Review

A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations

Nicola J. Nadeau and Chris D. Jiggins

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

Studies of the genetic basis of adaptive changes in natural populations are now addressing questions that date back to the beginning of evolutionary biology, such as whether evolution proceeds in a gradual or discontinuous manner, and whether convergent evolution involves convergent genetic changes. Studies that combine quantitative genetics and population genomics provide a powerful tool for identifying genes controlling recent adaptive change. Accumulating evidence shows that single loci, and in some cases single mutations, often have major effects on phenotype. This implies that discontinuous evolution, with rapid changes in phenotype, could occur frequently in natural populations. Furthermore, convergent evolution commonly involves the same genes. This implies a surprising predictability underlying the genetic basis of evolutionary changes. Nonetheless, most studies of recent evolution involve the loss of traits, and we still understand little of the genetic changes needed in the origin of novel traits.

New approaches to old questions

In the latter part of the 19th Century debate raged in the incipient field of evolutionary biology between the biometricians, who took a gradualist, incremental view of evolution, and the mutationists, who argued that evolutionary change could proceed in discontinuous leaps [1]. Such was the animosity that the leading biometricians, Karl Pearson and Walter Frank Raphael Weldon, were forced to establish their own journal rather than contend with the mutationist William Bateson who had a stranglehold on publication at the Royal Society. Although many of the early arguments were reconciled in the modern evolutionary synthesis, essentially the same debate was resurrected in different forms throughout the 20th Century, with the discontinuous evolutionary position represented with different emphasis by the 'hopeful monsters' of the developmental biologist Richard Goldschmidt in the 1940s [2], and to some extent by the 'punctuated equilibrium' hypothesis of the palaeontologist Stephen J. Gould in the 1970s [3]. In the 21st Century we now have the tools to study evolution on a genome-wide scale in a diverse array of natural populations, allowing us to uncover the mutational steps that characterise adaptive change. In this way, questions that date back to the early days of evolutionary biology are beginning to be answered.

Corresponding author: Jiggins, C.D. (c.jiggins@zoo.cam.ac.uk).

Indeed, there has been a recent explosion of studies attempting to identify the genetic basis of adaptive change in natural populations. These can be divided into forward genetics approaches, which assume no *a priori* knowledge of the genes involved and scan the genome for regions of interest, and those using reverse genetics or candidate-gene approaches in which genes are selected *a priori* (Box 1). Focusing on adaptation in wild animal populations under natural selective pressures, and primarily but not exclusively on morphological evolution, we look at three major questions: does evolution proceed by large-effect mutations or by the accumulation of many small-effect mutations? Are the same genes utilised when evolving similar phenotypes? And, what are the genetic changes required for the evolution of novel structures? Although our examples are restricted to animals, some similar patterns are emerging in plants [4–6] and microorganisms [7,8].

Is evolution jerky or gradual?

The question of whether evolution proceeds in major steps is really one of the relative contribution of different substitutions to an overall bout of adaptive change, sometimes termed an 'adaptive walk' [9,10]. Although an individual substitution will often correspond to a single novel mutation, this might not be the case if novel alleles are drawn from standing variation or from gene flow between different populations or species. Theoretical work has shown that the distribution of substitutions should follow an exponential distribution (with few substitutions of large effect and many of small effect) and large-effect substitutions are more likely in the early stages of adaptation to a novel environment; these adaptations are subsequently perfected by minor-effect substitutions [9,11]. Empirical tests of this theory in natural populations, however, are hampered by our inability to identify individual mutations and to track the timecourse of adaptive substitutions. Nonetheless, many studies have identified quantitative trait loci (QTL) across the genome that contribute to a particular adaptation. Importantly, these QTL can contain many individual mutations (as discussed below) but they do still provide some insight into the number of genes influencing a trait.

Before molecular studies, classical crossing experiments had already shown that some major phenotypic changes, such as the wing patterns of tropical *Heliconius* butterflies, or banding patterns of polymorphic *Cepaea* snails, could be controlled by a small number of Mendelian loci [12,13].

