doi:10.1111/evo.12081

THE LOCI OF REPEATED EVOLUTION: A CATALOG OF GENETIC HOTSPOTS OF PHENOTYPIC VARIATION

Arnaud Martin[1,2] and Virginie Orgogozo[3]

[1]*Department of Ecology and Evolutionary Biology, Cornell University, Corson Hall, 215 Tower Road, Ithaca, New York 14853*

[2]*E-mail: heliconiuswing@gmail.com*

[3]*CNRS UMR7592, Université Paris Diderot, Sorbonne Paris Cité, Institut Jacques Monod, 15 rue Hélène Brion, 75205, Paris cedex 13, France*

Received November 20, 2012
Accepted January 26, 2013

What is the nature of the genetic changes underlying phenotypic evolution? We have catalogued 1008 alleles described in the literature that cause phenotypic differences among animals, plants, and yeasts. Surprisingly, evolution of similar traits in distinct lineages often involves mutations in the same gene ("gene reuse"). This compilation yields three important qualitative implications about repeated evolution. First, the apparent evolution of similar traits by gene reuse can be traced back to two alternatives, either several independent causative mutations or a single original mutational event followed by sorting processes. Second, hotspots of evolution—defined as the repeated occurrence of de novo mutations at orthologous loci and causing similar phenotypic variation—are omnipresent in the literature with more than 100 examples covering various levels of analysis, including numerous gain-of-function events. Finally, several alleles of large effect have been shown to result from the aggregation of multiple small-effect mutations at the same hotspot locus, thus reconciling micromutationist theories of adaptation with the empirical observation of large-effect variants. Although data heterogeneity and experimental biases prevented us from extracting quantitative trends, our synthesis highlights the existence of genetic paths of least resistance leading to viable evolutionary change.

KEY WORDS: **Evolutionary genomics, genetic variation, morphological evolution, mutations, quantitative genetics.**

> On the one hand, we can point to a rich and formidable body of mathematical theory on phenotypic evolution, built largely on an infinitesimal foundation. On the other hand, we can point to a large and growing body of data on the genetic basis of adaptation. The problem, of course, is that the formidable theory says little or nothing about the formidable data.
>
> (H.A. Orr 2005)

Mutations are the nuts and bolts of evolution. Indeed, sequence modifications generate the heritable phenotypic variation upon which drift and selection can act, leading to the idea that mutations themselves are a primary force of evolution (Nei 2007). The "Quantitative Trait Gene" (QTG) program aims at identifying the genes and the mutations responsible for phenotypic variation between individuals, populations, and species—the so-called "Loci of Evolution" (Stern and Orgogozo 2008). Although this dataset is biased toward large effect variants and may thus be inappropriate to infer general statements about the mechanisms of evolution (Rockman 2011), we believe that shifts in concepts and questions might bypass these biases and lead to general and profound discoveries about the causes of phenotypic variation.

Important insights may come from work on repeated evolution, the independent evolution of similar features in two lineages. It is widely accepted that evolution produces similar phenotypic outcomes at many biological levels (McGhee 2011), suggesting that nature often "repeats itself" due to internal and external constraints on change (Brakefield 2006; Losos 2011; Wake et al.

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

2011). From the QTG program emerged the surprising finding that certain genes have repeatedly been major components of phenotypic variation of similar traits (Wood et al. 2005; Gompel and Prud'homme 2009; Kopp 2009), qualifying as *genetic hotspots of phenotypic variation* (Richardson and Brakefield 2003; McGregor et al. 2007; Papa et al. 2008; Stern and Orgogozo 2009) or as examples of *gene reuse* (Conte et al. 2012). However, these terms have been lacking a rigorous conceptualization.

To better comprehend the concept of hotspots, we updated and compiled from the literature an extensive catalog of evolutionarily relevant allelic variants. The data are publicly available (Martin and Orgogozo 2013), and we refer to its previous version (Stern and Orgogozo 2008) and to the Supplementary Notes of this article for information on its conception, structure, and caveats. Overall, our qualitative survey of the known Loci of Evolution uncovers several emerging patterns in the genetic basis of repeated evolution.

Part 1. Five Cases of Apparent Gene Reuse

The replicated identification of a gene as underlying repeated phenotypic evolution has been dubbed "gene reuse" (Conte et al. 2012). In this section, we decompose gene reuse into different mechanisms, based on whether repetition is on the *propagation* or on the *generation* of variation itself.

REVERSION TO THE ANCESTRAL STATE

An effect of repeated evolution can occur when a new phenotype evolved once and has subsequently been reverted to the ancestral state in a subset of the radiating branches in the phylogenetic tree (Fig. 1A). It is essential to stress that reversions produce a dual effect of repeated evolution, whether one considers the phenotypic level or the genotypic level. On the phenotypic side, the derived non-reverted traits may appear as independent repetitions although they are not (green boxes in Fig. 1A). On the genotypic side, in the case of gene reuse underlying a return to the ancestral phenotypic state, the "reverted" lineage harbors repeated mutations at the same locus (orange box in Fig. 1A, as in Fig. 1E). In the strict sense, reversions are different from trait losses, which do not always lead to ancestral phenotypic states (e.g., body hair loss in humans; Porter and Crandall 2003). Cases of phenotypic reversions whose underlying mutations have been narrowed down to a single gene are scarce. Production of diene courtship pheromones by both sexes in *Drosophila takahashi* evolved from a dimorphic diene-producing ancestor by a secondary reversion mutation in the *cis*-regulatory region of *desatF*, rather than from the direct inheritance of an ancestral monomorphic state (Shirangi et al. 2009). Reverse mutations have been observed in the photoreceptors of several fish lineages that have undergone multiple colonization events between bright- and dim-light environments, underlying adaptations to different water depths (Yokoyama et al. 2008; Nagai et al. 2011; Hofmann et al. 2012). Finally, gene "death and resurrection" (Bekpen et al. 2009) and "birth and death" (Malfavon-Borja et al. 2013) can also be seen as forms of reversion.

Reversions illustrate the need to treat the genotypic and phenotypic levels separately, as repetition is not in the lineages that first appear as "repeated." We explore below how this important distinction between levels enlightens fundamentally different modes of repetition.

SORTING OF ANCESTRAL POLYMORPHISMS

An effect of repeated phenotypic evolution can emerge from the spread in independent populations of alleles that exist as standing genetic variation in the ancestral population (Barrett and Schluter 2008; Fig. 1B). This is well illustrated by the repeated fixation of variants of the *EDA* gene that confer a reduction of armor plates in stickleback fishes (Colosimo et al. 2005). These alleles are associated with adaptations to recent and worldwide colonization events into freshwater environments, and also exist in the marine populations at cryptic levels under the form of recessive low-frequency variants. Armor plate reduction has thus occurred countless times in various lakes, but sequencing of the *EDA* locus in many freshwater populations pinpoints a single origin of the causative allele (Jones et al. 2012).

An equivalent phenomenon can be observed in absence of contemporary gene flow, when ancestral polymorphisms have been maintained during past speciation events (Avise and Robinson 2008; Fig. 1B). Known examples of incomplete lineage sorting of adaptive significance mostly include alleles under long-term balancing selection such as in plant / plant pathogen recognition systems (Stahl et al. 1999; Rose et al. 2007; Horger et al. 2012), major histocompatibility complexes in mammals (Edwards et al. 1997; Loisel et al. 2006), AB-blood type defining enzymes and viral response factors in primates (Newman et al. 2006; Ferrer-Admetlla et al. 2009; Segurel et al. 2012), photoreceptors sustaining color vision in New World monkeys (Hunt et al. 1998), and self-incompatibility genes in plants and fungi (Wu et al. 1998; Schierup et al. 2001; Charlesworth et al. 2006; Igic et al. 2006). The recent development of methods dedicated to the detection of ancestral polymorphisms should yield to better estimates of the frequency of this phenomenon (Scally et al. 2012; Segurel et al. 2012).

LATERAL TRANSFER

Another scenario of repeated phenotypic evolution by gene reuse concerns the introgression of causative alleles from a lineage to another in spite of barriers to gene flow (Fig. 1C). Such "lateral

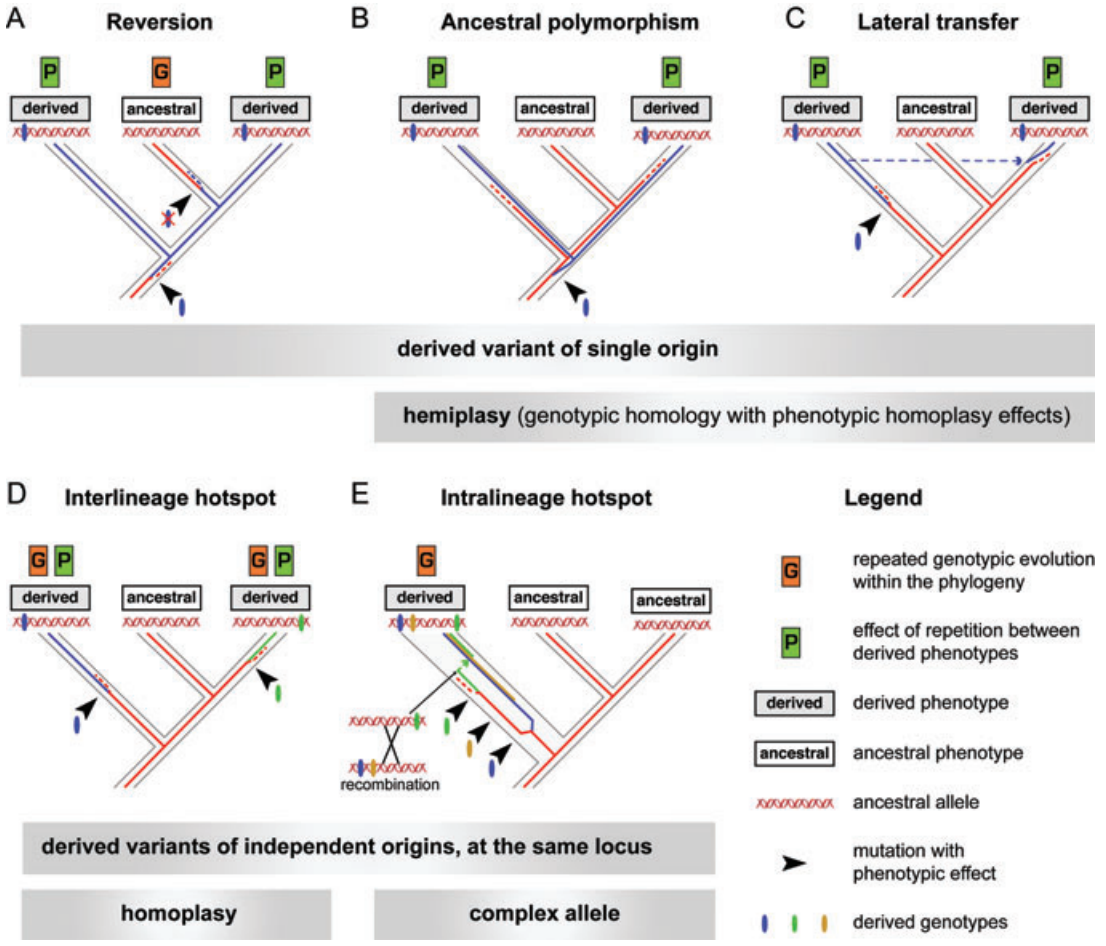

Figure 1. **Five cases of apparent repeated evolution by gene reuse. The discontinuous distribution of a character state on a phylogeny gives an effect of repeated evolution, including when this character state has appeared only once. Specifically, phenotypic reversions (A), propagations of ancestral polymorphisms (B, "adaptation from standing genetic variation" at the intraspecific level and "incomplete lineage sorting" at the interspecific level), and lateral transfers (C, introgression resulting from hybridization or vector-mediated horizontal transfer) generate contemporary effects of repetition of the derived phenotypes, even though these derived states originate from a single mutational event. In contrast, the concept of hotspot (D and E) involves at least two independent mutations, either at orthologous loci between lineages (interlineage hotspot, D) or at the same locus within a lineage (intralineage hotspot, E).**

