



HAL
open science

The differential view of genotype–phenotype relationships

Virginie Orgogozo, Baptiste Morizot, Arnaud Martin

► **To cite this version:**

Virginie Orgogozo, Baptiste Morizot, Arnaud Martin. The differential view of genotype–phenotype relationships. *Frontiers in Genetics*, 2015, 6, pp.179. <10.3389/fgene.2015.00179>. <hal-01354348>

HAL Id: hal-01354348

<https://hal.science/hal-01354348v1>

Submitted on 18 Aug 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

The differential view of genotype–phenotype relationships

Virginie Orgogozo^{1*}, Baptiste Morizot² and Arnaud Martin³

¹ CNRS, UMR 7592, Institut Jacques Monod, Université Paris Diderot, Paris, France, ² Aix Marseille Université, CNRS, CEPERC UMR 7304, Aix en Provence, France, ³ Department of Molecular Cell Biology, University of California, Berkeley, CA, USA

OPEN ACCESS

Edited by:

Sylvain Marcellini,
University of Concepcion, Chile

Reviewed by:

Jacob A. Tennesen,
Oregon State University, USA
Margarida Matos,
University of Lisbon, Portugal

*Correspondence:

Virginie Orgogozo,
CNRS, UMR 7592, Institut Jacques
Monod, Université Paris Diderot,
15 Rue Hélène Brion, 75013 Paris,
France
virginie.orgogozo@normalesup.org

Specialty section:

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

Received: 24 March 2015

Accepted: 28 April 2015

Published: 19 May 2015

Citation:

Orgogozo V, Morizot B and Martin A
(2015) The differential view
of genotype–phenotype relationships.
Front. Genet. 6:179.
doi: 10.3389/fgene.2015.00179

An integrative view of diversity and singularity in the living world requires a better understanding of the intricate link between genotypes and phenotypes. Here we re-emphasize the old standpoint that the genotype–phenotype (GP) relationship is best viewed as a connection between two differences, one at the genetic level and one at the phenotypic level. As of today, predominant thinking in biology research is that multiple genes interact with multiple environmental variables (such as abiotic factors, culture, or symbionts) to produce the phenotype. Often, the problem of linking genotypes and phenotypes is framed in terms of genotype and phenotype maps, and such graphical representations implicitly bring us away from the differential view of GP relationships. Here we show that the differential view of GP relationships is a useful explanatory framework in the context of pervasive pleiotropy, epistasis, and environmental effects. In such cases, it is relevant to view GP relationships as differences embedded into differences. Thinking in terms of differences clarifies the comparison between environmental and genetic effects on phenotypes and helps to further understand the connection between genotypes and phenotypes.

Keywords: genotype, phenotype, genetics, complex trait, GxE, GxG

Introduction

We sometimes seem to have forgotten that the original question in genetics was not what makes a protein but rather ‘what makes a dog a dog, a man a man.’

(Noble, 2006)

One fundamental question in biology is to understand what makes individuals, populations, and species different from each other. The concept of *phenotype*, which corresponds to the observable attributes of an individual, was coined in opposition to the *genotype*, the inherited material transmitted by gametes. Since the early proposal that genotypes and phenotypes form two fundamentally different levels of biological abstraction (Johannsen, 1911), the challenge has been to understand how they articulate with each other, how genotypes map onto phenotypes. In the last 15 years, more than 1,000 examples of DNA sequence changes have been linked to naturally occurring non-deleterious phenotypic differences between individuals or species in Eukaryotes (Martin and Orgogozo, 2013b). In human, the OMIM[®] catalog (Online Mendelian Inheritance in Man, <http://omim.org/>) compiling the genetic determinants of disease-related phenotypes totals more than 4,300 entries and a total of 2,493 published Genome-Wide Association Studies (GWAS) have been uncovering a wealth of sites in the genome that are statistically associated to complex traits

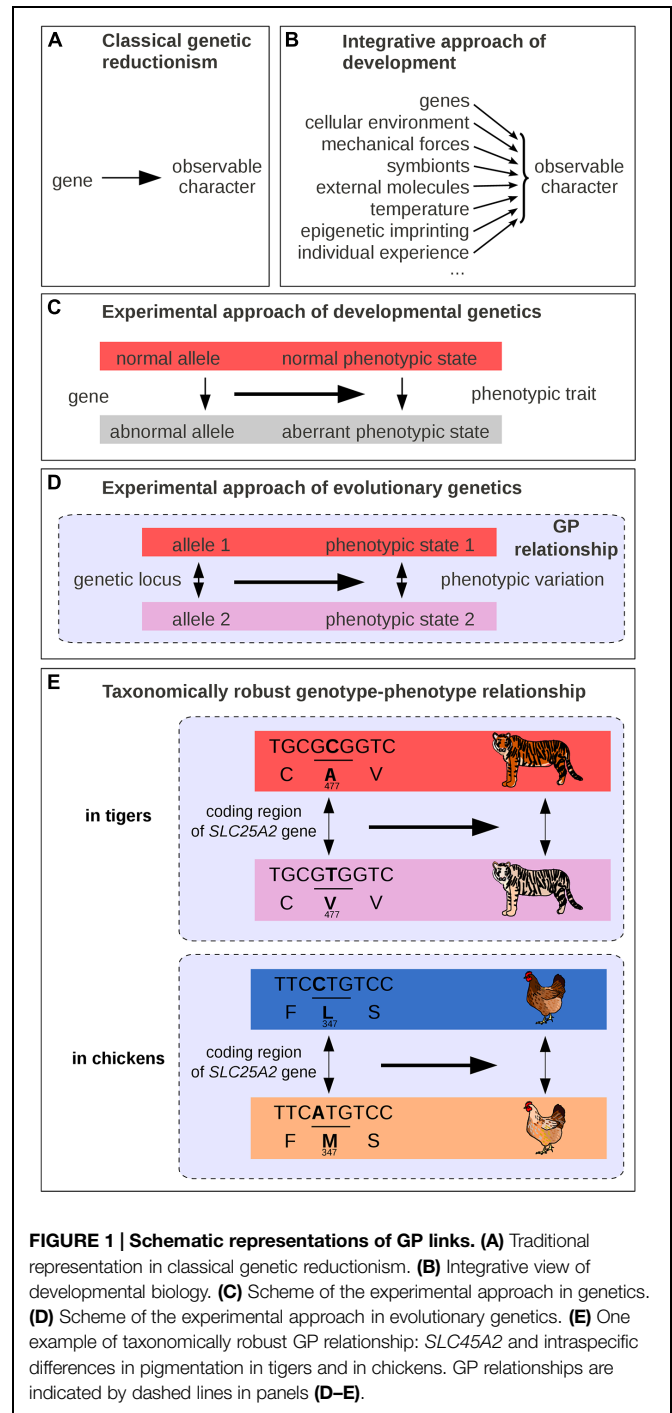
Abbreviations: GP, genotype–phenotype.

(Welter et al., 2014). As the detection of causal links between genetic and phenotypic variation is accelerating, a reexamination of our conceptual tools may help us in finding unifying principles within the swarm of data. Here we reflect on the relationship between genotypes and phenotypes and we address this essay to biologists who are willing to try to challenge their current understanding of phenotypes. We single out one useful point of view, the differential view. We then show that this simple framework remains insightful in the context of pervasive pleiotropy, epistasis, and environmental effects.

Genes as Difference Makers

Mutations isolated from laboratory strains have been instrumental to the understanding of the GP map. Under the classical scheme, a mutation is compared to a wild-type reference, and its phenotypic effects are used to infer gene function. This framework often leads to a semantic shortcut: from a genetic change causing a *variation* in phenotype, it is often convenient to assimilate the corresponding gene as a causal determinant of a trait (Keller, 2010; **Figure 1A**). It is common to find headlines expressing these simplifications, trumpeting to wide audiences the discovery of the “longevity” or “well-being” gene, that sacrifice scientific accuracy to psychological impact. Along these lines, should a gene whose mutation is lethal be called a “life gene”? What these over-simplified formulations truly mean is that *variation* at a given gene causes *variation* in a given phenotype (Dawkins, 1982; Schwartz, 2000; Waters, 2007). In fact, a gene alone can neither cause an observable phenotypic trait, nor can it be necessary and sufficient to the emergence of observable characteristics. Genes need a cellular environment, the combined action of multiple other genes, as well as certain physico-chemical conditions to have an observable effect on organisms (**Figure 1B**). For example, brown hair pigmentation in one human being is not just a product of the genes coding for pigment synthesizing enzymes but also of the presence of cells producing pigments of relevant substrate molecules (such as tyrosine for melanin), and of the amount of received sun light (Liu et al., 2013). Thus, the genetic reductionist approach, which only explores a few genetic parameters among the variety of causal factors, is vain to fully address the broad question of what makes hair brown, of what brings forth a particular biological structure, or process in its entirety. Nevertheless, genetic reductionism can be perfectly appropriate for identifying genetic loci where a change causes a phenotypic difference (**Figure 1C**). A *difference* in hair color between two individuals could be due in some cases to their genetic difference. We note, however, that not all phenotypic changes can be attributed to genetic changes. A difference in hair color could also be caused by non-genetic factors such as age, intensity of solar radiation or hair dyeing, or by a combination of both genetic and non-genetic differences.

While modern genetics was in its infancy, Alfred Sturtevant formulated the question of the GP map in simple terms: “one of the central problems of biology is that of differentiation – how does an egg develop into a complex many-celled organism?