0168-9525/$ – see front matter doi:10.1016/j.tig.2010.08.004 Trends in Genetics, November 2010, Vol. 26, No. 11

Box 1. Methods for identifying adaptive genes in natural populations

Quantitative genetics

QTL mapping: the study of the genetic architecture of quantitative traits. QTL are identified by genetic mapping either using cross(es) or natural pedigrees to look for regions of the genome that are preferentially inherited with a particular phenotype. Crosses can be between divergent populations or inbred laboratory strains with different trait values [69,39]. Quantitative genetics can also be carried out using natural populations with known pedigree [19]. The QTL thus identified could represent large genome regions and commonly contain multiple functionally important sites in one or more genes.

Association mapping: markers distributed over the genome are screened for an association with a variable phenotype. Commonly used in medical genetics, and for naturally varying traits especially in *Drosophila* [16,80], but less commonly applied in non-model systems because a sequenced genome and large sample sizes are required. Association mapping relies on 'historical' recombination over many generations in natural populations, potentially allowing the effects of much smaller regions of the genome to be characterised.

Population genomics

Divergence genomics: the remaining methods make no assumptions about the traits under selection but instead scan markers distributed over the genome [21] for evidence of selection. The genome is scanned for regions showing the greatest genetic divergence between locally adapted populations or species [often measured using an estimate of genetic variance among populations (F_{st})] compared to estimated or observed null distributions [81]. These so-called 'outlier' markers can indicate regions under selection. This approach carries a significant risk of false positives because all distributions will have extreme values that might not necessarily be due to selection.

Hitch-hiking mapping: the characteristic signatures of a selective sweep are reduced genetic diversity, a region of increased linkage disequilibrium (LD) and a distinctive haplotype structure. These patterns can be used to identify regions that have been subject to recent selection [82]. This approach suffers from difficulties in differentiating the effects of selection from those of population demography [83,84] because a population bottleneck can have very similar effects to those of a selective sweep. Demography can often stochastically influence different genome regions in different ways, giving rise to patterns similar to a selective sweep.

Subsequently, QTL analyses have shown that in fact a few loci of relatively major effect similarly contribute to adaptive differences between populations, even in continuous quantitative traits [14–18]. We arbitrarily consider a single substitution that contributes >10% of the difference between parental strains to be of 'large' effect (although in QTL analysis effect size is instead often estimated as a percentage of the phenotypic variance seen in an experimental cross, which is a less evolutionarily-relevant measure). A well-studied example of adaptation in wild populations is the stickleback, a marine fish that has repeatedly colonised and adapted to freshwater habitats. These freshwater populations have evolved reduced armour plating, pelvic spines and changes in body shape as well as physiological changes in response to the osmotic environment. This system has become an exemplar of convergent evolution, and QTL studies have looked at a range of these traits. Commonly, one or a few major-effect QTL explain a large proportion of the variance in such crosses [15–17].

QTL studies have their limitations, however, because the power to detect loci of small effect is limited by the number of individuals, either in laboratory crosses or in natural pedigrees [19], and a statistical artefact known as the 'Beavis effect' can inflate the apparent influence of single QTL due to sampling effects [20]. Population genomics therefore offers an alternative approach, and this takes advantage of molecular variation linked to selected sites in natural populations (Box 1). Such studies also often identify just a few loci associated with adaptive divergence between populations [21–25]. For example, 22% of the difference in wing size along a latitudinal cline in *Drosophila melanogaster* could be explained by allelic variation in the promotor sequence of the *cold acclimation* (*Dca*) gene in *Drosophila* [25]. On a larger scale, a recent stickleback study used high-throughput sequencing methods to screen over 20 000 sequence tags in natural populations of marine and freshwater sticklebacks, and this revealed several regions that were genetically divergent between populations, indicating habitat adaptation [21]. Some of these corresponded to previously identified QTL whereas others indicated novel regions that probably harbour previously unidentified genes involved in local adaptation. Population genomics can therefore be complementary to more established quantitative genetic techniques, with the two approaches not necessarily highlighting the same genomic regions (Box 2). The main drawbacks of the population genomics approach alone are that divergent regions cannot be directly related to particular traits under selection, and that the inference of selection can be confounded by demographic effects.

