Next generation genetics

Mogens Fenger *

Clinical Biochemistry, Molecular Biology, and Genetics, KBA339, Hvidovre, Denmark
Correspondence: mogens.fenger@hvh.regionh.dk

One of the goals in genetic research aims at identifying genes in biochemical and physiological processes to reveal genetic causes of rare and common diseases. Previous obstacles such as costly genotyping or sequencing have been reduced with the chip-based genomewide association studies (GWAS), now culminating with the latest toy—next generation sequencing methodologies (NGS). Concomitantly, computer technologies have evolved to an increasing use of multicore processors and distributed computing on large networks or grids. Although the technologies are not perfect, we now have unprecedented opportunities to perform genetic studies not possible just 10 years ago. The hype about these new technologies have been large, but all the promises have however not been fulfilled entirely as hoped for. Maybe because the hype has been more about the technologies as such, and less about their intended use. Millions of genetic variations have been detected by GWAS and NGS, but only a few have been linked to diseases—with almost no practical clinical significance. A major reason for this apparent deadlock is the inadequacy of the models used, which are based on the traditional "Mendelian" approach, in which one gene is supposed to have a main effect on a trait or a disease. However, most genes claimed to be associated with a disease have small effects and only a tiny fraction of the genetic variance has been captured.

In this short notice it is argued, why this traditional approach should be supplemented or even replaced by modeling approaches in accordance with the complexity of biological systems, if we shall have any reasonable hope to understand the genetics behind any trait and bring genetics into practical use in medicine for common diseases (Costanzo et al., 2010; Ramanan et al., 2012).

EVOLUTION, FITNESS, AND EPISTASIS

Evolution of phyla is a complex process governed by genomic as well as environmental factors (Marshall, 2006). Much theoretic and practical work about evolution are based on theories of adaptive landscapes of fitness and natural selection, as advocated by Fisher in his geometrical model of adaption (Fisher, 1930; Martin and Lenormand, 2006; Chevin et al., 2010; Weinreich and Knies, 2013). In this model fitness is determined in a multidimensional landscape of phenotypes or traits, on which a selection pressure is imposed that limits the number of viable phenotypes. Although the space of theoretically phenotypes increases with the complexity of an organism, this may come with a cost of decreasing adaptability (Fisher, 1930; Orr, 2000; Martin and Lenormand, 2006; Borenstein and Krakauer, 2008). The Fisher model(s) is not explicitly rooted in genetic models but rather considers the phenotypic pleiotropic effect of mutations, i.e., particular genes and loci are not formulated in the model. In contrast, Wrights formulation of evolution (Wright, 1920) can be described as a multidimensional mutational or genetic landscape in which each dimension corresponds to a specific locus. In "modern" terms these ideas can be formulated as the occurrence of stabilizing selection acting on the increasing mutational load possibly involving pleiotropic behavior of a given mutation, that is a mutation may affect several endophenotypes (Weinreich et al., 2006; Masel and Siegal, 2009). Pleiotropicity also means that a selection pressure imposed on one endophenotype may not only constrict the number of viable genotypes but also inflict a collateral selection on other endophenotypes and genes (Gavrilets and De Jong, 1993; Snitkin and Segre, 2011). The latter may be regarded as "innocent" bystanders and the preserved genotypes may just be those that happen to be around at the moment of selection.

Fitness is a measure of the capability of survival and reproduction of a species as the result of integrated action of many subprocesses conditional on the imposed selection pressure. However, less fitness may not necessarily result in an entire loss of a phenotype or trait, but may prevail and in fact increase fitness if the selection pressure changes. This scenario is supported by the long known fact that a mutation may have major effect in one genotypic background, but may only have a minor influence in another and hence escape purging by selection (Nijhout and Paulsen, 1997; Kouyos et al., 2006). The fitness landscape (or any other trait landscape) may thus be roughed with several local optima. This is clearly obvious in the landscape of species, but is also present within a species (Marshall, 2006).

For long the question has been if a mutation impose a pure additive effect on fitness or if epistasis (the effect of mutations in a gene or regulatory structure on the effects of other genes) is the prevalent driving force in evolution. The effect of any mutation (genic as well as exgenic including possible changes in epigenetic processes) may be increased, buffered, or ameliorated in particular

genetic backgrounds, while having negative (even lethal) effect in other genetic backgrounds. Buffering is the essence in evolutionary theory of canalization and organismal robustness, in which the phenotype appears robust to mutational perturbations. Mutations may accumulate and appear as cryptic or neutral variations as long as they are not selected against (Masel and Siegal, 2009). These cryptic genetic variations may be revealed if some genetic or environmental changes happen affecting the fitness and then contribute to evolvability (McGuigan and Sgró, 2009).