That is, of course, the traditional major problem of embryology; but it also appears in genetics in the form of the question, “How do genes produce their effects?” (Sturtevant, 1932). For long some geneticists may have thought that they were dissecting the morphogenetic mechanisms underlying the *formation* of phenotypic traits, while their experimental approach were in fact uncovering genes whose absence or alteration (mutations, deletions, duplications, rearrangements, etc.) leads to phenotypic *differences* (compare **Figure 1A** with **Figure 1C**). In fact, the

sentence “your hair is brown” can be interpreted either as an absolute observation (a description of a particular assemblage of molecules containing defined levels of the dark pigment eumelanin and of the pale pigment pheomelanin) or with implicit reference to other possibilities (it is brown and not of another color). Misconceptions arise because phenotypes are usually defined relative to possibilities that are not formulated explicitly. Our minds and our language often tend to confuse the objects whose variation is under consideration with the variation itself (Keller, 2010), and it is essential to remind that, in genetics, the objects of interest (e.g., a given genotype, an allele or a phenotype) deserve to be defined *relatively* to another reference state.

In summary, the classical genetic reductionist approach is inherently unable to elucidate all the factors responsible for observable characteristics in the living world (Stotz, 2012) but is a powerful and relevant method for dissecting the genetic levers of heritable phenotypic variation. Focusing on phenotypic variation between individuals rather than on absolute characters present in single organisms is key to better comprehend the genetic causes of phenotypic diversity.

The GP Relationship is between Two Levels of Variation

Thinking in terms of differences makes apparent an abstract entity that encapsulates both genetic and phenotypic levels. This entity is composed of a variation at a genetic *locus* (two alleles), its associated phenotypic change (two distinct phenotypic states), and their relationships (Figure 1D). The three of us name the assemblage of these elements a “gophe,” but here we simply call it a “genotype–phenotype relationship” (GP relationship). We will show that the GP relationship is much more than a simple and loosely defined interaction between two levels of organization: it is a cause-and-effect connection that facilitates our understanding of phenotypic diversity.

The Genetic Part of a GP Relationship

In current genome annotation databases, a *gene* is usually defined as a stretch of nucleic acids that is transcribed and codes for an RNA or a polypeptide with a known or presumed function (Gerstein et al., 2007). The genetic locus underlying a phenotypic difference is not necessarily a gene in the strict sense; it can span a particular base-pair, a coding region, a *cis*-regulatory region, or extend to an entire gene with its *cis*-regulatory regions, or even to a gene cluster (Table 1). As previously noted by others (Falk, 1984; Gilbert, 2000; Stern, 2000; Moss, 2003; Griffiths and Stotz, 2013), the concept of gene in developmental biology and in current genome annotation databases is distinct from the concept of gene in evolutionary biology. Here the emphasis is not on the gene itself as defined in genome databases, but rather on a case-by-case functional partitioning of the genome into difference-making loci. The genotypic part of a GP relationship can take the form of various alleles: distinct codons coding for different amino acids, insertions/deletions within a protein coding sequence, diverging versions of a particular *cis*-regulatory element, presence/absence of transposon insertions,

number of gene copies within a gene cluster prone to structural variation, etc. Within a genome not all nucleotide sites are associated with phenotypic variation. For instance, there are probably fragments of nucleotide sequences, including the so-called junk DNA (Graur et al., 2015), whose presence does not have any consequence on observable characteristics of the organism, besides being replicated, and possibly transcribed. There are also genetic loci that may have been associated with phenotypic variation in the past and that are no longer associated with phenotypic variation. For example, genetic variation in histone DNA binding coding regions may have been important during the early evolution of eukaryotic cells, but these genetic loci no longer harbor phenotypically relevant variation besides lethal mutations. Within a genome, there are thus nucleotide sites that are absolutely required for life, but that do not harbor viable phenotypically relevant variation themselves.

The Phenotypic Part of a GP Relationship

The phenotypic counterpart of the GP relationship refers to a kind of variation (hair color, level of toxin resistance, etc.) rather than to a state (blond hair, taster of phenylthiocarbamide, etc.; Table 1).

The phenotype associated with a genetic change is not necessarily confined to the organism that harbors the genetic mutation. For example, the difference between left- and right-coiled shells in the snail *Lymnaea peregra* is determined by a single genetic locus with maternal effect: the genotype of the mother, but not of the individual itself, is responsible for the direction of shell coiling (Boycott et al., 1931). In other cases, the causal genetic change lies within symbiont bacteria: aphid thermal tolerance can vary between individuals due to a point mutation in their bacterial symbiont (Dunbar et al., 2007). Certain phenotypic effects can also come up at a level higher than the organism harboring the genetic change (Dawkins, 1982), one exemplary case being the social organization of an ant colony (Wang et al., 2013).

The Differential Part of a GP Relationship

As defined above, the GP relationship encompasses a genetic difference and a phenotypic difference. The relationship of *difference* at both the genetic and the phenotypic level is quite abstract, and it can correspond to three distinct differences within the living world: (#1) a difference between two reproductively isolated taxa (living or extinct), (#2) a difference segregating within a population, and (#3) the difference that first appeared during evolution, between an organism harboring the ancestral allele/trait and its direct descendant which evolved the new allele/trait. Of note, the variation in phenotype does not always immediately follow the emergence of the new causing mutation, but can appear later from the singular assortment of alleles that are segregating in the population. For example, a new phenotype of reduced armor plates appeared in a freshwater stickleback population when a recessive *EDA* allele already present at cryptic levels ended up in a homozygous state in one individual (Colosimo et al., 2005; Jones et al., 2012). A major conceptual advance made by Charles Darwin was to

TABLE 1 | A few examples of GP relationships.

Genetic locus	Phenotypic trait	Organisms
Two coding sites in the <i>ABO HBG</i> gene	A/B blood group	Human, chimpanzee, gibbon
<i>Cis</i> -regulatory element in the <i>lactase</i> gene	Ability to digest milk	Various human populations
Number of duplications of the amylase genes <i>AMY1</i> or <i>AMY2B</i>	Ability to digest starch	Human, dog
Presence/absence of a complex of adjacent genes coding for carotenoid desaturases and carotenoid cyclase/synthases	Ability to produce carotenoids	Pea aphid, spider mite, gall midge fly, nematode
Coding sites in <i>opsin</i> genes	Color vision	Human, cetaceans, fishes, butterflies
Coding region of <i>F 3'5'H</i>	Flower pigmentation	Soybean, pea, annual phlox, potato, <i>lochroma</i>
Coding region of <i>FRIGIDA</i>	Flowering time	Thale cress, oilseed rape
Coding region of <i>BADH2</i>	Fragrance	Soybean, rice
Coding region of <i>HMA3</i>	Heavy metal tolerance	Thale cress, rice
Coding and <i>cis</i> -regulatory regions of <i>myostatin</i>	Muscle size	Cattle, sheep, dog, pig, horse, human
<i>Cis</i> -regulatory regions in the <i>achaete-scute</i> complex	Number and position of sensory bristles	Fruitflies
<i>Cis</i> -regulatory element in the <i>pitx1</i> gene	Pelvis morphology	Stickleback fish
Coding region of <i>SLC45A2</i>	Pigmentation of eye, hair, and skin	Human, tiger, chicken
Coding region of <i>Mc1R</i>	Pigmentation of hair and skin, but not eye	Human, mouse, cattle, chicken, guinea pig, horse, fox, pig, sheep, dog, rabbit, bear, jaguar, jaguarundi, squirrel, birds, sand lizard
<i>Cis</i> -regulatory regions of <i>tan</i> and <i>ebony</i>	Pigmentation pattern	Fruitflies
<i>Cis</i> -regulatory elements in the <i>shavenbaby</i> gene	Position and number of trichomes	Fruitflies
<i>Cis</i> -regulatory elements in the <i>optix</i> gene	Red color pattern on butterfly wings	Longwing butterflies
Coding region of <i>hemoglobin alpha</i> and <i>beta</i> chain genes	Resistance to hypoxia	Human, llama, crocodile, deer mouse, waterfowl
Coding region of <i>Ace</i>	Resistance to organophosphate insecticides	Potato beetle, aphids, mosquitoes, house fly, fruit flies, oriental fruitfly
Presence/absence of a <i>CypA</i> insertion within the <i>TRIM5a</i> gene	Resistance to retrovirus	Owl monkey, macaque, Old World monkeys
Coding sites in the <i>Nav1.4</i> gene	Resistance to tetrodotoxin or saxotoxin	Snakes, pufferfish, clam
Coding region of <i>TAS2R38</i>	Sensitivity to bitterness	Human, chimpanzee
<i>Cis</i> -regulatory and coding regions in <i>Agouti</i>	Skin and coat pigmentation	Human, deer mouse, cattle, pig, dog, cat, horse, fox, domesticated fox, quail, sheep
Number of duplications of the glucose transporter gene <i>HXT6</i>	Survival in low-glucose environment	Yeast
Number of duplications of the <i>CCL3L1</i> gene	Susceptibility to HIV infection and progression rate of AIDS after infection	Human
<i>Cis</i> -regulatory element in the <i>WntA</i> gene	Wing pigmentation pattern	Butterflies

See (Martin and Orgogozo, 2013a) for references and for additional cases.

relate variation among individuals within an interbreeding group (difference #2) with variation between taxonomic groups in space and time (difference #1; Lewontin, 1974a).

Note also that certain phenotypic changes may appear at the level of the entire organism when the “causative” mutation is accompanied by additional somatic mutations that are highly likely. For example, in women carrying a wild-type allele and a mutant allele of *BRCA1*, cells can produce wild-type *BRCA1* proteins since they carry one copy of the wild-type *BRCA1* allele. Nevertheless, these women have up to an 80% risk of developing breast or ovarian cancer by age 70 compared to women carrying two wild-type copies of *BRCA1*, due to the appearance of additional deleterious mutations within the wild-type *BRCA1* allele in their somatic breast cells (Narod and Foulkes, 2004).