Overall, there is therefore considerable evidence that single loci have major effects on adaptive phenotypes in the wild. However, this does not directly address the temporal question of whether evolutionary change is jerky or gradual. Genetic differences between current populations with divergent phenotypes might not be the same as the genetic changes that originally gave rise to these phenotypes, particularly if different species are being considered, and where multiple genetic changes affecting phenotype could have occurred [26]. As Ronald A. Fisher highlighted in the context of the evolution of butterfly mimicry, many small and sequential substitutions could lead to gradual evolution of a single major locus controlling adaptive differences [27]. Similarly, QTL of major effect could in fact harbour multiple linked substitutions of more minor effect [28–30]. Thus, we really need to identify the contribution of individual mutational steps to the differences between taxa, a much more difficult task. Directional, human-induced selection provides clear evidence for single mutations resulting in major phenotypic changes. Good examples include adaptation to insecticides in agricultural environments [31], laboratory evolution experiments with bacteria [7] and animal domestication [32], but the phenotypic effects of individual substitutions are less well documented for natural adaptations.

Nonetheless, numerous individual protein-coding mutations with major phenotypic effects have been described in natural populations, for example in the gain or loss of pigmentation [33–36]. Regulatory changes are harder to document, but detailed dissection of a promotor region of the *ebony* gene has recently shown that altitudinal variation in pigmentation in *D. melanogaster* is controlled by just five

Box 2. The benefits and drawbacks of quantitative genetic and population genomics approaches

QTL mapping. The QTL approach allows identification of genome regions influencing variation in a particular trait [14–18,29,34,36,53]. However, it usually relies on laboratory crosses and power is limited primarily by family size and recombination frequency [49]. The loci identified in captive populations might not be those responsible for trait differences in wild populations [85]. Genetic variation present in captive populations might not be representative of that present in the wild due to sampling, or could contain combinations of epistatically interacting alleles not found in the wild (as might be the case if a particular allele has been introgressed into a different genetic background). In particular, the influence of a QTL in the genetic background of a homogeneous captive strain might not be representative of the effect in a wild population. Genotype by environment interactions can also cause difficulties because the environment will normally differ and show less variation in the laboratory as compared to wild populations. Linkage disequilibrium will also tend to be higher in lab crosses than in the wild, resulting in QTLs often being large and encompassing many genes. The same QTL might not be detected by population genomics approaches where natural recombination could have broken up large regions of association.

Population Genomics. Population genomic approaches are therefore appealing because they do not require laboratory crosses [82,49] and can potentially take advantage of natural recombination to delimit QTL more finely. The major limitation of these approaches is that regions under selection cannot be directly related to adaptation at particular traits, hence the need to combine QTL and population genomics approaches. In addition, the power of these genomics approaches is limited by the vagaries of population demography and the history of selection [86]. Clear evidence of selection, in the form of a reduction in variation around a selected locus, is most likely under a 'strong' selective sweep when a novel advantageous mutation spreads rapidly to fixation. Signatures of selection are much weaker under 'soft' sweeps when selection acts on standing genetic variation, or on several mutations at a single locus [62], because there is greater opportunity for the advantageous allele to undergo recombination with surrounding variants before reaching fixation. Hence, selection is most difficult to detect when it acts on polygenic quantitative traits with standing genetic variation [84]. Finally, the signature of a selective sweep decays with time; in particular the characteristic reduction in nucleotide diversity will decay as new mutations arise and reach equilibrium.

mutations, with individual mutations responsible for up to 40% of the difference in expression between natural variants [37]. Similarly, individual deletion mutations can account for large expression differences at the *pituitary homeobox transcription factor 1* (*Pitx1*) gene between marine and freshwater sticklebacks, although the precise phenotypic effect of each mutation has not been measured independently of other sequence differences at this locus [38]; this highlights the difficulty of measuring the phenotypic effects of single mutational events in non-model systems.

