Cell PRESS

Causes and evolutionary significance of genetic convergence

Pascal-Antoine Christin[1], Daniel M. Weinreich[1] and Guillaume Besnard[2]

[1] Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02906, USA
[2] Imperial College Silwood Park Campus, Buckhurst Road Ascot, Berkshire SL5 7PY, UK

Convergent phenotypes provide extremely valuable systems for studying the genetics of new adaptations. Accumulating studies on this topic have reported surprising cases of convergent evolution at the molecular level, ranging from gene families being recurrently recruited to identical amino acid replacements in distant lineages. Together, these different examples of genetic convergence suggest that molecular evolution is in some cases strongly constrained by a combination of limited genetic material suitable for new functions and a restricted number of substitutions that can confer specific enzymatic properties. We discuss approaches for gaining further insights into the causes of genetic convergence and their potential contribution to our understanding of how the genetic background determines the evolvability of complex organismal traits.

Evolutionary convergence provides outstanding study systems

During the billions of years of evolution, similar selective pressures have occasionally led to the independent evolution of identical or similar traits in distantly related species, a phenomenon referred to as phenotypic convergence [1,2]. The recent wide use of genetic and/or phylogenetic approaches has uncovered diverse examples of repeated evolution of adaptive traits including the multiple appearances of eyes [3,4], echolocation in bats and dolphins [5,6], pigmentation modifications in vertebrates [7–10], mimicry in butterflies for mutualistic interactions [11], convergence of some flower traits in plants [12–15], and multiple independent evolution of particular protein properties [16,17]. The multiple origins of a trait represent exceptional replicates of evolutionary processes and can provide extremely valuable insights into the constraints and opportunities that govern evolution. In particular, comparing the genetic determinants of the independent origins of an adaptive phenotype can shed new light on the role of genomic background in restricting or opening new evolutionary trajectories towards adaptive innovations [18–22]. In this paper we discuss the potential causes of convergence at the genetic level together with their implications for our understanding of evolutionary biology in general.

When phenotypic convergence is caused by mutations in the same gene

In the numerous reports of phenotypic convergence the responsible genetic mechanisms remain largely unknown because their identification is often complicated by the involvement of complex biochemical cascades as well as epistatic interactions [19,23,24]. In some cases it has been shown that different loci are involved in phenotypic convergence (e.g. Refs [8,25,26]), demonstrating that similar phenotypes can be reached through alterations of distinct enzymes. However, other studies have traced phenotypic convergence to modifications of homologous genes (e.g. Refs [3,5,6,26,27]); in this paper such phenomena will be further referred to as convergent recruitment (Glossary).

The independent involvement of homologous genes in the emergence of a given phenotype probably results from strongly biased potential for a given phenotypic change as a consequence of mutations in different genes [28,29]. In cases where the new phenotype repeatedly occurs through a loss of enzymatic function, such as albinism or the absence of specific pigments [7,12], alterations of genes encoding elements involved in the biochemical cascade that cause the trait of interest are more likely to lead to the new phenotype. Silencing mutations also have a higher probability of being fixed when they occur in genes that can block the entire biochemical cascade without major deleterious pleiotropic effects on the organism. Therefore, genes involved in multiple functions are poor candidates for phenotype loss through gene silencing. However, repeated *cis*-regulatory changes involved in the recurrent loss (or gain) of organ-specific gene expression have been reported [30]. Such modifications allow silencing

Glossary

Convergence: independent appearance of the same trait in different lineages.
Convergent recruitment: the process of homologous gene becoming recurrently responsible for a novel function.
Convergent substitution: replacement of the same ancestral character (e.g. amino acid) by an identical character.
Epistatic interaction: influence of one gene on the expression of another gene.
Gene family: a group of homologous genes which are generally responsible for similar catalytic reactions. Multigene families contain several gene lineages, and these usually fulfill different functions.
Gene lineage: a gene family that arose via whole genome or gene duplication. Genes of the same lineage are orthologous, but more recent gene duplications can hamper the definition of orthology.
Homology: the relationship between genes that share a common ancestor. This includes orthologs as well as paralogs. This term is restricted to genes whose relationship can be deduced from sequence similarity.
Orthologs: genes in different species whose divergence is due to speciation.
Paralogs: genes whose divergence is due to single gene or whole genome duplication.
Phenotype: the observable characteristic that results from the expression of genes with the possibility of additional environmental effects. Phenotypes include organism traits as well as all measurable properties of enzymes.
Pleiotropic effect: the action of a single gene on apparently unrelated phenotypic traits.

