.74, $p < 0.05$). Linear regression analysis of expression level with exposure showed a highly significant ($p < 0.001$) correlation between CYP6N4v1 gene expression induction and increased exposure level (extract $R = 0.89$; soil $R = 0.92$), while the residuals did not significantly differ from normal distribution (data not shown). Subsequently, we decided to assay more genes related to biotransformation, previously identified by de Boer et al. (2011b) (expression data are provided in datasheet Table S2 in Supplementary Material). They are summarized in **Table 2**. Interestingly, transcriptional activation of three additional cytochrome P450s (CYP6N3v1, CYP2P3, CYP9/6) showed highly significant correlations with both extract and soil concentration. Particularly, CYP6N3v1 gene activation showed the highest correlation ($R = 0.99$, extract **Figure 6C**; $R = 0.96$, soil **Figure 6D**) with increased exposure levels (**Table 2**), despite the fact that this gene did not show dose-dependent transcriptional activation in the microarray experiment. This was probably due to detection limitations of the microarray technology, because the hybridization intensities were not above background levels. Moreover, alcohol dehydrogenase, deoxynucleoside kinase, transcription factor CCCTC-binding protein, phosphoserine amino transferase, and haloacid dehalogenase-like hydrolase showed highly significant dose-dependent transcriptional regulation (**Table 2**).

Finally we identified nine genes that showed a significant (one-way ANOVA, $p < 0.02$) sample type effect. Among these, the gene ubiquitin ligase E3 alpha (Fcc06380) indicated an increased general stress–response in soil samples. Also, developmental processes were increasingly affected due to the increased transcriptional activation of crossveinless-2 BMP binding protein (Fcc04834) and LMBR1 domain containing 2 (Fcc03839, associated with hedgehog transcriptional regulation). Although, direction of regulation (up- or down-regulation) was in concordance between the QPCR data and microarray data for most of the genes, we could be confirmed significant sample type-specific regulation for LMBR1 domain containing 2, ABC transporter (Fcc06002), Laminin A (Fcc00086). To conclude, we were able to indentify and confirm treatment specific and sample type specified gene expression assays by considering both (microarray and QPCR) gene expression analysis platforms.

DISCUSSION

Here we presented a full ecotoxicogenomics assessment of a complex environmental sample and identified large differences between the unprocessed, ecologically more relevant, raw soil, and an organic total extract. The extract samples showed a clearly different level of toxicity when compared with the original soil sample. As such, extract samples may have only weak predictive power to estimate the actual toxicity status in the field. The soil sample is highly contaminated with organic micropollutants such as polychlorinated biphenyls, organochlorine pesticides, and one brominated flame retardant (Legler et al., 2011). Heavy metals

Table 2 | QPCR assays and results of statistical analysis.

Code	Collembase	Function	Forward	Reverse	Efficiency	Linear regression$		*p* Value ANOVA$$
						Extract	Soil	
PHE1	Fcc05839	Alcohol dehydrogenase	GCGCGTAATCGATGCTAGTGT	GCCCAGGCCTCTTTAATGAAA	2.2	0.75***	0.83***	0.001
PHE13	Fcc06009	Putative deoxynucleoside kinase	TTGTCGGGAGGTCAGAAGTAGC	CACGGCACTCAATGGATCATC	1.9	0.47*	0.65***	0.008
PHE21	Fcc05124	CCCTC-binding factor-like	TGCATCACTACCTTGCCGAAG	CTGCGTAGTTTTCCCATGCTG	2.0	0.68***	0.83***	0.253
PHE23	Fcc04316	Phosphoserine aminotransferase	GCAGTGCCATTAAACCTTATGAGC	GTTGCCTTTCCGTATTTGGTTG	2.1	0.89***	0.94***	0.195
CYP5	Fcc00015	CYP9/6	TTCGATCACTTTTCGGACCG	TCATGTCTTTCCACTCTTGTCCC	1.9	0.69***	0.58**	0.035
CYP11	Fcc01780	CYP6N3v2	GCGTTAAAAGCGAGGCAAGA	GCGATATCCACGTTCGAATTGT	1.9	0.99***	0.96***	0.235
CYP19	Fcc00390	CYP2P3	CATTCCTTTACATGGCGAAATACC	CGATCATTTAAACATGGCTGTCTG	1.9	0.93***	0.89***	0.116
CYP23	Fcc01651	CYP6N4v1	TTCCATGCAAGTCATCACATCAG	CGGAAACACAAAGATTCGTTCTG	1.9	0.89***	0.92***	0.387
CD49	Fcc00343	Haloacid dehalogenase-like hydrolase	TATTTAATCCGCGACATGGGTC	AATAATGCCGTCCTGGTGTAGC	2.0	0.77***	0.56**	0.519
PHE5	Fcc04834	Crossveinless-2 BMP binding	AAGAGGTGGCCCAGTAAGTCATC	GGTTCAGGACCTTTCACTTCTTCTC	2.0	NS	NS	<0.001
PHE7	Fcc03839	LMBR1 domain containing 2	AAAAAGCGCAGACAGCCAGA	CTCCCAACATCCAGTCCTTTCA	2.0	NS	NS	<0.001
PHE11	Fcc06002	ABC transporter	TTGTCGGGAGGTCAGAAGTAGC	CACGGCACTCAATGGATCATC	1.9	0.48*	0.74***	<0.001
CPF3	Fcc06377	Exportin 7	CAGTGTTGTGAAATCGAGGAAAGC	CACCGACGAATCGTCAGATGA	1.9	0.49*	NS	<0.001
CPF5	Fcc06380	Ubiquitin ligase E3 alpha	TTCCACGACGAAGTCCTTGATC	CACGGTCAAAACAGAGGATGCT	1.9	NS	NS	0.004
CPF13	Fcc06453	NADH dehydrogenase subunit 5	TTGTCACACCCCAGAATCCCTA	TTAATTATCTCAGCCCCACCACCT	1.9	NS	0.47*	0.012
CPF27	Fcc06524	60S ribosomal protein L30	ATCGAGTTGGGCACTGCTTGT	GCCGAGGATCTAATGATATCGGA	1.9	0.78***	0.71***	<0.001
CPF51	Fcc03128	No significant hit	ACTATTTTTCATACGGCCTCCTGC	AATAGCGAGAAAATCTCCAAAGTCG	1.9	0.47*	NS	0.02
CD89	Fcc00086	Laminin A	GAGTGGAGCAGGACGCTGTAAT	GAACTTGGATTTAACTCCGTCGC	1.9	0.47**	0.67***	0.001

*$Linear regression analysis was performed to investigate dose-dependent transcriptional regulation at significance level *p < 0.05, ** p < 0.01, or *** p < 0.001.*

$$One way ANOVA was performed to investigate significant differential regulation between extract- or soil exposure.

Collembase, EST numbers from Collembase.org. Effiency, PCR effiency resulting from a standard curve analysis.

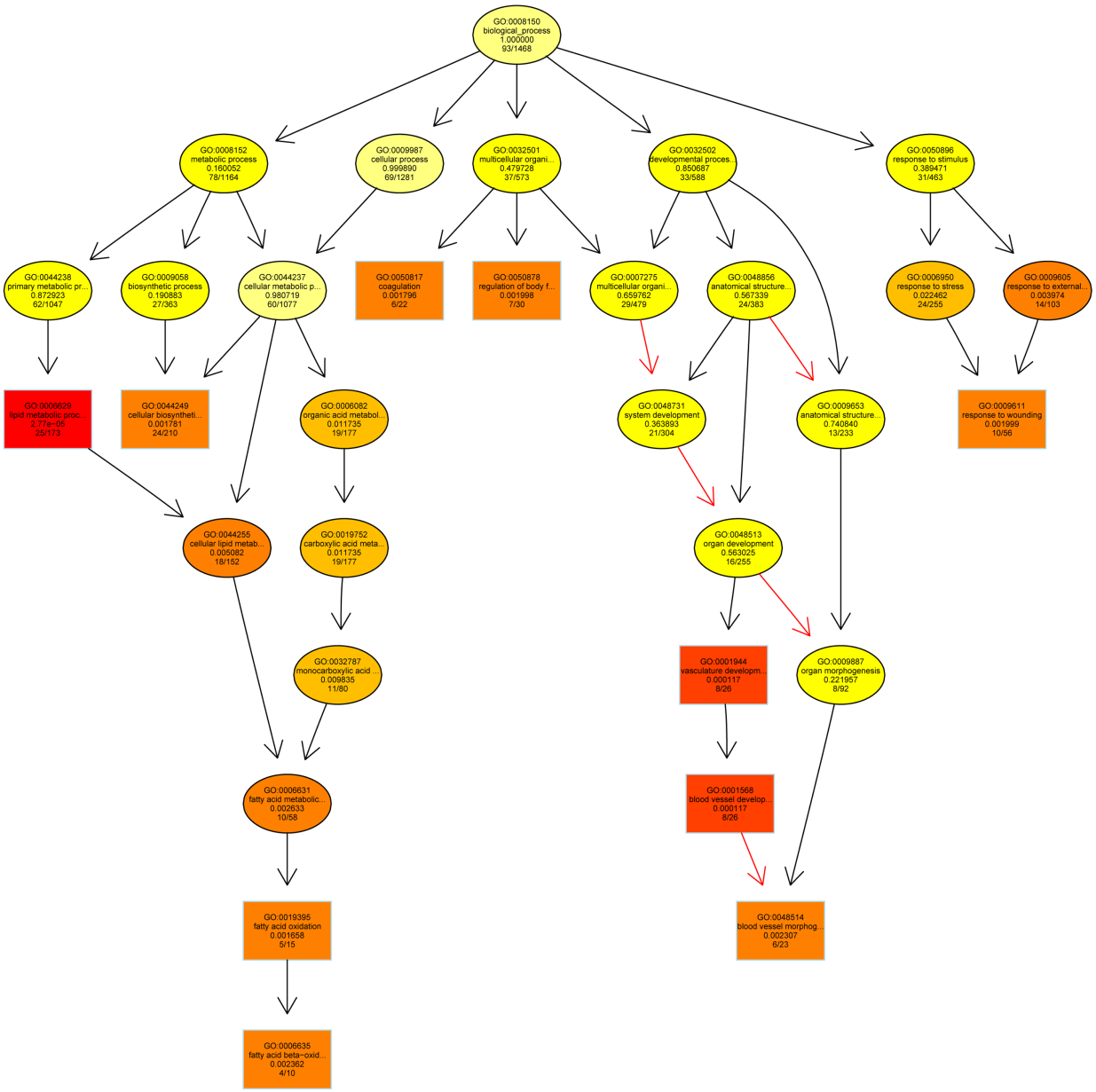


FIGURE 4 | Acyclic graph resulting from the Enrichment Analysis and showing the biological processes mostly affected by an interaction between sample type (Soil, Extract) and treatment (NOEC, DiR). Increasing coloring toward red represents increasing significance levels. Each sphere contains GO ID, description, significance level, and ratio regulated: total genes in GO ID.

are also elevated in the soil sample, including lead, copper, cadmium, and zinc. The exact chemical composition of the extract is not known, however due to the nature of the organic extraction procedure, metals are expected to be removed from the extract (Hubert et al., 2000). Moreover, microarray analysis generated important mechanistic information that can explain this discrepancy in toxicological effects. In combination with the QPCR data, we can conclude that developmental processes, fatty acid metabolism, and defense processes are adversely affected depending on which sample type was analyzed. Indeed, most of the genes (1929) were regulated in response to sample type. This is reflected in the PCA graph (**Figure 5**) where the first principle component explaining 42% of the variance divides the samples into either Extract or Soil. As much as 395 genes showed treatment-specific-regulation. Among them are genes (CYP6N4v1) that show a highly significant dose-dependent transcriptional activation, which was confirmed by QPCR analysis. Such genes may proof invaluable as genetic markers for soil pollution assessment.

ECOLOGICAL RELEVANCE

The ISO standardized *Folsomia* test resulted in a consistent difference in toxicity between soil samples and organic extracts, with

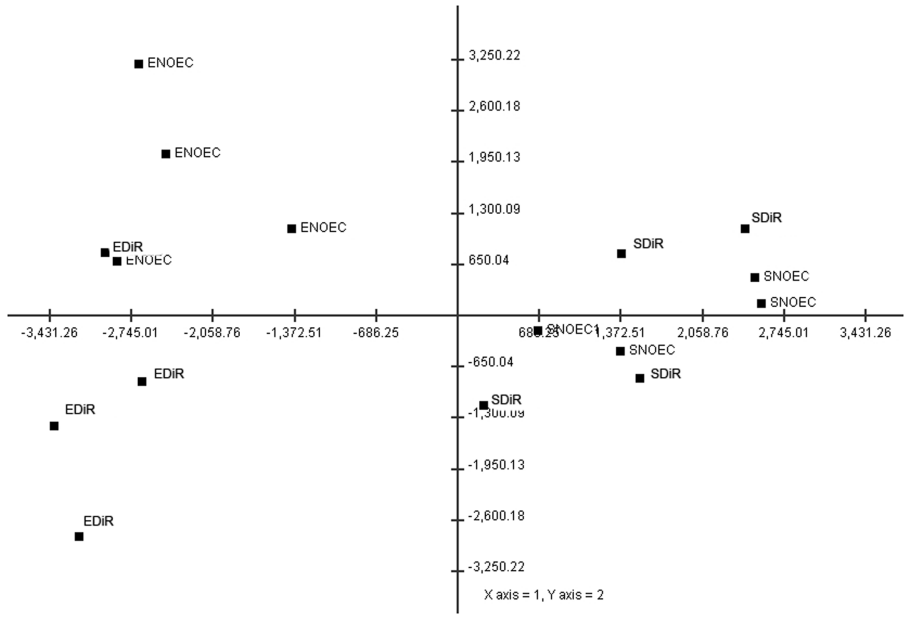


FIGURE 5 | Distribution of the samples in the space defined by two main components (axis) resulting from a Principal Component Analysis. The labels indicate the sample type; ENOEC, extract no-effect concentration; SNOEC, soil no-effect concentration; EDiR extract 50% decrease in reproduction; SDiR soil 50% decrease in reproduction. The localization of the points in the space suggests that these two factors are mainly responsible for the distribution of the data.

the former being much more toxic than the latter. This confirms the concern raised in earlier reports that toxicity tests based on extracts may generate uncertain levels of protection due to the fact that they do not address bioavailability of the original sample (Fent, 2004). The reason for this result could lie in differences in the availability of the toxic compounds between the original landfill soil and the natural soil spiked with the organic extracts, or in the loss of some of these substances during the extraction procedure. The extraction method using acetone/dichloromethane does not remove all toxic compounds; therefore it is likely that part of the soil toxicity is due to polar substances such as hydrophilic xenobiotic compounds and/or heavy metals.