transfers" (in the broad sense) can notably occur between interfertile species by hybridization, followed by generations of backcrosses in one species (Mallet 2005). For instance, in European populations of the house mouse *Mus musculus*, adaptive coding changes of the gene *vkorc1* that provide resistance to rodent poisons have been inherited by introgression of a derived, resistant allele from a partially sympatric species, the Algerian mouse *Mus spretus* (Song et al. 2011). In other words, poison resistance in two mouse species shares the same original mutational events: the resistant haplotype appeared once and then spread through introgression (Fig. 1C). The evolution of ray floret petals in *Senecio vulgaris* flowers (Kim et al. 2008), or the spread of melanism in coyotes and wolves due to admixture with domestic dogs (Anderson et al. 2009) are examples of introgression by recent contact between a wild species and its domesticated counterpart. Introgression also occurred by man-made hybridization between domesticated strains and thus facilitated the rapid transfer of traits of agricultural importance, as observed for genes that determine skin color in chickens (Eriksson et al. 2008), gustative quality in rice (Kovach et al. 2009), and resistance to mildew in barley (Piffanelli et al. 2004). Introgression of adaptive alleles may also shape patterns of convergence across multiple phylogenetic levels within complex radiations, and at least partially explains the phylogenetically scattered occurrence of red patterns in mimetic *Heliconius* butterflies (Hines et al. 2011; Pardo-Diaz et al. 2012; The Heliconius Genome Consortium 2012), as well as the repeated evolution of C4-type photosynthesis in the *Alloteropsis* grass lineage (Christin et al. 2012).

Horizontal gene transfers are a second mode of lateral transfer that does not involve hybridization between sibling species

(Syvanen 2012; Fig. 1C). This phenomenon can occur between extremely distant organisms (Ros and Hurst 2009), and is a primary force of genome evolution and adaptation in prokaryotes (Hehemann et al. 2010; Syvanen 2012). Between eukaryotes, the transfer of carotenoid synthesis enzymes from fungi to aphids (Moran and Jarvik 2010), and of an algal photosynthesis enzyme to a sea slug (Rumpho et al. 2008) offer spectacular examples of phenotypic evolution through gene transfer.

HOTSPOTS AND THE INDEPENDENT ORIGIN OF DERIVED VARIANTS

Ultimately, ancestral polymorphisms (adaptation from standing genetic variation; incomplete lineage sorting) and lateral transfers (introgression by secondary hybridization; gene transfer between distant species) rely on distinct allele histories (Fig. 1A–C), but result in the characteristic effects of hemiplasies where "the responsible lineage sorting processes have homoplasy-like consequences despite the fact that the character states themselves (i.e., the causative alleles) are genuinely homologous and apomorphic" (Avise and Robinson 2008). Most of these cases come from studies of the past 5 years, suggesting that hemiplasies may have a larger contribution to the pool of known evolutionarily relevant alleles than previously appreciated.

Loci that spread to distant lineages via lateral transfer or fixation of ancestral polymorphism (Fig. 1B,C) could certainly be considered as hotspots of evolution that inform us about how fit these variants are, but it is debatable whether they are genuine repetition (was a given trait "invented" twice?) or a genetic form of plagiarism. For instance, most biologists would agree that the molecular machinery that underlies carotenoid biosynthesis in aphids is not in itself a phenotypic novelty, as it was directly inherited from fungi (Moran and Jarvik 2010). For this reason we propose a definition of hotspots that restricts repetition to the independent appearance of causal mutations (Fig. 1D,E) rather than including the repeated propagation of existing variants (Fig. 1B,C). The exclusion of lateral transfers and ancestral polymorphisms also makes the hotspot definition more practical because the current literature usually does not allow to infer whether allele identity within a species reflects gene reuse or direct lineage descent.

We thus define the phenomenon of genetic hotspot of variation (thereafter abbreviated as hotspot), as *the repeated occurrence of de novo mutations at orthologous loci and directly causing similar phenotypic variations*. Repeated de novo mutations can occur at the same locus either in distinct lineages (interlineage hotspot, Fig. 1D) or within a single lineage (intralineage hotspot, Fig. 1E). Although the term "parallel evolution" is contentious due to contradictory usages and definitions (Arendt and Reznick 2008; Scotland 2011; Pearce 2012), interlineage hotspots are smoking guns of parallel evolution sensu Scotland (2011), where parallelisms are assimilated as homoplasies at the genotypic level. The criteria "orthologous loci" and "similar phenotypic variations" are discussed further below (end of Part 2). Next we qualitatively assess the body of experimental evidence for hotspots.

Part 2. Interlineage Hotspots are Found at Multiple Levels

An interesting property of the hotspot definition is that interlineage hotspots (Fig. 1D) can be detected as genes appearing several times in the Loci of Evolution catalog (Martin and Orgogozo 2013). In a second time, a variety of arguments such as phylogenetic distance, the identification of distinct causative mutations, and mapping studies with a multiple cross design allows to positively identify cases of hotspots based on de novo mutations (i.e., genes with several, distinct causative alleles). We detected 111 hotspot genes that have repeatedly caused the evolution of similar traits in independent lineages (Supporting Information). A total of 35.4% of the alleles (357 / 1008) were assigned to a gene based on linkage mapping for at least two independent evolutionary events. This number is probably inflated by ascertainment biases, as few rediscoveries are completely independent from a first discovery, even for linkage mapping studies. Nevertheless, this important level of replication in the dataset must reflect a propensity of certain genes to repeatedly underlie phenotypic variation. Genetic hotspots are found at multiple levels of analysis that we examine in this section.

Hotspots are Found at Various Genetic Resolutions

REPEATED EVOLUTION OF THE SAME AMINO ACIDS IN PROTEINS THAT INTERACT WITH ENVIRONMENTAL FACTORS

A striking feature that emerges from the evolutionary genetics literature is the existence of numerous cases of extreme hotspots, where identical amino acid changes have independently evolved in separate lineages (Wood et al. 2005; Christin et al. 2010).

Certain specialized effector genes encode proteins such as gas transporters, sensory proteins, and targets of xenobiotics that directly interact with extrinsic variables at the molecular level. It is intuitive that such highly specialized, environmental sensing genes are repeatedly the target of directional selection during adaptation to novel conditions. Resistance to xenobiotics has sometimes evolved independently via identical amino acid substitutions in single residues of a small set of genes (Table S2). In particular, the repeated appearances of insecticide resistance illustrate the rapidity of evolutionary change in

response to anthropogenic perturbations, and highlight a restricted adaptive landscape of accessible mutations. This is confirmed by short-term experimental evolution studies in yeast lines: resistance to a fungicide has evolved in 33 out of 35 cases by single mutations in the hotspot genes *ERG3* and *ERG6* (Gerstein et al. 2012) of which 4 are amino acid substitutions that appeared in multiple lines (Table S2). Of note, molecular convergence at the amino acid level is also widespread among drug-resistant malaria parasites and bacteria (e.g., Toprak et al. 2011; Manske et al. 2012).

Sensory specializations to similar environments also display many cases of extreme parallelism. For instance, analogous amino acid substitutions are found in the high-frequency hearing proteins of echolocating mammals (Davies et al. 2012aa; Shen et al. 2012), in vision-related opsins (Davies et al. 2012bb; Martin and Orgogozo 2013), and in hemoglobins (Perutz 1983; McCracken et al. 2009; Martin and Orgogozo 2013). The presence of such tuning sites, each displaying evolutionarily "hot" amino acid substitutions, unravels a limited number of effective mutations in these proteins where the structure–function relationship is relatively well understood.

In the current list of Loci of Evolution, independent evolution of the same amino acid was reported mostly in proteins that molecularly interact with environmental factors. Possible exceptions consist of the several key amino acid changes involved in the parallel evolution of the C4-photosynthetic pathway in grasses (Christin et al. 2007), and the Glu92Lys mutation in the constitutively active forms of the melanocortin receptor (Mc1R), which results in melanic forms of bananaquits, chicks, and mice (Theron et al. 2001). Another example of interest was captured during the experimental adaptation of yeast clones to low-glucose growth conditions, where three independent lines underwent an evolutionary reversion (Asp30Gly) in *Mkt1*, converting a derived laboratory allele to a wild-type state (Anderson et al. 2010; Parreiras et al. 2011). This case of extreme parallelism, in accordance with experimental evolution work in prokaryotes (Tenaillon et al. 2012), thus suggests that at least a fraction of the mutational trajectories that underlie the evolution of complex traits may be predictable at the nucleotide level.

Overall, the repeated evolution of specific amino acid residues are one of the best illustrations of the constraints that structure the protein adaptive landscape, with a finite number of substitutions allowing a functional and viable effect on the phenotype.

HOTSPOTS AT WIDER GENETIC RESOLUTIONS

Our hotspot definition is restricted to independent genetic variation at orthologous genes and applies to various genetic resolutions, from orthologous nucleotides or codons, to orthologous subgenetic regions (such as protein domains and *cis*-regulatory regions), to orthologous genes (Fig. 2). We did not include in hotspots cases of unlinked paralogues or different components of the same pathway, even though their insights into genotype-to-phenotype relationships must be appreciated equally.

Hotspots Span Various Phylogenetic Distances

HOTSPOTS OF EXPERIMENTAL EVOLUTION

Experimental evolution imposes controlled selective pressures during extremely short timescales. Genomic scans of experimentally evolved lines have uncovered unexpected patterns of repeated evolution between replicates. For instance, after selection of five yeast clones for growth in low-glucose conditions, three lines evolved null-alleles of the transcription repressor *MTH1*, and the two other lines evolved amplification of the glucose transporters *HXT6/7*, revealing two mutually exclusive but repeatable routes for adaptation to low-glucose conditions (Kvitek and Sherlock 2011). In another study comparing responses to limitations to various nutrients, five lines out of nine adapted to low-glucose conditions via amplification of *HXT6/7*, 15 out of 16 adapted to low-sulfate by amplification of the sulfate transporter *SUL1*, and two out of eight lines adapted to low-phosphate by structural variation at the *PHO5* gene (Gresham et al. 2008). The study of experimental evolution, particularly when starting from identical individuals (which removes standing genetic variation), thus offers an interesting proxy to understand mutational landscapes of adaptation, and already uncovers surprising patterns of repeatability within extremely short timescales.