Corresponding author: Christin, P.-A. (Pascal-Antoine_Christin@brown.edu)

0168-9525/$ – see front matter doi:10.1016/j.tig.2010.06.005 Trends in Genetics 26 (2010) 400–405

(or activation) of a specific gene function without pleiotropic effects.

The suitability of genes for new phenotypes that result in the addition or transformation of a pre-existing biochemical pathway is dictated by more subtle requirements. Novel genes can be created *de novo* from non-coding DNA [31,32], but new functions generally evolve through modifications of pre-existing genes. In this case, genes must meet two broad criteria to be eligible for a novel function: (i) they must have the possibility of being recruited for a new task without deleterious effect due to the loss or modification of the ancestral function, and (ii) their expression profiles and kinetics must be suitable for the new task.

If the new function can be acquired without altering its ancestral role (e.g. Refs [30,33]), the gene perfectly meets the first criterion. This is also true when genes become involved in a new function that replaces the ancestral role, for instance when the acquisition of new biochemical characteristics is a response to a switch of selection pressures after an environmental change or when the ancestral task became obsolete. The recruitment of genes can also be favored by genetic redundancy, which can result from whole genome or single gene duplications. One of the gene copies can acquire a new task while the other copy remains responsible for the ancestral function [34,35].

Regarding the second criterion, the pool of candidates for a new function is likely to be limited to genes encoding enzymes with compatible catalytic properties or to genes that can acquire them via successive substitutions without strongly deleterious transitional stages. The gene must also confer expression profiles that match the new function or be able to acquire the required expression patterns through few key mutations. In the case of C_4 photosynthesis, C_4-specific NADP-malic enzymes (NADP-me) evolved at least five times independently through modifications of NADP-me initially involved in non-photosynthetic functions [36]. The NADP-me gene family encompasses four lineages, but only one (*nadpme-IV*) was recruited in these five origins. Unlike the other NADP-me genes, *nadpme-IV* displays a transit peptide for expression in chloroplasts. This characteristic is essential for the C_4 pathway and probably accounts for the convergent recruitment of *nadpme-IV* for this new function.

Evolutionary significance of convergent recruitment

Convergent recruitment indicates that genes suitable for creating a given phenotype are rare [20,28,29]. Whereas the absence of appropriate genes in some lineages can hamper the acquisition of specific phenotypes, the presence of genes that are able to acquire a given function can enhance the probability that a given group of organisms evolves a new trait. During evolution coding sequences have reached different areas of protein space (Box 1) through the accumulation of amino acid replacements. Some genes have probably reached regions of the protein space where they became suitable for the emergence of a new phenotype through exaptation (i.e. a protein previously shaped for a specific function is coopted for a new trait). Testing these hypotheses about genetic evolvability and convergent recruitment might shed new lights on the genetic constraints and their impact on phenotype evolvability.

Box 1. Fitness landscapes over sequence space

Both proteins and DNA have an intrinsically digital nature – they consist of a finite number of sites, each of which can assume only a finite number of states. There is thus a finite (albeit very large) number of conceivable genes of a given length. John Maynard Smith [49] was perhaps the first to explicitly propose consideration of a molecular sequence space (but see Ref. [50] for the allelic analog) in which mutationally adjacent haploid genomes are also spatially adjacent. Can we model an evolving population moving through this high-dimensional space?