MOLECULAR MECHANISMS

A considerable difference was found between the two kinds of samples, concerning the number of genes affected by the exposure. This is in accordance with a recently published survey of ecotoxicogenomics studies by Van Straalen and Feder (2012), who showed that in most studies expression profiles at sublethal toxic effects of 10% DiR (EC10) and EC50 cluster together, while the main differences can be observed between sample type. Soil samples evoked gene regulatory changes in more than one order of magnitude fewer genes than the extract samples although we aimed to assess the two sample types at similar toxicity levels (NOEC and DiR of around 50% reduced reproduction). Due to the steepness of the logistic model through the soil toxicity data we may have chosen a soil sample that exerts more toxic effects than the DiR in Extact, although DiR in Soil did not significantly deviate from 50% DiR deduced from the logistic model. It is worth to mention that Nota et al. (2009) studied the effects of phenanthrene on *F. candida* and also found a smaller number of differentially expressed genes in response to high toxic concentration. Very recently, we obtained a similar result in a toxicogenomic study assessing stress–responses in *F. candida* exposed to the anti inflammatory drug Diclofenac (Roelofs et al. unpublished data). The explanation for this drop in transcriptional regulation needs further investigation. We speculate that higher toxic levels induce intense detoxification responses in the organisms, thus leaving less energy for other less essential processes, although Timmermans et al. (2009) showed that increased desiccation stress strongly increased the number of up- and down-regulated genes.

Some interesting biological processes were influenced by the exposure to contaminated samples. The GO term "lipid metabolic process" includes the chemical reaction and pathways involving all kinds of lipids. It is the biological process most significantly affected in this study. Within living systems, polar lipids have a fundamental structural role, being the main constituents of biological membranes. Furthermore, apolar lipids act as a reserve of energy. Lipid metabolism is regulated in order to ensure the correct balance between degradation and synthesis of lipids, according to the needs of the cells and of the whole organism. When facing environmental stress, an organism will activate a series of stress–response mechanisms in order to face the new conditions. These processes require energy and this might explain the changes in lipid metabolism suggested by the gene expression pattern. In fact, within this GO term category, we found a series of significant genes linked to lipid metabolism and transport. For example, Enoyl-CoA hydratase was found to be up-regulated in response to toxicant exposure: this enzyme is very efficient in metabolizing fatty acids to produce acetyl CoA and energy (Agnihotri and Liu, 2003), so

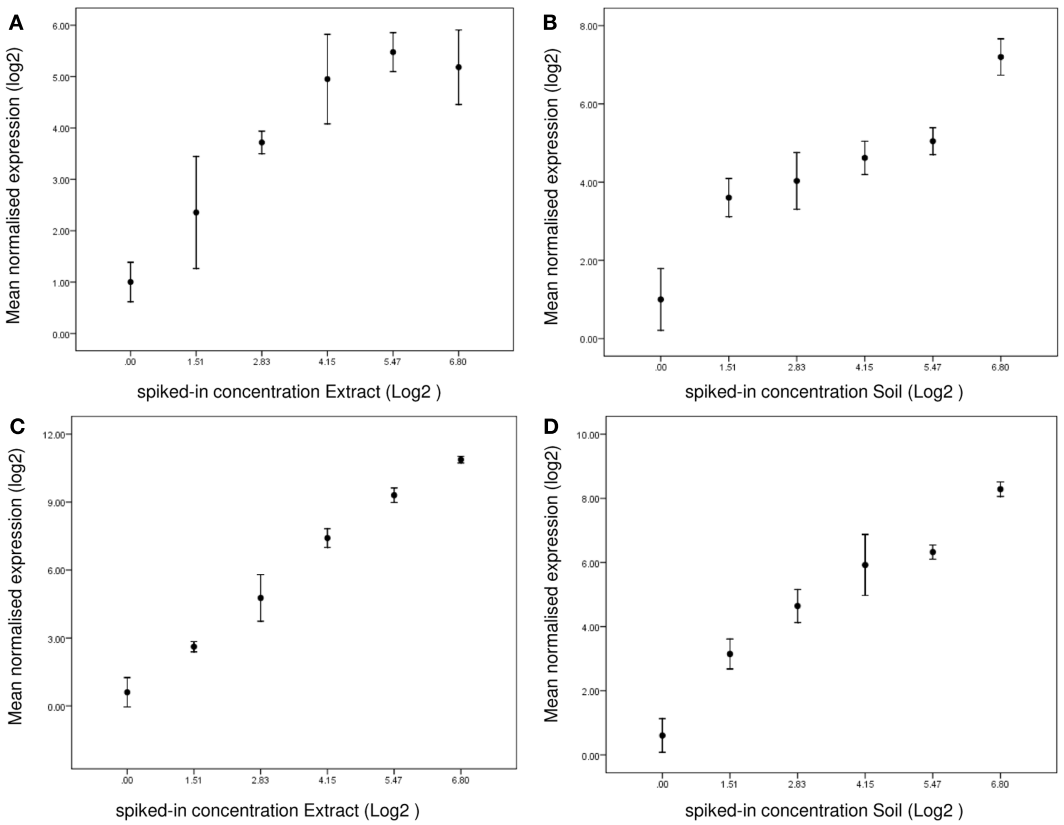


FIGURE 6 | Linear regression of gene expression as deduced from the QPCR measurements. (A) CYP6N4v1 expression in response to Extracts; **(B)** CYP6N4v1 expression in response to Soils; **(C)** CYP6N3v2 expression in response to Extracts. **(D)** CYP6N3v2 expression in response to Soils. X-axis, Log2 transformed spiked-in concentrations. Y-axis, Log2 normalized gene expression.

it might be up-regulated in order to sustain the energy-requiring processes. On the other side, fatty acid desaturase, that causes increase of unsaturated bonds in fatty acids of membranes (Los and Murata, 1998), was found to be down-regulated, probably because the double bonds in fatty acids are a target of oxidative stress. Another interesting biological process is "response to wounding," which includes any process resulting from a stimulus indicating damage of the organism. Within this group we found, for example, hedgehog, a developmentally active transcription factor that plays a vital role during early embryonic stages (Tabata and Kornberg, 1994). The hedgehog signaling pathway is intimately linked to cell growth and differentiation, so this protein could also be involved in the healing response. Vascular development is another significantly affected biological process. Springtails do not have a vascular system; thus it is difficult to translate this biological process to an invertebrate response. In fact, in this category we find hypoxia-inducible factor, a transcription factor that responds to changes in the level of available oxygen and mediates responses to hypoxia (Jiang et al., 1996). Furthermore, matrix metalloproteinase are represented in this category. This family of enzymes hydrolyze components of the extracellular matrix and play a central role in many biological processes, such as embryogenesis, normal tissue remodeling, wounding, etc (Nagase and Woessner, 1999). When investigating the single genes that fall into a GO term, it is interesting to notice that many of them are actually annotated to more than one term, and this in fact can be seen as a reflection of the number of interconnections that exist between different kinds of stress–responses. For example, within the context of lipid metabolism, a series of genes are annotated that are responding to different kinds of stress: glucosyl glucuronosyl transferases (involved in phase II metabolism of xenobiotics), glutathione *S* transferases, CYP450, catalases, dismutases, and other gene products found, by other authors, to be associated to oxidative stress or exposure to toxicants.

Some significant genes were found to both high and low levels of exposure in either kind of sample, soil and extract. Heat-shock proteins, that are part of the general stress–response and have a chaperone function (Feder and Hofmann, 1999), were always found to be up-regulated. A similar result was observed for ubiquitins, regulatory proteins that are involved in the degradation of damaged or unneeded proteins (Glickman and Ciechanover, 2002). An increase in expression of enzymes responsible for biotransformation and detoxification reactions was also observed in all cases of exposure. Mono-oxygenases such as cytochrome P450 for phase I, conjugation enzymes for phase II and ABC transporters for phase III were all found to be significantly up-regulated in contaminated samples compared to the control ones. These enzymes are responsible for the detoxification of organic

compounds, so their induction is expected in case of exposure to organic xenobiotic substances (Xu et al., 2005). Interestingly, two components of the antibiotic biosynthetic pathway were found to be up-regulated in soil samples: aminoadipyl–cysteinyl–valine synthetase and isopenicillin-*N*-synthetase. Cathepsins, a family of proteases that break apart other proteins and might play a role in apoptosis, also resulted overexpressed, in the extract samples. Hedgehog was down-regulated following exposure to the environmental samples, so the stress caused by toxic exposure is likely to influence important signaling pathways and therefore the correct development of the organism. Some antimicrobial genes were also found to be down-regulated after exposure, suggesting that pollution might adversely affect the insect immune system and increase its susceptibility to invading pathogens. Finally, down-regulation of genes coding for proteins with a folding function is likely to affect directly or indirectly important cellular structures and functions.

In conclusion the microarray experiment shows that transcriptomics data can add relevant information on the nature of the compounds that cannot be recovered by traditional bioassays. We showed this in a recent study on aged copper contamination in an agricultural field (de boer et al., 2012). In this case, the patterns of gene expression suggest that the adverse biological effect of Vlagheide soil is partly due to organic compounds inducing xenobiotic metabolism. These compounds remain active after solvent extraction and induce a similar set of genes compared to the intact soil samples. However, the dose-dependence of the extracts is less clear maybe due to altered bioavailability. In addition to organic compounds, toxicity of the field soils is also due to polar components. Heavy metals would be the most likely factor, because elevated levels were measured at the Vlagheide site. However, we did not recover gene expression profiles indicative of specific single metals, as in the study of Nota et al. (2010). This might be due to interactive effects of metal mixtures in field soils, as also described by Nota et al. (2010). Finally our study illustrates that it is essential to link transcriptomics bioassays to traditional ecotoxicity tests, because only in this way can the exposure levels applied in gene expression studies linked to defined phenotypic effects.

Finally, we demonstrate that a number of QPCR assays exert a wide dynamic range of transcript quantification activated in a dose-dependent manner. This dynamic range can be observed around important endpoint such as the NOEC and 50% DiR. This opens the possibility to link gene expression levels to adverse effects at the organismal level. Such molecular bioassays may become very useful in future soil quality testing, because they are fast and diagnostic for the type of toxicity. Future work will focus on thorough validation of selected gene expression assays using a wide range of environmental soil samples containing different classes of compounds.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/Toxicogenomics_/10.3389/fgene.2012.00085/abstract

Table S1 | Measurement of metals and soil parameters in three Vlagheide samples.

Table S2 | Log2 normalized gene expression values from biological replicates in the QPCR assays.

SampleTypeSign, heat map significant genes factor Sample Type (extract or soil).
TreatmentSign, heat map significant genes factor Exposure level (NOEC or DiR).
InteractionSign, heat map significant genes Sample Type × Treatment interaction.

REFERENCES

Agnihotri, G., and Liu, H. W. (2003). Enoyl-CoA hydratase: reaction, mechanism, and inhibition. *Bioorg. Med. Chem.* 11, 9–20.

Alexa, A., Rahnenfuehrer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22, 1600–1607.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Methodol.* 57, 289–300.

de Boer, M. E., de Boer, T. E., Marien, J., Timmermans, M., Nota, B., van Straalen, N. M., Ellers, J., and Roelofs, D. (2009). Reference genes for QRT-PCR tested under various stress conditions in *Folsomia candida* and *Orchesella cincta* (Insecta, Collembola). *BMC Mol. Biol.* 10, 54. doi:10.1186/1471-2199-10-54

de Boer, T. E., Birlutiu, A., Bochdanovits, Z., Timmermans, M., Dijkstra, T. M. H., Van Straalen, N. M., Ylstra, B., and Roelofs, D. (2011a). Transcriptional plasticity of a soil arthropod across different ecological conditions. *Mol. Ecol.* 20, 1144–1154.

de Boer, M. E., Berg, S., Timmermans, M. J. T. N., Den Dunnen, J. T., Van Straalen, N. M., Ellers, J., Roelofs, D. (2011b). High throughput nano-liter RT-qPCR to classify soil contamination using a soil arthropod. *BMC Mol. Biol.* 12, 11. doi:10.1186/1471-2199-12-11

de boer, T. E., Taş, N., Braster, M., Temminghoff, E. J., Röling, W. F., and Roelofs, D. (2012). The Influence of long-term copper contaminated agricultural soil at different pH levels on microbial communities and springtail transcriptional regulation. *Environ. Sci. Technol.* 46, 60–68.

Droge, S. T. J., Paumen, M. L., Bleeker, E. A. J., Kraak, M. H. S., and van Gestelt, C. A. M. (2006). Chronic toxicity of polycyclic aromatic compounds to the springtail *Folsomia candida* and the enchytraeid *Enchytraeus crypticus*. *Environ. Toxicol. Chem.* 25, 2423–2431.

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.

Fent, K. (2004). Ecotoxicological effects at contaminated sites. *Toxicology* 205, 223–240.

Glickman, M. H., and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* 82, 373–428.

Gong, P., Guan, X., Inouye, L. S., Deng, Y., Pirooznia, M., and Perkins, E. J. (2007). Transcriptomic analysis of RDX and TNT interactive sublethal effects in the earthworm *Eisenia fetida*. *BMC Genomics* 9, S15 doi:10.1186/1471-2164-9-S1-S15

Hubert, A., Wenzel, K. D., Manz, M., Weissflog, L., Engewald, W., and Schuurmann, G. (2000). High extraction efficiency for POPs in real contaminated soil samples using accelerated solvent extraction. *Anal. Chem.* 72, 1294–1300.

ISO, Soil Quality. (1999). *Inhibition of Reproduction of Collembola (Folsomia candida). ISO Guideline 11267*. Switzerland: International Standardization Organization, 1999–1916.

Jiang, B. H., Semenza, G. L., Bauer, C., and Marti, H. H. (1996). Hypoxia-inducible factor 1 levels vary exponentially over a physiologically relevant range of O-2 tension. *Am. J. Physiol.* 271, C1172–C1180.

Legler, J., van Velzen, M., Cenijn, P. H., Houtman, C. J., Lamoree, M. H., and Wegener, J. W. (2011). Effect-directed analysis of municipal landfill soil reveals novel developmental toxicants in the zebrafish *Danio rerio*. *Environ. Sci. Technol.* 45, 8552–8558. doi:10.1186/1471-2164-10-160

Los, D. A., and Murata, N. (1998). Structure and expression of fatty acid desaturases. *Biochim. Biophys. Acta* 1394, 3–15.

Menzel, R., Swain, S. C., Hoess, S., Claus, E., Menzel, S., Steinberg, C. E. W., Reifferscheid, G., and Stuerzenbaum, S. R. (2009). Gene expression profiling to characterize sediment toxicity – a pilot study using *Caenorhabditis elegans* whole genome microarrays. *BMC Genomics* 10, 160. doi:10.1186/1471-2164-10-160

Nagase, H., and Woessner, J. F. (1999). Matrix metalloproteinases. *J. Biol. Chem.* 274, 21491–21494.

Nota, B., Bosse, M., Ylstra, B., van Straalen, N. M., and Roelofs, D. (2009). Transcriptomics reveals extensive inducible biotransformation in the soil-dwelling invertebrate *Folsomia candida* exposed to phenanthrene. *BMC Genomics* 10, 326. doi:10.1186/1471-2164-10-236

Nota, B., Verweij, R. A., Molenaar, D., Ylstra, B., van Straalen, N. M., and Roelofs, D. (2010). Gene expression analysis reveals a gene set discriminatory to different metals in soil. *Toxicol. Sci.* 115, 34–40.