Canalization (or buffering) implies that the phenotypic mean tends to be preserved when a mutation occurs, but the cost is diminished variation of the phenotype, as new (deleterious) mutations are buffered leaving less degrees of freedom of variation compared to the premutational genotype. Thus, a particular phenotype representing a local maximum in the phenotypic landscape, is generated by an ensemble of genotypes, each depicting a path or trajectory of the genetic network. Generally, the probability of a given genetic path being accessible to generate a phenotype decreases with the number of mutations. However, as the number of paths increases exponentially with the number of mutations a large and increasing number of paths may eventually generate a phenotype. This hypothesis has been confirmed empirically (Dowell et al., 2010; Franke et al., 2011). These and many other studies have firmly established epistasis as a primary driving force in evolution and as a fundamental principle in governing biological processes (Rice, 1998; Segre et al., 2005; Weinreich et al., 2005, 2013; Bershtein et al., 2006; Borenstein and Krakauer, 2008; Pavlicev et al., 2008; Chevin et al., 2010; Lunzer et al., 2010; Breen et al., 2012; Huang et al., 2012; Hemani et al., 2013; Weinreich and Knies, 2013).

The mechanism behind interactions and epistasis has been extensively studied and includes concepts as sign epistasis (Weinreich et al., 2005), reciprocal epistasis (Poelwijk et al., 2011), and the expansion of the protein universe (Povolotskaya and Kondrashov, 2010) to mention just a few outstanding contributions. The reader is referred to the cited work and to the vast literature appearing now.

THE GENETIC AND PHENOTYPIC SPACES

Complex species like humans are organized in interacting and interdependent functional units called organs or multicellular tissues. This extends the complexity of the genetics to another level. Despite the constrains this impose, the phenotypic space is vast.

Suppose that a diploid organism like *Homo sapiens* with 23 sets of chromosomes only harbor one mutation in each chromosome. The theoretically number of gametes emerging by random segregation amounts to approximately 8.4 million. If all gametes are viable then the number of possible zygotes will be more than $7*10^{13}$ or more than 11.000 fold the number human beings ever lived on planet Earth. Most probably a vast amount of the gametes or zygotes are not viable, but nevertheless, genetic variations so far discovered runs in the millions. This maps to as many phenotypes and hence, two human beings will never be genetically identical.

Similarly, in a physiological process like blood pressure, which are regulated by say 100 interacting genes, more than 10^{30} networks with exactly the same topology can be constructed if just one mutation is present in each gene. This would map to as many physiological states and dynamics. Adding to the number of genes their alternative spliced forms, the vast number of posttranslational modifications of proteins, non-protein regulatory elements (metabolites, small regulatory RNAs), epigenetic modifications, non-genic regulatory and genome-organizing structures, and not the least interactions and communications between cells in a multicellular organisms like humans, the combinatorial space of interactions and hence phenotypes is (almost) infinite.