When mutation and recombination rates are high, populations occupy a cloud of points in sequence space (see for example the quasispecies literature; e.g. Ref. [51]). To simplify, many authors adopt the strong selection/weak mutation (SSWM) approximation [52]. In this regime the selective fate of each mutation (fixation or loss) will almost surely be resolved before the next mutation occurs. Thus under the SSWM approximation we can regard an evolving population as following a succession of spatially adjacent (or at least nearly so [53]) points in sequence space. Can we predict which such trajectory an evolving population is likely to follow?

Sewall Wright was the first to propose projecting fitness values for all genetic variants over sequence space [50]. This mapping from each point in sequence space to reproductive success of the corresponding genotype is commonly referred to as the fitness landscape. Subject to stochastic population genetic effects, an evolving population will 'climb' the steepest gradient to the nearest fitness peak on this landscape, and given data on this mapping one can compute the probability that an evolving population will follow each conceivable trajectory through sequence space to that point (e.g. Ref. [44]).

Note that although several authors have questioned the formal coherence of the fitness landscape (e.g. Refs [54–56]) we believe that little confusion exists when the landscape is defined as here. It is however important to acknowledge that, in addition to the SSWM assumptions, this predictive framework assumes constant haploid fitness values (i.e. invariant environment, no frequency-dependent selection, no dominance), and also disregards insertion/deletions, inversions and other gross mutational processes not readily represented in sequence space [23]. Finally, note that important theoretical [52,57,58] and empirical [59–62] questions remain regarding the evolutionary consequences of multi-peaked fitness landscapes.

Convergent recruitment of the same gene lineage from multigene families affords an ideal system for studying the predisposition of particular genes for a given novel function. Two types of phenotypic similarity can increase the likelihood of a gene lineage becoming involved in a new function: catalytic function and tissue expression pattern. We predict that, compared to the other members of the gene family, a recurrently recruited gene lineage will generally have a catalytic activity and an expression pattern closer to those needed for the novel reaction compared to other members of the same gene family. Both these predictions can be experimentally tested. We suggest that each paralog, identified through phylogenetics as being present in the ancestral species, might be assayed for its kinetic activity against the novel substrate or in the new reaction *in vitro*. Similarly, temporal and spatial expression patterns of each paralog can be assessed using quantitative PCR or similar technologies.

When phenotypic convergence is caused by identical substitutions

In addition to convergent recruitment, several studies have traced phenotypic convergence to identical genetic substitutions in different lineages [5,6,37–39]. Whereas

convergent substitutions can theoretically occur in coding and non-coding regions, most reports concern replacement of the same protein residue by an identical amino acid in independent lineages. After careful statistical consideration [37,40] or experimental demonstration [27,41], the adaptive value of repeated substitutions can be established. For instance, various combinations of the same five amino acid replacements on visual pigments (opsins) were responsible for independent changes of color vision in vertebrates, through similar shifts in the light absorption maximum of the encoded proteins [27]. Similarly, identical amino acid replacements at three sites within digestive RNases were shown to be driven by adaptive evolution in Asian and African leaf monkeys, and these independently adapted the enzyme to ruminant-like alimentary systems by lowering its optimal pH [41–43]. Amino acid replacements are more likely to occur by chance when caused by single nucleotide substitutions and those that confer a higher fitness will be preferentially fixed. Transitions to proteins with better suited kinetics but which involve transitional stages with lower fitness are less likely to take place, and evolutionary paths that lead to optimized enzymes via successively advantageous single nucleotide substitutions will be more frequently followed [44]. Thus, the observed adaptive convergent substitutions probably result from strong biases in the likelihood of different replacements in a limited number of proteins suitable for the emergence of the convergent phenotype [18].