Importantly, the GP difference is always defined relative to a population, or taxon, of interest (Sober, 1988). In less medically developed countries, humans carrying two defective

copies of the phenylalanine hydroxylase gene have serious medical problems including seizure and intellectual disabilities. In contrast, in most developed countries, such humans are diagnosed at birth and have a normal life span with normal mental development thanks to a phenylalanine-restricted diet (Armstrong and Tyler, 1955). Therefore the GP relationship involving the phenylalanine hydroxylase defective mutation is context-dependent: the mutation is associated with health problems in less medically developed countries but not in other countries. This example shows that the causal relationship between a genetic change and its associated phenotypic change can hide multiple embedded parameters (such as medical practices for the phenylalanine hydroxylase case) within the *ceteris paribus* assumption of “all other things being equal.”

In summary, the GP relationship is best viewed as a relationship between two variations, one at the genotypic level, and one at the phenotypic level. The human mind can elaborate concepts of increasing abstraction: concepts of things (e.g., a cell),

concepts of change (e.g., evolution), and concepts of relations (e.g., homology; Cassirer, 1910; Simondon, 1968). Here the concept of GP relationship establishes a relation between two *changes* (genetic and phenotypic). In the next paragraphs we will show that, compared to the usage of intuitive concepts of things, this detour through increased abstraction may prove more efficient to better understand phenotypic diversity.

Several Current Representations of the Connection between Genotype and Phenotype Implicitly Dismiss the Differential View

We argued above that the differential view should always be kept in mind when thinking about the connection between genotypes and phenotypes. GWAS, which represent the most popular method to detect genomic loci that are associated with complex traits in populations, are based on the analysis of differences (Visscher et al., 2012). Nevertheless, in current research the differential view is sometimes implicitly dismissed. When multiple factors are observed to influence phenotypic traits (Figure 1B), the differential view is considered as too simplistic and researchers often prefer to focus back on phenotypes of single individuals, without explicitly relating them to a phenotypic reference.

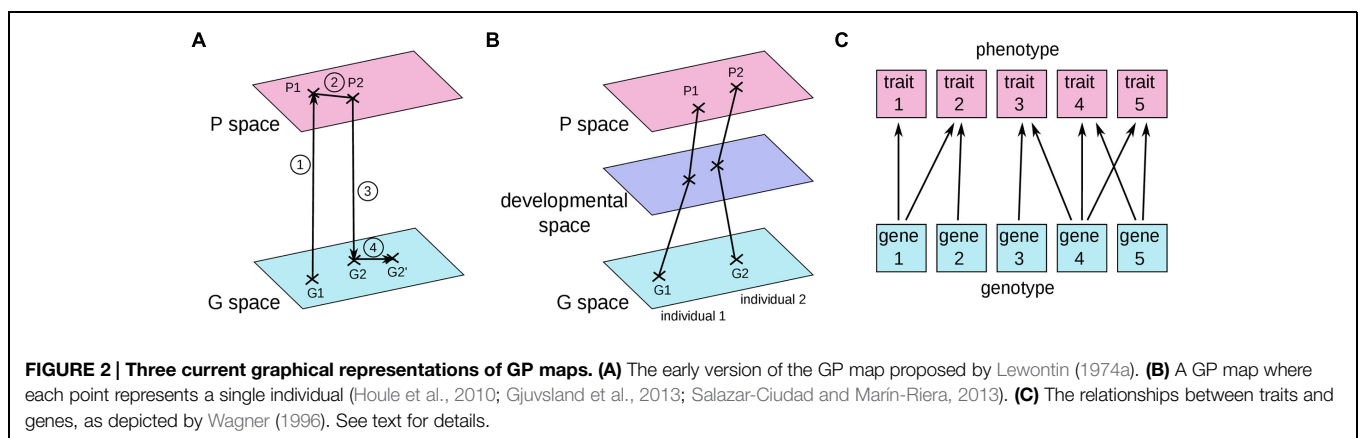
In most current articles, the problem of connecting the genotype to the phenotype is framed in terms of genotype and phenotype maps. The first GP map was introduced by Richard Lewontin in his book “The genetic basis of evolutionary change” (Lewontin, 1974a; Figure 2A). He indicated the average genotype of a population as a point in the space of all possible genotypes (G space) and the average phenotype of the same population as a corresponding point in the space of all possible phenotypes (P space). The evolutionary process was thus decomposed into four steps: (1) the average phenotype is derived from the development of the distinct genotypes in various environments; (2) migration, mating, and natural selection acts in P space to change the average phenotype of the initial population into the average phenotype of the individuals

which will have progeny; (3) the identity of successful parents determines which genotypes are preserved; and (4) genetic processes such as mutation and recombination modify position in G space.

In another common graphical representation (Figure 2B), a point in the G space and its corresponding point in the P space correspond to the genotype and the phenotype of a single individual (Fontana, 2002; Landry and Rifkin, 2012). Under such a representation, the abstract object that we defined above as the GP relationship would correspond to a “move” in genotype space associated with a “move” in phenotype space (or, better, a sum of several “moves” in genotype, and phenotype spaces because several distinct genomes can carry the two alternative alleles of a given GP relationship). In a third representation put forward by Wagner (1996; Figure 2C), individual genes are connected to individual traits.

Although these three graphical representations of GP maps may facilitate our understanding of certain aspects of biology, in all of them the GP relationship and the differential view are not easy to grasp. It is quite perplexing that the first person to draw such a GP map was Richard Lewontin, an eloquent advocate of the differential view (see for example his preface to Oyama, 2000, a masterpiece of persuasion). Because these graphics focus on individual rather than differential objects, we believe that these three representations implicitly incite us to go back to the more intuitive idea of one genotype associated with one phenotype. Losing sight of the differential view might also come from the molecular biology perspective, where proteins are viewed as having causal effects on their own, such as phosphorylation of a substrate or binding to a DNA sequence. Because of the two entangled definitions of the gene, either as encoding a protein, or as causing a phenotypic change (Griffiths and Stotz, 2013), it is easy to move from a differential view to a non-differential view of the GP relationship.

In summary, many current mental representations of the connection between genotype and phenotype implicitly dismiss the differential view. We will now show that the differential view is compatible with the fact that phenotypic traits are influenced by a complex combination of multiple factors and that we can find a relevant schematic representation of GP relationships.



The Problem of Pleiotropy

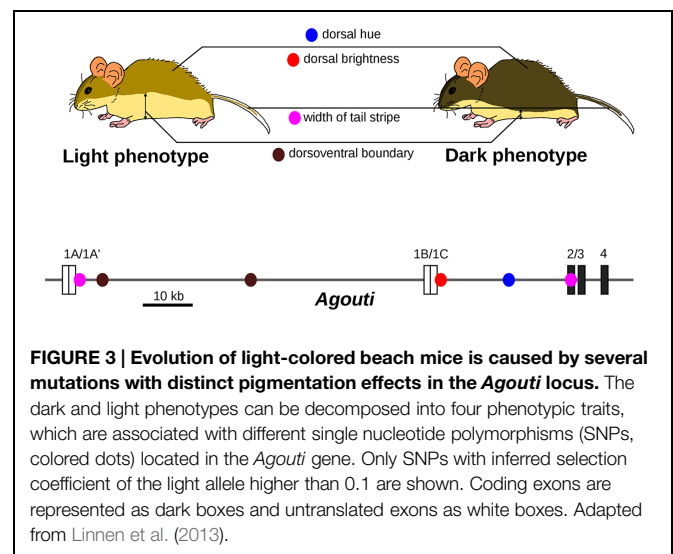
Decomposing an organism into elementary units such as anatomical structures has been instrumental in many biology disciplines such as physiology, paleontology and evolution. However, the issue is to identify the decomposition into characters that is most adequate for the question of interest. For questions related to relationships between organs of various individuals or species (such as homology), it might be appropriate to keep the traditional decomposition into anatomical structures (Wagner, 2014). Richard Lewontin and Günter Wagner defined characters as elements within an organism that answer to adaptive challenges and that represent quasi-independent units of evolutionary change (Lewontin, 1978; Wagner, 2000). Their definition deals with absolute traits observed in single organisms (for example the shape of a wing, or the number of digits in an individual) and is thus far from the differential view. Here, to better apprehend evolution and phenotypic diversity of the living world, we propose to decompose the observable attributes of an organism into multiple elementary GP variations that have accumulated through multiple generations, starting from an initial state. We insist that under this perspective, characters are not concrete objects (such as skin) but abstract entities defined by the existence of *differences* between two possible observable states (for example skin color). As an analogy, one can imagine two ways to produce a well-worn leather shoe of a particular shape. One can either assemble the different atoms into the same organization, or one can buy a shoe in a store and then subject it to a series of mechanical forces. We are naturally inclined to compare organisms to machines, and to think in terms of pieces that must be assembled to make a functional whole. However, the rampant metaphor of the designer or maker is inadequate for understanding the origin of present-day organisms (Coen, 2012). To understand the phenotypic features of a given organism it is more efficient to decompose it into abstract changes that occurred successively across evolutionary time, and not across developmental time. The initial state is a hypothetical ancestor of the organism under study.

Certain mutations (qualified as pleiotropic) are observed to affect several organs at once while others alter only one at a time (Paaby and Rockman, 2013; Zhang and Wagner, 2013). For pleiotropic mutations, we consider that the GP relationship should include all the phenotypic changes (in diverse organs, at various stages, etc.) associated with the genetic difference. For instance, the V370A mutation of the EDAR receptor is associated not only to hair thickness but also to changes in sweat gland and mammary gland density in Asian populations (Kamberov et al., 2013). The GP relationship is, in such cases, one-to-multi. Considering skin and eye as independent anatomical modules of the human body might seem appropriate for many evolutionary changes, but it is somewhat inadequate in cases where these two organs evolved a new pigmentation trait at once through a single mutation in the *SLC45A2* gene (Liu et al., 2013). Reasoning in terms of GP relationships strikes off the problem of finding a relevant decomposition into elementary anatomical structures. The elementary GP relationships themselves appear as adequate semi-independent

modules, whose combination can account for the observable characteristics of an organism.