Owen, J., Hedley, B. A., Svendsen, C., Wren, J., Jonker, M. J., Hankard, P. K., Lister, L. J., Sturzenbaum, S. R., Morgan, A. J., Spurgeon, D. J., Blaxter, M. L., and Kille, P. (2008). Transcriptome profiling of developmental and xenobiotic responses in a keystone soil animal, the oligochaete annelid *Lumbricus rubellus*. *BMC Genomics* 9, 266. doi:10.1186/1471-2164-9-266

Saeed, A. I., Hagabati, N. K., Braisted, J. C., Liang, W., Sharov, V., Howe, E. A., Li, J., Thiagarajan, M., White, J. A., and Quackenbush, J. (2006). TM4 microarray software suite. *Methods Enzymol.* 411, 134–193.

Tabata, T., and Kornberg, T. B. (1994). Hedgehog is a signaling protein with a key role in patterning drosophila imaginal disks. *Cell* 76, 89–102.

Timmermans, M. J. T. N., Roelofs, D., Nota, B., Ylstra, B., and Holmstrup, M. (2009). Sugar sweet springtails: on the transcriptional response of *Folsomia candida* (Collembola) to desiccation stress. *Insect Mol. Biol.* 18, 737–746.

Van Straalen, N. M., and Feder, M. E. (2012). Ecological and evolutionary functional genomics – how can it contribute to the risk assessment of chemicals? *Eviron Sci. Technol.* 46, 3–9.

van Straalen, N. M., and Roelofs, D. (2008). Genomics technology for assessing soil pollution. *J. Biol.* 7, 19.

Vandenbrouck, T., Jones, O. A. H., Dom, N., Griffin, J. L., and De Coen, W. (2010). Mixtures of similarly acting compounds in *Daphnia magna*: from gene to metabolite and beyond. *Environ. Int.* 36, 254–268.

Wettenhall, J. M., and Smyth, G. K. (2004). limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics* 20, 3705–3706.

Xu, C. J., Li, C. Y. T., and Kong, A. N. T. (2005). Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Arch. Pharm. Res.* 28, 249–268.

Conflict of Interest Statement: 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.

Received: 28 January 2012; paper pending published: 29 February 2012; accepted: 28 April 2012; published online: 16 May 2012.

Citation: Roelofs D, de Boer M, Agamennone V, Bouchier P, Legler J and van Straalen N (2012) Functional environmental genomics of a municipal landfill soil. Front. Gene. ***3****:85. doi: 10.3389/fgene.2012.00085*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

frontiers in
GENETICS

ORIGINAL RESEARCH ARTICLE
published: 15 March 2012
doi: 10.3389/fgene.2012.00035

Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift

Pablo Cingolani[1,2,3], Viral M. Patel[1], Melissa Coon[2], Tung Nguyen[4], Susan J. Land[2], Douglas M. Ruden[1,2]* and Xiangyi Lu[1]

[1] Institute of Environmental Health Sciences, Wayne State University, Detroit, MI, USA
[2] Department of Obstetrics and Gynecology, Wayne State University School of Medicine, C.S. Mott Center, Detroit, MI, USA
[3] School of Computer Science and Genome Quebec Innovation Centre, McGill University, Montreal, QC, Canada
[4] Department of Computer Sciences, Wayne State University, Detroit, MI, USA

Edited by:
Michael Aschner, Vanderbilt University Medical Center, USA

Reviewed by:
Michael Aschner, Vanderbilt University Medical Center, USA
Stephen Sturzenbaum, King's College London, UK

***Correspondence:**
Douglas M. Ruden, Department of Obstetrics and Gynecology, Wayne State University School of Medicine, C.S. Mott Center, Detroit, MI 48201, USA.
e-mail: douglasr@wayne.edu

This paper describes a new program SnpSift for filtering differential DNA sequence variants between two or more experimental genomes after genotoxic chemical exposure. Here, we illustrate how SnpSift can be used to identify candidate phenotype-relevant variants including single nucleotide polymorphisms, multiple nucleotide polymorphisms, insertions, and deletions (InDels) in mutant strains isolated from genome-wide chemical mutagenesis of *Drosophila melanogaster*. First, the genomes of two independently isolated mutant fly strains that are allelic for a novel recessive male-sterile locus generated by genotoxic chemical exposure were sequenced using the Illumina next-generation DNA sequencer to obtain 20- to 29-fold coverage of the euchromatic sequences. The sequencing reads were processed and variants were called using standard bioinformatic tools. Next, SnpEff was used to annotate all sequence variants and their potential mutational effects on associated genes. Then, SnpSift was used to filter and select differential variants that potentially disrupt a common gene in the two allelic mutant strains. The potential causative DNA lesions were partially validated by capillary sequencing of polymerase chain reaction-amplified DNA in the genetic interval as defined by meiotic mapping and deletions that remove defined regions of the chromosome. Of the five candidate genes located in the genetic interval, the *Pka-like* gene *CG12069* was found to carry a separate pre-mature stop codon mutation in each of the two allelic mutants whereas the other four candidate genes within the interval have wild-type sequences. The *Pka-like* gene is therefore a strong candidate gene for the male-sterile locus. These results demonstrate that combining SnpEff and SnpSift can expedite the identification of candidate phenotype-causative mutations in chemically mutagenized *Drosophila* strains. This technique can also be used to characterize the variety of mutations generated by genotoxic chemicals.

Keywords: personal genomes, *Drosophila melanogaster*, whole-genome SNP analysis, next-generation DNA sequencing

INTRODUCTION

There are two types of chemicals that cause developmental abnormities in organisms – genotoxic chemicals and non-genotoxic chemicals. Genotoxic chemicals directly alkylate or oxidize the DNA and cause inappropriate base pairing. This causes permanent genetic mutations after exposing germline cells to genotoxic chemicals. Non-genotoxic chemicals are thought to cause epigenetic changes in the DNA that cause developmental abnormalities. Most non-genotoxic chemicals only affect development or the health of the organism exposed, but some non-genotoxic chemicals such as the estrogenic chemical diethylstilbestrol (DES) can cause developmental abnormalities and increased susceptibility to cancer for several generations (reviewed in Ruden et al., 2005).

Random mutagenesis such as chemical mutagenesis with the genotoxic chemical ethyl methane sulfonate (EMS) is an incredibly powerful tool for generating mutant strains of cells or organisms for purposes of studying all types of biological processes. In mutant bacteria or yeast, identification of the mutated genes is often done by transforming wild-type DNA into the cells and screening for rescue of the mutant phenotype. One could then sequence the DNA that rescues the phenotype to find the gene mutated. In *Drosophila melanogaster*, a causative DNA lesion for an observable phenotype is traditionally done by meiotic mapping of the mutant locus using a series of visible genetic markers that span the chromosome (Anderson, 1992). Deficiencies that delete defined regions of the chromosome, typically tens to hundreds of kilobases long, can then be used to further refine the boundaries of the mutated gene locus (Parks et al., 2004; Ryder et al., 2007). However, these positional cloning techniques are not only labor-intensive and time consuming, but also without a guarantee of success. This frequently leads to inevitable delays in molecular and functional characterization of the gene involved, even in the post genomic era.

With the development of next-generation DNA sequencing instruments, whole-genome sequencing is becoming feasible to replace labor-intensive positional cloning methods. However, we are limited by the capacity of the current bioinformatic programs to rapidly and reliably process sequence variants including single nucleotide polymorphisms (SNPs), multiple nucleotide polymorphisms (MNPs), insertions, and deletions (InDels) between the wild-type control and the mutant genomes. This is especially the case in dealing with mutant strains isolated from random chemical mutagenesis that typically introduces quite large numbers of background sequence variants and SNPs into the mutant genome, only one of which is likely responsible for the mutant phenotype.

Furthermore, all current next-generation sequencers produce frequent errors, especially when approaching the 3′-ends of each short read. Using current technologies, a short read is typically 70–150 bp long. As the euchromatic genome of *D. melanogaster* is 117 million base pairs (Mbp), machine-generated errors by themselves are sufficient to produce thousands of false SNPs in whole-genome sequencing data. To expedite the analyses of whole-genome sequencing data and to reduce number of false positives, we have developed the programs SnpEff (Pablo Cingolani and Douglas M. Ruden; Fly, in press; Platts et al., 2009) and SnpSift. These programs can categorize and filter thousands of variants per second, based on their locations in the transcriptional unit and potential mutational effects on transcription or translation. By comparing several sequencing experiments, the number of false positives can be reduced.

Whole-genome sequencing to identify a causative SNP has not been established for *D. melanogaster* mutants (Hillier et al., 2008; Wang et al., 2010). Here, we describe how SnpEff[1] and SnpSift[2] can be used together to identify causative gene candidate using just two alleles of a male-sterile *Drosophila* locus. Both programs have web based interfaces available via the Galaxy project[3].

RESULTS

WHOLE-GENOME SEQUENCING OF MALE-STERILE MUTANTS X1 AND X2

Two allelic male-sterile mutations, X1 and X2, were identified in a F_3 genetic screen (Yang et al., 2011). Briefly, males isogenic for the third chromosome were fed the chemical mutagen ethyl methane sulfonate (EMS) for 12 h (10 mM in 1% sucrose solution; Ruden et al., 1997) and then mated with virgin females of the genotype w^{1118}; *TM2/TM6,Sb*. Approximately 10,000 of the F_1 males (w^{1118}; */TM2 or w^{1118}: */TM6, Sb; * represents the mutagenized third chromosome) were then mated individually to w^{1118}; *TM2/TM6,Sb* virgin females to generate ∼6,000 lines, each carrying a mutagenized third chromosome. From the F_3 flies, males homozygous for the mutagenized chromosome (*/*) were tested for low fertility by crossing to virgin females from a wild-type stock ($y^1 w^1$). From this genetic screen, approximately 50 lines were saved that have low male fertility. They were placed into complementation groups by crossing to each other in ∼1,275 crosses (i.e., $1{,}275 = N(N+1)/2$, where $N = 50$). The characterization of two alleles of the same complementation group that we call X1 and X2 are presented. Details of the other male-sterile mutations isolated in the screen and phenotypic analyses of X1 and X2 will be presented elsewhere.

Males homozygous for X1 and X2 were sequenced (see Materials and Methods), producing over 90 million combined sequencing reads (∼76 bp per read), ∼10% of which are of insufficient quality and discarded. The remaining sequence reads represent approximately 20- to 29-fold coverage of the euchromatic DNA (**Figure 1**). These unique sequence reads were aligned to the reference genome (y^1; cn^1 bw^1 sp^1 strain, dm5.30), variant calls were performed, and 204,250 homozygous SNPs were found with a quality score greater than 70 (**Figure 2**). There were also 97,574 heterozygous SNPs, but they were not analyzed further because the sequenced genomic DNA samples were purified from the X1/X1 and X2/X2 homozygous flies. We found that greater than 99.99% of the homozygous SNPs were identical for X1 and X2 and these have to be common background variants because X1 and X2 were derived from the same parental strain. The remaining SNPs differ between X1 and X2 and they are associated with 141 genes, which were examined further (**Figure 3**, see below).

FINDING PHENOTYPE-CAUSATIVE CANDIDATE SNPs IN X1 AND X2

Figure 3 shows a flowchart of how the causative SNPs in X1 and X2 were identified. In order to identify the phenotype-causative candidate SNPs, we first assumed that they change an amino acid, splice site, reading frame, start or stop codon since these types of SNPs potentially alter the activity of the protein produced (we call these class 1 SNPs). Other types of SNPs such as intronic, intergenic, 5′UTR, 3′UTR, upstream, and downstream are less likely to affect gene function and they are considered secondarily only if no candidate genes could be identified from the first category of SNPs (we call these class 2 SNPs). Second, we considered the differential SNPs that are unique to either X1 or X2, but not common for X1 and X2 (**Figure 3A**). The way that the male-sterile screen was conducted ensured that X1 and X2 carried independently mutagenized chromosomes, so it is very unlikely that they have identical phenotype-causative SNPs (see Materials and Methods). Out of the 16,921 class 1 SNPs in X1 and X2, we found that 558 SNPs

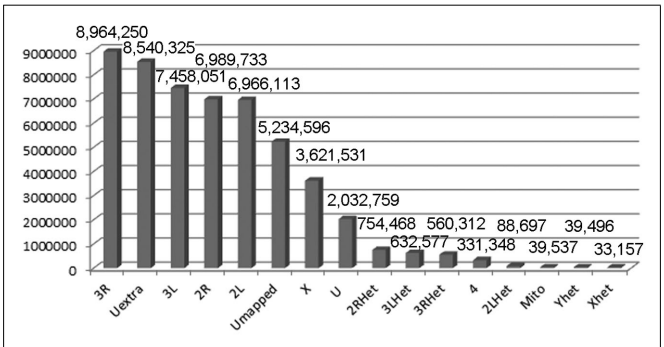


FIGURE 1 | Mapping X1 to the reference genome. The reference genome used was the latest FlyBase version (dm5.30). The quality score was arbitrarily set at 70 and above for this table. The numbers indicate the numbers of reads mapped to the indicated genomic region. U, unmapped regions. Het, heterochromatic regions.

[1]snpeff.sourceforge.net
[2]snpeff.sourceforge.net/SnpSift.html
[3]www.galaxy.psu.edu

are uniquely present in X1 and 447 SNPs are uniquely present in X2 (**Figure 3A**). For this analysis, thresholds above a certain level, such as 70, were not used because we did not want to eliminate a candidate SNP because it fell below an arbitrary threshold. For **Figure 2**, for illustrative purposes, we used a threshold score of 70, based on the quality score distribution for this sequencing run (McCarthy, 2010). Quality score, is defined by SAMtools as the probability of error in decibels, that is $q = -10\ \log(p)$, where p is the error probability and the logarithm is in base 10. Typically range for quality scores is from 1 to 100 with the higher score having a greater probability of being a real SNP and, therefore, not a sequencing artifact (McCarthy, 2010).

Next, we analyzed only the class 1 SNPs on the chromosome 3 since the X1 and X2 mutant strains were generated by using the third chromosome balancer (**Figure 3B**). As a general exercise, we did not begin our analysis by focusing on the third chromosome alone because this may not be applicable to other experimental settings. Considering just the third chromosome, there are 81 class 1 SNPs associating with 81 genes in X1, and 68 class 1 SNPs in 68 genes in X2. Of most interest are the eight genes that are commonly affected in both X1 and X2; i.e., the SNPs differ, but these SNPs associate with the same eight genes. Since the male-sterile phenotypes of X1 and X2 are presumably caused by two different SNPs affecting the same gene, we focused on these eight genes, which are A*nk2*, *Hsromega*, *CG12069*, *prc*, *CG13826*, *Muc68Ca*, *Rgl*, and *sls* (**Figure 3C**; **Table 1**). However, *CG12069* has SNPs with scores of 102 in X1 and 66 in X2 (**Table 1**). The score of 66 can be considered significant and it is substantially higher than the scores for the other seven candidate genes which have scores ranging from 1 to 36 with the majority having scores less than 5 (**Table 1**). *CG12069* was named as Pka-like in the Flybase because it encodes a protein with 51% amino acid identity to the adjacent *Pka-C2* which encodes a cAMP-dependent protein kinase A catalytic subunit (**Figure 4A**).