Adaptive convergent substitutions are expected to arise when the mutated genes are similar because genetic similarity implies that the probability and effects of amino acid replacements are more likely to be comparable, increasing the likelihood of fixation of identical amino acid substitutions. Thus, the more similar the genes are the more likely convergent adaptive substitutions will be. Pesticide resistance was acquired several times through the same

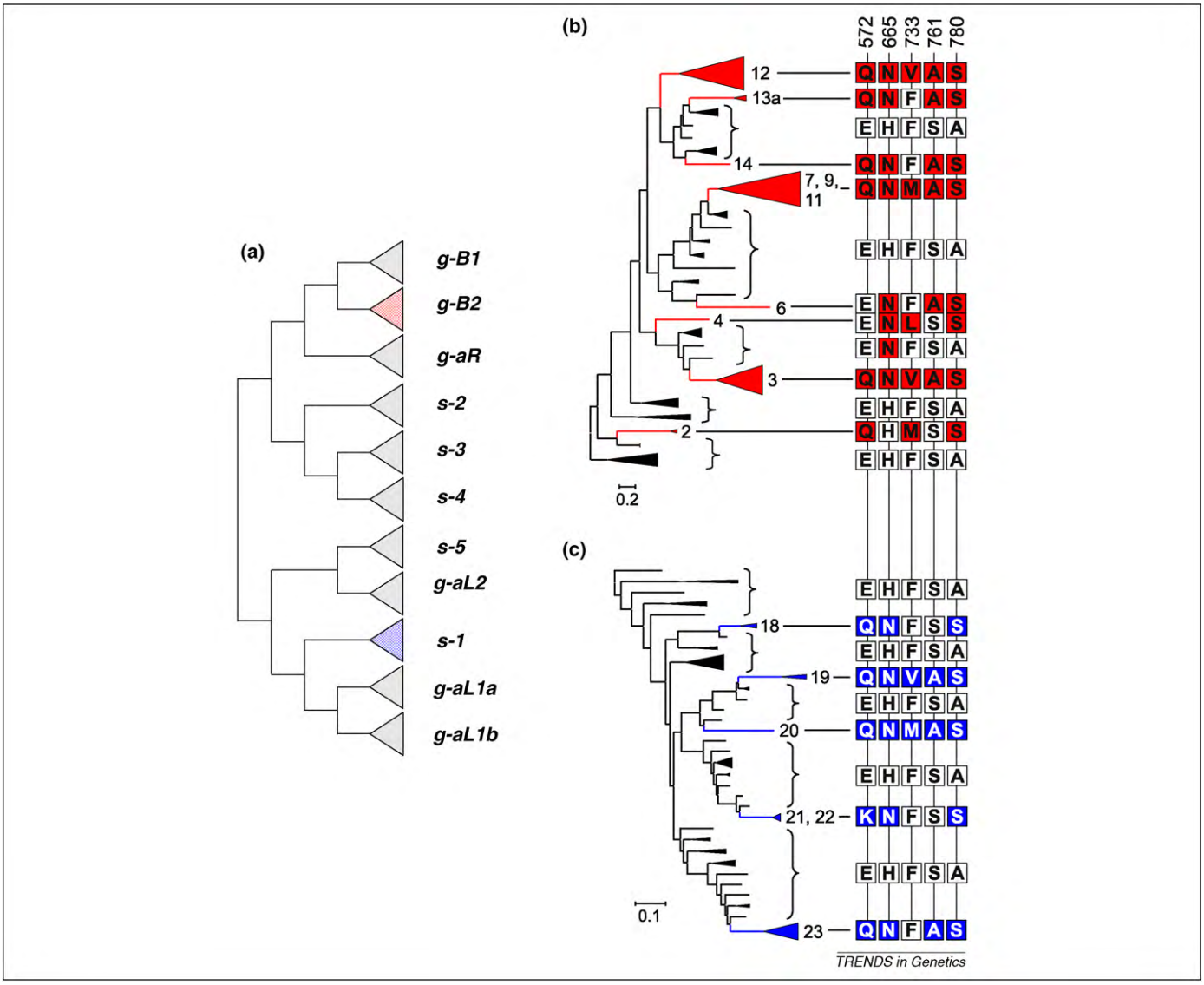

Figure 1. Genetic convergence of C_4 PEPC. **(a)** Phylogenetic tree illustrating the evolutionary relationships between the different PEPC gene lineages (*ppc*) of grasses (g) and sedges (s), following Refs [47,48]. Each gene lineage is named according to Ref. [48]. The gene lineage convergently recruited for C_4 photosynthesis is hatched in red for grasses and in blue for sedges. **(b)** Detail of the phylogenetic tree of the grass gene lineage *g-B1*, according to Ref. [47]. **(c)** Detail of the phylogenetic tree of the sedge gene lineage *s-1*, according to Ref. [48]. C_4-specific genes are in red for grasses and blue for sedges. C_4 lineages are numbered according to Ref. [39]. Scale bars represent expected substitutions per site. On the right, most frequent amino acids at positions shown to be under C_4-specific positive selection in both grasses and sedges are indicated. Putative C_4-adaptive residues are highlighted in red for grasses and blue for sedges.