The Problem of Continuous Complex Traits

Under the differential concept of GP relationships, one crucial point is to decompose observable traits into a series of semi-independent phenotypic variations, that is to identify the elementary changes that have occurred during evolution. Experimental approaches are available to decompose a given phenotypic difference into appropriate finer sub-variations. For example, crossing plants with different leaf shapes yields a progeny that exhibits a composite range of intermediate leaf shapes. Principal component analysis uncovered elementary leaf shape changes that can together account for the difference in shape between parental lines and that appear to be caused by distinct genomic regions (Langlade et al., 2005). This suggests to some extent that “the sum obscures the parts.” What we traditionally consider as complex traits can be made of simpler traits, more amenable to genetic analysis. Another illuminating example is the abdominal pigmentation in the *Drosophila dunni* group. Taken as a single variable, the levels of pigmentation show a complex genetic architecture, but decomposing adult patterns into anatomical sub-units unravels discrete genetic control for each sub-trait (Hollocher et al., 2000). A better known case is the evolution of body color in beach mice. The difference in color between light-colored beach mice and dark mice can be decomposed into distinct phenotypes (dorsal hue, dorsal brightness, width of tail stripe, and dorsoventral boundary), which are all associated with distinct mutations in the *Agouti* gene (Linnen et al., 2013; Figure 3). Each *Agouti* genetic locus appears to be dedicated to the specification of pigmentation in a given body part. Together, they form a group of tightly linked loci that are associated with changes in coat pigmentation.



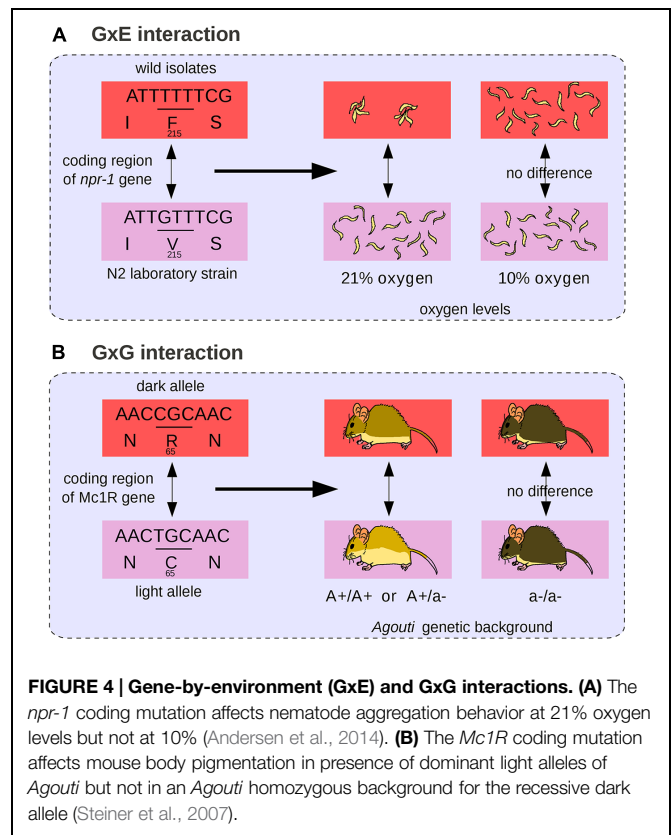
While complex traits may not always be reducible to a suite of simple GP relationships, it is possible that traits such as adult human height, the most emblematic quantitative trait predicted to consist of many genetic effects of small size (Fisher, 1930), might also be decomposed into elementary variations, each explaining more discrete sub-traits. While some determinants of human height such as *LIN28B* have been associated to adult height at different ages, other genes have only reached statistical significance in stage-specific studies focusing on fetal growth and height velocity at puberty (Lettre, 2011). In other words, these data suggest that human height may be a composite trait that is modulated by several GP relationships, each acting at different phases of developmental growth.

The Problem of Epistasis and GxE

Gene-by-Environment (GxE) interaction occurs when the phenotypic effect of a given genetic change depends on environmental parameters. Similarly, epistasis, or GxG interaction, occurs when the phenotypic effect of a given genetic change depends on the allelic state of at least one other locus (Phillips, 2008; Hansen, 2013). There is increasing evidence that GxG and GxE interactions are of fundamental importance to understand evolution and inheritance of complex traits (Gilbert and Epel, 2009; Hansen, 2013). We propose that both phenomena can be integrated into the basic GP differential framework, where both GxG and GxE interactions inject a layer of context-dependence, and result in differences embedded within differences.

The difference in color pigmentation between dark and light-colored beach mice mentioned previously (Figure 3) is not only due to mutations in *Agouti* but also to a coding mutation in the *MC1R* gene that decreases pigmentation (Steiner et al., 2007; Figure 4B). The effect of the *MC1R* mutation is visible only in presence of the light-colored-associated derived *Agouti* haplotype. Here the *Mc1R* locus is considered to interact epistatically with the *Agouti* locus. In this case, we propose that the GP relationship does not comprise a single phenotypic difference but two possible phenotypic differences (a change in coat pigmentation or no change at all). The choice between these two phenotypic differences is determined by the genetic background (here at the *Agouti* locus). The differential view thus remains relatively straightforward for two-loci interactions: the context-dependence of the phenotype is translated into a choice between two possible phenotypic differences. We propose that a GP relationship involving a mutation subjected to multiple epistatic interactions should comprise all possible phenotypic differences that can result from the mutation in all genetic backgrounds. Among all possible phenotypic variations, the phenotypic difference that will be observed is determined by other genetic loci. In general, GxG interactions involve multiple sites that are dispersed across the genome (Bloom et al., 2013).

An example of GxE interaction (see also Figure 4A) is the naturally occurring loss-of-function allele of *brx* in *Arabidopsis*



plants, which is associated with accelerated growth and increased fitness in acidic soils, and with severely reduced root growth compared to wild-type in normal soils (Gujas et al., 2012). GxE interactions are usually analyzed in the form of a *norm of reaction*, which represents all the observable traits of a *single* genotype across a range of environments (Johannsen, 1911; Sarkar, 1999). In the case of GxE interactions, we propose that the GP relationship should comprise all the possible phenotypic changes that can be caused by the associated genetic change across various experimental conditions. The associated phenotypic change is thus a difference between two norms of reaction. A textbook example is the variation in temperature-size rule in *C. elegans*. Like most other animals, *C. elegans* nematodes grow larger at low temperature, but a wild-type laboratory strain of *C. elegans* originating from Hawaii shows no variation in body size across various temperatures. An amino acid change in a calcium-binding protein is responsible for the decreased ability of the Hawaiian strain to grow larger at low temperature (Kammenga et al., 2007). Here the norm of reaction (representing nematode body size across a range of temperatures) differs between nematodes and the associated GP relationship encompasses the difference between these two slopes.

The range of phenotypic variations embodied within GP relationships subjected to GxG and GxE interactions can be quite overwhelming, especially in cases when several tissues are affected by the same mutation, and when the phenotypic variation of each tissue is influenced by other genomic loci and by environmental conditions. In fact, the phenotypic effects of a mutation always

rely on other pieces of DNA from the same genome, so that any GP relationship can be considered to experience epistasis. In other words, a genetic locus affecting a phenotype never acts independently of other DNA sequences. For instance, a given opsin allele will only lead to particular color vision properties if an eye is formed and if this eye receives light during its development, allowing effective vision neural circuits to form. For the differential view to be tractable, we advise not to consider all possible genetic backgrounds and environmental conditions, but to restrict possibilities to potential environments, and segregating alleles that are relevant to the population of interest (Sober, 1988).

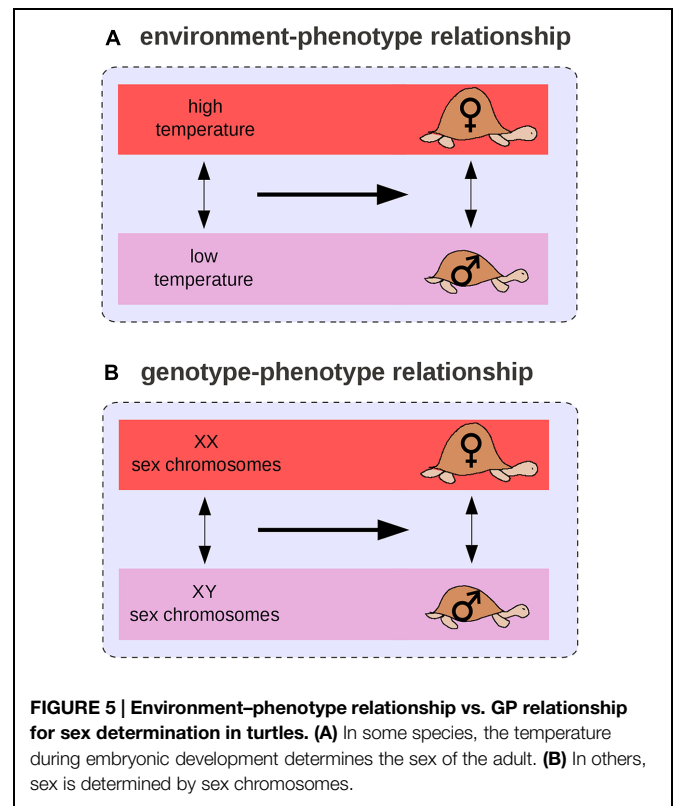
In summary, in presence of epistasis or GxE interactions, a genetic change is not associated with a single phenotypic difference but with multiple possible phenotypic differences, among which one will be achieved, depending on the environment and the genetic background. The context-dependence can be represented schematically as GP differences embedded into other genotype and environment differences.

The Differential View of Genetic and Environmental Effects on Phenotypes

As underlined by multiple authors (most notably Waddington, 1957; Oyama, 2000; Keller, 2010), genes and environment act jointly on the phenotype, and in most cases it is impossible to disentangle the effect of one from the other. Here we show that reasoning in terms of differences helps to clarify the comparison between genetic and environmental effects on phenotypes. However, we identify certain cases where the comparison remains difficult.