POPULATION STRUCTURE AND GENETIC NETWORKS

Two basic aspects must be addressed in population genetics: (1) biological processes even in its most simple forms are blue-printed in the genome of interacting networks of genes; (2) the exp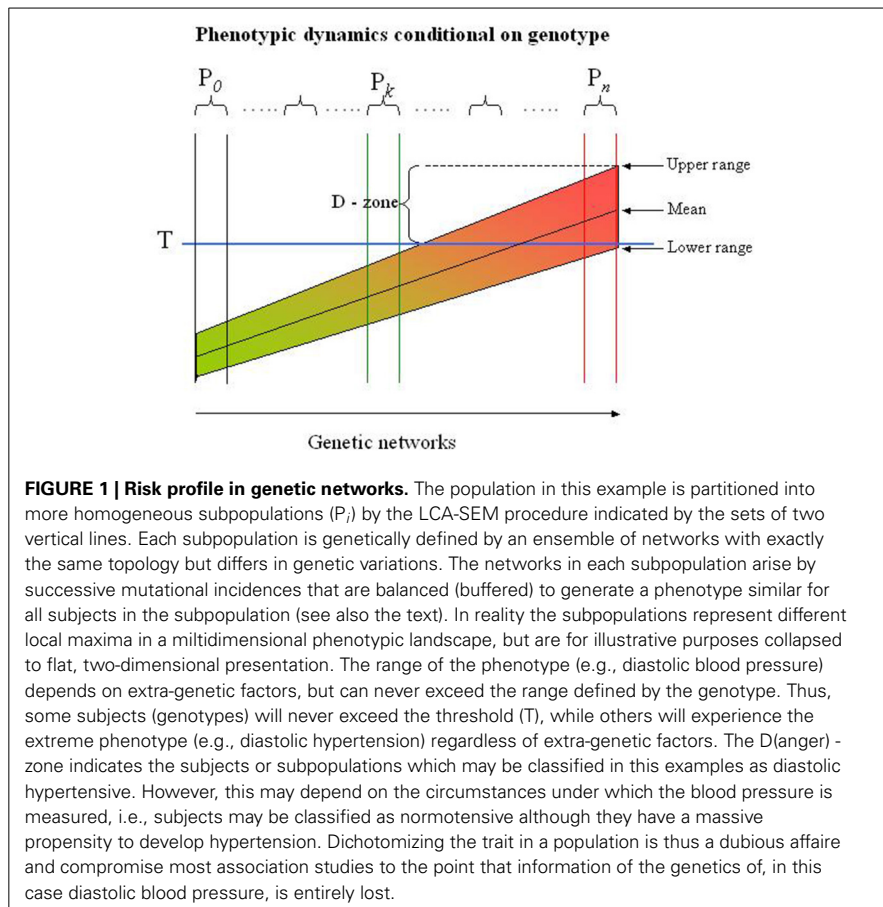ression of the biological processes and phenotypes are conditional on the genetic variations and their inherent epistasis.

The genetic networks coding a trait can be mapped as a "continuum" reflecting the physiological states they define (see the **Figure 1**). Neighboring networks can be distinguished by variations in one or several genes or non-genetic regulatory structures, but may appear physiological similar as most genetic variations have small effects. The sensitivity of a network to external factors is encoded in the genome, and it is the variations in the process-specific genes and regulatory structures that determines the range of the response to an external perturbation. Identifying genetic networks are neither simple, nor transparent: functional networks are multipartite structures and are not secluded entities but rather interconnected with other networks (e.g., the glucose and the fat metabolisms are highly intermingled processes). Nevertheless, it may be possible to define a reasonable number of sub-ensembles of networks to be interpretable.

HETEROGENEITY

Population heterogeneity refers to the mixture of phenotypically homogeneous subpopulations, although in the extreme no subpopulation is truly homogenous as each subject harbors a unique genomewide genotype. A more or less well-defined phenotype thus comprise an ensemble of genotypes in the population. The initial task is then to cluster subjects into more physiological homogeneous subpopulations to increase the accuracy and power of the genetic analysis (Fenger et al., 2008, 2011). The application of appropriate cluster algorithms is generally ill posed, as no universal formal criteria for the best clustering is available. Many of the well-known classification procedures implement some data-reduction or feature selection (Saeys et al., 2007), but any manipulations of the data space are likely to result in loss of information and should be avoided.

Allocating subjects to subpopulations is an art of modeling hidden or latent variables as the number of subpopulations is not known *a priori*. A way to resolve this is by applying the concept of latent class (LCA) in a

FIGURE 1 | Risk profile in genetic networks. The population in this example is partitioned into more homogeneous subpopulations (P$_i$) by the LCA-SEM procedure indicated by the sets of two vertical lines. Each subpopulation is genetically defined by an ensemble of networks with exactly the same topology but differs in genetic variations. The networks in each subpopulation arise by successive mutational incidences that are balanced (buffered) to generate a phenotype similar for all subjects in the subpopulation (see also the text). In reality the subpopulations represent different local maxima in a miltidimensional phenotypic landscape, but are for illustrative purposes collapsed to flat, two-dimensional presentation. The range of the phenotype (e.g., diastolic blood pressure) depends on extra-genetic factors, but can never exceed the range defined by the genotype. Thus, some subjects (genotypes) will never exceed the threshold (T), while others will experience the extreme phenotype (e.g., diastolic hypertension) regardless of extra-genetic factors. The D(anger) - zone indicates the subjects or subpopulations which may be classified in this examples as diastolic hypertensive. However, this may depend on the circumstances under which the blood pressure is measured, i.e., subjects may be classified as normotensive although they have a massive propensity to develop hypertension. Dichotomizing the trait in a population is thus a dubious affaire and compromise most association studies to the point that information of the genetics of, in this case diastolic blood pressure, is entirely lost.