HOTSPOTS OF DOMESTICATION IN PLANTS AND ANIMALS

The study of the genetic basis of quantitative traits has traditionally focused on agricultural traits, revealing the malleability of phenotypes over relatively short timescales. Entries related to domestication QTGs in plants and animals constitute a third of the total dataset (Table S1). Domestication QTGs are extremely useful for addressing the repeatability of evolution, because they correspond to replicated experiments where the same traits have been selected in diverse plants (grain properties, color, resistance to abiotic stress, flowering time) and animals (coat color, meat, or milk yield). Domestication hotspot genes cover an unexpectedly wide range of species and traits (Tables S3 and S4), suggesting that the domestication process may have largely recruited variation at the same genes in different species. Of note, new genome-wide methods of genotype–phenotype associations have started to provide a comparative template for the study of repeated evolution. In a genome scan of selection conducted on 74 worldwide breeds of sheep (Kijas et al. 2012), 8 of the top 31 genomic regions that showed a significant signature of differentiation contain candidate

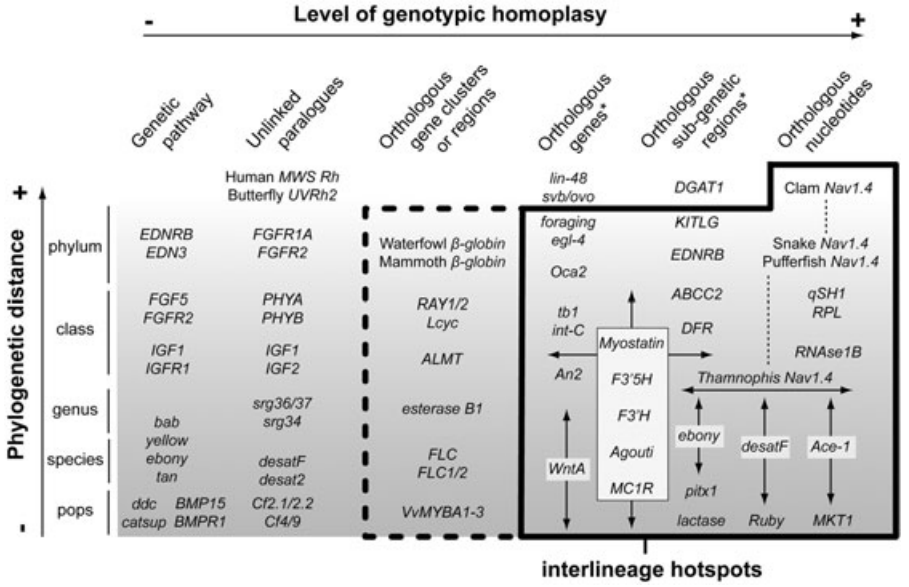

Figure 2. **A space for "Quantitative Trait Gene" (QTG)-based cases of repeated evolution. This abstract space features known QTGs as referred to in the text and in the loci of evolution catalog (Martin and Orgogozo 2013). Even when a single gene name is given, each entry corresponds to comparisons in at least two lineages. The vertical scale approximates the phylogenetic distance between these comparisons. Asterisks: the horizontal position of QTGs between these categories is nondefinitive and depends both on the genetic resolution of the corresponding studies, and on subjective criteria used to define orthology at the subgenic level, especially at noncoding sites.**

genes that were previously identified as domestication genes in other species. Genome scans on selected lines will greatly benefit from increased resolution, replication, and sample sizes. It will be interesting to see if such approaches confirm the existence of a small toolkit of domestication genes (as seen in Tables S3 and S4), potentially overlapping with hotspots of adaptation in the wild.

HOTSPOTS ACROSS LONG DISTANCES AND AT THE LIMITS OF PHENOTYPIC SIMILARITY

At the other extreme of the evolutionary timescale, hotspots can also span very long phylogenetic distances (Fig. 2). This is the case of melanin biosynthetic pathway genes such as *OCA2*, *MC1R*, and *KITLG* that are determinants of pigmentation differences between population of both humans and fishes (Table S4), or of an amino acid substitution in the Nav1.4 sodium channel that independently evolved in mollusks and vertebrates (Table S2). The *Drosophila* gene *foraging* and its nematode orthologue *egl-4* not only provide an example of hotspot across a long phylogenetic distance, but are also rare examples of natural variation underlying behavioral traits (Martin and Orgogozo 2013). Natural *cis*-regulatory alleles of *foraging* determine food-search activity in flies (Osborne et al. 1997; Mery et al. 2007). Strikingly, natural variation at the *egl-4* locus in the nematode *Pristionchus pacificus* also affects food-search behavior: these organisms are facultative parasites of beetles, and the *egl-4* variants influence host species preference based on pheromone cues (Hong et al. 2008). Whether "food-search behavior" represents a legitimate phenotypic similarity is debatable, but it is also plausible that *for/egl-4* has been at the core of a genetic module involved in similar traits before the fly / worm split.

Furthermore, *cis*-regulatory variation in *lin-48* and its orthologue *ovo/shavenbaby* (*ovo/svb*) is linked to evolution of excretory duct position in nematodes and larval cuticle patterning in flies, respectively (Sucena et al. 2003; Wang and Chamberlin 2004; Frankel et al. 2012). At an even more dramatic phylogenetic distance, artificial selection for oil content and milk yield during domestication of maize and dairy cattle both reflect single amino acid changes in a type I acyl-CoA : diacylglycerol acyltransferase (DGAT1) enzyme (Grisart et al. 2004; Zheng et al. 2008). The *lin-48/ovo-svb* and cattle *DGAT1* / maize *DGAT1* genes show clear orthology in each case (Wang and Chamberlin 2004; Turchetto-Zolet et al. 2011), and they may be considered to underlie similar trait phenotypic variations if our trait definition moves from observable levels ("excretory duct position" vs. "cuticle patterning"; "oil content" vs. "milk yield") toward molecular or developmental mechanisms (variations in "epithelial patterning" and "lipid biosynthesis"). Whether or not these examples are treated as hotspots, they provide a potential reflection on how to categorize phenotypes.

In conclusion, our compilation of existing data shows that hotspots exist at many levels: from orthologous nucleotides to

genetic pathways, and from individuals of the same population to distantly related species (Fig. 2).

Hotspots Can Involve Loss-of-Function or Gain-of-Function Mutations

CODING LOSSES-OF-FUNCTION AS HOTSPOTS FOR RAPID ADAPTATION

Stop, frameshift, and certain missense mutations, as well as the deletion of coding sequences or entire genes, tend to eliminate protein production and activity. Here again, cases of genetic replication illustrate how coding mutations of large effect can provide selective advantages in a given environment (Table S5). This is exemplified by the widespread evolution of rapid cycling in the wild plant model species *Arabidopsis thaliana*: whereas five genetically defined pathways control the timing of flowering (Srikanth and Schmid 2011), wild-populations of *A. thaliana* have repeatedly taken the same two genetic paths—null mutations in *FRI* or *FLC* (Table S5)—to accelerate their life cycle, a strategy that provides a selective advantage apparently sufficient to overcome the possible deleterious effects of these null alleles.

Such opportunistic and replicated use of knockout mutations in specific environments is well illustrated by cases of responses to extreme selective pressure. For instance, several mutations in the red blood cell enzyme G6PD confer resistance to malaria and show strong signatures of positive selection in spite of causing human pathologies (Tishkoff et al. 2001; Louicharoen et al. 2009; Timmann et al. 2012). Alternatively, loss-of-function may simply be viable when a novel environment abrogates the necessity for a gene, as seen in the independent loss-of-functions alleles of the *Oca2* and *Mc1R* pigmentation genes in subterranean populations of cavefish (Protas et al. 2006; Gross et al. 2009).

Although mutations that impair gene function are common, it is reasonable to see them as rather short-term solutions to special environmental conditions.

CIS-REGULATORY ALTERATIONS ARE HOTSPOTS FOR TISSUE-SPECIFIC TRAIT LOSS

The cases of coding loss-of-function mutations mentioned earlier concern genes of relatively low phenotypic pleiotropy, meaning that their inactivation affects a small number of biological structures and processes. A corollary is that the more pleiotropic a gene is, the less likely it is to yield a viable phenotype when inactivated, except perhaps when existing in multiple and functionally redundant copies. It is thus worth mentioning here that regulators of morphogenesis, in spite of being generally more pleiotropic than genes that affect physiological traits (Liao et al. 2010), can in some cases bear large deletions in the wild. However, these are typically restricted to regulatory regions that direct tissue-specific expression, resulting in local loss-of-functions without known pleiotropic effects, consistent with previous arguments on the importance of *cis*-regulatory evolution (Stern and Orgogozo 2008). For example, the repeated loss of pelvic spines in distant populations of freshwater sticklebacks has involved independent deletions of a pelvic-specific regulatory region of the *pitx1* gene in the last 10,000 years (Chan et al. 2010); loss of abdominal pigmentation in the young *Drosophila santomea* lineage is explained by three different loss-of-function mutations in an enhancer of the *tan* locus (Jeong et al. 2008; Rebeiz et al. 2009b); and adaptations to saline soils in *Arabidopsis* coastal populations involves small deletions in the promoter of the sodium transporter *AtHKT1*, that evolved independently in Japan and Europe, and in each case reduce gene expression in roots (Rus et al. 2006).

Overall, we can deduce from these empirical findings that gene inactivations, be they of coding or *cis*-regulatory nature, are genuine components of the mutational landscape that can yield adaptive responses in the wild. When reaching fixation in a species, null and hypomorphic alleles result in removal of genetic information, which can restrict their evolutionary potential (Zufall and Rausher 2004). Phenotypic losses may be reversed in the presence of low-frequency or introgressing functional alleles, as proposed in the reversed evolution of low-plated stickleback fishes in answer to recent shifts in predation regime in an urban freshwater environment (Kitano et al. 2008). In such cases, null and hypomorphic alleles can simply be seen as large effect switches that transiently contribute to evolutionary changes. Permanent losses of functionally important genes can alternatively be maintained over long evolutionary timescales. For example, insects lost the ability to synthesize cholesterol through independent losses of several genes of the cholesterol biosynthetic pathway (Clayton 1964; Vinci et al. 2008). Although these gene losses created a complete metabolic dependence toward dietary sterols, which could be detrimental in the long term, insects represent more than half of all known living animal species.

HOTSPOTS OF CONSTRUCTIVE EVOLUTION

Clearly evolution does not always occur by removing bits of information. We consider here as constructive evolution the mutations that result in an increase in complexity (this includes the gain of an interaction with a trait inhibitor, resulting in an apparent loss-of-function).

Whole-gene duplications are an obvious way to add information at the sequence level, and such changes create redundancy that can become a template for repeated phenotypic evolution (Rohner et al. 2009).