By analogy with the GP relationship, we can define the environment-phenotype relationship as an environmental variation (two environments), its associated phenotypic change (distinct phenotypic states), and their relationships. For example, in many turtle species, a change in temperature during egg development is associated with the male/female sex difference (Figure 5A) and at least six transitions from environmental to genetic sex determination (Figure 5B) occurred across the turtle phylogeny (Pokorná and Kratochvíl, 2009). In this case, environmental and genetic effects can be compared: sex chromosomes and temperature have the same phenotypic effect on turtles. Such observations led West-Eberhard (2003, 2005) to propose the “genes as followers” hypothesis, which suggests that novel phenotypic states are more likely to arise first from a change in the environment than from a genetic mutation, and that mutations occur only later, in modifying the threshold for expression of the novel trait. West-Eberhard (2003, 2005) extrapolated from differences segregating within populations (difference #2) to differences that arose temporally during the evolution of a population (difference #1).

The independent evolution of directional left–right asymmetry from symmetrical ancestors in multiple lineages has provided a major argument supporting the “gene as follower” hypothesis (Palmer, 2004). Under this framework, it is stipulated that directional asymmetry, where all individuals are same-sided, has often evolved from a “random asymmetry” state, where the directionality will depend on environmental factors and thus vary between genetically identical individuals. For instance,



the strongest claw of a lobster will develop based on usage and has a priori equal probabilities to develop on the left or on the right side. We can see how the “genes as followers” formula applies here: the environment triggers an asymmetry, and later in evolution some genetic effects can bias its directionality on one side or the other. But while asymmetry “occurs before genetic variation exists to control it,” the differential view makes it clear that the genetic effect on directionality is not comparable to the environmental effect that triggers the asymmetry. The genetic change makes a switch between the final 100% same-sided condition and an initial condition where 50% of the cases are dextral and 50% sinistral. In contrast, the two alternative phenotypic states resulting from variation in the environment are considered to be 100% dextral and 100% sinistral. This example shows that for the sake of accuracy it is important to explicitly state the differences that are being considered within a GP relationship.

The differential view provides a theoretical framework that can help in designing experiments to investigate the proper variables: one can compare different genotypes in a fixed environment (classic GP relationship), compare the response of a fixed genotype to two different environments (phenotypic plasticity), or compare the sensitivity of two different genotypes to two different environments (wherein the phenotypic variation becomes a *difference in a difference*; see for example Engelman et al., 2009; Thomas, 2010).

Various quantitative methods have been developed to disentangle genetic from environmental effects and to quantify GxE interactions (Lynch and Walsh, 1998). Yet in certain

situations it can be impossible to separate genetic from environmental effects in a biologically meaningful way, even when reasoning in terms of differences (Lewontin, 1974b). Populations of the beetle *Calathus melanocephalus* comprise two morphs, long-winged, and short-winged (Schwander and Leimar, 2011). The long-winged morph only develop from homozygous individuals for a recessive allele segregating in the population, and only when food conditions are good. In this case, the genetic and environmental effects are intermingled (Figures 6A,B). In the theoretical case of a population comprising only short-winged heterozygous animal that have been raised in starving conditions and long-winged ones, both genes and environment are responsible for the wing difference between individuals and it is impossible to estimate the proportion of environmental and genetic effects because genes and environment act on distinct levels along the complex causal link between genotypes and phenotypes.

Another case that questions the classical environment/genetic distinction is when the addition of certain symbiotic bacteria modifies the host phenotype. Mice fed with a *Lactobacillus* strain of bacteria show reduced anxiety-related behaviors compared to control mice fed with broth without bacteria (Bravo et al., 2011). Here the behavioral difference is caused by a switch between presence or absence of a particular gut symbiont. The cause of the phenotypic difference is not a simple change in a DNA sequence,

nor a simple environmental change disconnected from genetic changes, but a switch between presence and absence of a factor that can be considered as an environmental factor – the bacteria – which contains DNA whose mutations may also change the host phenotype.

In conclusion, reasoning in terms of differences can help to clarify the comparison between genetic and environmental effects on phenotypes. However, the issues are nothing but simple. Since genes and environment act on distinct levels along the complex causal link between genotypes and phenotypes, in certain cases it is impossible to disentangle both causes.

A Clarification on the Terminology Gain/Loss and Permissive/Instructive

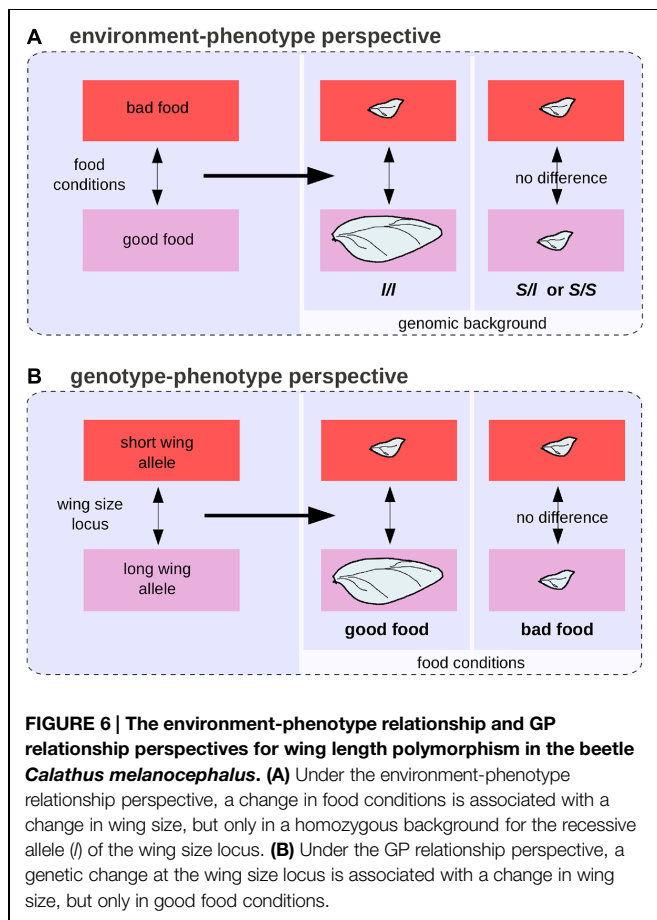
Phenotypic differences appear to fall under two major categories, either the presence/absence of something (for example body hair or the ability to digest milk), or the shift between two alternatives that are both present (for example two hair colors). Similarly, on the genotype side, a mutation can correspond to the presence/absence of a relevant DNA sequence, or to a nucleotide polymorphism. The differential perspective makes it evident that a loss of phenotype is not necessarily associated with a loss of genetic material, and vice versa. For example, the evolutionary gain of dark pigments covering the entire coat of animals has often been associated with a loss of the *Mc1R* gene (Gompel and Prud'homme, 2009). Furthermore, as one of us noted previously (appendix of Stern and Orgogozo, 2008), gain or loss for a phenotype is subjective. For example, loss of hair might also be considered as gain of naked epidermis. Most insect epidermal cells differentiate into one of these two alternative states and both states involve large gene regulatory networks. It is not clear which phenotypic state represents a gain or loss relative to the other. Even on the genotypic side, defining losses and gains can be difficult. The insertion of a transposable element can knock down a gene, whereas a deletion can sometimes creates a new binding site for an activator of transcription. As a matter of fact, the evolutionary gain of *desatF* expression in *D. melanogaster* occurred through a series of three deletions, each creating an hexamer motif that is required for *desatF* expression (Shirangi et al., 2009).

Similarly, the differential perspective on environmental effects highlights the fallacy of the distinction between permissive and instructive signals. A permissive signal is associated with the presence/absence of a phenotype and an instructive signal with the shift between two alternatives that are both present. As argued above, these distinctions at the phenotypic level are not clear-cut.

In conclusion, we suggest that the gain/loss and instructive/permissive terminology should be used with caution.

Taxonomically Robust GP Relationships

A mutation is expected to produce a somewhat reproducible phenotypic variation within a population. Such reproducibility



in phenotypic outcome is required to allow genetic evolution and adaptation by natural selection (Lewontin, 1974a; Kirschner and Gerhart, 1998). Indeed, a newly formed allele that would generate yet another phenotype each time it ends up in a different organism would not be subjected to natural selection. Reasoning in terms of variation, rather than considering alleles as isolated entities, makes it clear that competition occurs between alleles that span the same genetic locus. Natural selection acts directly on the allelic variation that is consistently associated with a given phenotypic variation, which is the GP relationship itself. The GP relationship is thus a basic unit of evolutionary change, on which natural selection acts.