structural equation modeling framework (SEM) (Bollen, 1989; Muthen, 2002; Skrondal and Rabe-Hesketh, 2007; Fenger et al., 2008, 2011). The idea of the LCA-SEM approach is to model a physiological process, and therefore the most appropriate study population would be a random selection of subjects as each subject provide information of the physiological process. Genetic structures and variations are not necessarily modeled directly, but are embedded as latent variables in the SEM structure and are mapped or reflected by the measured variables. Modeling in this framework addresses two pivotal issues in complex data: resolving the heterogeneity in the population, and simultaneously evaluating the data structure within the sub-populations. This approach outperforms most other classification methods in almost all aspects (Magidson and Vermunt, 2002).

An emerging line of methods implement ensembles of classification functions (Polikar, 2006). These approaches are particularly attractive when features in multi-source or distributed data sets are partly or completely disjoint, or if access to data in data set is limited to a subset of objects. Thus, the problem of missing data and hence reduced power may be circumvented to some extent and represent a potential alternative to imputing missing genetic data.

Undoubtedly, new and promising methods will merge, in particular as theoretical ideas mentioned below are integrated in future developments.

INHERITANCE: GENES OR INFORMATION?

Understanding and integrating the wealth of genetic data in a biological and medical context requires new approaches and techniques. Fortunately many new approaches are emerging increasingly embracing the nature of biological systems. In particularly the recent developments in network theory (Dorogovtsev and Mendes, 2002; Newman, 2009) including concepts of modularity (Newman and Girvan, 2004), stochastic block modeling (Karrer and Newman, 2011), statistical mechanics (Reichardt and Bornholdt, 2006; Ronhovde and Nussinov, 2009), and information theory (Anand and Bianconi, 2010) are promising.

All these methods comprise passing of information that have its corollary in genetics. Stabilizing selection may give rise to prevailing linkage disequilibrium of genes within and across chromosomes (Fenger et al., 2011). Such linkage disequilibrium arise as a consequence of preserving physiological processes regardless of the physical structure of the genome. Thus, inheritance is not a simple matter of passing on genetic material, but rather to combine information harbored in the genome into a viable organism. Information theoretic approaches in genetics may therefore be more promising (how abstract it may be) than traditional association methods, although transformation of these theories to biological structures may not be always straightforward.

IS VALIDATION IN GENETICS ACTUALLY POSSIBLE?

Validation is a standard requirements in genetic association studies today. However, validation of an association of a genetic variant to a trait or disease is often not or only sporadically obtained and for that reason a gene may be dismissed as disease related (Ioannidis, 2007; Shriner and Vaughan, 2011), or even to be pivotal in a physiological process (Fenger et al., 2011; Spijkers et al., 2011). It should hopefully be clear from the discussions above that validation should be expected to be an exception. A non-validated association may simply indicate a local optimum in the phenotypic landscape that happens to be detected because the genotypes in a population are permissive for expressing an apparent main effect. It has indeed been demonstrated that the lack of validation actually suggest more complex genetic structures governs a trait (Greene et al., 2009; Fenger et al., 2011), and including epistasis in the analysis may eventual confirm the importance of non-validated associations.

In the end the importance of genetic variations should be confirmed by cellular experiments. If possible, studies should be done in the cells where the genes in a network has its effect (which we often

do not know). A gargantuan endeavor and at the moment maybe wild-fetched, but eventually any genetic variation has to be substantiated in a real biological context - not just as a statistical phenomenon.

REFERENCES

Anand, K., and Bianconi, G. (2010). Gibbs entropy of network ensembles by cavity methods. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 82:011116. doi: 10.1103/PhysRevE.82.011116

Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N., and Tawfik, D. S. (2006). Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature* 444, 929–932. doi: 10.1038/nature05385

Bollen, K. A. (1989). *Structural Equations with Latent Variables.* Hobroken, NJ: John Wiley and Sons.