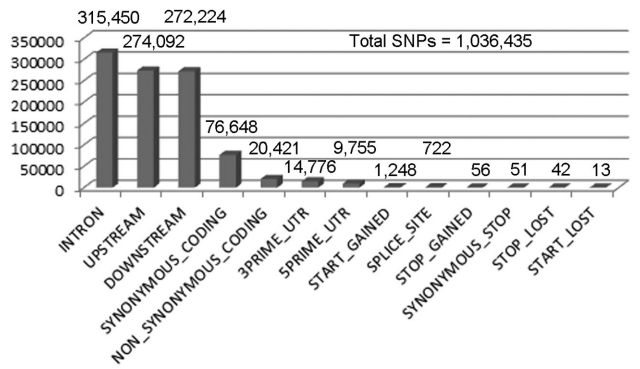


FIGURE 2 | Single nucleotide polymorphism calling for X1 SNPs with a quality score greater than or equal to 70. We performed SNP calling using Samtools, which produced 1,943,047 SNPs with a quality score > 1. Out of these, 1,036,435 are homozygous SNPs. The low quality SNPs were filtered out using an arbitrary threshold of 70 (the peak of the distribution) leaving 204,205 homozygous SNPs. A summary of the remaining homozygous SNPs found in each category is shown in the numbers above the bars.

VALIDATING X1 AND X2 AS NONSENSE ALLELES OF *CG12069*

Further analysis of the two SNPs in *CG12069* of X1 and X2 indicated that both of them are nonsense mutations causing premature translational termination at different amino acid residues of the Pka-like protein. X1 contains a TGG/TGA SNP that converts the tryptophan (W) residue 308 to a stop codon whereas X2 contains a CAG/TAG SNP that converts the glutamine (Q) residue 9 to a stop codon (**Figure 4B**). X1 will make the first 308 out of 356 amino acids of Pka-like. However, the Pka-like function is likely diminished because the conserved region of Pka-like with *Drosophila virilis* extends beyond amino acid 308. Also, the conserved ATP-binding domain of Pka-like extends beyond amino acid 308 (**Figure 4C**). X2 will only make the first eight amino acids of Pka-like, but there is another in-frame ATG codon at amino acid 10 that, if it supports translation initiation, would make a functional protein. However, there is a poor match to the Kozak consensus sequence, 5′-ACC-ATG-G-3′, flanking the downstream ATG site, 5′-CAG-ATG-C-3′. Since a good match to

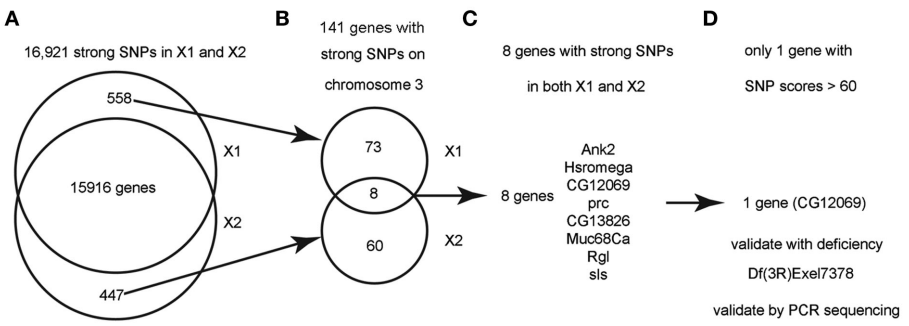


FIGURE 3 | Flowchart for finding the causative SNPs in X1 and X2. **(A)** SnpEeff identified 16,921 "class 1" SNPs (see text) with a quality score > 1 in both X1 and X2 (zero quality scores are usually resulted from reads mapping to multiple genomic regions). There are 558 SNPs that are only present in X1 and 447 SNPs that are only present in X2. **(B)** Since we know that X1 and X2 are on chromosome 3, we focused on the 141 strong SNPs on chromosome 3 that are present in X1 or X2 but not both. There are only eight genes that are commonly affected by unique SNPs in both X1 and X2 (note that the eight genes have at least two SNPs at different bases). **(C)** List of the eight genes with SNPs in both X1 and X2. See **Table 1** for more details. **(D)** Only one gene, *CG12069/Pka*-like, contained SNPs with scores > 60. These SNPs were validated by capillary sequencing of PCR-amplified DNA from the genetic interval of the male-sterile locus as defined by meiotic and deletion mapping data (see text). ca.

Table 1 | Gene candidates for X1 and X2.

Gene Name	X1 SNPs	Score	X2 SNPs	Score
Ank2	15	All < 5	14	All < 5
Hsromega	4	All < 5	4	All < 5
CG12069 (Pka-like)	1	102 (W308/*)	1	66 (Q9/*)
prc	2	1, 10	2	2, 21
CG13826	1	36 (I70/F)	1	30 (I70/L)
Muc68Ca	1	1	1	2
Rgl	1	30 (N8/T)	1	33 (N8/S)
sls	1	1	1	1

X1 SNPs and X2 SNPs, the number of SNPs in the indicated gene in X1 and X2. Score, the SNP quality score produced by the alignment and variant call software (e.g., SamTools and BcfTools).

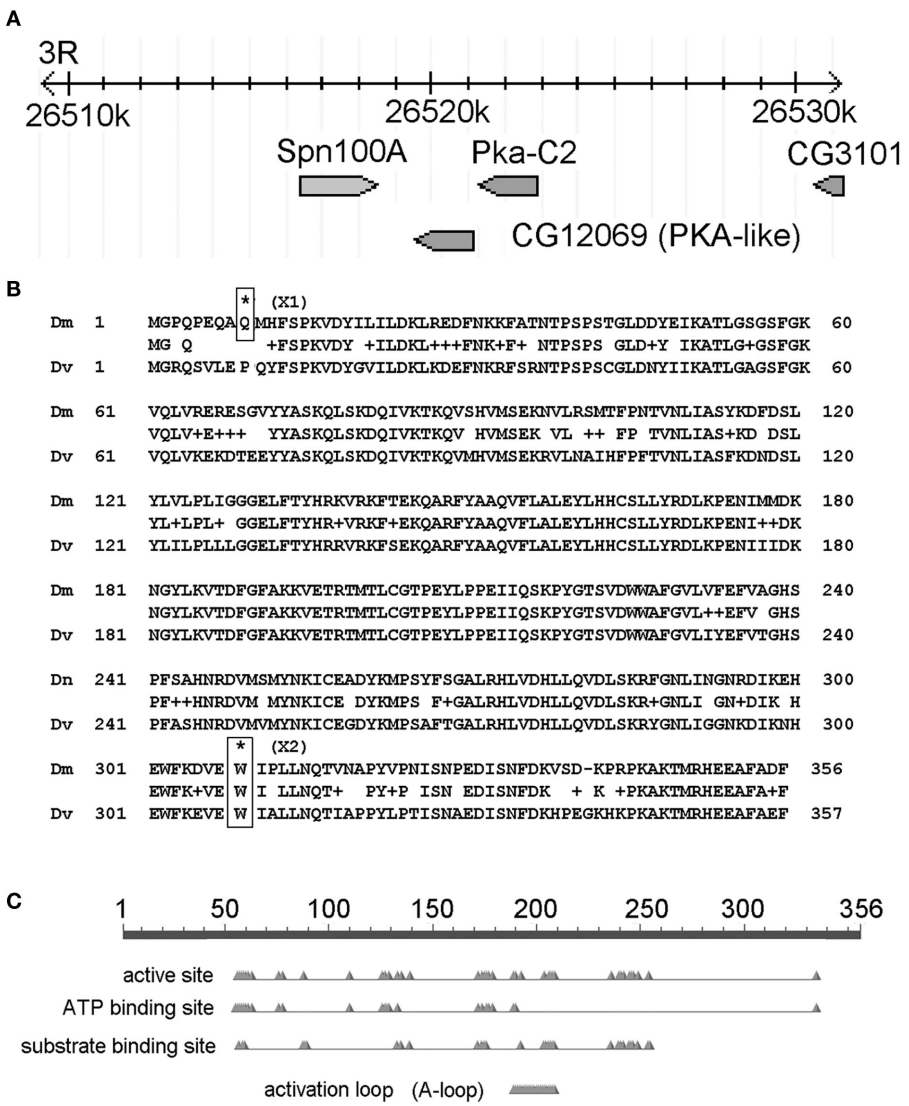


FIGURE 4 | The candidate gene mutated in X1 and X2 is *CG12069/Pka-like*. (A) Map of the *CG12069/Pka-like* region on chromosome 3R. The image is adapted from the FlyBase genome browser. The genomic location (26,520 k) is indicated in kilobase pairs. **(B)** Location of X1 and X2 SNPs. **(C)** Conserved domains in *CG12069/Pka-like*.

the Kozak sequence is generally required for efficient translation, (Kozak, 1987) it is possible that the downstream ATG is not used for translation. We note that the correct translation start sequence, 5′-GCA-ATG-C-3′, has a slightly better match to the Kozak sequence.

Since the male-sterile phenotypes of X1 and X2 homozygotes are nearly as strong as that of the males of the mutation over *Df(3R)Exel7378* that deletes *CG12069*, it is likely that the pre-mature stop codon mutations in *CG12069* are the causative loss-of-function mutations. To confirm this, we crossed X1 or X2 with chromosomal deletions that overlap with *Df(3R)Exel7378*. We found that the male-sterile phenotypes of X1 and X2 failed to complement *Df(3R)Exel7378* (3R:26388946;26620677), but complemented *Df(3R)BSC504* (3R:26253789;26512985) and *Df(3R)Exel8194* (3R:26582117;26713967). These localize the genetic boundary of X1 and X2 to a 69,132-bp of DNA interval from 26,512,985 to 26,582,117[4]. The ∼69 kb of DNA encodes 10 annotated genes, of which five are highly expressed in the testis, including *CG12069*. No SNPs were found in the remaining four candidate genes expressed in the testes (*CG12066*, *CG31010*, *CG1340*, *CG15543*), suggesting that *CG12069* is a strong candidate gene for the sperm storage defects of X1 and X2.

To further confirm the SNPs identified by SnpEff and SnpSift, genomic DNA samples were isolated from X1 and X2 homozygous mutant males and regions containing exons were amplified by polymerase chain reaction (PCR), cloned into pGEMT (Promega), and sequenced by capillary DNA sequencing (Applied Biosystems, Inc.). Sequencing confirmed the presence of stop codon SNPs in *CG12069* in both X1 and X2 at the expected locations. Thus, we conclude that the male-sterile alleles of X1 and X2 probably contain mutations in the *CG12069* gene. Complete validation will require a *CG12069* rescue transgene that is expressed in the male testes. However, phenotypic rescue of the male-sterile and sperm motility phenotypes of X1 and X2 is beyond the scope of this paper and will be presented elsewhere.

DISCUSSION

In this paper, we show that SnpEff and SnpSift can be used to identify causative SNPs in EMS-generated alleles of a new male-sterile mutant locus that we isolated from random chemical mutagenesis screens. We performed whole-genome shotgun sequencing of the two non-complementing alleles, X1 and X2, and showed that only a single gene, *CG12069/Pka-like*, was affected by SNPs at two different places, generating two different truncated proteins. The SNPs were confirmed by PCR amplification and capillary sequencing and further genetic mapping of the mutant locus using overlapping chromosomal deletions. From these, we conclude that a single lane of next-generation sequencing on the GAIIx instrument is probably sufficient for identifying homozygous causative SNP candidates in *Drosophila*. It should be emphasized that, in this case, we sequenced the DNA from homozygous flies. We were also able to use this technique to identify heterozygous SNPs isolated in a separate genetic screen (data not shown; Ruden et al., 1999). It was lucky that X1 and X2 were both nonsense mutations that designate strong SNPs and these occurred at two different codon positions in the same gene. Nevertheless, SnpEff and SnpSift can also analyze weak SNPs such as those located in the 5′ UTR or promoter regions and it should be possible to use a similar strategy to identify mutations that contain SNPs at regulatory regions of the genes, such as in many examples of population studies.

Recently, the Bellen laboratory developed rapid meiotic mapping techniques to map a recessive-lethal mutation to within a few kilobases to transposons containing easily visualized marker genes such as mini-w^+ or y^+ (Zhai et al., 2003). Meiotic mapping can be used to further delimit the regions of the genome and facilitate identification of candidate genes by whole-genome sequencing approach. We know of at least one other laboratory that has used next-generation sequencing to identify chemically induced mutations in *Drosophila*, but this was done with PCR-amplified DNA fragment from the ∼1-Mbp region of interest (Wang et al., 2010). Deficiencies, such as in the Exelixis and DrosDel collections that have known breakpoints, (Parks et al., 2004; Ryder et al., 2007) can be used to fine map the mutant locus further, often to a region small enough to PCR amplify and sequence with conventional capillary sequencing techniques.

Although we sequenced homozygous DNA, it is conceivable that larger fold of sequence coverage should overcome complication of data resulting from sequencing heterozygous DNA when the mutation is lethal. Langley et al. (2011) have recently shown that one can "circumvent heterozygosity" by sequencing the genome of a single haploid *D. melanogaster* embryo. The haploid embryo is gynogenetically produced by mating females with males homozygous for the recessive male-sterile mutation *ms(3)K81*, which jumps start embryogenesis without incorporating the sperm DNA in the developing embryo (Langley et al., 2011). Another alternative method to circumvent heterozygosity for recessive-lethal mutations is to use "green balancers" that carry, for example, *Kr-Gal4* driving GFP expression in the embryo and thus allowing the enrichment of homozygous mutant embryos prior DNA sequencing (Casso et al., 1999, 2000). The Bloomington stock center has green balancer stocks for the X chromosome (*FM7*), the second chromosome (*CyO*), and the third chromosome (*TM3,Sb*[5]). When a recessive-lethal allele is balanced with a green balancer, one needs only to select for non-GFP expressing embryos to ensure that the flies are homozygous in genotypes (Casso et al., 1999, 2000).

In summary, we describe a new tool, SnpSift that can be used to help identify causative SNPs in mutants derived from random chemical mutagenesis screens. This tool, along with SnpEff, has currently set to analyze and identify SNPs associated with phenotypes of not only *Drosophila* mutant strains but also other organisms including humans.

MATERIALS AND METHODS

PREPARING GENOMIC DNA LIBRARY FOR PAIRED-END SEQUENCING

Drosophila genomic DNA from the strains X1 and X2 was prepared using an AutoPure LS (Qiagen) Kit. A genomic DNA library was prepared from 5 μg purified *Drosophila* DNA according to

[4]flystocks.bio.indiana.edu

[5]www.flybase.org

the standard protocol using a Paired-End Sample Prep Kit for the GAIIx (Illumina). The DNA library was then used for cluster generation and sequencing analysis using the Genome Analyzer IIx using Illumina standard protocols. Methods for DNA manipulation, including sample preparation, formation of single-molecule arrays, cluster growth, and sequencing were all done by the standard protocols from Illumina, Inc. All sequencing was performed using two lanes (one for X1 and one for X2) in paired-end sequencing mode on an Illumina Genome Analyzer version 2 (GA2X) that was equipped with a 1-megapixel camera. The Illumina sequencing kits used allowed for 76 base single-end reads. Each lane of DNA sequencing had over 90 million reads.