A major discovery of the past 20 years is that variation at certain genetic loci produce comparable phenotypic variation not only in various individuals of one population, but also in extremely diverse taxa (Martin and Orgogozo, 2013b). In other words, certain GP relationships are taxonomically robust and present across a large range of species. This implies that the genetic and environmental backgrounds have remained relatively constant or have appeared repeatedly throughout evolution to allow for genetic loci to generate similar phenotypic changes in various taxonomic groups. This important finding was quite unsuspected some 50 years ago. For a long time the singularity observed in the living world was expected to reflect a comparable singularity at the genetic level, implicating disparate and non-conserved genes, specific to each lineage (Mayr, 1963). As Mayr (1963) once proposed in 1963, “Much that has been learned about gene physiology makes it evident that the search for homologous genes is quite futile except in very close relatives [. . .]. The saying “Many roads lead to Rome” is as true in evolution as in daily affairs” (Mayr, 1963). In other words, the genetic loci that make a man a man were expected to be different from the ones that make a dog, or a fish. Later, in the 80–90s, a few researchers suggested quite the contrary that evolution proceeds through mutations in conserved protein-coding genes (Romero-Herrera et al., 1978; Perutz, 1983; Stewart et al., 1987; Carroll et al., 2005) – but they had little experimental data at hand to support their view (Tautz and Schmid, 1998). As of today, the accumulating data on the mutations responsible for natural variation make it clear that the diversity in living organisms share a common genetic basis on at least three points. First, comparative developmental biology revealed that animals share common sets of key regulatory genes with conserved functions (Wilkins, 2002, 2014; Carroll et al., 2005). Second, most interspecific differences in animals and plants for which the underlying genetic basis has been at least partly identified (154 cases out of 160) are due to mutations at homologous genes, and very few (6/160) are due to new genes, which nevertheless represent duplicates of existing genes (Martin and Orgogozo, 2013b). Third, multiple cases of similar phenotypic changes have been shown to involve mutations of the same homologous genes in independent lineages (Table 1), sometimes across large phylogenetic distances. For instance, the difference in pigmentation between white and orange Bengal tigers has recently been mapped to a single mutation in the transporter protein gene *SLC45A2* (Xu et al., 2013), and this gene has also been associated with hypopigmented eyes, skin, hair, and feathers in humans and chickens (Xu et al., 2013;

Figure 1E). A more dramatic example is the recent evolution of a toxin resistance in three species that diverged more than 500 million years ago – a clam, a snake and a pufferfish – via the same amino acid substitution in a conserved gene (Bricelj et al., 2005; Geffeny et al., 2005; Venkatesh et al., 2005; Feldman et al., 2012). Such striking patterns of genetic repetition have now been found for more than 100 genes in animals and plants (Martin and Orgogozo, 2013b). Despite existing methodological biases favoring conserved genes in the search for quantitative trait loci (Rockman, 2012; Martin and Orgogozo, 2013b), the level of genetic repetition remains astounding and suggests that for the evolution of at least certain phenotypic differences, relatively few genetic roads lead to Rome (Stern, 2013). Nowadays, one should not be surprised that a piece of DNA associated with a complex wing color pattern in one *Heliconius* butterfly species provides similar wings and collective protection from the same predators when introduced into the genome of other butterflies (Supple et al., 2014). What makes a dog a dog or a man a man is now partly explained by singular assortments of taxonomically robust GP relationships, which are found in multiple lineage branches.

Certain environment–phenotype relationships are also taxonomically robust. For example, across most taxa body size is affected by nutrition; iron deficiency can cause anemia and certain toxic compounds can be lethal. In ectotherms the temperature of the organism depends on the environmental temperature. Given the daunting number of environmental conditions that can be conceived, it is probably impossible to determine whether taxonomically robust GP relationships or taxonomically robust environment–phenotype relationships are more prevalent. Furthermore, whether taxonomically robust GP relationships represent an exceptional and small fraction, or a significant proportion, of all GP relationships is a matter of debate. In any case, the existence of taxonomically robust GP relationships is now clear and should be broadly accepted by the research community.

Some of the most striking teachings of modern biology include the discovery that living beings share the same genetic material (DNA or RNA), the same genetic code (with few exceptions), and the same basic cellular machinery. It is thus far from paradoxical that individual differences are built upon similarities, and the finding that certain GP relationships persists over long evolutionary times completes the picture.

The precise predictive power resulting from the existence of taxonomically robust GP relationships is rare in biology, and is only starting to be exploited at its full potential. Long-range conservations of GP links now fully justify the use of comparative genetics approaches to tackle pragmatic problems. For instance, crop domestication took the form of similar selective pressures in many species, and we now have experimental evidence that this process has repeatedly involved mutations in the same set of conserved genes (Paterson et al., 1995; Martin and Orgogozo, 2013b). This observation opens up interesting applications, as we can use this emerging body of genetic expertise to assist the domestication of future crops, or to use marker-assisted strategies to produce and

maintain crop biodiversity (Lenser and Theißen, 2013). GP predictability is already used in the identification of strains that evolved resistance to different pest control strategies, with extreme cases targeting anti-malarial drugs tolerance in *Plasmodium* parasites (Manske et al., 2012), antibiotic resistance in bacteria and yeasts (Fischbach, 2009; MacCallum et al., 2010), or even more dramatically, the anthropogenic evolution of insecticide-resistance in diverse cohorts of insects, regardless of their pest status (Ffrench-Constant et al., 2004; Martin and Orgogozo, 2013b).

Furthermore, repeatability in the genetic basis of phenotypic variation suggests that clinical research is also likely to benefit from genetic studies of a large range of model species (Robinson and Webber, 2014). For instance, natural variation in the tolerance to methotrexate, a chemotherapy drug, was mapped in *Drosophila* fruitflies to genes whose human orthologs are also associated with the response of patients to this drug (Kislukhin et al., 2013), thus extending the use of model organisms as disease models.

Toward a Gene-Based Classification of Phenotypes

One original aspect of framing the GP connection in terms of individual GP relationships is that it allows to classify phenotypes according to their underlying genetic basis. On a first level, GP relationships implicating different regions within the same gene and producing comparable phenotypic outcomes can be grouped together. Simple cases of GxG interactions have been found between tightly linked mutations, generally within a coding sequence or within a *cis*-regulatory element, when they generate a non-additive effect on the phenotype. For example, a particular mutation in an enhancer was observed to produce different shifts in expression pattern of the downstream coding gene, depending on neighboring DNA sequence (Frankel et al., 2011; Rogers et al., 2013). Similarly, amino acid mutations in a hemoglobin gene was found to increase or decrease affinity to oxygen, depending on the allelic state of other sites (Natarajan et al., 2013). In such cases, it is intuitive to group such genetically linked sites together as they all affect the same kind of phenotypic trait.

Absence of melanin pigments in animals has been associated with mutations in several genes, including *OCA2*, *kit ligand* or *Mc1R* (reviewed in Gompel and Prud'homme, 2009; Liu et al., 2013). Whereas the absence of melanin is traditionally considered as one character state, albinism, irrespective of the underlying genetic basis, we propose here to distinguish *OCA2*-associated albinism from *Mc1R*-associated albinism, or from albinism associated with any other gene. One interest of decomposing the variation within the living world into these multiple elementary GP relationships is that these elements can then be grouped together into successively larger groups. Elementary phenotypic changes involving different genes that are part of the same genetic pathway could also be grouped together as concomitant components of the same phenotype-modulating mechanism. This is clearly the case for the TGF- β signaling molecules BMP15, GDF9, and the TGF- β receptor BMPRI1B,

that have all been repeatedly associated to variations in ovarian function in humans and in domestic sheep breeds (reviewed in Luong et al., 2011).

Another important consequence of the GP relationship perspective is that apparently distinct phenotypic changes caused by similar genetic loci in various organisms can be examined further to uncover what might be a common basic phenotypic change (Deans et al., 2015). For example, fly larvae and nematode worms have distinctive food search behaviors but mutations in the same orthologous gene (*for/egl-4*) have been shown to alter the intensity of food search behaviors in both organisms (Osborne et al., 1997; Mery et al., 2007; Hong et al., 2008). It is thus plausible that a basic behavioral change, which underlies seemingly distinctive fly and nematode food search changes, represents a conserved GP relationship across nematodes and flies. This somewhat borderline example illustrates the challenge to incorporate widespread comparative thinking into our global understanding of biology. Is a mutation in a mouse model relevant to human disease? Can we consider that a mouse phenotype is similar to a human condition if its genetic basis is different? We and others predict that the search for orthologous phenotypes, or “phenologs” (McGary et al., 2010), will represent a major task for modern genetics and will require a fruitful alliance between applied and evolutionary biology.

Conclusion

In this paper, we bring back the differential concept of gene (Schwartz, 2000) into our framework for understanding the GP map. The differential view of the GP relationship helps to clarify the genetic and environmental effects on phenotypes and their connection. It also opens up new avenues of thinking, in particular regarding the decomposition of observable features within an organism and the representation of GP maps. Furthermore, the existence of taxonomically robust GP relationships encourages an unabashed use of comparative genetics to predict the genetic basis of phenotypic variation in diverse groups of organisms, and this predictive power has an important potential for translational research in agronomy and clinical research.

Acknowledgments

We deeply acknowledge Marie-Anne Félix for enlightening discussions and Thomas Pradeu for bringing several relevant papers to our attention. We also thank Giuseppe Baldacci, Marie-Anne Félix, Pierre-Henri Gouyon, Alexandre Peluffo, Mark Siegal, David Stern, and the reviewers for their insightful comments on the manuscript. The research leading to this paper has received funding from the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant Agreement no. 337579) and from the John Templeton Foundation (Grant no. 43903).