Borenstein, E., and Krakauer, D. C. (2008). An end to endless forms: epistasis, phenotype distribution bias, and nonuniform evolution. *PLoS Comput. Biol.* 4:e1000202. doi: 10.1371/journal.pcbi.1000202

Breen, M. S., Kemena, C., Vlasov, P. K., Notredame, C., and Kondrashov, F. A. (2012). Epistasis as the primary factor in molecular evolution. *Nature* 490, 535–538. doi: 10.1038/nature11510

Chevin, L. M., Martin, G., and Lenormand, T. (2010). Fisher's model and the genomics of adaption: restricted pleiotropy, heterogenous mutation, and parallel evolution. *Evolution* 64, 3213–3231. doi: 10.1111/j.1558-5646.2010.01058.x

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., et al. (2010). The genetic landscape of a cell. *Science* 327, 425–431. doi: 10.1126/science.1180823

Dorogovtsev, S. N., and Mendes, J. F. (2002). Evolution of networks. *Adv. Phys.* 51, 1079–1187. doi: 10.1080/00018730110112519

Dowell, R. D., Ryan, O., Jansen, A., Cheung, D., Agarwala, S., Danford, T., et al. (2010). Genotype to phenotype: a complex problem. *Science* 328, 469. doi: 10.1126/science.1189015

Fenger, M., Linneberg, A., Jorgensen, T., Madsbad, S., Sobye, K., Eugen-Olsen, J., et al. (2011). Genetics of the ceramide/sphingosine-1-phosphate rheostat in blood pressure regulation and hypertension. *BMC Genet.* 12:44. doi: 10.1186/1471-2156-12-44

Fenger, M., Linneberg, A., Werge, T., and Jorgensen, T. (2008). Analysis of heterogeneity and epistasis in physiological mixed populations by combined structural equation modelling and latent class analysis. *BMC Genet.* 9:43. doi: 10.1186/1471-2156-9-43

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection.* London: Nature Publishing Group.

Franke, J., Klözer, A., de Visser, J. A., and Krug, J. (2011). Evolutionary accessibility of mutational pathways. *PLoS Comput. Biol.* 7:e1002134. doi: 10.1371/journal.pcbi.1002134

Gavrilets, S., and De Jong, G. (1993). Pleiotropic models of polygenic variation, stabilizing selection, and epistasis. *Genetics* 134, 609–625.

Greene, C. S., Penrod, N. M., Williams, S. M., and Moore, J. H. (2009). Failure to replicate a genetic association may provide important clues about genetic architecture. *PLoS ONE* 4:e5639. doi: 10.1371/journal.pone.0005639

Hemani, G., Knott, S., and Haley, C. (2013). An Evolutionary Perspective on epistasis and the missing heritability. *PLoS Genet.* 9:e1003295. doi: 10.1371/journal.pgen.1003295

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

Ioannidis, J. P. (2007). Non-replication and inconsistency in the genome-wide association setting. *Hum. Hered.* 64, 203–213. doi: 10.1159/000103512

Karrer, B., and Newman, M. E. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 83:016107. doi: 10.1103/PhysRevE.83.016107

Kouyos, R. D., Otto, S. P., and Bonhoeffer, S. (2006). Effect of varying epistasis on the evolution of recombination. *Genetics* 173, 589–597. doi: 10.1534/genetics.105.053108

Lunzer, M., Golding, G. B., and Dean, A. M. (2010). Pervasive cryptic epistasis in molecular evolution. *PLoS Genet.* 6:e1001162. doi: 10.1371/journal.pgen.1001162

Magidson, J., and Vermunt, J. K. (2002). Latent class models for clustering: a comparison with K-means. *Can. J. Market. Res.* 20, 37–44. Available online at: http://statisticalinnovations.com/technicalsupport/cjmr.pdf

Marshall, C. R. (2006). Explaining the Cambrian "Explosion" of animals. *Annu. Rev. Earth Planet. Sci.* 34, 355–384. doi: 10.1146/annurev.earth.33.031504.103001

Martin, G., and Lenormand, T. (2006). A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* 60, 893–907. doi: 10.1111/j.0014-3820.2006.tb01169.x

Masel, J., and Siegal, M. L. (2009). Robustness: mechanisms and consequences. *Trends Genet.* 25, 395–403. doi: 10.1016/j.tig.2009.07.005

McGuigan, K., and Sgró C. M. (2009). Evolutionary consequences of cryptic genetic variation. *Trends Ecol. Evol.* 24, 305–311. doi: 10.1016/j.tree.2009.02.001

Muthen, B. O. (2002). Beyond SEM: general latent variabel modeling. *Behaviormetrika* 29, 81–117. doi: 10.2333/bhmk.29.81