The growing evidence for the importance of large-effect mutations in adaptation contrasts with data on the genetic architecture of quantitative traits in flies, mice and humans [39]. In these systems, quantitative traits are generally controlled by many loci of small effect, commonly with extensive pleiotropic effects on different aspects of phenotype. Early QTL studies did identify single QTL of large effect, and in many cases these were later found to consist of many linked sites of smaller effect [39]. Therefore, this discrepancy might partly reflect the relative infancy of studies addressing the genetic basis of adaptive traits. Nonetheless, there does appear to be a genuine discrepancy between the results from studies of background quantitative variation within populations and those of adaptive divergence between populations. We would argue that this is unsurprising because standing variation in quantitative traits is largely controlled by neutral or 'nearly neutral' variants and recessive deleterious alleles; by contrast, the adaptive traits described here have been subject to strong directional selection. As a population adapts to a novel environment, those rare mutations of large phenotypic effect that avoid deleterious side-effects will be most strongly favoured. Thus, almost by definition, the mutations that contribute to adaptation will be a highly non-random subset of those that contribute to standing variation in quantitative traits [40,41]. Furthermore, as highlighted by a recent study of insecticide resistance in *Drosophila*, the effective population size relevant for strong adaptation can be considerably larger than that estimated for neutral variation, such that extremely rare mutational events are available for adaptive change [42]. In summary, we argue that the genetic architecture of adaptive traits is different to that for standing variation in quantitative traits, and shows a greater role for substitutions of large effect. Nonetheless, a more precise quantification of this difference will require more examples in which the influence of mutational effects on natural adaptations is quantified.

Those same old genes again?

A related question is the predictability of evolutionary change. If selection preferentially acts on genes and mutations of large effect, are certain genes more likely to be the target of selection than others? Convergent evolution, whereby the same phenotype has evolved independently in two or more lineages, is a ubiquitous phenomenon and offers an excellent opportunity to test for the repeatability of genetic changes during adaptation (following Ref. [43], we use the term convergent rather than parallel evolution). As we gain more insight into the genetic basis of adaptive traits it is hard to escape the conclusion that certain genes have been repeatedly and preferentially used in multiple species and populations (Table 1).

A recent study showed that two haemoglobin genes in eight species of Andean waterfowl have multiple amino acid differences between high and low altitude populations, with a high incidence of the same substitutions occurring independently [44]. Similarly, and now widely cited, coding changes in the *melanocortin 1 receptor* (*MC1R*) gene are associated with both increases and decreases in pigmentation in many wild vertebrate populations, often with the same amino acid substitutions implicated in the production of similar phenotypes [35,45–47]. More broadly, there is evidence that *MC1R* has had an increased rate of evolutionary change in response to sexual selection on multiple independent lineages in the pheasant family [48]. A recent QTL study identified *MC1R* as being independently responsible for pigmentation loss in two cavefish populations, supporting the view that *MC1R* is a key gene in the evolution of pigmentation across vertebrates [14].

However, the 'candidate gene' approach suffers from an inevitable ascertainment bias. It would be preferable to determine how many of the loci involved in convergence are

Table 1. Examples of adaptive phenotypic convergence resulting from convergence at the molecular level[a]

Organism(s)	Gene(s)[b]	Phenotypic effect	Level of convergence	Technique(s) used to identify genes	Refs
Banaquits, snow geese, skuas, Flycatchers Jaguars, jaguarundi, rock pocket mice	*MC1R/Mc1r*	Pigmentation darkening	Orders, classes	Candidate gene	[35][c] [87] [46][c]
Beach mice Lizards, (possibly mammoths) Cavefish	*MC1R/Mc1r*	Pigmentation lightening	Classes	QTL, candidate gene Candidate gene QTL	[47][c] [14]
Human Beach/oldfield mice	*ASIP/Agouti*	Pigmentation lightening	Populations, orders	Association mapping, candidate gene Candidate gene	[88] [60]
Cavefish Human	*oca2/OCA2*	Pigmentation lightening	Populations, classes	QTL, candidate gene Association mapping, candidate gene	[36] [88]
Threespine stickleback Human	*Kitlg*	Pigmentation lightening	Populations, classes	QTL Candidate gene, association mapping	[16]
Heliconius melpomene and *H. erato*	Two genes mapping to homologous regions	Wing colour pattern	Species	Genetic mapping, population genetics	[50–52]
Drosophila gunungcola and *D. mimetica*	*yellow* (loss of *cis*-regulatory element)	Loss of pigmentation spot	Species	Candidate gene, expression and transgenics	[64]
Four *Drosophila* species	*yellow*	Gain of wing spots	Species	Candidate genes, expression, transgenics	[71,64]
Threespine stickleback Ninespine stickleback (Possibly manatees)	*Pitx1*	Loss of pelvic spines/ structures	Populations Genera (Classes)	QTL, transgenics Interspecific crosses (developmental similarities)	[38,89,17] [90]
Threespine stickleback	*Eda*	Loss of armour plating	Populations	QTL, population genomics, transgenics	[15,21]
Pelagic antarctic silverfish and blackfin icefish	Expression patterns: *col1a1*, *col2a1b*, *col10a1*	reduction of skeletal mineralisation – increased buoyancy	Genera	Expression of candidate genes	[91]
Six *Drosophila* species	*svb/ovo*	Loss of trichomes	Species	Genetic mapping, interspecific crosses, expression	[24]
Eight waterfowl species	αA and βA haemoglobin genes	Adaptation to high altitude	Genera	Candidate genes	[44]
Several bird species	*SWS1* opsin gene	Ultraviolet sensitivity	Species, orders	Candidate gene	[92]
Old-world primates, howler monkeys Cavefish	*LWS* opsin duplication	Trichromacy	Infra-order, classes	Candidate gene	[93][c] [94]
Asian and African leaf monkeys	*RNASE1* duplication	Adaptation to foregut fermentation	Genera	Candidate gene	[95]
Human	Lactase-phlorizin hydrolase (*LCT*)	Digestion of lactose from milk in adulthood	Populations	Candidate gene	[63]
Colias butterflies, Glanville fritillary butterfly	*Pgi*	Dispersal behaviour	Families	Candidate gene	[96][c]
Colias butterflies Copper butterfly, Sierra willow leaf beetle, Sand crab	*Pgi*	Thermal adaptation	Families, classes	Allozyme scan Candidate gene	[96][c]
D. melanogaster	*adh* *period* *IR* Transposable elements in introns of *sra* and *ago2*	Adaptation to latitude	Populations	Candidate genes Population genomics	[97] [98] [99] [22]
Arthropods, annelids, chordates	Expression patterns: *Notch, engrailed, wingless* (*Wnt1*)	Segmentation	Phyla	Expression	[66]