Gene copy number variation within and between species can also result in increases in gene dosage of adaptive relevance, as illustrated by repeated amplification of the metal transporters

MTP1, *HMA3*, and *HMA4* in plants tolerant to contaminated soils (Hanikenne et al. 2008; Ueno et al. 2011; Shahzad et al. 2012), or of the *amylase1* gene in humans in association with starch-rich diets (Perry et al. 2007). This mechanism can sustain rapid adaptive bouts, as seen in populations of *Culex* mosquitoes where multiple amplifications of the *esterase B1* genes have evolved in response to the recent use of organophosphorous insecticides (Guillemaud et al. 1999).

Alternatively, gain-of-function alleles can result from coding changes that confer increased affinity of a protein to another molecule. This is the case of the esterase E3 enzymes, which instead of acquiring resistance to organophosphorous compounds via copy increase, resulted in similar insecticide-catalytic properties via the same amino acid change that appeared independently in three fly species (Table S2; Hartley et al. 2006). Interestingly, both modes of quantitative increase in genetic activity are visible in the parasitic yeast *Candida albicans*: drug-resistant strains independently evolved via both gene amplification and gain-of-function coding mutations of the transcription factor *TAC1* (Coste et al. 2007).

HOTSPOTS OF TISSUE-SPECIFIC NOVELTIES BY *CIS*-REGULATORY TINKERING

As already discussed, coding and copy number variations can be constrained by their pleiotropic effects, whereas *cis*-regulatory mutations have the potential to modulate gene expression in specific tissue types and developmental stages, providing precision and flexibility to phenotypic change. We know that a few nucleotide changes can suffice to create novel, albeit complex enhancer activities (Eichenlaub and Ettwiller 2011). Regardless of whether these mutations have a quantitative, qualitative, negative or positive effect on transcription, mutations that increase the number or stability of molecular interactions of a given *cis*-element with *trans*-regulatory factors can be regarded as "constructive," in the sense that they increase the complexity of the regulatory mechanism. For instance, cases where small deletions actually result in novel binding sites (Shirangi et al. 2009) or where a loss of expression is due to the novel binding of an inhibitor of transcription are examples of constructive evolution.

A few cases are known where derived phenotypes arose from single *cis*-regulatory changes. The evolution of lactase persistence in humans is perhaps the best illustration of not only how a novel trait can arise from a noncoding mutation, but also of adaptive evolution that recidivates, pinpointing strong biases in the genomic loci that can yield to adaptive variation. Lactase persistence, the ability to consume milk during adult life, has independently evolved in pastoral cultures of Europe and Sub-Saharan Africa less than 8,000 years ago (Tishkoff et al. 2006; Dunne et al. 2012). A single European allele and three alleles of Sub-Saharan origin each arose independently through a single mutation within a 100-bp wide long-range enhancer of the gene encoding the lactase enzyme that cleaves lactose (Enattah et al. 2002; Tishkoff et al. 2006). These single base-pair substitutions all individually increase in vitro gene expression and *trans*-regulatory factor affinity compared to the ancestral haplotypes, and correlate with increases in lactase activity in adults (Lewinsky et al. 2005; Tishkoff et al. 2006). The *lactase* enhancer is thus a genuine hotspot of "gain-of-function" evolution in modern humans over short evolutionary scales.

In conclusion, hotspots include cases of both constructive evolution and losses, and cover many types of mutations (coding, structural, and *cis*-regulatory).

Part 3. Intralineage Hotspots Form Large-Effect QTGs

LARGE-EFFECT ALLELES AS THE LOW-HANGING FRUITS OF THE QTG PROGRAM

The QTG approach has been condemned because it has an inescapable bias toward large-effect QTGs as a consequence of top-down approaches, which fairly survey the entire genome, but can only pick up the low-hanging fruits with largest phenotypic effects (Rockman 2011). Even though great advances are still to come in mapping techniques, there will always be a bias toward large-effect QTGs because identification of the refractory small-effect QTGs requires an incomparable amount of genotyping and phenotyping efforts. As a matter of fact, virtually all QTGs listed here have rather large phenotypic effects (Martin and Orgogozo 2013). However, this bias might not be as problematic as it seems. Indeed, recent QTG empirical data clarify the nature of large-effect alleles, and argue that alleles of large phenotypic effects are not necessarily the result of large-effect mutations. A recent review article compares the race for QTG identification to a Gold Rush: large-effect alleles can be compared to golden nuggets that are immediately visible to the prospector, whereas low-grade ore (small-effect alleles) remains undetectable (Rockman 2011). But as we will see next, this metaphor can be pushed further: actual golden nuggets, rather than spontaneously forming large chunks of metal, are made of smaller particles that gradually agglomerate by thermal or mechanical accretion in a water stream (Hough et al. 2007).

COMPLEX ALLELES ARE CLUSTERS OF TIGHTLY LINKED VARIANTS

> Large effects due to a single locus may be due to multiple associated polymorphisms (or sequential fixations in isolated populations) rather than individual mutations of large effect.
>
> L.F. Stam and C.C. Laurie (1996)

A complex allele results from several mutations underlying phenotypic variation that accumulate in the same gene. It is

thus similar to the hotspot concept, with the difference that here replication occurs within a single lineage (Fig. 1E). This idea is not new. Early studies of the genetic architecture of alcohol dehydrogenase (Adh) activity in *Drosophila* already pinpointed that at least three mutations contribute to variation in Adh protein levels (Stam and Laurie 1996), and recent analyses have since confirmed that gene expression levels are often shaped by the combined effects of multiple sites (e.g., Tao et al. 2006). We examine below how the evidence for this phenomenon has extended to phenotypes above the molecular level.

PHENOTYPIC SYNERGY BETWEEN MULTIPLE CODING SITES DURING PROTEIN EVOLUTION

Detectable shifts in protein function are not always reducible to single mutations, in which cases they can be considered as examples of intralineage hotspots (Fig. 1E). Photoreceptor evolution provides canonical examples of such colocalized changes. For instance, rhodopsin from deepwater loosejaw fishes shows a light wavelength absorbance shift of +40 nm that confers sensitivity to the red bioluminescence of their own forehead "lantern" (O'Day and Fernandez 1974), and at least six amino acid substitutions at effective spectral tuning residues explain this red shift (Yokoyama et al. 2008). Similar studies that measure phenotypes "substitution by substitution" have also uncovered synergistic effects where the phenotype of combined point mutations is more than the sum of their individual effects. To illustrate this principle, point mutations of the hotspot tuning sites 86 and 93 of the vertebrate violet/UV-sensitive opsin SWS1 show individual infinitesimal shifts of 2 and 0 nm, but trigger a 25-nm shift when combined (Takahashi and Yokoyama 2005). Combinations of certain other mutations are strictly additive, rather than synergistic. Yet, we now begin to see how individual mutations can have small effects, although their association can result in large, synergistic effects. In other cases known as examples of "compensatory" evolution, a phenotypically silent mutation can influence the effect of an other mutation within the same protein, for instance countering its deleterious effects (Campbell et al. 2010). In summary, our current knowledge of protein evolution argues in favor of synergistic models in which individual mutations of small-effects generate large-effects when clustered together.

COMPLEX *CIS*-REGULATORY VARIATION SHAPES NOVEL GENE EXPRESSION DOMAINS

Most of the evo-devo literature reports divergent gene expression patterns associated with phenotypic changes. Such spatial (heterotopies) or temporal (hetechronies) shifts in gene expression are caused either by *cis*-regulatory mutations in the gene of interest or by coding or *cis*-regulatory changes in an upstream regulator gene. Recent data suggest that subtle changes in gene expression patterns evolve through the accumulation of several fine-tuning mutations in the same *cis*-regulatory elements (Wittkopp and Kalay 2011). For instance, at least two mutations in the noncoding region *HACNS1* in humans appear to be responsible for a novel expression pattern in developing limbs (Prabhakar et al. 2008). In *Drosophila santomea*, the gene *Nep1* evolved a novel optic lobe expression domain in less than 500,000 years through four mutations within a 923-bp *cis*-regulatory element, and they are all necessary to produce this novel expression pattern (Rebeiz et al. 2011).

COMPLEX *CIS*-REGULATORY VARIATION SHAPES TISSUE-SPECIFIC PHENOTYPIC VARIATION

Although the phenotypic effects of the *HACNS1*- and *Nep1*-derived variants remain unclear, these studies inform us that other *cis*-regulatory variants may have emerged from several mutations, in particular when the derived phenotypes are linked to spatial shifts in gene expression (for examples of known adaptive relevance, see Mery et al. 2007; Wittkopp et al. 2009; Manceau et al. 2011; Reed et al. 2011; Roberts et al. 2011; Martin et al. 2012). A few studies have narrowed derived gene expression activities to distinct regions of the same genes (Wang and Chamberlin 2004; Bickel et al. 2011; Loehlin and Werren 2012), but they do not reach sufficient resolution to test the effects of individual mutations.

This caveat was overridden in a study that deciphered the respective effects of point mutations that collectively contribute to the loss of larval trichomes in *Drosophila sechellia* (Frankel et al. 2011). First, the causative variation responsible for the loss of dorsolateral trichomes was localized to a 500 bp-region, within an enhancer of the aforementioned *ovo/svb* gene. The same region drives robust *ovo/svb* expression in the corresponding epidermal cells of *Drosophila melanogaster*, and the authors identified 13 point mutations in this region that are specific to *D. sechellia* and could thus account for the "naked" phenotype in this species. Engineering of individual *D. sechellia*-specific point mutations has minute effects on the number of trichomes in transgenic *D. melanogaster* larvae, whereas the combined modification of all causative sites results in a dramatic reduction of trichomes. Opposite effects are observed in larvae carrying a *D. sechellia* enhancer that has been reverse engineered site by site to a *D. melanogaster* state. This study illustrates how a derived phenotype can arise from the combination of many *cis*-regulatory micromutations, rather than from a single mutation of large effect. Whether the divergent *D. sechellia* morphology emerged gradually through the successive appearance and fixation of causing mutations or from the opportunistic recombination of standing genetic variation is unknown, due to limited nucleotide diversity in present populations of *D. sechellia*.

This question was tackled in a study that dissected the effects of point mutations that segregate in natural populations (Rebeiz

et al. 2009a). *Cis*-regulatory variation of the *ebony* melanin synthesis inhibitor accounts for most of the variation in abdominal pigmentation of western African populations of *D. melanogaster* (Rebeiz et al. 2009a). Local differences in pigmentation between high-altitude dark-colored and low-altitude light-colored flies are explained by quantitative differences in *ebony* expression that are under the control of at least five single nucleotide polymorphisms. These five substitutions are clustered in two haplotype blocks, both within an enhancer of about 1 kb that drives abdominal expression, and each block accounts for about half of this enhancer activity in a reporter assay. Interestingly, these two haplotype blocks have different evolutionary histories. One was first assembled from the combination of three ancient mutations, whereas the other arose more recently by de novo mutation. This study illustrates with exquisite details how a complex allele can arise during evolution by both recombination between standing genetic variation and repeated de novo mutations within one locus (Fig. 1E).