Newman, M. E. J. (2009). The structure and functions of complex networks. *SIAM Rev.* 45, 167–256. doi: 10.1137/S003614450342480

Newman, M. E. J., and Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 69:026113. doi: 10.1103/PhysRevE.69.026113

Nijhout, H. F., and Paulsen, S. M. (1997). Develomental models and polygenic characters. *Am. Nat.* 149, 394–405. doi: 10.1086/285996

Orr, H. A. (2000). Adaptation and the cost of complexity. *Evolution* 54, 13–20. doi: 10.1111/j.0014-3820.2000.tb00002.x

Pavlicev, M., Kenney-Hunt, J. P., Norgard, E. A., Roseman, C. C., Wolf, J. B., and Cheverud, J. M. (2008). Genetic variation in pleiotropy: differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. *Evolution* 62, 199–213. doi: 10.1111/j.1558-5646.2007.00255.x

Poelwijk, F. J., T-ânase-Nicola, S., Kiviet, D. J., and Tans, S. J. (2011). Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J. Theor. Biol.* 272, 141–144. doi: 10.1016/j.jtbi.2010.12.015

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits Syst. Mag.* 21–45. doi: 10.1109/MCAS.2006.1688199

Povolotskaya, I. S., and Kondrashov, F. A. (2010). Sequence space and the ongoing expansion of the protein universe. *Nature* 465, 922–926. doi: 10.1038/nature09105

Ramanan, V. K., Shen, L., Moore, J. H., and Saykin, A. J. (2012). Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends Genet.* 28, 323–332. doi: 10.1016/j.tig.2012.03.004

Reichardt, J., and Bornholdt, S. (2006). Statistical mechanics of community detection. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 74:016110. doi: 10.1103/PhysRevE.74.016110

Rice, S. H. (1998). The evolution of canalization and the breaking of von Baer's laws: modeling the evolution of development with epistasis. *Evolution* 52, 647–656. doi: 10.2307/2411260

Ronhovde, P., and Nussinov, Z. (2009). Multiresolution community detection for megascale networks by information-based replica correlations. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 80:016109. doi: 10.1103/PhysRevE.80.016109

Saeys, Y., Inza, I., and Larranaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics* 23, 2507–2517. doi: 10.1093/bioinformatics/btm344

Segre, D., Deluna, A., Church, G. M., and Kishony, R. (2005). Modular epistasis in yeast metabolism. *Nat. Genet.* 37, 77–83. doi: 10.1038/ng1489

Shriner, D., and Vaughan, L. K. (2011). A unified framework for multi-locus association analysis of both common and rare variants. *BMC Genomics* 12:89. doi: 10.1186/1471-2164-12-89

Skrondal, A., and Rabe-Hesketh, S. (2007). Latent variable modelling: a survey. *Scand. J. Stat.* 34, 712–745. doi: 10.1111/j.1467-9469.2007.00573.x

Snitkin, E. S., and Segre, D. (2011). Epistatic interaction maps relative to multiple metabolic phenotypes. *PLoS Genet.* 7:e1001294. doi: 10.1371/journal.pgen.1001294

Spijkers, L. J., van den Akker, R. F., Janssen, B. J., Debets, J. J., De Mey, J. G., Stroes, E. S., et al. (2011). Hypertension is associated with marked alterations in sphingolipid biology: a potential role for ceramide. *PLoS ONE* 6:e21817. doi: 10.1371/journal.pone.0021817

Weinreich, D. M., Delaney, N. F., Depristo, M. A., and Hartl, D. L. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312, 111–114. doi: 10.1126/science.1123539

Weinreich, D. M., and Knies, J. L. (2013). Fishers's geometric model of adaption meets the functional synthesis: data on pairwise epistasis fro fitness yields insights into the shape and size of

phenotypic space. *Evolution* 67, 2957–2972. doi: 10.1111/evo.12156

Weinreich, D. M., Lan, Y., Wylie, C. S., and Heckendorn, R. B. (2013). Should evolutionary geneticists worry about higher-order epistasis? *Curr. Opin. Genet. Dev.* 23, 700–707. doi: 10.1016/j.gde.2013.10.007

Weinreich, D. M., Watson, R. A., and Chao, L. (2005). Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59, 1165–1174. doi: 10.1554/04-272

Wright, S. (1920). The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. *Proc. Natl. Acad. Sci. U.S.A.* 6, 320–332. doi: 10.1073/pnas.6.6.320

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.