[a]This table contains all the animal examples that we are aware of in which phenotypic convergence results from changes in the same genes.
[b]Abbreviations: *adh*, alcohol dehydrogenase; *ago2, argonaute 2*; *col1a1*, collagen, type I, alpha 1; *col2a1b*, collagen, type II, alpha-1b; *col10a1*, collagen, type X, alpha 1; *IR, Insulin-like Receptor*; Kitlg, *Kit ligand*; *LWS*, long wavelength sensitive; *oca2/OCA2*, oculocutaneous albinism II; *Pgi, Phosphoglucose isomerase*; *sra, sarah*; *SWS1*, short wavelength sensitive 1; *svb/ovo, shavenbaby/ovo*.
[c]Reviews containing original references.

shared between the species across the whole genome. For example, in northern and southern hemisphere latitudinal clines in *D. melanogaster*, a survey of all transposable elements across the genome identified ten elements showing evidence of selection, of which two were shared between hemispheres [22]. Similar parallel patterns were seen in a genome-wide survey of genetic divergence across the same clines [23]. Although such studies provide clear evidence for a degree of parallelism in the genetic basis of adaptation, where different loci are identified it is unclear whether this represents selection on different traits, or different genes affecting convergent traits. Again, a disadvantage of the population genomic approach is therefore that genetic signatures of selection cannot be linked directly to particular traits.

Thus, studies ideally need to combine both QTL and population genomic approaches [49]. Again, the stickleback system provides a good example, and population genomic data have provided striking evidence for repeated divergence of the same genomic regions in the freshwater populations [21], with most regions identified corresponding to previously identified QTL for traits relating to morphology and osmoregulation, for example the *Ectodysplasin* (*Eda*) gene [15]. Nonetheless, some of the regions identified in the population survey had not been previously described in QTL analysis of individual traits, and some major QTL were not highlighted in the population survey (see Box 2 for reasons why the two approaches could differ). Detailed study of the genes underlying differences in armour plating and pelvic spines, *Eda* and *Pitx1* respectively, has shown that convergence can result from either selection of the same allele from standing variation [15] in the case of *Eda*, or repeated independent mutation of the same *cis*-regulatory region in multiple freshwater populations in the case of *Pitx1* [38].