Analysis software

Image analysis software was provided as part of the Genome Analyzer analysis pipeline and configured for fully automatic parameter selection. Single-end reads were 76 bases in total length. Quality control was performed using FastQC, showing overall low error rates. The reference genome used was the latest FlyBase version at the time (y^1; cn^1 bw^1 sp^1 strain, Dm5.30). The data was aligned using the BWA algorithm (Li and Durbin, 2009). A total of 5,234,506 reads were NOT mapped to the genome (i.e., 10.01%). This is usually due to low quality reads or reads have missing base calling information (i.e., "B" in the quality stream). The rest of the reads for X1 and X2 were mapped as indicated. Gap estimation: according to the mapping software, the gap between pair-end reads is 360 ± 20 bp. The distribution percentiles are 345 (25%), 360 (50%), and 375 (75%). The set of[6] and to the NCBI's map of RefSeq and candidate *Drosophila* genes[7].

Reads were filtered using a minimum mapping quality of 20 (MAPQ). Variant calling was performed using SamTools (Li et al., 2009) and BcfTools. When using individual calls without base alignment quality (BAQ) model, (Li, 2011) a total of 1,036,435 homozygous SNPs were detected. Using multi-sample calling methods and BAQ model, (Li, 2011) the number of homozygous SNPs was reduced to 204,250. Variant annotation and filtering was performed using the software SnpEff (Cingolani et al., Fly, in press) and SnpSift, described below.

SnpSift

Variant filtering was performed using an in-house development tool set called SnpSift[8]. This tool set works almost exclusively on variant call format (VCF) files according to the specification for versions 4 or 4.1 (Danecek et al., 2011). The two main components used in this work were "SnpSift caseControl" and "SnpSift filter." Frequently asked questions (FAQs) are addressed on our web site.

SnpSift caseControl

This tool counts the number of genotypes present in two user-defined groups ("case" and "control"), and then it calculates a *p*-value based on Fisher exact test. For each group, either homozygous, heterozygous, or both kinds of variants can be used.

SnpSift filter

This module performs filtering based on arbitrary expressions. In order to be able to parse arbitrary expressions, we created a top-down recursive grammar [also known as LL(*) grammar] using ANTLR (Parr, 2007). Using the lexer and parser created by ANTLR we are able to parse expressions by creating an abstract syntax tree (AST) for the expression. An AST is a well-known structure, very common in compiler design, that is used to represent the arbitrary input expressions from the user. The AST tree is converted into an *interpreter syntax tree (IST)*, which is a tree composed of objects capable of interpreting conditions, expressions, and functions. This means that the IST is like AST, but it is also capable of performing expression evaluation. The result of the filter expression is the value of the root node in the IST.

There are well-known variables pre-defined according to the VCF format specification. Other additional variables and their respective data types are parsed from VCF meta-information in the file header. As specified in the norm, INFO meta-information lines define the type and the number of values (e.g., an array) in each INFO sub-field. Automatic variable conversion is implemented (e.g., INT is automatically converted to FLOAT whenever required). Genotype fields are similarly parsed by using FORMAT meta-information header lines.

Each VCF entry (i.e., each non-header line in a VCF file) is converted into a set of "variable = value" tuples, which are feed into the interpreter tree. The IST, created using the user expression, interprets the user-defined expression from top to bottom trying to assign a Boolean value to the root node. If the result from evaluating the IST is "true" then the VCF line is either printed to standard output or marked as PASS in the FILTER field; likewise, if it is "false," the line is filtered out (i.e., not printed) or marked as failed in the FILTER field. **Table A1** in Appendix shows a list of allowed operators used in SnpSift and **Table A2** in Appendix shows some functions commonly used in SnpSift expressions. Language definition and examples are shown in Appendix.

SnpSift is platform independent and available as an open source as part of the SnpEff project[9]. A web based interface is available via the Galaxy project (see text foot note 1).

DATA ACCESS

SnpEff and SnpSift Data can be accessed from the data file for X1 and X2 by contacting Douglas M. Ruden.

ACKNOWLEDGMENTS

This work was supported by a Michigan Core Technology grant from the State of Michigan's 21st Century Fund Program to the Wayne State University Applied Genomics Technology Center. This work was also supported by the Environmental Health Sciences Center in Molecular and Cellular Toxicology with Human Applications Grant P30 ES06639 at Wayne State University, NIH R01 grants (ES012933) to Douglas M. Ruden, and DK071073 to Xiangyi Lu.

[6] ftp://ftp.flybase.net/genomes/dmel/dmel_r5.12_FB2008_09/gff/

[7] ftp://ftp.ncbi.nih.gov/genomes/Drosophila_melanogaster/mapview/seq_gene.md.gz

[8] SnpEff.sourceforge.net/SnpSift.html

[9] SnpEff.sourceforge.net/SnpSift.html

REFERENCES

Anderson, K. (1992). The Making of a fly – the genetics of animal design – Lawrence, Pa. *Science* 256, 1053–1054.

Casso, D., Ramirez-Weber, F., and Kornberg, T. B. (2000). GFP-tagged balancer chromosomes for *Drosophila melanogaster*. *Mech. Dev.* 91, 451–454.

Casso, D., Ramirez-Weber, F. A., and Kornberg, T. B. (1999). GFP-tagged balancer chromosomes for *Drosophila melanogaster*. *Mech. Dev.* 88, 229–232.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., and 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158.

Hillier, L. W., Marth, G. T., Quinlan, A. R., Dooling, D., Fewell, G., Barnett, D., Fox, P., Glasscock, J. I., Hickenbotham, M., Huang, W., Magrini, V. J., Richt, R. J., Sander, S. N., Stewart, D. A., Stromberg, M., Tsung, E. F., Wylie, T., Schedl, T., Wilson, R. K., and Mardis, E. R. (2008). Whole-genome sequencing and variant discovery in *C. elegans*. *Nat. Methods* 5, 183–188.

Kozak, M. (1987). An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15, 8125–8148.

Langley, C. H., Crepeau, M., Cardeno, C., Corbett-Detig, R., and Stevens, K. (2011). Circumventing heterozygosity: sequencing the amplified genome of a single haploid *Drosophila melanogaster* embryo. *Genetics* 188, 239–246.

Li, H. (2011). Improving SNP discovery by base alignment quality. *Bioinformatics* 27, 1157–1158.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078.

McCarthy, A. (2010). Third generation DNA sequencing: pacific biosciences' single molecule real time technology. *Chem. Biol.* 17, 675–676.

Parks, A. L., Cook, K. R., Belvin, M., Dompe, N. A., Fawcett, R., Huppert, K., Tan, L. R., Winter, C. G., Bogart, K. P., Deal, J. E., Deal-Herr, M. E., Grant, D., Marcinko, M., Miyazaki, W. Y., Robertson, S., Shaw, K. J., Tabios, M., Vysotskaia, V., Zhao, L., Andrade, R. S., Edgar, K. A., Howie, E., Killpack, K., Milash, B., Norton, A., Thao, D., Whittaker, K., Winner, M. A., Friedman, L., Margolis, J., Singer, M. A., Kopczynski, C., Curtis, D., Kaufman, T. C., Plowman, G. D., Duyk, G., and Francis-Lang, H. L. (2004). Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. *Nat. Genet.* 36, 288–292.

Parr, T. (2007). *The Definitive ANTLR Reference: Building Domain-Specific Languages*. Lewisville: Pragmatic Bookshelf.

Platts, A. E., Land, S. J., Chen, L., Page, G. P., Rasouli, P., Wang, L., Lu, X., and Ruden, D. M. (2009). Massively parallel resequencing of the isogenic *Drosophila melanogaster* strain w(1118); iso-2; iso-3 identifies hotspots for mutations in sensory perception genes. *Fly (Austin)* 3, 192–203.

Ruden, D. M., Cui, W., Sollars, V., and Alterman, M. (1997). A *Drosophila* kinesin-like protein, Klp38B, functions during meiosis, mitosis, and segmentation. *Dev. Biol.* 191, 284–296.

Ruden, D. M., Wang, X., Cui, W., Mori, D., and Alterman, M. (1999). A novel follicle-cell-dependent dominant female sterile allele, StarKojak, alters receptor tyrosine kinase signaling in *Drosophila*. *Dev. Biol.* 207, 393–407.

Ruden, D. M., Xiao, L., Garfinkel, M. D., and Lu, X. (2005). Hsp90 and environmental impacts on epigenetic states: a model for the trans-generational effects of diethylstibesterol on uterine development and cancer. *Hum. Mol. Genet.* 14, R149–R155.

Ryder, E., Ashburner, M., Bautista-Llacer, R., Drummond, J., Webster, J., Johnson, G., Morley, T., Chan, Y. S., Blows, F., Coulson, D., Reuter, G., Baisch, H., Apelt, C., Kauk, A., Rudolph, T., Kube, M., Klimm, M., Nickel, C., Szidonya, J., Maróy, P., Pal, M., Rasmuson-Lestander, A., Ekström, K., Stocker, H., Hugentobler, C., Hafen, E., Gubb, D., Pflugfelder, G., Dorner, C., Mechler, B., Schenkel, H., Marhold, J., Serras, F., Corominas, M., Punset, A., Roote, J., and Russell, S. (2007). The DrosDel deletion collection: a *Drosophila* genomewide chromosomal deficiency resource. *Genetics* 177, 615–629.

Wang, H., Chattopadhyay, A., Li, Z., Daines, B., Li, Y., Gao, C., Gibbs, R., Zhang, K., and Chen, R. (2010). Rapid identification of heterozygous mutations in *Drosophila melanogaster* using genomic capture sequencing. *Genome Res.* 20, 981–988.

Yang, Y., Cochran, D. A., Gargano, M. D., King, I., Samhat, N. K., Burger, B. P., Sabourin, K. R., Hou, Y., Awata, J., Parry, D. A., Marshall, W. F., Witman, G. B., and Lu, X. (2011). Regulation of flagellar motility by the conserved flagellar protein CG34110/Ccdc135/FAP50. *Mol. Biol. Cell* 22, 976–987.

Zhai, R. G., Hiesinger, P. R., Koh, T. W., Verstreken, P., Schulze, K. L., Cao, Y., Jafar-Nejad, H., Norga, K. K., Pan, H., Bayat, V., Greenbaum, M. P., and Bellen, H. J. (2003). Mapping *Drosophila* mutations with molecularly defined P element insertions. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10860–10865.

Conflict of Interest Statement: 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.

Received: 20 February 2012; accepted: 24 February 2012; published online: 15 March 2012.

*Citation: Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM and Lu X (2012) Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. Front. Gene. **3**:35. doi: 10.3389/fgene.2012.00035*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.

APPENDIX

SnpSIFT FILTER: LANGUAGE DEFINITION

This section shows the language definition for SnpSift filter. Operators (see **Table A1**) and functions (see **Table A2**) can be used to create arbitrary expressions that are evaluated using the information in each VCF line.

SnpSIFT FILTER: LANGUAGE DEFINITION AND USAGE EXAMPLES

Using the SnpSift filter, arbitrary expressions can be evaluated. Since an arbitrary number of conditions can be combined using Boolean operators, the expressions can be complex, allowing significant flexibility.

Some examples:

1-) Filter out variants with quality less than 30:

```
cat variants.vcf | java -jar SnpSift.jar " ( QUAL >= 30 )" > filtered.vcf
```

2-)Filter out variants with quality less than 30 but keep InDels that have quality 20 or more:

```
cat variants.vcf | java -jar SnpSift.jar "(( exists INDEL ) & (QUAL >= 20)) | (QUAL >= 30 )" > filtered.vcf
```

3-)Same as example 2, but keeping also any homozygous variant present in more than 3 samples:

```
cat variants.vcf | java -jar SnpSift.jar "(countHom > 3) | (( exists INDEL ) & (QUAL >= 20)) | (QUAL >= 30 )" > filtered.vcf
```

4-)Same as example 3, but keeping also heterozygous variants with coverage 25 or more:

```
cat variants.vcf | java -jar SnpSift.jar "((countHet > 0) && (DP >= 25)) | (countHom > 3) | (( exists INDEL ) & (QUAL >= 20)) | (QUAL >= 30 )" > filtered.vcf
```

Table A1 | Operators allowed in SnpSift filter.

Operand	Description	Data type	Example
=	Equality test	FLOAT, INT or STRING	(REF = 'A')
>	Greater than	FLOAT or INT	(DP > 20)
≥	Greater or equal than	FLOAT or INT	(DP ≥ 20)
<	Less than	FLOAT or INT	(DP < 20)
≤	Less or equal than	FLOAT or INT	(DP ≤ 20)
=∼	Match regular expression	STRING	(REL =∼ 'AC')
!∼	Does not match regular expression	STRING	(REL!∼ 'AC')
&	AND operator	Boolean	(DP > 20) & (REF = 'A')
|	OR operator	Boolean	(DP > 20) | (REF = 'A')
!	NOT operator	Boolean	! (DP > 20)
exists	The variable exists (not missing)	Any	(exists INDEL)

Table A2 | Functions implemented in SnpSift filter.

Function	Description	Data type	Example
countHom	Count number of homozygous genotypes	No arguments	(**countHom()** > 0)
countHet	Count number of heterozygous genotypes	No arguments	(**countHet()** > 2)
countVariant	Count number of genotypes that are variants (i.e., not reference 0/0)	No arguments	(**countVariants()** > 5)
countRef	Count number of genotypes that are NOT variants (i.e., reference 0/0)	No arguments	(**countRef()** < 1)

SNPSIFT FILTER: VARIABLES

For each VCF entry, the variables are populated and made available in the analyzed expressions. The values used to populate the variables are obtained from different fields of the VCF entry. There are four main groups of variables:

- Fields: these are mandatory valued from the VCF specification and are the first columns in a VCF file ("CHROM, POS, ID, REF, ALT, QUAL, or FILTER").
- INFO field: each value defined in the info field is made available using the type specified according to the VCF meta-information lines in the header section. Some "well-known" variables are predefined and do not need corresponding header entries (see VCF specification for a list of well-known INFO fields).
- Genotype fields: each genotype field is available using the GEN[] array. Subfields of this array include all variables in each genotype field. Types are casted according to the VCF meta-information lines in the header section.
- Effect fields: the "EFF" sub-field from the INFO field (created by SnpEff program) is further parsed and made available. This is parsed as an array since one variant can be annotated with more than one effect.
- Sets: expressions can test if a value belongs to a set. Sets are defined in files having one value per line. This files are parsed when using the "–set" command line option. Values from sets can be used in expressions by using the "in" operator.

Fields

Available variable names are: "CHROM, POS, ID, REF, ALT, QUAL, or FILTER."