HOTSPOTS AS ACCUMULATORS OF VARIATION

The repeated occurrence within a gene of independent mutations that cause similar phenotypic changes closely resembles the *interlineage hotspot* phenomenon described in Part 2 (Fig. 1D). Here, the repetition does not occur in distinct evolutionary branches, but within a single branch, and is thus considered as an *intralineage hotspot* (Fig. 1E). Importantly, *ebony* is not only involved in intraspecific variation in *D. melanogaster* but is also responsible for both intra- and interspecific variation in another clade of flies (Wittkopp et al. 2009). The *ebony* case thus reveals that intralineage hotspot genes may also be interlineage hotspot genes for the same phenotype. This might also be the case for other interlineage hotspots, where elaborate phenotypic variations are unlikely to be based on a single mutation (e.g., Mery et al. 2007; Manceau et al. 2011; Reed et al. 2011; Martin et al. 2012). In fact the difference between inter- and intralineage hotspots is only a matter of level of observation, because haplotypes undergoing recurrent mutation and recombination are lineages in their own right (Pennings and Hermisson 2006). In any case, the general concept of hotspot ultimately refers to the fact that for a given type of phenotypic change, the relevant genetic variation is often highly localized in the genome.

Discussion

HOW FREQUENT ARE HOTSPOTS? CHALLENGES FOR QUANTITATIVE META-ANALYSES

The large number of hotspots in the data shows that some of the evolutionarily viable regions of the phenotypic landscape can be reached by genetic changes at a restricted set of loci. Yet the multiple biases in the catalog construction (Supplementary Notes) or in the experimental detection of Loci of Evolution itself (Rockman 2011), together with the heterogeneity in sources and experimental approaches challenge rigorous quantitative hypothesis testing with the present dataset. The "Linkage Mapping" sub-dataset is relatively comprehensive and contains studies that often lack a priori expectations on genetic features, but still carries a strong sampling bias toward certain phenotypic categories and shallow phylogenetic nodes (Table S1). The prevalence of hotspots we observed is also inflated by detection biases toward large-effect alleles (Rockman 2011). Elusive and dispersed small-effect mutations can fine tune phenotypes and collectively explain large fractions of the observed variation but are experimentally masked by linkage with large-effect loci and pervasive epistasis (Huang et al. 2012; Lorenz and Cohen 2012; Zuk et al. 2012). Therefore, one cannot conclude from current results that genetic paths are limited for *all* types of phenotypic changes. Independent evolution of similar traits has sometimes been shown to occur through mutations in different genes (e.g., Gruber et al. 2007; Aminetzach et al. 2009; Shapiro et al. 2009; Anderson et al. 2010; Roelants et al. 2010; Kvitek and Sherlock 2011; Protas et al. 2011; Green II and Extavour 2012) indicating that large-effect variations are not always caused by the same loci.

Although we encourage future quantitative meta-analysis over all the Loci of Evolution, we urge that this must be done with caution. A reasonable alternative may reside in the extraction of datasets restricted to a specific phenotypic class (Streisfeld and Rausher 2011). This "metamodel" approach (Kopp 2009) corrects for data heterogeneity while dealing with more manageable and meaningful variables, and would at least provide a desirable estimate of the proportion of large effect QTLs that are caused by hotspots for a given trait.

WHAT MAKES A HOTSPOT "Hot"

> It strains one's faith in the laws of chance to imagine that identical changes should crop out again and again if the possibilities are endless and the probabilities equal.
>
> A.F. Shull (1935)

Interlineage hotspots suggest that even with different genetic backgrounds and population histories, similar evolutionary changes are often caused by mutations at orthologous loci. What intrinsic properties of a given gene make it a path of least resistance for driving phenotypic evolution? Multiple papers recently discussed the various intrinsic properties of a gene that might make it a hotspot for evolution (Papa et al. 2008; Gompel and Prud'homme 2009; Kopp 2009; Stern and Orgogozo 2009). Our purpose is not to summarize them all again here, but to discuss potential avenues of research to gain better insight into this important yet unresolved question.

A first explanation is mutational bias. For instance, complexes of gene clusters as well as repeat-rich regions are prone to

ectopic exchange, explaining the spectacular repeatability of gene amplifications, including within the time frame of experimental evolution studies (Gresham et al. 2008; Kvitek and Sherlock 2011). Colocalized occurrences of narrower mutations testify to mutational biases. For instance, the recurrent deletion of the stickleback "pelvic spine" enhancer of *pitx1* correlates with a genome-wide hotspot of chromosomal instability (Chan et al. 2010). Similarly, the independent occurrence of three functional gene fusions between the gene *CypA* and the antiviral factor *TRIM5alpha* that conferred resistance to retroviruses during primate evolution first defies our imagination (Malfavon-Borja et al. 2013). But in retrospect, the probability for such chimeras to appear might be relatively substantial because *CypA* exists in multiple copies that bombard the genome by retrotransposition (Zhang et al. 2003).

A second category of explanation deals with the concept of *optimal pleiotropy* (Kopp 2009), which we interpret as the capacity of a locus to generate variation in a given phenotypic trait without affecting other functions. The problem is different depending on the position of a gene in the differentiation cascade (Stern and Orgogozo 2009). For instance, highly specialized "effector" genes expressed in a limited number of cell types are intuitively major modulators of the relevant phenotypes. After all, it is not surprising that "Metal tolerance" repeatedly maps to genes specialized in metallic ion transport in plants, or that the ability to digest lactose repeatedly maps to a lactose-cleaving enzyme (Martin and Orgogozo 2013). Hence the truism: *specialized genes drive the evolution of specialized traits*.

Conversely, regulator genes that form a hub in a regulatory network between a series of upstream activators (the input) and a battery of downstream effector genes have been proposed to be hotspot genes, with *ovo/svb* incarnating the typical hub gene (Stern and Orgogozo 2009; Chanut-Delalande et al. 2012). DNA-binding domain transcription factors form 20.8% (103/494) of the Linkage Mapping entries in our dataset, whereas these genes form only 5.5–7% of the protein coding genes annotated in the human, *Drosophila*, *and Arabidopsis* genomes (Guo et al. 2005; Adryan and Teichmann 2006; Lee et al. 2007). Because the "Linkage Mapping" sub-dataset has limited ascertainment biases on gene function, this enrichment suggests that the ability to regulate batteries of downstream target genes is common among hotspot genes. Finally, hotspots of *cis*-regulatory evolution necessitate complex and modular regulatory regions, which led to the proposal that genes with large intergenic or intronic regions have more potential for *cis*-regulatory evolution (Knecht et al. 2007). Here again, future investigations including a detailed understanding of mutational biases and regulatory regions should allow more rigorous tests of these hypotheses.

HOTSPOTS RECONCILE MICROMUTATIONISM AND LARGE QTL

> [. . .] the conflicting views of micromutationism and macromutationism can actually reflect observations of the same molecular mechanisms at different levels of resolution.
>
> A.P. McGregor et al. (2007)

> The Hopeful Monster is dead. Long live the Hotspot!
>
> D.L. Stern (2010)

At first glance, the empirical discovery of large-effect QTLs seems to refute the widely accepted infinitesimal model (Rockman 2011), which stipulates that adaptive evolution involves genetic variants of small phenotypic effect that drive minute steps onto the phenotypic landscape (Orr 2005). In other words, the QTG program seems to have resurrected a mild form of saltationism, where large-effect alleles are assimilated to the molecular underpinnings of impossible "Hopeful Monsters." However, one can notice here a terminological conundrum where large-effect alleles are assumed to result from large-effect mutations (Akam 1998). In an attempt to justify the existence of discrete variation with gradualism, one of the founders of Neo-Darwinism already proposed that "Hopeful Monsters," if they exist in nature, "may be due to the accumulation of small gene mutations" (Huxley 1942; Stern 2000), but this intuition has been largely overlooked due to the lack of empirical data.

Recent discoveries of intralineage hotspots illustrate how alleles with large phenotypic effects can actually be aggregates of multiple small-effect mutations that have arisen independently. Hotspots of variation and large-effect alleles composed of multiple micromutations thus resolve the paradox between the theoretical prediction that many mutations of infinitesimal size predominate in evolution, and the empirical finding that few loci of relatively large effects underlie many cases of phenotypic variation in nature (as foreseen in McGregor et al. 2007; Stern 2010). Because they aggregate into large-effect QTGs, small-effect QTNs in intralineage hotspot genes make themselves discernible to current approaches. Even though small-effect QTNs remain high-hanging fruits when dispersed over the genome (Rockman 2011), we can rejoice that the ones underlying large phenotypic changes are within reach. Finally, the finding that multiple mutations underlie large effects revives the need to incorporate this scenario in current theoretical models on the distribution and detection of mutational effects (Pennings and Hermisson 2006; Hermisson and McGregor 2008).

Conclusion

> To collect and codify the facts of Variation is, I submit, the first duty of the naturalist. This work should be undertaken if only

> to rid our science of that excessive burden of contradictory assumptions by which it is now oppressed. [. . .] The only way in which we may hope to get at the truth is by the organization of systematic experiments in breeding, a class of research that calls perhaps for more patience and more resources than any other form of biological enquiry. Sooner or later such an investigation will be undertaken and then we shall begin to know.
>
> W. Bateson (1894)

The QTG program is nothing else than a modern realization of the vision of William Bateson, who defended a decade before inventing the term "genetics" that the material basis of variation was central for understanding evolution. The empirical approach he envisioned continues to relieve us from the burden of old paradoxes. The data allowed us to derive here three qualitative statements about repeated evolution. First, we delineated five scenarios of repeated evolution and gene reuse by separating their genotypic and phenotypic components (Fig. 1). Second, we found that hotspots cover a wide range of processes and exist at previously unsuspected levels. Third, hotspots broadly testify for a tendency of genetic variation with specific phenotypic effects to cluster in the genome. This is an important step in unifying theory and experimental evidence in evolutionary genetics as it erases former discrepancies between gradualism and saltationism: alleles of large effect can be explained by the accumulation of small effect mutations at hotspot loci.

Now that data are becoming available at an accelerating cadence, it is time to deploy collective efforts in completing and perfecting the list of "Loci of Evolution" as a community resource to continue to gain insights into the genetic basis of evolution.

ACKNOWLEDGMENTS

We are indebted to M.-A. Félix, D. Stern, T. Schilling, A. McGregor, M. Rockman, and two anonymous referees for their input on the manuscript, and to R. Reed for stimulating its genesis. This work was supported by the National Science Foundation grant IOS-1052541 and an ATIP (Action Thématique et Incitative sur Programme)-Avenir grant. The authors declare no conflict of interest. AM conceived the project and compiled the Loci of Evolution catalog. AM and VO wrote the manuscript.

LITERATURE CITED

Adryan, B., and S. A. Teichmann. 2006. FlyTF: a systematic review of site-specific transcription factors in the fruit fly *Drosophila melanogaster*. Bioinformatics 22:1532–1533.