Mimetic butterflies offer another striking example because different species converge on the same wing pattern (Figure 1). For example, *Heliconius melpomene* and *Heliconius erato* exhibit convergent colour patterns wherever they co-occur, driven by Müllerian mimicry (convergence in appearance of two or more distasteful species in order to minimise the cost of educating predators), and much of the colour pattern variation in both species can be explained by a small number of Mendelian loci [13]. Mapping of these switch genes has revealed that similar phenotypic changes map to the same genomic regions in both species, making it highly likely that the similarities in colour pattern have evolved through changes in orthologous genes in the two species [50]. Population surveys complement the mapping approach, and have highlighted narrow gene regions with parallel divergence between forms of the mimetic species on a much finer genomic scale [51,52].

A complementary approach, which has been employed in studies of lake whitefish (*Coregonus clupeaformis*), is to study the response of the transcriptome to divergent adaptation [54,55]. Multiple populations of these fish have independently evolved a dwarf form that exploits limnetic habitats and also shows phenotypic convergence with the related cisco (*Coregonus artedi*), which it replaces ecologically across their range. At the intraspecific level a signifi-

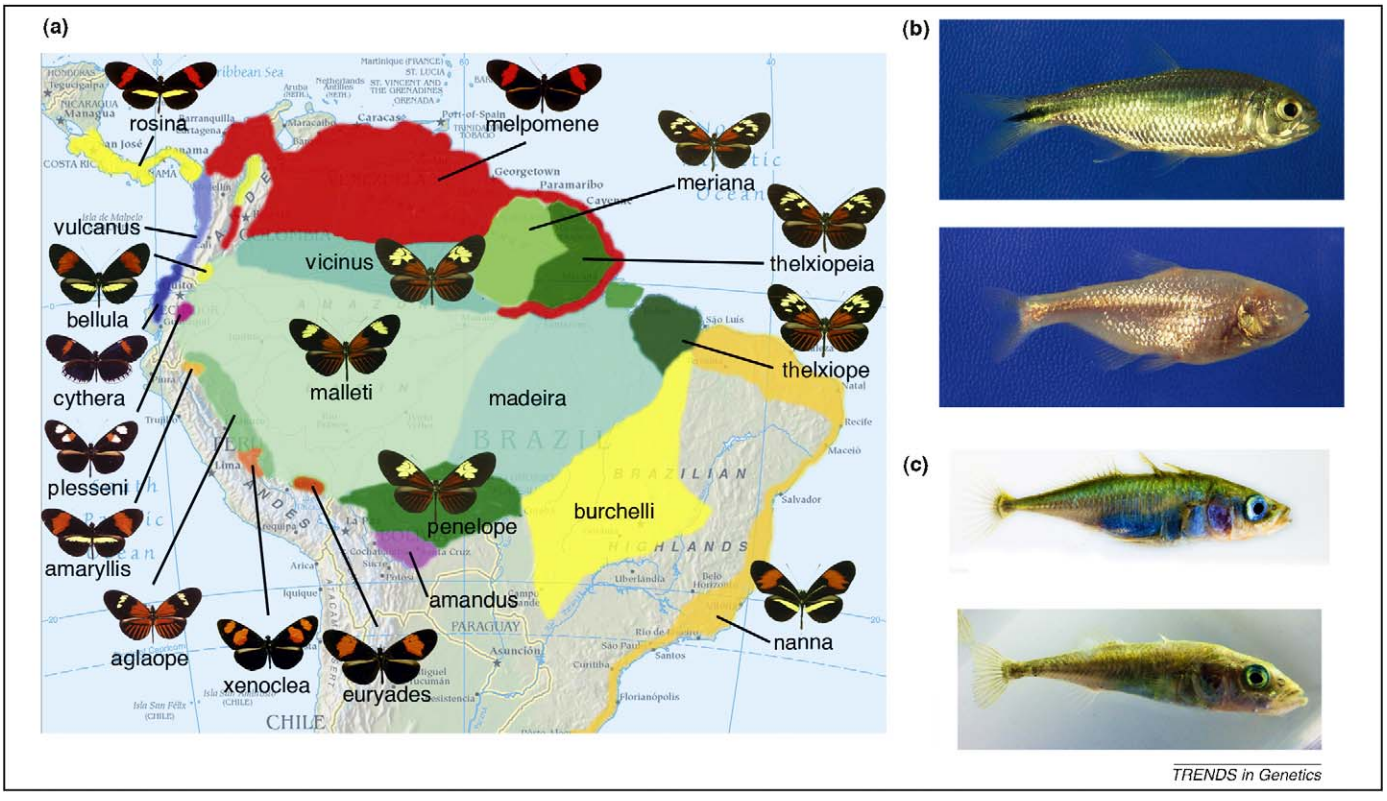

Figure 1. Examples of intra-specific adaptive change that have been studied genetically. **(a)** Geographic races of *Heliconius melpomene* across South and Central America. Each race is a near-perfect mimic of a race in the parallel radiation of *Heliconius erato*. Reproduced with permission from Bioscience (UC Press). **(b)** The Mexican tetra, *Astyanax mexicanus,* showing eyed surface form (above) and eyeless cave form (below). **(c)** The three-spine stickleback, *Gasterosteus aculeatus* showing marine (above) and freshwater (below) forms. Note in particular the loss of dorsal spines in freshwater form. Photo credits to Yoshiyuki Yamamoto and Frank Chan.

cant number of the same genes were differentially expressed in independent pairs of dwarf and normal populations; many of these were expressed in opposite directions between populations, suggesting that this is not a simple case of convergence via the same genetic mechanisms [54]. However, at the interspecific level, six of the same candidate genes were upregulated in both dwarf whitefish and ciscos, and overall there was greater similarity of transcriptional profiles between the dwarf whitefish and ciscos than there was between the normal whitefish and ciscos [55]. Thus, parallel evolutionary changes can also involve parallel differences in gene expression. Of course, expression changes might not be regulated by the same underlying genetic changes; so this work implies that the same pathways are involved in convergent adaptation but the genetic basis of divergence could be very different.

Some clear counterexamples have also emerged recently, such as the remarkable evolution of identical frog toxins using different genes [56], or convergent abdominal pigmentation in *Drosophila* species that has evolved via a diversity of genetic mechanisms [57]. Similarly, there are several counterexamples in which *MC1R* is not involved in changes in vertebrate pigmentation [58–60]. Pigmentation and eye reduction in cavefish [61] and pale coat colour in different populations of beach mice [34,45] have both been cited as examples of convergent phenotypes that can result from more than one genetic mechanism [43]. However, in these cases there appear to be a small number of genes, any one of which can be deployed in any particular situation, but some of which are also repeatedly used in different populations or species, supporting the idea that repeated use of the same genes is a general rule rather than the exception.