amino acid replacement in a given gene, either within the same species or in taxa of the same family [45,46], and the possibility of reaching a new adaptive phenotype through a single substitution strongly favored this evolutionary endpoint. The pancreatic digestive RNases of some Asian and African monkeys evolved new kinetics through convergent adaptive substitutions [42,43]. However, the distantly related RNases of ruminants evolved similar kinetics but via divergent genetic changes [41]. Furthermore, it was shown that the amino acid replacements that adapted leaf monkey RNases would have opposite effects on the ruminant enzymes [41]. This suggests that the repeatability of genetic adaptive changes can depend on the evolutionary distance separating the organisms with convergent phenotypes and thus on the genetic background, although some evidence demonstrates that recurrent recruitment is not perfectly predicted by the divergence of the organisms [19]. For instance, changes in coloration have been caused by convergent recruitment in species ranging from reptiles to very diverse mammals [19,21], but two populations of the same mouse species acquired pale pelages by mutations of different genes [8]. Thus, the sequence divergence of recruited genes rather than the phylogenetic distance between organisms must be considered when studying the genetics of phenotypic convergence, although these are often correlated. The repeated evolution of phosphoenolpyruvate carboxylase (PEPC) optimized for C_4 photosynthesis in grasses recruited the same gene lineage (out of a total of six gene duplicates) at least eight times and involved similar or identical amino acid changes at a relatively high proportion of sites (21/450; ~5%) [47]. Another PEPC gene lineage was independently recruited five times in C_4 sedges (Figure 1) and, similarly, C_4 optimized characteristics were reached through repeated mutations at 16 amino acid sites [48]. Interestingly, only five of these amino acid replacements were shared between grasses and sedges (Figure 1). This suggests that different starting points can open new evolutionary paths [41]. Nevertheless, some changes were necessary whatever gene lineage is recruited, as shown by the same five amino acid replacements involved in multiple evolution of C_4 PEPC from genes that diverged more than 120 million years ago [48].

In summary, the biased probability of different replacements leads in some cases to adaptive convergent substitutions, and the extent of this phenomenon depends on the similarity of the recruited genes. However, convergent substitutions represent only a small portion of the adaptive genetic changes putatively involved in new phenotypes and are often accompanied by non-convergent substitutions [37,43]. The importance of convergent versus divergent amino acid replacements in convergent phenotype switches will probably differ greatly between genes. Its quantification would give strong insights into the role of the genetic background in determining the novel enzymatic functions that can emerge under natural selection.

Concluding remarks

Whereas convergent recruitment suggest that only a few genes have the potential to create a specific phenotypic change, the occurrence of convergent adaptive substitutions at diverse taxonomic scales tells us that some substitutions are more likely to be involved in the emergence of a novel adaptation. Studies of natural selection at the genetic level can benefit greatly from the information provided by convergent adaptive substitutions because these provide naturally occurring genetic points under a shared selection pressure (Box 2). Reciprocally, assessing the fitness landscapes of different properties of the encoded proteins is required to reveal the causes of convergent adaptive substitutions. Comparison of the areas of the fitness landscape crossed during the independent emergence of a convergent phenotype would help to estimate the constraints operating on evolutionary transitions at the genetic level. The possibility of considering convergent modifications in similar as well as distantly related genes would moreover shed light on the importance of the ancestral genetic architecture in opening roads to different evolutionary optima.