Examples:

1-) Any variant in chromosome 1:

```
"( CHROM = 'chr1' )"
```

2-) Variants between two positions:

```
"( POS > 123456 ) & ( POS < 654321 )"
```

3-) Variants having an ID and it matches the regular expression "rs":

```
"(exists ID) & ( ID =  'rs' )"
```

4-) Variants having reference "A":

```
"( REF = 'A' )"
```

5-) Variants having an alternative "T":

```
"( ALT = 'T' )"
```

6-) Variants having quality over 30:

```
"( QUAL > 30 )"
```

6-) Variants having Filter value is either "PASS" or it is missing:

```
"( na FILTER ) | (FILTER = 'PASS')"
```

INFO field

Variable names from INFO field. E.g., if the info field has "DP=48;AF1=0;. . ." e.g.,:

(DP > 10) & (AF1 = 0)

Multiple value

Info field variables can have multiple values (comma separated). These multiple valued fields are represented as an array. Individual values can be accessed using an index. E.g., If the INFO field has "CI95=0.04167,0.5417," then the following expression is valid:

"(CI95[0] > 0.1) & (CI95[1] <= 0.3)"

An asterisk may be used to represent "ANY" variable index. So the following example is "true" if any of the values in the CI95 field is more than 0.1:

"(CI95[∗] > 0.1)"

Genotype fields

Variables from genotype fields are represented as an array. The individual values are accessed using an index (sample number) followed by a variable name. E.g., If the genotypes are "GT:PL:GQ 1/1:255,66,0:63 0/1:245,0,255:99," then the following expression is "true":

"(GEN[0].GQ > 60) & (GEN[1].GQ > 90)"

An asterisk may be used to represent "ANY" variable index

"(GEN[∗].GQ > 60)"

Genotype having multiple fields

These are represented as arrays, so individual values can be accessed using an index (sample number) followed by a variable name and then another index. E.g., If the genotypes are "GT:PL:GQ 1/1:255,66,0:63 0/1:245,0,255:99," then the following expression is valid:

"(GEN[0].PL[2] = 0)"

Also in this case, an asterisk may be used to represent "ANY" variable index, e.g.,:

"(GEN[0].PL[∗] = 0)"

And another asterisk may be used to represent "ANY" genotype index, e.g.,:

"(GEN[∗].PL[∗] = 0)"

Sets

are defined by the "-s" (or "–set") command line option. Each file must have one string per line. They are named based on the order used in the command line (e.g., the first one is "SET[0]," the second one is "SET[1]," etc.) An example of the set expression (assuming your command line was "-s set1.txt -s set2.txt -s set3.txt"):

"(ID in SET[2])"

Effect fields

Effect fields created by SnpEff are accessed using an index (effect number) followed by a sub-field name. Available sub-field are:

- EFFECT: effect (e.g., SYNONYMOUS_CODING, NON_SYNONYMOUS_CODING, FRAME_SHIFT, etc.)
- IMPACT: [HIGH, MODERATE, LOW, MODIFIER]
- FUNCLASS: [NONE, SILENT, MISSENSE, NONSENSE]
- CODON: codon change (e.g., "ggT/ggG")
- AA: amino acid change (e.g., "G156")
- GENE: gene name (e.g., "PSD3")
- BIOTYPE: gene biotype, as described by the annotations (e.g., "protein_coding")
- CODING: gene is [CODING, NON_CODING]
- TRID: transcript ID
- EXID: exon ID

Examples:

1-) The following expression is true if the first effect is NON_SYNONYMOUS:

"(EFF[0].EFFECT = 'NON_SYNONYMOUS_CODING')"

2-) This expression is true if ANY effect is NON_SYNONYMOUS:

"(EFF[∗].EFFECT = 'NON_SYNONYMOUS_CODING')"

3-) This expression is true if ANY effect is NON_SYNONYMOUS on gene TCF7L2:

"(EFF[∗].EFFECT = 'NON_SYNONYMOUS_CODING') & (EFF[∗].GENE = 'TCF7L2')"

frontiers in
GENETICS

ORIGINAL RESEARCH ARTICLE
published: 04 April 2012
doi: 10.3389/fgene.2012.00051

Influence of nitrate and nitrite on thyroid hormone responsive and stress-associated gene expression in cultured *Rana catesbeiana* tadpole tail fin tissue

Ashley Hinther[1], Thea M. Edwards[2], Louis J. Guillette Jr.[3] and Caren C. Helbing[1]*

[1] Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC, Canada
[2] School of Biological Sciences, Louisiana Tech University, Ruston, LA, USA
[3] Department of Obstetrics and Gynecology, Medical University of South Carolina, Hollings Marine Laboratory, Charleston, SC, USA

Edited by:
Jonathan Freedman, National Institute of Environmental Health Science, National Institute of Health, USA

Reviewed by:
Michael J. Liguori, Abbott Laboratories, USA
Jason Unrine, University of Kentucky, USA

***Correspondence:**
Caren C. Helbing, Department of Biochemistry and Microbiology, University of Victoria, P.O. Box 3055 STN CSC, Victoria, BC, Canada V8W 3P6.
e-mail: chelbing@uvic.ca

Nitrate and nitrite are common aqueous pollutants that are known to disrupt the thyroid axis. In amphibians, thyroid hormone (TH)-dependent metamorphosis is affected, although whether the effect is acceleration or deceleration of this developmental process varies from study to study. One mechanism of action of these nitrogenous compounds is through alteration of TH synthesis. However, direct target tissue effects on TH signaling are hypothesized. The present study uses the recently developed cultured tail fin biopsy (C-fin) assay to study possible direct tissue effects of nitrate and nitrite. Tail biopsies obtained from premetamorphic *Rana catesbeiana* tadpoles were exposed to 5 and 50 mg/L nitrate (NO_3–N) and 0.5 and 5 mg/L nitrite (NO_2–N) in the absence and presence of 10 nM T_3. Thyroid hormone receptor β (TRβ) and Rana larval keratin type I (RLKI), both of which are TH-responsive gene transcripts, were measured using quantitative real time polymerase chain reaction. To assess cellular stress which could affect TH signaling and metamorphosis, heat shock protein 30, and catalase (CAT) transcript levels were also measured. We found that nitrate and nitrite did not significantly change the level of any of the four transcripts tested. However, nitrate exposure significantly increased the heteroscedasticity in response of TRβ and RLKI transcripts to T_3. Alteration in population variation in such a way could contribute to the previously observed alterations of metamorphosis in frog tadpoles, but may not represent a major mechanism of action.

Keywords: nitrate, nitrite, frog, thyroid hormone, metamorphosis, C-fin, organ culture assay, quantitative real time polymerase chain reaction

INTRODUCTION

Currently over 100,000 manufactured chemicals are produced in the marketplace (European Union Commission, 2006). Many of these chemicals have endocrine disrupting abilities and more specifically, are disruptors of the thyroid axis. Most endocrine disruptors can be classified as plasticizers, pesticides, industrial chemicals, heavy metals, or plant and fungal compounds; however, ions such as nitrate (NO_3^-) and nitrite (NO_2^-) have endocrine disrupting abilities as well (Crain, 2000; Sampat, 2000; Gray et al., 2001). Environmental nitrate can come from many sources including agricultural fertilizer, waste from animal production, and burning fossil fuels, industrial effluent, and wastewater treatment plant discharges (Rouse et al., 1999; Camargo et al., 2004; De Groef et al., 2006).

In the US, the current public health maximal level for safe drinking water is 10 mg/L nitrate (measured as NO_3–N) and 1 mg/L nitrite (NO_2–N; US EPA, 2006, 2009). In Canada, the Canadian Council of Ministers of the Environment (CCME) guideline for the protection of aquatic life has set the maximum level of nitrate at 13 mg/L(NO_3^-/L) in freshwater and 16 mg/L (NO_3^-/L) in marine water; the level for nitrite in freshwater is 60 μg/L (NO_3^-/L) and there is no level set for marine water (CCME, 2007). Health Canada has set the maximal allowable concentration in drinking water at 10 mg/L nitrate (NO_3–N) and 3.2 mg/L nitrite (NO_2–N; Health Canada, 2008). Nitrate concentrations have been found as high as 25 mg/L NO_3–N in surface waters and 100 mg/L NO_3–N in ground water, yet there is currently no guideline for the protection of wildlife (Rouse et al., 1999; Camargo et al., 2004).

In aquatic environments, nitrogen exists in four forms in descending order of toxicity: ammonium ion, ammonia, nitrite, and nitrate. Although nitrate is the least toxic form of the four, it is the most stable and therefore the most abundant. Under aerobic conditions, ammonia and ammonium can be oxidized to nitrite by *Nitrosomonas* bacteria, and then to nitrate by *Nitrobacter and Nitrospira bacteria* (Sharma and Ahlert, 1977). When oxygen is low, denitrifying bacteria can use nitrate as a terminal electron acceptor and make nitrogen gas (N_2; reviewed in Camargo et al., 2005).

Aquatic animals are exposed to nitrate and nitrite through ingestion or epithelial absorption across skin or gills (Onken et al., 2003). High levels of these contaminants cause methemoglobinemia, also called "brown blood" disease in fish and amphibians and "blue baby" syndrome in humans. Methemoglobin is formed from nitrate/nitrite-induced oxidation of hemoglobin, which prevents

normal oxygen binding and leads to hypoxia (Porter et al., 1999). Toxicity of nitrite and nitrate depends on body size and developmental stage, increases with increasing concentration and exposure time, and decreases with water salinity and environmental adaptation (Rouse et al., 1999; Camargo et al., 2004).

In addition to the toxic effects of nitrate and nitrite, exposure to these chemicals adversely affects the thyroid axis in multiple vertebrate species (reviewed in Edwards et al., 2006). For example, high doses of nitrate caused goiter and depressed serum thyroxine (T_4) and 3,5,3′-triiodothyronine (T_3) in rats and sheep (Zaki et al., 2004) and nitrite decreased serum T_4 while T_3 levels were unchanged in the sea bream (Deane et al., 2007). Toad and frog tadpoles exposed to nitrate exhibit altered metamorphic development; a TH-dependent process (Wyngaarden et al., 1952, 1953; Xu and Oldham, 1997; Edwards et al., 2006; Ortiz-Santaliestra and Sparling, 2007). Although these observations can be explained, in part, by competition of nitrate and nitrite with iodine uptake, transport, and retention in the thyroid gland that impairs TH synthesis (Crow et al., 2001; Hampel and Zollner, 2004), the contribution of nitrate and nitrite to alteration of TH signaling pathways at the cellular level in amphibian target tissues is not known.

The present study uses the recently developed "C-fin" assay to expose *Rana catesbeiana* premetamorphic tadpole tail fin biopsies to nitrate and nitrite with or without T_3 to determine if nitrate and nitrite affect TH-signaling within a TH-responsive tissue directly. We assessed TH-signaling by quantifying the levels of TH-responsive gene transcripts, thyroid hormone receptor β (*TR*β) and *Rana* larval type I keratin (*RLKI*), as well as cellular stress markers, heat shock protein (*HSP30*), and catalase (*CAT*). Alteration of the transcriptome is an essential component in TH-mediated tadpole metamorphosis (Shi, 2000) and part of the change in the tail transcriptome includes an increase in *TR*β transcripts and a decrease in *RLKI* transcripts (Domanski and Helbing, 2007). There is considerable precedent linking *TR*β transcript levels to progression through TH-dependent metamorphosis where perturbations from expected levels are indicative of altered postembryonic development (Crump et al., 2002; Opitz et al., 2006; Veldhoen et al., 2006a; Zhang et al., 2006; Helbing et al., 2007a,b; Ji et al., 2007; Skirrow et al., 2008).

MATERIALS AND METHODS

EXPERIMENTAL ANIMALS

Premetamorphic *R. catesbeiana* tadpoles were caught locally (Victoria, BC, Canada) or purchased from Ward's Natural Science Ltd. (St. Catherines, ON, Canada). Taylor and Kollros (TK; Taylor and Kollros, 1946) stage VI–VIII animals were used. Animals were housed in the University of Victoria aquatics facility and maintained in 100 gallon fiberglass tanks containing recirculating water at 12°C with exposure to natural daylight. Tadpoles were fed daily with spirulina (Aquatic ELO-Systems, Inc., FL, USA). Animals used in this study were treated and maintained in accordance with the guidelines of the Canadian Council on Animal Care.

ORGAN CULTURE OF TAIL FIN BIOPSIES

Preparation of the tail fin biopsy cultures was adapted from conditions described previously (Veldhoen et al., 2006b; Ji et al., 2007). Premetamorphic (TK stage VI–VIII; Taylor and Kollros, 1946) *R. catesbeiana* tadpoles were euthanized in 0.1% tricaine methanesulfonate (Syndel Laboratories, Vancouver, BC, Canada) in 25 mM sodium bicarbonate, and subsequently washed four times in 125 mL per tadpole of sterile magnesium-free (MFM) solution (7.5 mM Tris–HCl pH 7.6, 88 mM NaCl, 1 mM KCl, 2.4 mM $NaHCO_3$, 0.88 mM $CaCl_2$). Eight biopsies were obtained per animal ($n = 7$–16 animals), from the dorsal and ventral tail fins using a 6 mm dermal biopsy punch (Miltex, Inc., York, PA, USA), to allow the assessment of eight treatments per animal. Sixteen animals were used for each exposure.

Test chemicals were prepared in water as 1,000× concentrates and stored at −20°C. They were applied in equal volumes at 1 μL chemical stock/mL of media. Treatments included: a vehicle control (NaOH), sodium gluconate control (Na-G, used as a control for sodium; CAS S-2054, Sigma-Aldrich), sodium nitrate ($NaNO_3$; CAS BP360-500g, Fisher; measured as 5 and 50 mg/L NO_3–N), or sodium nitrite ($NaNO_2$; CAS S2252-500g, >99.5% purity, Sigma-Aldrich; measured as 0.5 and 5 mg/L NO_2–N), in the absence and presence of 10 nM T_3 (prepared as a 10^{-5} M stock in 400 μM NaOH), as well as a 10 nM T_3 treatment alone. Where treatments did not include T_3, an equal volume of NaOH vehicle was applied to a final concentration of 400 nM. This concentration did not affect the medium pH. Biopsies were cultured individually in 1 mL 70% strength Leibovitz's L15 medium (Gibco, Invitrogen) supplemented with 10 mM HEPES pH 7.5, 50 units/mL penicillin G sodium, 50 μg/mL streptomycin sulfate (Gibco, Invitrogen), and 50 μg/mL neomycin (Sigma-Aldrich), using 24-well culture plates (Primaria, BD Biosciences) at 25°C in air for 48 h.

The biopsies were pretreated with 0.5 mL of the appropriate concentration of the test chemical or NaOH control in culture media for 2 h prior to the addition of T_3. After the 2 h incubation, 0.5 mL of the appropriate concentration of the test chemical plus 20 nM T_3 (in 800 μM NaOH) were added into the wells giving a final concentration of 10 nM T_3 (in 400 nM NaOH). For the wells not containing T_3, 0.5 mL of the appropriate concentration of the test chemical plus 800 μM NaOH (for a final concentration of 400 nM NaOH) were added. At the end of the 48 h incubation period for each treatment, the biopsy was stored in 100 μL of RNA*later* (Ambion Inc., Austin, TX, USA) for 24 h at 4°C and then transferred to −20°C until it was processed for RNA.

ISOLATION OF RNA AND QUANTIFICATION OF GENE EXPRESSION

RNA was isolated using TRIzol reagent as described previously (Hinther et al., 2010a,b). cDNA was synthesized from 5 μL (∼0.5 μg) total RNA as per manufacturer's protocol using the RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas) as described in (Hinther et al., 2010b). The cDNA products were diluted fivefold prior to PCR amplification and stored at −20°C.

The levels of mRNAs encoding *TR*β, *RLKI*, *HSP30*, *CAT*, and ribosomal protein L8 (*rpL8*) were determined using a MX3005P real time quantitative PCR system (Stratagene, La Jolla, CA, USA) using gene-specific primers as described previously (Hinther et al., 2010b). Expression profiles of the *rpL8* transcript normalizer were invariant ($p = 0.998$ and 0.950, nitrate and nitrite data sets, respectively). The amplified DNA signals for all QPCR reactions were

evaluated for specificity based upon their thermodenaturation profiles. Data that failed to produce a profile indicative of gene target-specific detection were removed before analysis. If a control condition did not pass the quality measure above for a given animal, then the data for all conditions associated with that animal were removed for that gene transcript due to the repeated measures nature of the data set.