In some instances (such as the *Eda* locus in sticklebacks [15]) the common genetic basis of convergent traits is most likely to be due to independent selective events operating on the same pool of standing genetic variation. In which case, independent utilisation of the same genetic variation is not particularly surprising given that this variation will be most accessible to selection [62]. However, in many cases both the mutations and selective events have occurred independently in multiple populations or species [14,24,35,38,44,63,64]. Again, these results contrast strikingly with what is known about quantitative genetic variation in similar phenotypes, where tens or hundreds of genes can influence traits such as bristle number in flies or coat colour in mice [65,39]. Thus, it seems clear that the bias towards utilisation of the same loci in convergent evolution is a result of selection rather than an underlying mutational bias. Recent attempts to explain the predictability of genetic evolution have focused on the lack of negative side-effects at some loci due to their position in regulatory pathways, meaning that they can evolve more freely without altering traits other then those under selection [40,41]. This predictability must be aided by shared developmental systems [2]. For example, segmentation has been proposed to have evolved independently in several animal phyla through co-option of ancient conserved regulatory networks [66]. As we move beyond the simple observation of a shared genetic basis for convergence, the detection of general rules governing the genetic basis of adaptive changes promises to be an exciting area for future research.

Origins of novelty

Many of the well-studied examples of adaptive evolution have involved trait loss, such as the loss of bony structures in freshwater stickleback populations [15,38] and the reduction of pigmentation and eyes in cavefish [61]. However, over the broad sweep of evolutionary time what we would really like to explain is the gain of complexity and the origins of novel adaptations. It has previously been suggested that certain kinds of mutations might be more liable to cause trait loss [33] (but see Ref. [40]). Is it possible that trait loss could involve different evolutionary processes to trait gain?

In some cases loss-of-function mutations could result simply from relaxed constraints following a change in environment or ecology. For example, *Drosophila sechellia* is a host-plant specialist in which all 136 olfactory and gustatory receptors show high rates of change as compared to other *Drosophila* species [68]. This includes a large number of loss-of-function mutations leading to many pseudogenes. This pattern is significantly different to that found at other genes in the genome, and in some cases probably represents adaptation to the novel host, but much of the change is probably due to a reduction in the number of olfactory and gustatory receptors needed for a specialised life style. Similarly, reduction of eyes and pigmentation in cavefish involves loss of complexity, although such losses can be due to genetic drift in the case of pigmentation or positive natural selection for eye reduction [69].