Box 2. Convergent adaptive substitutions and the functional synthesis

Since the 1960s the study of natural selection at the molecular level has been largely confined to the statistical analysis of within- and between-species nucleotide variation (but see Ref. [63] for a notable exception). In this research program, patterns of genetic variation are commonly compared to expectations under the null model of selective neutrality (e.g. Ref. [64]), and many examples of genes under Darwinian selection have now been observed (e.g. Refs [37,38,65,66]). However the statistical nature of this work leaves us largely ignorant of the mechanistic determinants of natural selection [67]. Recently, a number of workers have begun to address this gap in understanding at the level of individual protein-coding genes. First, given particular evolutionary starting and ending alleles, many or all mutational intermediates are constructed using methods of reverse genetics. Next, each allele's contributions to whole-organism fitness are assessed. Together these data provide a picture of the fitness landscape (Box 1) over which evolution between endpoints must travel (e.g. Refs [44,68–71]). Moreover, at the level of protein-coding genes individual mutational effects on organismal fitness must be mediated by their effect on mechanistically more proximal traits such as protein-folding stability, substrate-binding affinity, and aggregation and degradation potential [72]. Thus, by combining well-established techniques from protein biology, biochemistry and biophysics it is becoming possible to dissect the exact mechanistic determinants of natural selection at this level of organization (e.g. Ref. [68]). This research program has been termed the functional synthesis [67].

One important limitation thus far in work of this sort is the explicit assumption of a particular genetic endpoint of evolution, whereas natural selection is concerned with phenotypic endpoints. Thus, characterization of the topography and mechanistic determinants of the fitness landscape as described might not fully explain the molecular basis of natural selection if evolving populations discover alternative mutations that yield the favored phenotype. Applications of this experimental program to cases of convergent adaptive substitution offer a unique opportunity to overcome this limitation because the repeated observation across lineages of the same substitutions under positive selection strongly suggests that there are only a few alternative genetic endpoints that can yield the required phenotype.

Acknowledgements

Funding for this work was provided by the Swiss National Science Foundation grant PBLAP3-129423 to P.A.C. and the Intra-European fellowship PIEF-GA-2008-220813 to G.B. The authors thank Kaya Schmandt and three anonymous reviewers for helpful comments on earlier versions of the manuscript.