STATISTICAL ANALYSES

Statistical analyses were performed using PASW 18.0 (Chicago, IL, USA) software. The C-fin data were not normally distributed based upon the Shapiro–Wilk test. We used the Friedman and paired Wilcoxon tests since these data were generated from a repeated measures type of experimental design. Homogeneity of variance was determined using the Levene's test. Correlation coefficients were generated using Spearman's rho.

The data were analyzed in two ways: First, the test chemical results in the absence of T_3 were examined relative to the vehicle control. Second, the test chemicals in combination with T_3 results were compared relative to T_3 alone. In the latter case, the response to a test chemical in the presence of T_3 was expressed as a fold change relative to the response to T_3 alone for each individual. This approach reduces the effect of inter-animal variation, enabling us to better identify chemical-induced perturbations relative to each individual's ability to respond to T_3. Therefore the T_3 values in this comparison were given a value of one and the graphs show the fold change relative to the T_3-induced response. Statistical significance was identified when $p < 0.05$.

RESULTS AND DISCUSSION

In order to validate the assay, we first examined the biopsy responses to T_3 treatment alone. **Figure 1A** shows the biological variation of the controls and the relative variation in T_3 response before taking repeated measures into account. **Figure 1B** shows the data after normalizing the data to the individual animal's baseline transcript levels into such that every control animal was assigned a value of 1. T_3 treatment alone increased the *TR*β transcript levels by a median 7.7-fold ($p = 0.0001$, Wilcoxon, $n = 23$; **Figure 1**), decreased *RLKI* transcript levels by 2.9-fold ($p = 0.0001$, Wilcoxon, $n = 26$; **Figure 1**), increased *HSP30* transcript levels by 1.7-fold ($p = 0.024$, Wilcoxon, n = 31; **Figure 1**), and reduced *CAT* transcript levels by a median 1.3-fold ($p = 0.024$, Wilcoxon, $n = 32$; **Figure 1**). All transcript responses were similar to previous observations (Hinther et al., 2010a,b, 2011).

The C-fin experimental design allows for the determination of relationship between the individual animal's baseline transcript levels and the extent of change in transcript level in response to chemical treatment. We examined the correlation between the baseline (control) levels of each transcript to each other and to the individual's level of fold induction in response to T_3 exposure (**Table 1**). A strong negative correlation between baseline transcript levels and the degree of fold response to T_3 for all four transcripts was observed (**Table 1**). This observation was consistent with previously reported observations for *TR*β and *RLKI* transcripts (Hinther et al., 2010a). A strong negative correlation was observed between the baseline levels of *CAT* and the

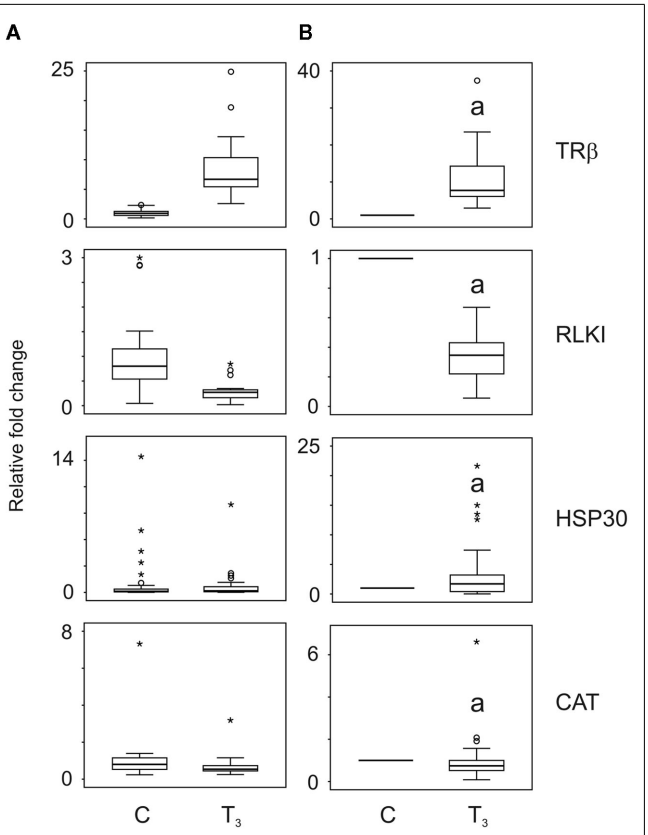


FIGURE 1 | The effect of application of repeated measures analysis on QPCR data generated for thyroid hormone receptor β (*TR*β), Rana larval keratin I (*RLKI*), heat shock protein 30 (*HSP30*), and catalase (*CAT*) transcripts. Tail fin biopsies were exposed to vehicle control (C) or 10 nM T_3. The data ($n = 23$–32) are plotted as independent measures in **(A)** and then as repeated measures in **(B)**. Taking the individual's baseline expression levels greatly reduces overall variation and enables the identification of perturbation of a response to T_3 relative to each individual's normal T_3 response. Box plots show medians ± first and third quartiles. The whiskers indicate minimum and maximum values excluding outliers and extreme values. Outlier (cases between 1.5 and 3.0 box lengths from the upper or lower edge of the box) and extreme values (cases >3.0 box lengths from the upper or lower edge of the box) are indicated by an open circle and asterisk, respectively. Statistical significance is indicated with the letter "a" for $p < 0.05$.

T_3-dependent reduction of *RLKI* transcripts and *RLKI* and the T_3-dependent decrease of the *CAT* transcript (**Table 1**). A positive correlation was observed between the baseline levels of *HSP30* and the T_3-dependent increase of *TR*β mRNAs (**Table 1**).

Exposure to 0.5 and 5 mg/L NO_2–N (in the form of $NaNO_2$) for 48 h did not have any effect on the TH-responsive gene transcripts, *TR*β and *RLKI*, in the absence ($p = 0.683$ and 0.257, respectively; Friedman) or presence of T_3 ($p = 0.282$ and 0.751, respectively, Friedman; **Figure 2**). Exposure to 5 mg/L sodium gluconate (Na-G; as a control for sodium) also did not result in a significant effect compared to the control ($p = 0.300$–0.875, Wilcoxon; **Figure 2**).

Exposure to 5 and 50 mg/L NO_3–N (in the form of $NaNO_3$) did not result in a change in *TR*β and *RLKI* transcript levels

Table 1 | Spearman's rho correlation analysis comparing baseline transcript levels with extent of (fold) induction in response to T_3 treatment.

			Fold induction by T_3			
			TRβ	**RLKI**	**HSP30**	**CAT**
Baseline	TRβ	Correlation coefficient	−0.645	−0.065	0.159	0.018
		p Value	0.000*	0.396	0.240	0.468
		N	23	19	22	23
	RLKI	Correlation coefficient	0.058	−0.570	−0.124	−0.350
		p Value	0.407	0.001*	0.278	0.040*
		N	19	26	25	26
	HSP30	Correlation coefficient	0.452	−0.015	−0.552	0.025
		p Value	0.017*	0.471	0.001*	0.448
		N	22	25	31	31
	CAT	Correlation coefficient	0.136	−0.465	−0.006	−0.618
		p Value	0.267	0.008*	0.487	0.000*
		N	23	26	31	32

Significance is indicated with an asterisk.

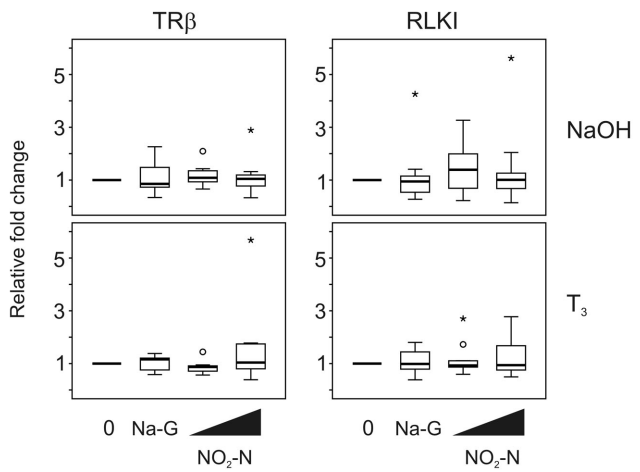


FIGURE 2 | QPCR analysis of thyroid hormone receptor β (*TR*β) and Rana larval keratin I (*RLKI*) transcript levels in the C-fin assay after exposure to nitrite in the absence or presence of 10 nM T_3. Tail fin biopsies were exposed to vehicle control (water; 0) and the indicated test chemicals for 48 h in the presence of 400 nM NaOH or 10 nM T_3 in 400 nM NaOH solvent. Test chemical concentrations were 5 mg/L sodium control (in the form of sodium gluconate, Na-G), 0.5 and 5 mg/L NO_2–N (in the form of $NaNO_2$). The results are expressed as fold change relative to the vehicle control (NaOH; upper panels) or to the vehicle +T_3-induced levels (lower panels) and represent QPCR data from $n = 7$–12 animals. Increasing concentrations of test chemicals are represented by bevels. See **Figure 1** legend for more details.

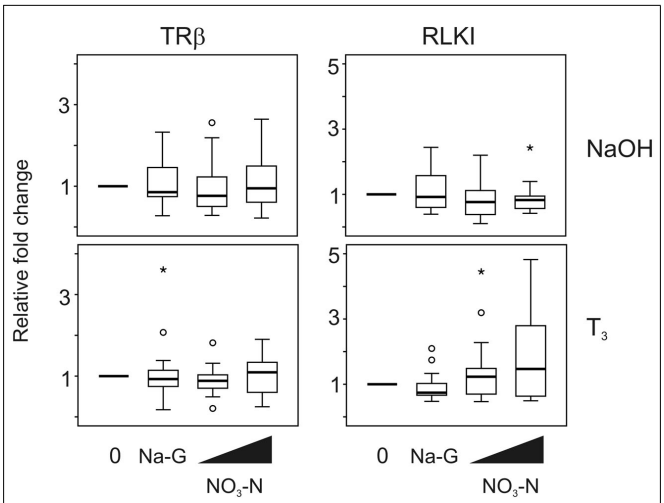


FIGURE 3 | QPCR analysis of thyroid hormone receptor β (*TR*β) and Rana larval keratin I (*RLKI*) transcript levels in the C-fin assay after exposure to nitrate in the absence or presence of 10 nM T_3. Tail fin biopsies were exposed to vehicle control (water; 0) and the indicated test chemicals for 48 h in the presence of 400 nM NaOH or 10 nM T_3 in 400 nM NaOH solvent. Test chemical concentrations were 50 mg/L sodium control (in the form of sodium gluconate, Na-G), 5 and 50 mg/L NO_3–N (in the form of $NaNO_3$). The results are expressed as fold change relative to the vehicle control (NaOH; upper panels) or to the vehicle +T_3-induced levels (lower panels) and represent QPCR data from $n = 14$–16 animals. Increasing concentrations of test chemicals are represented by bevels. See **Figure 1** legend for more details.

in the absence ($p = 0.565$ and 0.913, respectively, Friedman; **Figure 3**) or presence of T_3 ($p = 0.066$ and 0.529, respectively, Friedman; **Figure 3**). The 50 mg/L sodium control (in the form of sodium gluconate, Na-G) in this experiment had no effect as well ($p = 0.480$–1.000, Wilcoxon; **Figure 3**).

Exposure to 0.5 and 5 mg/L NO_2–N did not affect *HSP30* and *CAT* transcript levels in the absence ($p = 0.444$ and 0.185, respectively, Friedman; **Figure 4**) or presence of T_3 ($p = 0.570$ and 0.779, respectively, Friedman; **Figure 4**). Exposure to 5 mg/L Na-G also did not result in a significant effect ($p = 0.438$–0.717, Wilcoxon; **Figure 4**).

Exposure to 5 and 50 mg/L NO_3–N did not result in a change in stress-responsive transcript levels in the absence ($p = 0.282$ and 0.819, *HSP30* and *CAT* transcripts respectively, Friedman) or presence of T_3 ($p = 0.074$ and 0.819, respectively, Friedman; **Figure 5**). Exposure to 50 mg/L Na-G had no effect on the stress-responsive transcripts ($p = 0.796$–1.000, Wilcoxon; **Figure 5**).

Changes in population variation have been associated with endocrine disruptive events and exposure to pollutants (Orlando and Guillette, 2001). An increase in variance is often found in contaminant-exposed sites compared with reference site populations, in part, due to varying individual responses to the environmental stressor. Since more individuals are at the perimeter of a population range away from the more homogeneous central part of the range, the contaminant-exposed population is less able to adapt to environmental stress and may require additional energy budget expenditures to survive (Orlando and Guillette, 2001). Thus, variation can represent an additional indicator of population health not necessarily captured by measures of central

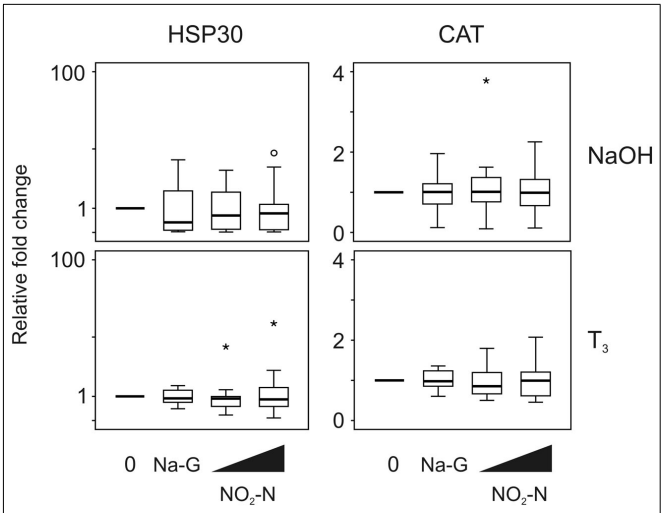


FIGURE 4 | QPCR analysis of heat shock protein 30 (*HSP30*) and catalase (*CAT*) transcript levels in the C-fin assay after exposure to nitrite in the absence or presence of 10 nM T_3. Tail fin biopsies were exposed to vehicle control (water; 0) and the indicated test chemicals for 48 h in the presence of 400 nM NaOH or 10 nM T_3 in 400 nM NaOH solvent. Test chemical concentrations were 5 mg/L sodium control (in the form of sodium gluconate, Na-G), 0.5 and 5 mg/L NO_2–N (in the form of $NaNO_2$). The results are expressed as fold change relative to the vehicle control (NaOH; upper panels) or to the vehicle +T_3-induced levels (lower panels) and represent QPCR data from $n = 15$–16 animals. Increasing concentrations of test chemicals are represented by bevels. See **Figure 1** legend for more details.