References

- Andersen, E. C., Bloom, J. S., Gerke, J. P., and Kruglyak, L. (2014). A variant in the neuropeptide receptor npr-1 is a major determinant of *Caenorhabditis elegans* growth and physiology. *PLoS Genet.* 10:e1004156. doi: 10.1371/journal.pgen.1004156
- Armstrong, M. D., and Tyler, F. H. (1955). Studies on phenylketonuria. I. Restricted phenylalanine intake in phenylketonuria. *J. Clin. Invest.* 34, 565–580. doi: 10.1172/JCI103105
- Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V., and Kruglyak, L. (2013). Finding the sources of missing heritability in a yeast cross. *Nature* 494, 234–237. doi: 10.1038/nature11867
- Boycott, A. E., Diver, C., Garstang, S. L., and Turner, F. M. (1931). The inheritance of sinistrality in *Limnaea peregra* (Mollusca, Pulmonata). *Philos. Trans. R. Soc. Lond. Ser. B* 219, 51–131. doi: 10.1098/rstb.1931.0002
- Bravo, J. A., Forsythe, P., Chew, M. V., Escaravage, E., Savignac, H. M., Dinan, T. G., et al. (2011). Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16050–16055. doi: 10.1073/pnas.1102999108
- Bricelj, V. M., Connell, L., Konoki, K., Macquarrie, S. P., Scheuer, T., Catterall, W. A., et al. (2005). Sodium channel mutation leading to saxitoxin resistance in clams increases risk of PSP. *Nature* 434, 763–767. doi: 10.1038/nature03415
- Carroll, S. B., Grenier, J., and Weatherbee, S. (2005). *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*. Malden, MA: John Wiley & Sons.
- Cassirer, E. (1910). *Substance et Fonction*, 2010 Edn. Paris: Nabu Press.
- Coen, E. (2012). *Cells to Civilizations: The Principles of Change That Shape Life*. Princeton, NJ: Princeton University Press. doi: 10.1515/9781400841653
- Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G. Jr., Dickson, M., Grimwood, J., et al. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* 307, 1928–1933. doi: 10.1126/science.1107239
- Dawkins, R. (1982). *The Extended Phenotype: The Long Reach of the Gene*. Oxford: Oxford University Press.
- Deans, A. R., Lewis, S. E., Huala, E., Anzaldo, S. S., Ashburner, M., Balhoff, J. P., et al. (2015). Finding our way through phenotypes. *PLoS Biol.* 13:e1002033. doi: 10.1371/journal.pbio.1002033
- Dunbar, H. E., Wilson, A. C. C., Ferguson, N. R., and Moran, N. A. (2007). Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS Biol.* 5:e96. doi: 10.1371/journal.pbio.0050096
- Engelman, C. D., Baurley, J. W., Chiu, Y.-F., Joubert, B. R., Lewinger, J. P., Maenner, M. J., et al. (2009). Detecting gene-environment interactions in genome-wide association data. *Genet. Epidemiol.* 33(Suppl. 1), S68–S73. doi: 10.1002/gepi.20475
- Falk, R. (1984). The gene in search of an identity. *Hum. Genet.* 68, 195–204. doi: 10.1007/BF00418388
- Feldman, C. R., Brodie, E. D. Jr., Brodie, E. D. III, and Pfrender, M. E. (2012). Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4556–4561. doi: 10.1073/pnas.1113468109
- Ffrench-Constant, R. H., Daborn, P. J., and Le Goff, G. (2004). The genetics and genomics of insecticide resistance. *Trends Genet.* 20, 163–170. doi: 10.1016/j.tig.2004.01.003
- Fischbach, M. A. (2009). Antibiotics from microbes: converging to kill. *Curr. Opin. Microbiol.* 12, 520–527. doi: 10.1016/j.mib.2009.07.002
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon. doi: 10.5962/bhl.title.27468
- Fontana, W. (2002). Modelling “evo-devo” with RNA. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 24, 1164–1177. doi: 10.1002/bies.10190
- Frankel, N., Erezylmaz, D. F., McGregor, A. P., Wang, S., Payre, F., and Stern, D. L. (2011). Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* 474, 598–603. doi: 10.1038/nature10200
- Geffeney, S. L., Fujimoto, E., Brodie, E. D. III, Brodie, E. D. Jr., and Ruben, P. C. (2005). Evolutionary diversification of TTX-resistant sodium channels in a predator-prey interaction. *Nature* 434, 759–763. doi: 10.1038/nature03444
- Gerstein, M. B., Bruce, C., Rozowsky, J. S., Zheng, D., Du, J., Korbel, J. O., et al. (2007). What is a gene, post-ENCODE? History and updated definition. *Genome Res.* 17, 669–681. doi: 10.1101/gr.6339607
- Gilbert, S. F. (2000). “Genes classical and genes developmental: the different uses of genes in evolutionary syntheses,” in *The Concept of the Gene in Development and Evolution*, eds P. Beurton, R. Falk, and H.-J. Rheinberger (Cambridge, NY: Cambridge University Press).
- Gilbert, S. F., and Epel, D. (2009). *Ecological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution*. Sunderland, MA: Sinauer Associates.
- Gjuvsland, A. B., Vik, J. O., Beard, D. A., Hunter, P. J., and Omholt, S. W. (2013). Bridging the genotype-phenotype gap: what does it take? *J. Physiol.* 591, 2055–2066. doi: 10.1113/jphysiol.2012.248864
- Gompel, N., and Prud'homme, B. (2009). The causes of repeated genetic evolution. *Dev. Biol.* 332, 36–47. doi: 10.1016/j.ydbio.2009.04.040
- Graur, D., Zheng, Y., and Azevedo, R. B. R. (2015). An Evolutionary Classification of Genomic Function. *Genome Biol. Evol.* 7, 642–645. doi: 10.1093/gbe/evv021
- Griffiths, P., and Stotz, K. (2013). *Genetics and Philosophy: An Introduction*. Cambridge: Cambridge University Press. doi: 10.1017/CBO978051744082
- Gujas, B., Alonso-Blanco, C., and Hardtke, C. S. (2012). Natural *Arabidopsis* brx loss-of-function alleles confer root adaptation to acidic soil. *Curr. Biol.* 22, 1962–1968. doi: 10.1016/j.cub.2012.08.026
- Hansen, T. F. (2013). Why epistasis is important for selection and adaptation. *Evolution* 67, 3501–3511. doi: 10.1111/evo.12214
- Hollocher, H., Hatcher, J. L., and Dyreson, E. G. (2000). Genetic and developmental analysis of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. *Evol. Int. J. Org. Evol.* 54, 2057–2071. doi: 10.1111/j.0014-3820.2000.tb01249.x
- Hong, R. L., Witte, H., and Sommer, R. J. (2008). Natural variation in *Pristionchus pacificus* insect pheromone attraction involves the protein kinase EGL-4. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7779–7784. doi: 10.1073/pnas.0708406105
- Houle, D., Govindaraju, D. R., and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* 11, 855–866. doi: 10.1038/nrg2897
- Johannsen, W. (1911). The genotype conception of heredity. *Am. Nat.* 45, 129–159. doi: 10.1086/279202
- Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., et al. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484, 55–61. doi: 10.1038/nature10944
- Kamberov, Y. G., Wang, S., Tan, J., Gerbault, P., Wark, A., Tan, L., et al. (2013). Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell* 152, 691–702. doi: 10.1016/j.cell.2013.01.016
- Kammenga, J. E., Doroszuk, A., Riksen, J. A. G., Hazendonk, E., Spiridon, L., Petrescu, A.-J., et al. (2007). A *Caenorhabditis elegans* wild type defies the temperature–size rule owing to a single nucleotide polymorphism in tra-3. *PLoS Genet* 3:e34. doi: 10.1371/journal.pgen.0030034
- Keller, E. F. (2010). *The Mirage of a Space between Nature and Nurture*. Durham, NC: Duke University Press. doi: 10.1215/9780822392811
- Kirschner, M., and Gerhart, J. (1998). Evolvability. *Proc. Natl. Acad. Sci. U.S.A.* 95, 8420–8427. doi: 10.1073/pnas.95.15.8420
- Kislukhin, G., King, E. G., Walters, K. N., Macdonald, S. J., and Long, A. D. (2013). The genetic architecture of methotrexate toxicity is similar in *Drosophila melanogaster* and humans. *PLoS One* 8, 1301–1310. doi: 10.1371/journal.pone.0066619
- Landry, C. R., and Rifkin, S. A. (2012). The genotype-phenotype maps of systems biology and quantitative genetics: distinct and complementary. *Adv. Exp. Med. Biol.* 751, 371–398. doi: 10.1007/978-1-4614-3567-9_17
- Langlade, N. B., Feng, X., Dransfield, T., Copsey, L., Hanna, A. I., Thébaud, C., et al. (2005). Evolution through genetically controlled allometry space. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10221–10226. doi: 10.1073/pnas.0504210102
- Lenser, T., and Theissen, G. (2013). Molecular mechanisms involved in convergent crop domestication. *Trends Plant Sci.* 18, 704–714. doi: 10.1016/j.tplants.2013.08.007
- Lette, G. (2011). Recent progress in the study of the genetics of height. *Hum. Genet.* 129, 465–472. doi: 10.1007/s00439-011-0969-x
- Lewontin, R. (1974a). *The Genetic Basis of Evolutionary Change (Columbia Biological Series)*. Columbia: Columbia University Press.
- Lewontin, R. C. (1974b). The analysis of variance and the analysis of causes. *Am. J. Hum. Genet.* 26, 400–411.