Akam, M. 1998. Hox genes, homeosis and the evolution of segment identity: no need for hopeless monsters. Int. J. Dev. Biol. 42:445–451.

Aminetzach, Y. T., J. R. Srouji, C. Y. Kong, and H. E. Hoekstra. 2009. Convergent evolution of novel protein function in shrew and lizard venom. Curr. Biol. 19:1925–1931.

Anderson, J. B., J. Funt, D. A. Thompson, S. Prabhu, A. Socha, C. Sirjusingh, J. R. Dettman, L. Parreiras, D. S. Guttman, A. Regev, L. M. Koh, et al. 2010. Determinants of divergent adaptation and Dobzhansky-Muller interaction in experimental yeast populations. Curr. Biol. 20:1383–1388.

Anderson, T. M., B. M. vonHoldt, S. I. Candille, M. Musiani, C. Greco, D. R. Stahler, D. W. Smith, B. Padhukasahasram, E. Randi, J. A. Leonard, et al. 2009. Molecular and evolutionary history of melanism in North American Gray Wolves. Science 323:1339–1343.

Arendt, J., and D. Reznick. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. 23:26–32.

Avise, J. C., and T. J. Robinson. 2008. Hemiplasy: a new term in the lexicon of phylogenetics. Syst. Biol. 57:503–507.

Barrett, R. D. H., and D. Schluter. 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23:38–44.

Bateson, W. 1894. Materials for the study of variation, treated with especial regard to discontinuity in the origin of species. Macmillan, London.

Bekpen, C., T. Marques-Bonet, C. Alkan, F. Antonacci, M. B. Leogrande, M. Ventura, J. M. Kidd, P. Siswara, J. C. Howard, and E. E. Eichler. 2009. Death and resurrection of the human IRGM Gene. PLoS Genet. 5:e1000403.

Bickel, R. D., A. Kopp, and S. V. Nuzhdin. 2011. Composite effects of polymorphisms near multiple regulatory elements create a major-effect QTL. PLoS Genet. 7:e1001275.

Brakefield, P. M. 2006. Evo-devo and constraints on selection. Trends Ecol. Evol. 21:362–368.

Campbell, K. L., J. E. E. Roberts, L. N. Watson, J. Stetefeld, A. M. Sloan, A. V. Signore, J. W. Howatt, J. R. H. Tame, N. Rohland, and T. J. Shen. 2010. Substitutions in woolly mammoth hemoglobin confer biochemical properties adaptive for cold tolerance. Nat. Genet. 42:536–540.

Chan, Y. F., M. E. Marks, F. C. Jones, G. Villarreal Jr., M. D. Shapiro, S. D. Brady, A. M. Southwick, D. M. Absher, J. Grimwood, J. Schmutz, et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science 327:302–305.

Chanut-Delalande, H., P. Ferrer, F. Payre, and S. Plaza. 2012. Effectors of tridimensional cell morphogenesis and their evolution. Semin. Cell Dev. Biol. 23:341–349.

Charlesworth, D., E. Kamau, J. Hagenblad, and C. Tang. 2006. Trans-specificity at loci near the self-incompatibility loci in Arabidopsis. Genetics 172:2699–2704.

Christin, P.-A., N. Salamin, V. Savolainen, M. R. Duvall, and G. Besnard. 2007. C-4 photosynthesis evolved in grasses via parallel adaptive genetic changes. Curr. Biol. 17:1241–1247.

Christin, P.-A., D. M. Weinreich, and G. Besnard. 2010. Causes and evolutionary significance of genetic convergence. Trends Genet. 26:400–405.

Christin, P.-A., E. J. Edwards, G. Besnard, S. F. Boxall, R. Gregory, E. A. Kellogg, J. Hartwell, and C. P. Osborne. 2012. Adaptive evolution of C4 photosynthesis through recurrent lateral gene transfer. Curr. Biol. 22:445–449.

Clayton, R. B. 1964. The utilization of sterols by insects. J. Lipid Res. 5:3–19.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal, M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, D. M. Kingsley, et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307:1928–1933.

Conte, G. L., M. E. Arnegard, C. L. Peichel, and D. Schluter. 2012. The probability of genetic parallelism and convergence in natural populations. Proc. R. Soc. Lond. B. 279:5039–5047.

Coste, A., A. Selmecki, A. Forche, D. Diogo, M. E. Bougnoux, C. d'Enfert, J. Berman, and D. Sanglard. 2007. Genotypic evolution of azole resistance mechanisms in sequential *Candida albicans* isolates. Eukaryot. Cell 6:1889–1904.

Davies, K. T. J., J. A. Cotton, J. D. Kirwan, E. C. Teeling, and S. J. Rossiter. 2012a. Parallel signatures of sequence evolution among hearing genes in echolocating mammals: an emerging model of genetic convergence. Heredity 108:480–489.

Davies, W. I. L., S. P. Collin, and D. M. Hunt. 2012b. Molecular ecology and adaptation of visual photopigments in craniates. Mol. Ecol. 21:3121–3158.

Dunne, J., R. P. Evershed, M. Salque, L. Cramp, S. Bruni, K. Ryan, S. Biagetti, and S. di Lernia. 2012. First dairying in green Saharan Africa in the fifth millennium B.C. Nature 486:390–394.

Edwards, S. V., K. Chesnut, Y. Satta, and E. K. Wakeland. 1997. Ancestral polymorphism of Mhc class II genes in mice: implications for balancing selection and the mammalian molecular clock. Genetics 146:655–668.

Eichenlaub, M. P., and L. Ettwiller. 2011. De novo genesis of enhancers in vertebrates. PLoS Biol. 9:e1001188.

Enattah, N. S., T. Sahi, E. Savilahti, J. D. Terwilliger, L. Peltonen, and I. Jarvela. 2002. Identification of a variant associated with adult-type hypolactasia. Nat. Genet. 30:233–237.

Eriksson, J., G. Larson, U. Gunnarsson, B. Bed'hom, M. Tixier-Boichard, L. Stromstedt, D. Wright, A. Jungerius, A. Vereijken, E. Randi, et al. 2008. Identification of the *Yellow Skin* gene reveals a hybrid origin of the domestic chicken. PLoS Genet. 4:e1000010.

Ferrer-Admetlla, A., M. Sikora, H. Laayouni, A. Esteve, F. Roubinet, A. Blancher, F. Calafell, J. Bertranpetit, and F. Casals. 2009. A natural history of FUT2 polymorphism in humans. Mol. Biol. Evol. 26:1993–2003.

Frankel, N., D. F. Erezyilmaz, A. P. McGregor, S. Wang, F. Payre, and D. L. Stern. 2011. Morphological evolution caused by many subtle-effect substitutions in a transcriptional enhancer. Nature 474:598–603.

Frankel, N., S. Wang, and D. L. Stern. 2012. Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. Proc. Nat. Acad. Sci. USA 109:20975–20979.

Gerstein, A. C., D. S. Lo, and S. P. Otto. 2012. Parallel genetic changes and non-parallel gene-environment interactions characterize the evolution of drug resistance in yeast. Genetics 192:241–252.

Gompel, N., and B. Prud'homme. 2009. The causes of repeated genetic evolution. Dev. Biol. 332:36–47.

Green II, D. A., and C. G. Extavour. 2012. Convergent evolution of a reproductive trait through distinct developmental mechanisms in *Drosophila*. Dev. Biol. 372:120–130.

Gresham, D., M. M. Desai, C. M. Tucker, H. T. Jenq, D. A. Pai, A. Ward, C. G. DeSevo, D. Botstein, and M. J. Dunham. 2008. The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet. 4:e1000303.

Grisart, B., F. Farnir, L. Karim, N. Cambisano, J.-J. Kim, A. Kvasz, M. Mni, P. Simon, J.-M. Frere, W. Coppieters, et al. 2004. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proc. Nat. Acad. Sci. USA 101:2398–2403.

Gross, J. B., R. Borowsky, and C. J. Tabin. 2009. A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. PLoS Biol. 5:e1000326.

Gruber, J. D., A. Genissel, S. J. Macdonald, and A. D. Long. 2007. How repeatable are associations between polymorphisms in achaete-scute and bristle number variation in *Drosophila*? Genetics 175:1987–1997.

Guillemaud, T., M. Raymond, A. Tsagkarakou, C. Bernard, P. Rochard, and N. Pasteur. 1999. Quantitative variation and selection of esterase gene amplification in *Culex pipiens*. Heredity 83:87–99.

Guo, A., K. He, D. Liu, S. Bai, X. Gu, L. Wei, and J. Luo. 2005. DATF: a database of Arabidopsis transcription factors. Bioinformatics 21:2568–2569.

Hanikenne, M., I. N. Talke, M. J. Haydon, C. Lanz, A. Nolte, P. Motte, J. Kroymann, D. Weigel, and U. Kramer. 2008. Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. Nature 453:391–395.

Hartley, C. J., R. D. Newcomb, R. J. Russell, C. G. Yong, J. R. Stevens, D. K. Yeates, J. La Salle, and J. G. Oakeshott. 2006. Amplification of DNA from preserved specimens shows blowflies were preadapted for the rapid evolution of insecticide resistance. Proc. Nat. Acad. Sci. USA 103:8757–8762.

Hehemann, J. H., G. Correc, T. Barbeyron, W. Helbert, M. Czjzek, and G. Michel. 2010. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. Nature 464:908–912.

Hermisson, J., and A. P. McGregor. 2008. Pleiotropic scaling and QTL data. Nature 456:E3–E3.

Hines, H. M., B. A. Counterman, R. Papa, P. A. de Moura, M. A. Cardoso, M. Linares, J. Mallet, R. D. Reed, C. D. Jiggins, M. R. Kronforst, et al. 2011. A wing patterning gene redefines the mimetic history of *Heliconius* butterflies. Proc. Natl. Acad. Sci. USA 108:19666–19671.

Hofmann, C. M., N. J. Marshall, K. Abdilleh, Z. Patel, U. E. Siebeck, and K. L. Carleton. 2012. Opsin evolution in Damselfish: convergence, reversal, and parallel evolution across tuning sites. J. Mol. Evol.:1–13.

Hong, R. L., H. Witte, and R. J. Sommer. 2008. Natural variation in *Pristionchus pacificus* insect pheromone attraction involves the protein kinase EGL-4. Proc. Nat. Acad. Sci. USA 105:7779–7784.

Horger, A. C., M. Ilyas, W. Stephan, A. Tellier, R. A. L. van der Hoorn, and L. E. Rose. 2012. Balancing selection at the tomato RCR3 guardee gene family maintains variation in strength of pathogen defense. PLoS Genet. 8:e1002813.

Hough, R. M., C. R. M. Butt, S. M. Reddy, and M. Verrall. 2007. Gold nuggets: supergene or hypogene? Aust. J. Earth Sci. 54:959–964.

Huang, W., S. Richards, M. A. Carbone, D. Zhu, R. R. H. Anholt, J. F. Ayroles, L. Duncan, K. W. Jordan, F. Lawrence, M. M. Magwire, et al. 2012. Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. Proc. Nat. Acad. Sci. USA 109:15553–15559.