References

1 Wake, D.B. (1991) Homoplasy: the result of natural selection or evidence of design limitations? *Am. Nat.* 138, 543–567
2 Conway Morris, S. (2003) *Life's solution: Inevitable Humans in a Lonely Universe*, Cambridge University Press
3 Fernald, R.D. (2006) Casting a genetic light on the evolution of eyes. *Science* 313, 1914–1918
4 Kozmik, Z. (2008) Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8989–8993
5 Liu, Y. *et al.* (2010) Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* 20, R53–R54
6 Li, Y. *et al.* (2010) The hearing gene Prestin unites echolocating bats and whales. *Curr. Biol.* 20, R55–R56
7 Protas, M.E. *et al.* (2006) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat. Genet.* 38, 107–111
8 Steiner, C.C. *et al.* (2009) The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. *Mol. Biol. Evol.* 26, 35–45
9 Rosenblum, E.B. *et al.* (2010) Molecular and functional basis of phenotypic convergence in white lizards at white sands. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2113–2117
10 Hubbard, J.K. *et al.* (2010) Vertebrate pigmentation: from underlying genes to adaptive function. *Trends Genet.* 26, 231–239
11 Baxter, S.W. *et al.* (2008) Convergent evolution in the genetic basis of Mullerian mimicry in *Heliconius* butterflies. *Genetics* 180, 1567–1577
12 Whittall, J.B. *et al.* (2006) Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in *Aquilegia*. *Mol. Ecol.* 15, 4645–4657
13 Hoballah, M.E. *et al.* (2007) Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* 19, 779–790
14 Preston, J.C. and Hileman, L.C. (2009) Developmental genetics of floral symmetry evolution. *Trends Plant Sci.* 14, 147–154
15 Rosin, F.M. and Kramer, E.M. (2009) Old dogs, new tricks: regulatory evolution in conserved genetic modules leads to novel morphologies in plants. *Dev. Biol.* 332, 25–36
16 Doolittle, R.F. (1994) Convergent evolution: the need to be explicit. *Trends Biochem. Sci.* 19, 15–18
17 Gherardini, P.F. *et al.* (2007) Convergent evolution of enzyme active sites is not a rare phenomenon. *J. Mol. Biol.* 372, 817–845
18 Hodin, J. (2000) Plasticity and constraints in development and evolution. *J. Exp. Zool.* 288, 1–20
19 Arendt, J. and Reznick, D. (2008) Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* 23, 26–32
20 Kalisz, S. and Kramer, E.M. (2008) Variation and constraints in plant evolution and development. *Heredity* 100, 171–177
21 Gompel, N. and Prud'homme, B. (2009) The causes of repeated genetic evolution. *Dev. Biol.* 332, 36–47
22 Stern, D.L. and Orgogozo, V. (2009) Is genetic evolution predictable? *Science* 323, 746–751
23 Weinreich, D.M. *et al.* (2005) Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59, 1165–1174
24 Poelwijk, F.J. *et al.* (2007) Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* 445, 383–386
25 Wittkopp, P.J. *et al.* (2003) *Drosophila* pigmentation evolution: divergent genotypes underlying convergent phenotypes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1808–1813
26 Fry, B.G. *et al.* (2009) The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu. Rev. Genomics Hum. Genet.* 10, 483–511
27 Yokoyama, S. and Radlwimmer, F.B. (2001) The molecular genetics and evolution of red and green color vision in vertebrates. *Genetics* 158, 1697–1710
28 Woods, R.D. *et al.* (2006) Test of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 9107–9112
29 Chevin, L.M. *et al.* (2010) Fisher's model and the genomics of adaptation: restricted pleiotropy, heterogenous mutation and parallel evolution. *Evolution* 64, DOI: 10.1111/j.1558-5646.2010.01058.x
30 Prud'homme, B. *et al.* (2006) Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. *Nature* 440, 1050–1053
31 Zhou, Q. *et al.* (2008) On the origin of new genes in *Drosophila*. *Genome Res.* 18, 1446–1455
32 Knowles, D.G. and McLysaght, A. (2009) Recent *de novo* origin of human protein-coding genes. *Genome Res.* 19, 1752–1759
33 Aharoni, A. *et al.* (2004) The 'evolvability' of promiscuous protein functions. *Nat. Genet.* 37, 73–76
34 Taylor, J.S. and Raes, J. (2004) Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.* 38, 615–643
35 Hahn, M.W. (2009) Distinguishing among evolutionary models for the maintenance of gene duplicates. *J. Hered.* 100, 605–617
36 Christin, P.A. *et al.* (2009) Insights on C4 photosynthesis evolution from genomics and phylogenetics of the NADP-malic enzyme multigene family in grasses. *Genome Biol. Evol.* 2009, 221–230
37 Castoe, T.A. *et al.* (2008) Adaptive evolution and functional redesign of core metabolic proteins in snakes. *PLoS One* 3, e2201
38 Castoe, T.A. *et al.* (2009) Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* 106, 8986–8991
39 Christin, P.A. *et al.* (2008) Evolutionary switch and genetic convergence on *rbcL* following the evolution of C_4 photosynthesis. *Mol. Biol. Evol.* 25, 2361–2368
40 Zhang, J.Z. and Kumar, S. (1997) Detection of convergent and parallel evolution at the amino acid sequence level. *Mol. Biol. Evol.* 14, 527–536
41 Zhang, J. (2006) Parallel functional changes in the digestive RNases of ruminants and colobines by divergent amino acid substitutions. *Mol. Biol. Evol.* 20, 1310–1317
42 Zhang, J.Z. (2006) Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nat. Genet.* 38, 819–823
43 Yu, L. *et al.* (2010) Adaptive evolution of digestive RNASE1 genes in leaf-eating monkeys revisited: new insights from ten additional colobines. *Mol. Biol. Evol.* 27, 121–131
44 Weinreich, D.M. *et al.* (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312, 111–114
45 ffrench-Constant, R.H. *et al.* (2004) The genetics and genomics of insecticide resistance. *Trends Genet.* 20, 163–170
46 Liu, W.J. *et al.* (2007) Single-site mutations in the carboxyltransferase domain of plastid acetyl-CoA carboxylase confer resistance to grass-specific herbicides. *Proc. Natl. Acad. Sci. U. S. A.* 104, 3627–3632
47 Christin, P.A. *et al.* (2007) C_4 photosynthesis evolved in grasses via parallel adaptive genetic changes. *Curr. Biol.* 17, 1241–1247
48 Besnard, G. *et al.* (2009) Phylogenomics of C_4 photosynthesis in sedges (Cyperaceae): multiple appearances and genetic convergence. *Mol. Biol. Evol.* 26, 1909–1919
49 Smith, J.M. (1970) Natural selection and the concept of a protein space. *Nature* 225, 563–564
50 Wright, S. (1932) The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Proceedings of the Sixth International Congress of Genetics* (Jones, D.F., ed.), pp. 356–366, Brooklyn Botanic Garden
51 Eigen, M. *et al.* (1988) Molecular quasi-species. *J Phys. Chem.* 92, 6881–6891
52 Gillespie, J.H. (1984) Molecular evolution over the mutational landscape. *Evolution* 38, 1116–1129
53 Weinreich, D.M. and Chao, L. (2005) Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution* 59, 1175–1182
54 Provine, W.B. (1986) *Sewall Wright and Evolutionary Biology*, University of Chicago Press
55 Kaplan, J. (2008) The end of the adaptive landscape metaphor? *Biol. Philos.* 23, 625–638
56 Pigliucci, M. (2008) Sewall Wright's adaptive landscapes: 1932 vs. 1988. *Biol. Philos.* 23, 591–603
57 Gavrilets, S. (1997) Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.* 12, 307–312
58 Gavrilets, S. (ed.) (2004) *Fitness landscapes and the origin of species*, Princeton University Press
59 Wade, M.J. (1992) Sewall Wright: gene interaction and the shifting balance theory. *Oxf. Surv. Evol. Biol.* 8, 35–62
60 Coyne, J.A. *et al.* (1997) Perspective: a critique of Sewall Wright's shifting balance theory of evolution. *Evolution* 51, 643–671