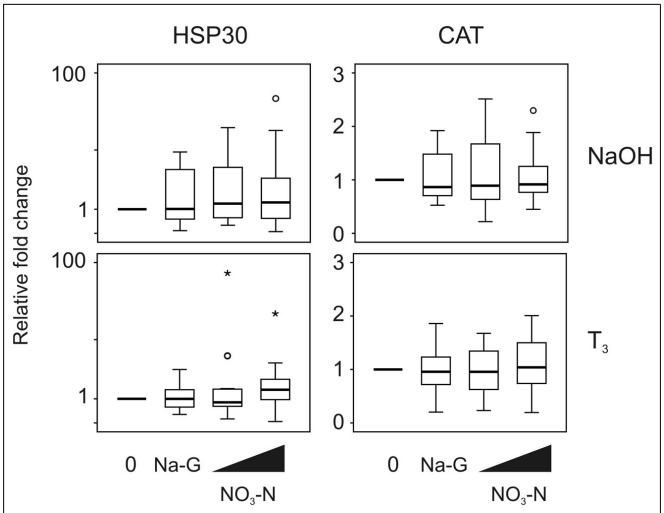


FIGURE 5 | QPCR analysis of heat shock protein 30 (*HSP30*) and catalase (*CAT*) transcript levels in the C-fin assay after exposure to nitrate in the absence or presence of 10 nM T_3. Tail fin biopsies were exposed to vehicle control (water; 0) and the indicated test chemicals for 48 h in the presence of 400 nM NaOH or 10 nM T_3 in 400 nM NaOH solvent. Test chemical concentrations were 50 mg/L sodium control (in the form of sodium gluconate, Na-G), 5 and 50 mg/L NO_3–N (in the form of $NaNO_3$). The results are expressed as fold change relative to the vehicle control (NaOH; upper panels) or to the vehicle +T_3-induced levels (lower panels) and represent QPCR data from $n = 16$ animals. Increasing concentrations of test chemicals are represented by bevels. See **Figure 1** legend for more details.

tendency (Orlando and Guillette, 2001). No alterations in heteroscedasticity were observed for any transcripts between the Na-G controls or the nitrite treatments (**Table 2**). This was also the case for nitrate in the absence of hormone. However, when T_3 was present, nitrate exposure affected the degree of heteroscedasticity in both *TR*β and *RLKI* mRNAs, but not *HSP30* or *CAT* transcripts (**Table 2**). A change in heteroscedasticity, as observed with nitrate exposure, suggests an alteration in the response to TH at the tissue level that is consistent with the conflicting accelerator and inhibitory effects on TH-dependent processes that have previously been observed (Xu and Oldham, 1997; Edwards et al., 2006; Ortiz-Santaliestra and Sparling, 2007). The data in the present study suggest that nitrate and nitrite differ in cellular effects on TH signaling while not eliciting stress responses in the TH-responsive tail fin tissue. Moreover, direct cellular effects of nitrate on peripheral tissues as a mechanism in influencing metamorphosis still remains a possibility but that this effect is not straightforward. Examination of additional time points would be useful to evaluate whether TH-mediated response kinetics are altered.

It has been postulated that nitrite and nitrate could act as nitric oxide donors through a non-genomic mechanism (Guillette and Edwards, 2005; Hannas et al., 2010). Nitric oxide donors have been shown to mimic the ability of T_4 to suppress catalase enzyme activity associated with tail shortening and apoptosis *in vitro* (Kashiwagi et al., 1999). However, a definitive connection between nitrate and nitrite and nitric oxide production in amphibian tissues has not been established, although this relationship has been shown in *Daphnia* (Hannas et al., 2010). The fact that nitrate elicited some response whereas nitrite did not on cultured tail fin suggest that this influence could be limited. Direct effects of nitrate and nitrite upon other amphibian tissues (such as the thyroid gland) through genomic and/or non-genomic

Table 2 | Analysis of variation using Levine's test.

Treatment	Transcript	Levene statistic	*p* Value
NO_2–N	*TR*β	0.484	0.624
	RLKI	0.261	0.772
	HSP30	0.563	0.574
	CAT	0.526	0.595
NO_3–N	*TR*β	0.183	0.834
	RLKI	1.073	0.352
	HSP30	2.891	0.066
	CAT	0.348	0.708
NO_2–N + T_3	*TR*β	2.786	0.088
	RLKI	0.174	0.841
	HSP30	1.698	0.196
	CAT	0.205	0.815
NO_3–N + T_3	*TR*β	3.449	0.041*
	RLKI	7.542	0.002*
	HSP30	1.254	0.296
	CAT	0.102	0.903

Groups with significant heteroscedasticity are indicated with an asterisk.

methods from tissue culture experiments and comparison to molecular responses elicited from whole animal exposures remain to be determined.

ACKNOWLEDGMENTS

This work was funded through a grant from the Natural Sciences and Engineering Research Council to CCH.

REFERENCES

Camargo, J. A., Alonso, A., and de la Puente, M. (2004). Multimetric assessment of nutrient enrichment in impounded rivers based on benthic macroinvertebrates. *Environ. Monit. Assess.* 96, 233–249.

Camargo, J. A., Alonso, A., and Salamanca, A. (2005). Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates. *Chemosphere* 58, 1255–1267.

CCME. (2007). *Canadian Water Quality Guidelines for the Protection of Aquatic Life, Environment Canada.* Ottawa: Canadian Council of Ministers of the Environment.

Crain, E. F. (2000). Environmental threats to children's health: a challenge for pediatrics: 2000 ambulatory pediatric association (APA) presidential address. *Pediatrics* 106, 871–875.

Crow, G. L., Luer, W. H., and Harshbarger, J. C. (2001). Histological assessment of goiters in elasmobranch fishes. *J. Aquat. Anim. Health* 13, 1–7.

Crump, D., Werry, K., Veldhoen, N., Van Aggelen, G., and Helbing, C. C. (2002). Exposure to the herbicide acetochlor alters thyroid hormone-dependent gene expression and metamorphosis in *Xenopus laevis*. *Environ. Health Perspect.* 110, 1199–1205.

Deane, E. E., Zhou, L., and Woo, N. Y. (2007). Effects of mitogenic hormones on HSP70 expression in a silver sea bream fibroblast cell line and a primary macrophage preparation. *Gen. Comp. Endocrinol.* 152, 183–188.

De Groef, B., Decallonne, B. R., Van der Geyten, S., Darras, V. M., and Bouillon, R. (2006). Perchlorate versus other environmental sodium/iodide symporter inhibitors: potential thyroid-related health effects. *Eur. J. Endocrinol.* 155, 17–25.

Domanski, D., and Helbing, C. C. (2007). Analysis of the *Rana catesbeiana* tadpole tail fin proteome and phosphoproteome during T3-induced apoptosis: identification of a novel type I keratin. *BMC Dev. Biol.* 7, 94. doi:10.1186/1471-213X-7-94

Edwards, T. M., McCoy, K. A., Barbeau, T., McCoy, M. W., Thro, J. M., Guillette, L. J. Jr. (2006). Environmental context determines nitrate toxicity in Southern toad (Bufo terrestris) tadpoles. *Aquat. Toxicol.* 78, 50–58.

European Union Commission. (2006). *Environmental Fact Sheet, "REACH – A New Chemicals Policy for the EU."* Brussels: European Commission.

Gray, B. R., Hill, W. R., and Stewart, A. J. (2001). Effects of development time, biomass and ferromanganese oxides on nickel sorption by stream periphyton. *Environ. Pollut.* 112, 61–71.

Guillette, L. J., and Edwards, T. M. (2005). Is nitrate an ecologically relevant endocrine disruptor in vertebrates? *Integr. Comp. Biol.* 45, 19–27.

Hampel, R., and Zollner, H. (2004). Iodine supply and incidence of goitrogenics in Germany. *Ernahrungs Umschau* 51, 132.

Hannas, B. R., Das, P. C., Li, H., and LeBlanc, G. A. (2010). Intracellular conversion of environmental nitrate and nitrite to nitric oxide with resulting developmental toxicity to the crustacean *Daphnia magna*. *PLoS ONE* 5, e12453. doi:10.1371/journal.pone.0012453

Health Canada. (2008). *Guidelines for Canadian Drinking Water Quality Summary Table.* Ottawa: Government of Canada, 1–14.

Helbing, C., Bailey, C., Ji, L., Gunderson, M., Zhang, F., Veldhoen, N., Mu, R., Lesperance, M., Holcombe, G., Kosian, P., Tietge, J., Korte, J. J., and Degitz, S. (2007a). Identification of gene expression indicators for thyroid axis disruption in a *Xenopus laevis* metamorphosis screening assay Part 1: effects on the brain. *Aquat. Toxicol.* 82, 227–241.

Helbing, C. C., Ji, L., Bailey, C. M., Veldhoen, N., Zhang, F., Holcombe, G. W., Kosian, P. A., Tietge, J., Korte, J. J., and Degitz, S. J. (2007b). Identification of gene expression indicators for thyroid axis disruption in a *Xenopus laevis* metamorphosis screening assay Part 2. Effects on the tail and hindlimb. *Aquat. Toxicol.* 82, 215–226.

Hinther, A., Bromba, C., Wulff, J., and Helbing, C. (2011). Effects of triclocarban, triclosan, and methyl triclosan on thyroid hormone action and stress in frog and mammalian culture systems. *Environ. Sci. Technol.* 45, 5395–5402.

Hinther, A., Domanski, D., Vawda, S., and Helbing, C. C. (2010a). C-fin: a cultured frog tadpole tail fin biopsy approach for detection of thyroid hormone-disrupting chemicals. *Environ. Toxicol. Chem.* 29, 380–388.

Hinther, A., Vawda, S., Skirrow, R. C., Veldhoen, N., Collins, P., Cullen, J. T., van Aggelen, G., and Helbing, C. C. (2010b). Nanometals induce stress and alter thyroid hormone action in amphibia at or below North American water quality guidelines. *Environ. Sci. Technol.* 44, 8314–8321.

Ji, L., Domanski, D., Skirrow, R. C., and Helbing, C. C. (2007). Genistein prevents thyroid hormone-dependent tail regression of *Rana catesbeiana* tadpoles by targetting protein kinase C and thyroid hormone receptor alpha. *Dev. Dyn.* 236, 777–790.

Kashiwagi, A., Hanada, H., Yabuki, M., Kanno, T., Ishisaka, R., Sasaki, J., Inoue, M., and Utsumi, K. (1999). Thyroxine enhancement and the role of reactive oxygen species in tadpole tail apoptosis. *Free Radic. Biol. Med.* 26, 1001–1009.

Onken, H., Tresguerres, M., and Luquet, C. M. (2003). Active NaCl absorption across posterior gills of hyperosmoregulating *Chasmagnathus granulatus*. *J. Exp. Biol.* 206, 1017–1023.

Opitz, R., Lutz, I., Nguyen, N. H., Scanlan, T. S., and Kloas, W. (2006). Analysis of thyroid hormone receptor beta A mRNA expression in *Xenopus laevis* tadpoles as a means to detect agonism and antagonism of thyroid hormone action. *Toxicol. Appl. Pharmacol.* 212, 1–13.

Orlando, E. F., and Guillette, L. J. Jr. (2001). A re-examination of variation associated with environmentally stressed organisms. *Hum. Reprod. Update* 7, 265–272.

Ortiz-Santaliestra, M. E., and Sparling, D. W. (2007). Alteration of larval development and metamorphosis by nitrate and perchlorate in southern leopard frogs (*Rana sphenocephala*). *Arch. Environ. Contam. Toxicol.* 53, 639–646.

Porter, W. P., Jaeger, J. W., and Carlson, I. H. (1999). Endocrine, immune, and behavioral effects of aldicarb (carbamate), atrazine (triazine) and nitrate (fertilizer) mixtures at groundwater concentrations. *Toxicol. Ind. Health* 15, 133–150.

Rouse, J. D., Bishop, C. A., and Struger, J. (1999). Nitrogen pollution: an assessment of its threat to amphibian survival. *Environ. Health Perspect.* 107, 799–803.

Sampat, P. (2000). Groundwater shock: the polluting of the world's major freshwater stores. *World Watch* 10–22.

Sharma, B., and Ahlert, R. C. (1977). Nitrification and nitrogen removal. *Water Res.* 11, 897–925.

Shi, Y.-B. (2000). *Amphibian Metamorphosis: From Morphology to Molecular Biology.* New York, NY: Wiley-Liss.

Skirrow, R. C., Veldhoen, N., Domanski, D., and Helbing, C. C. (2008). Roscovitine inhibits thyroid hormone-induced tail regression of the frog tadpole and reveals a role for cyclin C/Cdk8 in the establishment of the metamorphic gene expression program. *Dev. Dyn.* 237, 3787–3797.

Taylor, A. C., and Kollros, J. J. (1946). Stages in the normal development of *Rana pipiens* larvae. *Anat. Rec.* 94, 7–24.

US EPA. (2006). *National Recommended Water Quality Criteria. Office of Water, Office of Science and Technology.* Washington, DC: United States Environment Protection Agency.

US EPA. (2009). *National Drinking Water Regulations. Office of Water, Office of Science and Technology.* Washington, DC: United States Environment Protection Agency.

Veldhoen, N., Skirrow, R., Osachoff, H., Wigmore, H., Clapson, D., Gunderson, M., van Aggelen, G., and Helbing, C. (2006a). The bactericidal agent triclosan modulates thyroid hormone-associated gene expression and disrupts postembryonic anuran development. *Aquat. Toxicol.* 80, 217–227.

Veldhoen, N., Skirrow, R. C., Ji, L., Domanski, D., Bonfield, E. R., Bailey, C. M., and Helbing, C. C. (2006b). Use of heterologous cDNA arrays and organ culture in the detection of thyroid hormone-dependent responses in a sentinel frog, *Rana catesbeiana*. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 1, 187–199.

Wyngaarden, J. B., Stanbury, J. B., and Rapp, B. (1953). The effects of iodine, perchlorate, thiocyanate, and nitrate administration upon the iodide concentrating mechanism of the rat thyroid. *Endocrinology* 52, 568–574.

Wyngaarden, J. B., Wright, B. M., and Ways, P. (1952). The effect of certain anions upon the accumulation and retention of iodide by the thyroid gland. *Endocrinology* 50, 537–549.

Xu, Q., and Oldham, R. S. (1997). Lethal and sublethal effects of nitrogen fertilizer ammonium nitrate on common toad (*Bufo bufo*) tadpoles. *Arch. Environ. Contam. Toxicol.* 32, 298–303.

Zaki, A., Chaoui, A. A., Talibi, A., Derouiche, A. F., Aboussaouira, T., Zarrouck, K., Chait, A., and Himmi, T. (2004). Impact of nitrate intake in drinking water on the thyroid gland activity in male rat. *Toxicol. Lett.* 147, 27–33.

Zhang, F., Degitz, S. J., Holcombe, G. W., Kosian, P. A., Tietge, J., Veldhoen, N., and Helbing, C. C. (2006). Evaluation of gene expression endpoints in the context of a *Xenopus laevis* metamorphosis-based bioassay to detect thyroid hormone disruptors. *Aquat. Toxicol.* 76, 24–36.

Conflict of Interest Statement: 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.

Received: 23 January 2012; paper pending published: 26 February 2012; accepted: 20 March 2012; published online: 04 April 2012.

Citation: Hinther A, Edwards TM, Guillette Jr. LJ and Helbing CC (2012) Influence of nitrate and nitrite on thyroid hormone responsive and stress-associated gene expression in cultured Rana catesbeiana tadpole tail fin tissue. Front. Gene. ***3****:51. doi: 10.3389/fgene.2012.00051*

This article was submitted to Frontiers in Toxicogenomics, a specialty of Frontiers in Genetics.