Hunt, D. M., K. S. Dulai, J. A. Cowing, C. Julliot, J. D. Mollon, J. K. Bowmaker, W. H. Li, and D. Hewett-Emmett. 1998. Molecular evolution of trichromacy in primates. Vis. Res. 38:3299–3306.

Huxley, J. 1942. Evolution. Modern synthesis. John Wiley & Sons, New York.

Igic, B., L. Bohs, and J. R. Kohn. 2006. Ancient polymorphism reveals unidirectional breeding system shifts. Proc. Nat. Acad. Sci. USA 103:1359–1363.

Jeong, S., M. Rebeiz, P. Andolfatto, T. Werner, J. True, and S. B. Carroll. 2008. The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. Cell 132:783–793.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484:55–61.

Kijas, J. W., J. A. Lenstra, B. Hayes, S. Boitard, L. R. Porto Neto, M. San Cristobal, B. Servin, R. McCulloch, V. Whan, K. Gietzen, et al. 2012. Genome-wide analysis of the World's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS Biol. 10:e1001258.

Kim, M., M.-L. Cui, P. Cubas, A. Gillies, K. Lee, M. A. Chapman, R. J. Abbott, and E. Coen. 2008. Regulatory genes control a key morphological and ecological trait transferred between species. Science 322:1116–1119.

Kitano, J., D. I. Bolnick, D. A. Beauchamp, M. M. Mazur, S. Mori, T. Nakano, and C. L. Peichel. 2008. Reverse evolution of armor plates in the threespine stickleback. Curr. Biol. 18:769–774.

Knecht, A. K., K. E. Hosemann, and D. M. Kingsley. 2007. Constraints on utilization of the EDA-signaling pathway in threespine stickleback evolution. Evol. Dev. 9:141–154.

Kopp, A. 2009. Metamodels and phylogenetic replication: a systematic approach to the evolution of developmental pathways. Evolution 63:2771–2789.

Kovach, M. J., M. N. Calingacion, M. A. Fitzgerald, and S. R. McCouch. 2009. The origin and evolution of fragrance in rice (Oryza sativa L.). Proc. Nat. Acad. Sci. USA 106:14444–14449.

Kvitek, D. J., and G. Sherlock. 2011. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. PLoS Genet. 7:e1002056.

Lee, A. P., Y. Yang, S. Brenner, and B. Venkatesh. 2007. TFCONES: a database of vertebrate transcription factor-encoding genes and their associated conserved noncoding elements. BMC Genomics 8:441.

Lewinsky, R. H., T. G. K. Jensen, J. Moller, A. Stensballe, J. Olsen, and J. T. Troelsen. 2005. T-13910 DNA variant associated with lactase persistence interacts with Oct-1 and stimulates lactase promoter activity in vitro. Hum. Mol. Gen. 14:3945–3953.

Liao, B. Y., M. P. Weng, and J. Zhang. 2010. Contrasting genetic paths to morphological and physiological evolution. Proc. Nat. Acad. Sci. USA 107:7353–7758.

Loehlin, D. W., and J. H. Werren. 2012. Evolution of shape by multiple regulatory changes to a growth gene. Science 335:943–947.

Loisel, D. A., M. V. Rockman, G. A. Wray, J. Altmann, and S. C. Alberts. 2006. Ancient polymorphism and functional variation in the primate MHC-DQA1 5' cis-regulatory region. Proc. Nat. Acad. Sci. USA 103:16331–16336.

Lorenz, K., and B. A. Cohen. 2012. Small and large effect QTL interactions underlie variation in yeast sporulation efficiency. Genetics 192:1123–1132.

Losos, J. B. 2011. Convergence, adaptation, and constraint. Evolution 65:1827–1840.

Louicharoen, C., E. Patin, R. Paul, I. Nuchprayoon, B. Witoonpanich, C. Peerapittayamongkol, I. Casademont, T. Sura, N. M. Laird, and P. Singhasivanon. 2009. Positively selected G6PD-Mahidol mutation reduces *Plasmodium vivax* density in Southeast Asians. Science 326:1546–1549.

Malfavon-Borja, R., L. I. Wu, M. Emerman, and H. S. Malik. 2013. Birth, decay, and reconstruction of an ancient TRIMCyp gene fusion in primate genomes. Proc. Nat. Acad. Sci. USA 110:E583–592.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20:229–237.

Manceau, M., V. S. Domingues, R. Mallarino, and H. E. Hoekstra. 2011. The developmental role of Agouti in color pattern evolution. Science 331:1062–1065.

Manske, M., O. Miotto, S. Campino, S. Auburn, J. Almagro-Garcia, G. Maslen, J. O'Brien, A. Djimde, O. Doumbo, I. Zongo, et al. 2012. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. Nature 487:375–379.

Martin, A., and V. Orgogozo. 2013. Data from: the loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. Dryad Digital Repository. doi:10.5061/dryad.v66p0.

Martin, A., R. Papa, N. J. Nadeau, R. I. Hill, B. A. Counterman, G. Halder, C. D. Jiggins, M. R. Kronforst, A. D. Long, W. O. McMillan, et al. 2012. Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. Proc. Nat. Acad. Sci. USA 109:12632–12637.

McCracken, K. G., C. P. Barger, M. Bulgarella, K. P. Johnson, S. A. Sonsthagen, J. Trucco, T. H. Valqui, R. E. Wilson, K. Winker, and M. D. Sorenson. 2009. Parallel evolution in the major haemoglobin genes of eight species of Andean waterfowl. Mol. Ecol. 18:3992–4005.

McGhee, G. 2011. Convergent evolution: limited forms most beautiful. The MIT Press, Cambridge, Massachussets.

McGregor, A. P., V. Orgogozo, I. Delon, J. Zanet, D. G. Srinivasan, F. Payre, and D. L. Stern. 2007. Morphological evolution through multiple cis-regulatory mutations at a single gene. Nature 448:587-U586.

Mery, F., A. T. Belay, A. K. C. So, M. B. Sokolowski, and T. J. Kawecki. 2007. Natural polymorphism affecting learning and memory in *Drosophila*. Proc. Nat. Acad. Sci. USA 104:13051–13055.

Moran, N. A., and T. Jarvik. 2010. Lateral transfer of genes from fungi underlies carotenoid production in aphids. Science 328:624–627.

Nagai, H., Y. Terai, T. Sugawara, H. Imai, H. Nishihara, M. Hori, and N. Okada. 2011. Reverse evolution in RH1 for adaptation of cichlids to water depth in lake Tanganyika. Mol. Biol. Evol. 28:1769–1776.

Nei, M. 2007. The new mutation theory of phenotypic evolution. Proc. Nat. Acad. Sci. USA 104:12235–12242.

Newman, R. M., L. Hall, M. Connole, G. L. Chen, S. Sato, E. Yuste, W. Diehl, E. Hunter, A. Kaur, and G. M. Miller. 2006. Balancing selection and the evolution of functional polymorphism in Old World monkey TRIM5alpha. Proc. Nat. Acad. Sci. USA 103:19134–19139.

O'Day, W. T., and H. R. Fernandez. 1974. *Aristostomias scintillans* (Malacosteidae): a deep-sea fish with visual pigments apparently adapted to its own bioluminescence. Vis. Res. 14:545–550.

Orr, H. A. 2005. The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6:119–127.

Osborne, K. A., A. Robichon, E. Burgess, S. Butland, R. A. Shaw, A. Coulthard, H. S. Pereira, R. J. Greenspan, and M. B. Sokolowski. 1997. Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. Science 277:834–836.

Papa, R., A. Martin, and R. D. Reed. 2008. Genomic hotspots of adaptation in butterfly wing pattern evolution. Curr. Opin. Genet. Dev. 18:559–564.

Pardo-Diaz, C., C. Salazar, S. W. Baxter, C. Merot, W. Figueiredo-Ready, M. Joron, W. O. McMillan, and C. D. Jiggins. 2012. Adaptive introgression across species boundaries in heliconius butterflies. PLoS Genet. 8:e1002752.

Parreiras, L. S., L. M. Kohn, and J. B. Anderson. 2011. Cellular effects and epistasis among three determinants of adaptation in experimental populations of *Saccharomyces cerevisiae*. Eukaryot Cell 10:1348–1356.

Pearce, T. 2012. Convergence and parallelism in evolution: a neo-gouldian account. Brit. J. Phil. Sci. 63:429–448.

Pennings, P. S., and J. Hermisson. 2006. Soft sweeps III: the signature of positive selection from recurrent mutation. PLoS Genet. 2:e186.

Perry, G. H., N. J. Dominy, K. G. Claw, A. S. Lee, H. Fiegler, R. Redon, J. Werner, F. A. Villanea, J. L. Mountain, and R. Misra. 2007. Diet and the evolution of human amylase gene copy number variation. Nat. Genet. 39:1256–1260.

Perutz, M. F. 1983. Species adaptation in a protein molecule. Mol. Biol. Evol. 1:1–28.

Piffanelli, P., L. Ramsay, R. Waugh, A. Benabdelmouna, A. D'Hont, K. Hollricher, J. H. Jorgensen, P. Schulze-Lefert, and R. Panstruga. 2004. A barley cultivation-associated polymorphism conveys resistance to powdery mildew. Nature 430:887–891.

Porter, M. L., and K. A. Crandall. 2003. Lost along the way: the significance of evolution in reverse. Trends Ecol. Evol. 18:541–547.

Prabhakar, S., A. Visel, J. A. Akiyama, M. Shoukry, K. D. Lewis, A. Holt, I. Plajzer-Frick, H. Morrison, D. R. FitzPatrick, V. Afzal, et al. 2008. Human-specific gain of function in a developmental enhancer. Science 321:1346–1350.

Protas, M. E., C. Hersey, D. Kochanek, Y. Zhou, H. Wilkens, W. R. Jeffery, L. I. Zon, R. Borowsky, and C. J. Tabin. 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nat. Genet. 38:107–111.

Protas, M. E., P. Trontelj, and N. H. Patel. 2011. Genetic basis of eye and pigment loss in the cave crustacean, *Asellus aquaticus*. Proc. Nat. Acad. Sci. USA 108:5702–5707.

Rebeiz, M., J. E. Pool, V. A. Kassner, C. F. Aquadro, and S. B. Carroll. 2009a. Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. Science 326:1663–1667.

Rebeiz, M., M. Ramos-Womack, S. Jeong, P. Andolfatto, T. Werner, J. True, D. L. Stern, and S. B. Carroll. 2009b. Evolution of the tan locus contributed to pigment loss in *Drosophila santomea*: a response to Matute et al. Cell 139:1189–1196.

Rebeiz, M., N. Jikomes, V. A. Kassner, and S. B. Carroll. 2011. Evolutionary origin of a novel gene expression pattern through co-option of the latent activities of existing regulatory sequences. Proc. Nat. Acad. Sci. USA 108:10036–10043.

Reed, R. D., R. Papa, A. Martin, H. M. Hines, B. A. Counterman, C. Pardo-Diaz, C. D. Jiggins, N. L. Chamberlain, M. R. Kronforst, R. Chen, et al. 2011. optix drives the repeated convergent evolution of butterfly wing pattern mimicry. Science 333:1137–1141.