61 Coyne, J.A. *et al.* (2000) Is Wright's shifting balance process important in evolution? *Evolution* 54, 306–317
62 Goodnight, C.J. and Wade, M.J. (2000) The ongoing synthesis: a reply to Coyne, Barton and Turelli. *Evolution* 54, 317–324
63 Perutz, M.F. (1983) Species adaptation in a protein molecule. *Mol. Biol. Evol.* 1, 1–28
64 Nielsen, R. (2001) Statistical tests of selective neutrality in the age of genomics. *Heredity* 86, 641–647
65 Kosakovsky Pond, S.L. *et al.* (2006) Adaptation to different human populations by HIV-1 revealed by codon-based analysis. *PLoS Comput. Biol.* 2, e62
66 Studer, R.A. *et al.* (2008) Pervasive positive selection on duplicated and nonduplicated vertebrate protein coding genes. *Genome Res.* 18, 1393–1402
67 Dean, A.M. and Thornton, J.W. (2007) Mechanistic approaches to the study of evolution: the functional synthesis. *Nat. Rev. Genet.* 8, 675–688
68 Lunzer, M. *et al.* (2005) The biochemical architecture of an ancient adaptive landscape. *Science* 310, 499–501
69 Kogenaru, M. *et al.* (2009) Revealing evolutionary pathways by fitness landscape reconstruction. *Crit. Rev. Biochem. Mol. Biol.* 44, 169–174
70 Lozovsky, E.R. *et al.* (2009) Stepwise acquisition of pyrimethamine resistance in the malaria parasite. *Proc. Natl. Acad. Sci. U. S. A.* 106, 12025–12030
71 Romero, P.A. and Arnold, F.H. (2009) Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* 10, 866–876
72 DePristo, M.A. *et al.* (2005) Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat. Rev. Genet.* 6, 678–687