- Lewontin, R. (1978). Adaptation. *Sci. Am.* 293, 212–228. doi: 10.1038/scientificamerican0978-212
- Linnen, C. R., Poh, Y.-P., Peterson, B. K., Barrett, R. D. H., Larson, J. G., Jensen, J. D., et al. (2013). Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* 339, 1312–1316. doi: 10.1126/science.1233213
- Liu, F., Wen, B., and Kayser, M. (2013). Colorful DNA polymorphisms in humans. *Semin. Cell Dev. Biol.* 24, 562–575. doi: 10.1016/j.semcdb.2013.03.013
- Luong, H. T. T., Chaplin, J., McRae, A. F., Medland, S. E., Willemsen, G., Nyholt, D. R., et al. (2011). Variation in BMPR1B, TGFRB1 and BMPR2 and control of dizygotic twinning. *Twin Res. Hum. Genet. Off. J. Int. Soc. Twin Stud.* 14, 408–416. doi: 10.1375/twin.14.5.408
- Lynch, M., and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer Associates.
- MacCallum, D. M., Coste, A., Ischer, F., Jacobsen, M. D., Odds, F. C., and Sanglard, D. (2010). Genetic dissection of azole resistance mechanisms in *Candida albicans* and their validation in a mouse model of disseminated infection. *Antimicrob. Agents Chemother.* 54, 1476–1483. doi: 10.1128/AAC.01645-1649
- Manske, M., Miotto, O., Campino, S., Auburn, S., Almagro-Garcia, J., Maslen, G., et al. (2012). Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature* 487, 375–379. doi: 10.1038/nature11174
- Martin, A., and Orgogozo, V. (2013a). Data from: *The Loci of Repeated Evolution: A Catalogue of Genetic Hotspots of Phenotypic Variation*. Available at: <http://datadryad.org/resource/> doi: 10.5061/dryad.v66p0
- Martin, A., and Orgogozo, V. (2013b). The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evol. Int. J. Orgn. Evol.* 67, 1235–1250. doi: 10.1111/evo.12081
- Mayr, E. (1963). *Animal Species and Evolution*. Harvard: Harvard University Press. doi: 10.4159/harvard.9780674865327
- McGary, K. L., Park, T. J., Woods, J. O., Cha, H. J., Wallingford, J. B., and Marcotte, E. M. (2010). Systematic discovery of nonobvious human disease models through orthologous phenotypes. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6544–6549. doi: 10.1073/pnas.0910200107
- Mery, F., Belay, A. T., So, A. K.-C., Sokolowski, M. B., and Kawecki, T. J. (2007). Natural polymorphism affecting learning and memory in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13051–13055. doi: 10.1073/pnas.0702923104
- Moss, L. (2003). One, Two (Too?), Many Genes? *Q. Rev. Biol.* 78, 57–67. doi: 10.1086/367581
- Narod, S. A., and Foulkes, W. D. (2004). BRCA1 and BRCA2: 1994 and beyond. *Nat. Rev. Cancer* 4, 665–676. doi: 10.1038/nrc1431
- Natarajan, C., Inoguchi, N., Weber, R. E., Fago, A., Moriyama, H., and Storz, J. F. (2013). Epistasis among adaptive mutations in deer mouse hemoglobin. *Science* 340, 1324–1327. doi: 10.1126/science.1236862
- Noble, D. (2006). *The Music of Life: Biology beyond the Genome*. Oxford: Oxford University Press.
- Osborne, K. A., Robichon, A., Burgess, E., Butland, S., Shaw, R. A., Coulthard, A., et al. (1997). Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science* 277, 834–836. doi: 10.1126/science.277.5327.834
- Oyama, S. (2000). *The Ontogeny of Information: Developmental Systems and Evolution*, 2nd Edn. Cambridge, NY: Duke University Press. doi: 10.1215/9780822380665
- Paaby, A. B., and Rockman, M. V. (2013). The many faces of pleiotropy. *Trends Genet.* 29, 66–73. doi: 10.1016/j.tig.2012.10.010
- Palmer, A. R. (2004). Symmetry breaking and the evolution of development. *Science* 306, 828–833. doi: 10.1126/science.1103707
- Paterson, A. H., Lin, Y. R., Li, Z., Schertz, K. F., Doebley, J. F., Pinson, S. R., et al. (1995). Convergent domestication of cereal crops by independent mutations at corresponding genetic Loci. *Science* 269, 1714–1718. doi: 10.1126/science.269.5231.1714
- Perutz, M. F. (1983). Species adaptation in a protein molecule. *Mol. Biol. Evol.* 1, 1–28.
- Phillips, P. C. (2008). Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* 9, 855–867. doi: 10.1038/nrg2452
- Pokorná, M., and Kratochvíl, L. (2009). Phylogeny of sex-determining mechanisms in squamate reptiles: are sex chromosomes an evolutionary trap? *Zool. J. Linn. Soc.* 156, 168–183. doi: 10.1111/j.1096-3642.2008.00481.x
- Robinson, P. N., and Webber, C. (2014). Phenotype ontologies and cross-species analysis for translational research. *PLoS Genet* 10:e1004268. doi: 10.1371/journal.pgen.1004268
- Rockman, M. V. (2012). The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evol. Int. J. Orgn. Evol.* 66, 1–17. doi: 10.1111/j.1558-5646.2011.01486.x
- Rogers, W. A., Salomone, J. R., Tacy, D. J., Camino, E. M., Davis, K. A., Rebeiz, M., et al. (2013). Recurrent modification of a conserved cis-regulatory element underlies fruit fly pigmentation diversity. *PLoS Genet* 9:e1003740. doi: 10.1371/journal.pgen.1003740
- Romero-Herrera, A. E., Lehmann, H., Joysey, K. A., and Friday, A. E. (1978). On the evolution of myoglobin. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 283, 61–163. doi: 10.1098/rstb.1978.0018
- Salazar-Ciudad, I., and Marín-Riera, M. (2013). Adaptive dynamics under development-based genotype-phenotype maps. *Nature* 497, 361–364. doi: 10.1038/nature12142
- Sarkar, S. (1999). From the reaktionsnorm to the adaptive norm: the norm of reaction, 1909–1960. *Biol. Philos.* 14, 235–252. doi: 10.1023/A:1006690502648
- Schwander, T., and Leimar, O. (2011). Genes as leaders and followers in evolution. *Trends Ecol. Evol.* 26, 143–151. doi: 10.1016/j.tree.2010.12.010
- Schwartz, S. (2000). “The differential concept of the gene: past and present,” in *The Concept of the Gene in Development and Evolution*, eds P. J. Beurton, R. Falk, and H. J. Rheinberger (Cambridge, NY: Cambridge University Press), 26–39. doi: 10.1017/CBO9780511527296.004
- Shirangi, T. R., Dufour, H. D., Williams, T. M., and Carroll, S. B. (2009). Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. *PLoS Biol.* 7:e1000168. doi: 10.1371/journal.pbio.1000168
- Simondon, G. (1968). *L'individuation à la Lumière des Notions de Forme et d'Information*, 2005th Edn. Grenoble: Millon.
- Sober, E. (1988). Apportioning causal responsibility. *J. Philos.* 85, 303–318. doi: 10.2307/2026721
- Steiner, C. C., Weber, J. N., and Hoekstra, H. E. (2007). Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol.* 5:e219. doi: 10.1371/journal.pbio.0050219
- Stern, D. L. (2000). Evolutionary developmental biology and the problem of variation. *Evol. Int. J. Orgn. Evol.* 54, 1079–1091. doi: 10.1111/j.0014-3820.2000.tb00544.x
- Stern, D. L. (2013). The genetic causes of convergent evolution. *Nat. Rev. Genet.* 14, 751–764. doi: 10.1038/nrg3483
- Stern, D., and Orgogozo, V. (2008). The loci of evolution: How predictable is genetic evolution? *Evol. Int. J. Orgn. Evol.* 62, 2155–2177. doi: 10.1111/j.1558-5646.2008.00450.x
- Stewart, C. B., Schilling, J. W., and Wilson, A. C. (1987). Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature* 330, 401–404. doi: 10.1038/330401a0
- Stotz, K. (2012). Murder on the development express: who killed nature/nurture? *Biol. Philos.* 27, 919–929. doi: 10.1007/s10539-012-9343-1
- Sturtevant, A. H. (1932). The use of mosaics in the study of the developmental effects of genes. *Proc. Sixth Int. Congr. Genet. Ithaca N. Y.* 1, 304–307.
- Supple, M., Papa, R., Counterman, B., and McMillan, W. O. (2014). The genomics of an adaptive radiation: insights across the heliconius speciation continuum. *Adv. Exp. Med. Biol.* 781, 249–271. doi: 10.1007/978-94-007-7347-9_13
- Tautz, D., and Schmid, K. J. (1998). From genes to individuals: developmental genes and the generation of the phenotype. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 353, 231–240. doi: 10.1098/rstb.1998.0205
- Thomas, D. (2010). Gene–environment-wide association studies: emerging approaches. *Nat. Rev. Genet.* 11, 259–272. doi: 10.1038/nrg2764
- Venkatesh, B., Lu, S. Q., Dandona, N., See, S. L., Brenner, S., and Soong, T. W. (2005). Genetic basis of tetrodotoxin resistance in pufferfishes. *Curr. Biol.* 15, 2069–2072. doi: 10.1016/j.cub.2005.10.068
- Visscher, P. M., Brown, M. A., McCarthy, M. I., and Yang, J. (2012). Five years of GWAS discovery. *Am. J. Hum. Genet.* 90, 7–24. doi: 10.1016/j.ajhg.2011.11.029

- Waddington, C. H. (1957). *The Strategy of the Genes*, 2014 Edn Reprint. London: Routledge.
- Wagner, G. P. (1996). Homologues, natural kinds and the evolution of modularity. *Am. Zool.* 36, 36–43. doi: 10.1093/icb/36.1.36
- Wagner, G. P. (2000). *The Character Concept in Evolutionary Biology*. San Diego, CA: Academic Press.
- Wagner, G. P. (2014). *Homology, Genes, and Evolutionary Innovation*. Princeton, NJ: Princeton University Press. doi: 10.1515/9781400851461
- Wang, J., Wurm, Y., Nipitwattanaphon, M., Riba-Grognuz, O., Huang, Y.-C., Shoemaker, D., et al. (2013). A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* 493, 664–668. doi: 10.1038/nature11832
- Waters, C. K. (2007). Causes that make a difference. *J. Philos.* 104, 551–579.
- Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., et al. (2014). The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 42, D1001–D1006. doi: 10.1093/nar/gkt1229
- West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution*. Oxford: Oxford University Press.
- West-Eberhard, M. J. (2005). Developmental plasticity and the origin of species differences. *Proc. Natl. Acad. Sci. U.S.A.* 102(Suppl. 1), 6543–6549. doi: 10.1073/pnas.0501844102
- Wilkins, A. (2014). “The genetic tool-kit: the life-history of an important metaphor,” in *Advances in Evolutionary Developmental Biology*, ed. J. Todd Streebman (Hoboken, NJ: John Wiley & Sons), 1–14.
- Wilkins, A. S. (2002). *The Evolution of Developmental Pathways*. Sunderland, MA: Sinauer Associates.
- Xu, X., Dong, G.-X., Hu, X.-S., Miao, L., Zhang, X.-L., Zhang, D.-L., et al. (2013). The genetic basis of white tigers. *Curr. Biol.* 23, 1031–1035. doi: 10.1016/j.cub.2013.04.054
- Zhang, J., and Wagner, G. P. (2013). On the definition and measurement of pleiotropy. *Trends Genet.* 29, 383–384. doi: 10.1016/j.tig.2013.05.002

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.

Copyright © 2015 Orgogozo, Morizot and Martin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.