Richardson, M. K., and P. M. Brakefield. 2003. Developmental biology: hotspots for evolution. Nature 424:894–895.

Roberts, R. B., Y. Hu, R. C. Albertson, and T. D. Kocher. 2011. Craniofacial divergence and ongoing adaptation via the hedgehog pathway. Proc. Nat. Acad. Sci. USA 108:13194–13199.

Rockman, M. V. 2011. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. Evolution 66:1–17.

Roelants, K., B. G. Fry, J. A. Norman, E. Clynen, L. Schoofs, and F. Bossuyt. 2010. Identical skin toxins by convergent molecular adaptation in frogs. Curr. Biol. 20:125–130.

Rohner, N., M. Bercsenyi, L. Orban, M. E. Kolanczyk, D. Linke, M. Brand, C. Nusslein-Volhard, and M. P. Harris. 2009. Duplication of fgfr1 permits Fgf signaling to serve as a target for selection during domestication. Curr. Biol. 19:1642–1647.

Ros, V. I., and G. D. Hurst. 2009. Lateral gene transfer between prokaryotes and multicellular eukaryotes: ongoing and significant? BMC Biol. 7:20.

Rose, L. E., R. W. Michelmore, and C. H. Langley. 2007. Natural variation in the Pto disease resistance gene within species of wild tomato (Lycopersicon). II. Population genetics of Pto. Genetics 175:1307–1319.

Rumpho, M. E., J. M. Worful, J. Lee, K. Kannan, M. S. Tyler, D. Bhattacharya, A. Moustafa, and J. R. Manhart. 2008. Horizontal gene transfer of the algal nuclear gene psbO to the photosynthetic sea slug *Elysia chlorotica*. Proc. Nat. Acad. Sci. USA 105:17867–17871.

Rus, A., I. Baxter, B. Muthukumar, J. Gustin, B. Lahner, E. Yakubova, and D. E. Salt. 2006. Natural variants of At*HKT1* enhance Na+ accumulation in two wild populations of *Arabidopsis*. PLoS Genet. 2:e210.

Scally, A., J. Y. Dutheil, L. W. Hillier, G. E. Jordan, I. Goodhead, J. Herrero, A. Hobolth, T. Lappalainen, T. Mailund, T. Marques-Bonet, et al. 2012. Insights into hominid evolution from the gorilla genome sequence. Nature 483:169–175.

Schierup, M. H., A. M. Mikkelsen, and J. Hein. 2001. Recombination, balancing selection and phylogenies in MHC and self-incompatibility genes. Genetics 159:1833–1844.

Scotland, R. W. 2011. What is parallelism? Evol. Dev. 13:214–227.

Segurel, L., E. E. Thompson, T. Flutre, J. Lovstad, A. Venkat, S. W. Margulis, J. Moyse, S. Ross, K. Gamble, and G. Sella. 2012. The ABO blood group is a trans-species polymorphism in primates. Proc. Nat. Acad. Sci. USA 109:18493–18498.

Shahzad, Z., F. Gosti, H. Frerot, E. Lacombe, N. Roosens, P. Saumitou-Laprade, and P. Berthomieu. 2012. The five AhMTP1 zinc transporters undergo different evolutionary fates towards adaptive evolution to zinc tolerance in *Arabidopsis halleri*. PLoS Genet. 6:e1000911.

Shapiro, M. D., B. R. Summers, S. Balabhadra, J. T. Aldenhoven, A. L. Miller, C. B. Cunningham, M. A. Bell, and D. M. Kingsley. 2009. The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks. Curr. Biol. 19:1140–1145.

Shen, Y.-Y., L. Liang, G.-S. Li, R. W. Murphy, and Y.-P. Zhang. 2012. Parallel evolution of auditory genes for echolocation in bats and toothed whales. PLoS Genet. 8:e1002788.

Shirangi, T. R., H. D. Dufour, T. M. Williams, and S. B. Carroll. 2009. Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. PLoS Biol. 7:e1000168.

Shull, A. F. 1935. Weismann and Haeckel: one hundred years. Science 81: 443–452.

Song, Y., S. Endepols, N. Klemann, D. Richter, F.-R. Matuschka, C.-H. Shih, M. W. Nachman, and M. H. Kohn. 2011. Adaptive Introgression of Anticoagulant Rodent Poison Resistance by Hybridization between Old World mice. Curr. Biol. 21:1296–1301.

Srikanth, A., and M. Schmid. 2011. Regulation of flowering time: all roads lead to Rome. Cell. Mol. Life Sci. 68:2013–2037.

Stahl, E. A., G. Dwyer, R. Mauricio, M. Kreitman, and J. Bergelson. 1999. Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. Nature 400:667–671.

Stam, L. F., and C. C. Laurie. 1996. Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. Genetics 144:1559–1564.

Stern, D. 2000. Evolutionary developmental biology and the problem of variation. Evolution 54:1079–1091.

Stern, D. L. 2010. Evolution, development, and the predictable genome. Roberts & Company, Greenwood Village, CO.

Stern, D. L., and V. Orgogozo. 2008. The loci of evolution: how predictable is genetic evolution? Evolution 62:2155–2177.

———. 2009. Is genetic evolution predictable? Science 323:746–751.

Streisfeld, M. A., and M. D. Rausher. 2011. Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. Evolution 65:629–642.

Sucena, E., I. Delon, I. Jones, F. Payre, and D. L. Stern. 2003. Regulatory evolution of shavenbaby/ovo underlies multiple cases of morphological parallelism. Nature 424:935–938.

Syvanen, M. 2012. Evolutionary implications of horizontal gene transfer. Annu. Rev. Genet. 46:341–358.

Takahashi, Y., and S. Yokoyama. 2005. Genetic basis of spectral tuning in the violet-sensitive visual pigment of African clawed frog, *Xenopus laevis*. Genetics 171:1153–1160.

Tao, H., D. R. Cox, and K. A. Frazer. 2006. Allele-specific KRT1 expression is a complex trait. PLoS Genet. 2:e93.

Tenaillon, O., A. Rodriguez-Verdugo, R. L. Gaut, P. McDonald, A. F. Bennett, A. D. Long, and B. S. Gaut. 2012. The molecular diversity of adaptive convergence. Science 335:457–461.

The Heliconius Genome Consortium. 2012. A butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487:94–98.

Theron, E., K. Hawkins, E. Bermingham, R. E. Ricklefs, and N. I. Mundy. 2001. The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, *Coereba flaveola*. Curr. Biol. 11:550–557.

Timmann, C., T. Thye, M. Vens, J. Evans, J. May, C. Ehmen, J. Sievertsen, B. Muntau, G. Ruge, and W. Loag. 2012. Genome-wide association study indicates two novel resistance loci for severe malaria. Nature 489:443–446.

Tishkoff, S. A., R. Varkonyi, N. Cahinhinan, S. Abbes, G. Argyropoulos, G. Destro-Bisol, A. Drousiotou, B. Dangerfield, G. Lefranc, J. Loiselet et al. 2001. Haplotype diversity and linkage disequilibrium at human

G6PD: recent origin of alleles that confer malarial resistance. Science 293:455–462.

Tishkoff, S. A., F. A. Reed, A. Ranciaro, B. F. Voight, C. C. Babbitt, J. S. Silverman, K. Powell, H. M. Mortensen, J. B. Hirbo, and M. Osman. 2006. Convergent adaptation of human lactase persistence in Africa and Europe. Nat. Genet. 39:31–40.

Toprak, E., A. Veres, J. B. Michel, R. Chait, D. L. Hartl, and R. Kishony. 2011. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat. Genet. 44:101–105.

Turchetto-Zolet, A., F. Maraschin, G. de Morais, A. Cagliari, C. Andrade, M. Margis-Pinheiro, and R. Margis. 2011. Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. BMC Evol. Biol. 11:263.

Ueno, D., M. J. Milner, N. Yamaji, K. Yokosho, E. Koyama, M. Clemencia Zambrano, M. Kaskie, S. Ebbs, L. V. Kochian, and J. F. Ma. 2011. Elevated expression of *TcHMA3* plays a key role in the extreme Cd tolerance in a Cd-hyperaccumulating ecotype of *Thlaspi caerulescens*. Plant J. 66:852–862.

Vinci, G., X. Xia, and R. A. Veitia. 2008. Preservation of genes involved in sterol metabolism in cholesterol auxotrophs: facts and hypotheses. PLoS One 3:e2883.

Wake, D. B., M. H. Wake, and C. D. Specht. 2011. Homoplasy: from detecting pattern to determining process and mechanism of evolution. Science 331:1032–1035.

Wang, X. D., and H. M. Chamberlin. 2004. Evolutionary innovation of the excretory system in *Caenorhabditis elegans*. Nat. Genet. 36:231–232.

Wittkopp, P. J., and G. Kalay. 2011. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat. Rev. Genet. 13:59–69.

Wittkopp, P. J., E. E. Stewart, L. L. Arnold, A. H. Neidert, B. K. Haerum, E. M. Thompson, S. Akhras, G. Smith-Winberry, and L. Shefner. 2009. Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in *Drosophila*. Science 326:540–544.

Wood, T., J. Burke, and L. Rieseberg. 2005. Parallel genotypic adaptation: when evolution repeats itself. Genetica:157–170.

Wu, J., S. J. Saupe, and N. L. Glass. 1998. Evidence for balancing selection operating at the het-c heterokaryon incompatibility locus in a group of filamentous fungi. Proc. Nat. Acad. Sci. USA 95:12398–12403.

Yokoyama, S., T. Tada, H. Zhang, and L. Britt. 2008. Elucidation of phenotypic adaptations: molecular analyses of dim-light vision proteins in vertebrates. Proc. Nat. Acad. Sci. USA 105:13480–13485.

Zhang, Z., P. M. Harrison, Y. Liu, and M. Gerstein. 2003. Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. Genome Res. 13:2541–2558.

Zheng, P., W. B. Allen, K. Roesler, M. E. Williams, S. Zhang, J. Li, K. Glassman, J. Ranch, D. Nubel, W. Solawetz, et al. 2008. A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat. Genet. 40:367–372.

Zufall, R. A., and M. D. Rausher. 2004. Genetic changes associated with floral adaptation restrict future evolutionary potential. Nature 428:847–850.

Zuk, O., E. Hechter, S. R. Sunyaev, and E. S. Lander. 2012. The mystery of missing heritability: genetic interactions create phantom heritability. Proc. Nat. Acad. Sci. USA 109:1193–1198.

Associate Editor: A. Kopp

Supporting Information

Additional Supporting information may be found in the online version of this article at the publisher's website:

Figure S1. Progress in the discovery of loci of evolution by linkage mapping.

Table S1. Number of loci of evolution according to their taxon level.

Table S2. Codon-level hotspots of evolution of resistance to xenobiotics.

Table S3. Hotspots of variation in domesticated plants.

Table S4. Hotspots of variation in domesticated animals.

Table S5. Hotspots of evolution linked to coding loss-of-function mutations.