In some cases gains of phenotypic complexity have been studied at the genetic level, such as the gain of wing spots in *Drosophila* species [70,71], elaborated beetle horns [72], or butterfly wing patterns [73]. These commonly involve co-option of existing genes and pathways, but the observation that existing genes are involved in the development of such structures is in itself unsurprising. Presumably this co-option generally occurs through the evolution of novel tissue-specific *cis*-regulatory regions [37], but such changes have only rarely been documented [71]. In general, gains of complexity have been studied on deeper timescales than losses, making it difficult to infer the precise sequence of events that has led to the origin of the novel structure. Of course, to some extent the difference between loss and gain could be a question of semantics, so for example the loss of trichomes could be called gain of naked cuticle [40]. However there clearly are complex structures that are gained during evolution, such as butterfly wing patterns, and we currently know little about how this process takes place. Promising systems include rapidly evolving ornaments such as the colour patterns of butterflies and cichlid fish [51,52,58,74], the elongated eyes stalks of stalk-eyed flies and the swords of swordtail fish, all of which are being characterised at the molecular level [75,76], and progress is likely to be rapid in the near future.

Concluding remarks and future perspectives

Over the last decade or so the advances in our understanding of the link between genotype and phenotype have been profound. It is clear that evolution can often proceed by

means of single mutations with major effects on phenotype, a result that would have been a surprise to Darwin and some of his earliest followers. This is consistent with theory that predicts a distribution of effect sizes, with a few substitutions making a major contribution to evolutionary change during an 'adaptive walk' [9,11]. Our understanding of how such major shifts can remain beneficial without incurring deleterious side-effects is illuminated by studies of convergent evolution. These have shown that, of the many genes likely to influence variation in a particular trait, it is common for only one or a few to be recruited by natural selection. This implies that particular points in genetic pathways, or particular kinds of mutations, can avoid deleterious side-effects and move phenotypes in predictable ways towards local adaptive peaks [41]. To comprehend these results requires a realisation of the vast size of natural populations, especially compared to laboratory stocks of model organisms. Only then can we understand how what appear to be extraordinarily unlikely events, such as the repeated deletion of precisely the same promotor sequence [38], can actually become predictable evolutionary outcomes.

Much of the recent progress has been facilitated by new technologies, making this an exciting time for evolutionary genetics. Advances in sequencing and other technologies mean that large numbers of individuals can be genotyped at genome-wide scales, facilitating both quantitative genetic and population genomics approaches [77] and dramatically increasing the diversity of systems in which adaptive divergence can be analysed genetically. A particularly exciting development is the restriction-site-associated DNA (RAD) sequencing method used for population genomic studies of sticklebacks [21], which allows large amounts of sequence data to be generated for mapping or population genetic studies, but without the need for prior identification of SNP sites [78]. Further, techniques are now available for target enrichment that allow next-generation sequencing of particular genome regions of interest [79]. For the future, there is a clear need to diversify the study systems and investigate the generality of results inferred from a handful of candidate genes and model systems, but fortunately the tools are now available to do just that.

Acknowledgement

This work was funded by the Leverhulme Trust, the BBSRC and the Royal Society.

References

1 Provine, W.B. (2001) *The Origins of Theoretical Population Genetics*, (1st edn), University of Chicago Press
2 Goldschmidt, R.B. (1940) *The Material Basis of Evolution*, Yale University Press
3 Eldredge, N. and Gould, S. (1972) Punctuated equilibria: an alternative to phyletic gradualism. In *Models in Paleobiology* (Schopf, T.J.M., ed.), pp. 82–115, Freeman, Cooper and Company
4 Johanson, U. *et al.* (2000) Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* 290, 344–347
5 Harrison, C.J. *et al.* (2005) Independent recruitment of a conserved developmental mechanism during leaf evolution. *Nature* 434, 509–514
6 Bradshaw, H.D. and Schemske, D.W. (2003) Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* 426, 176–178
7 Barrick, J.E. *et al.* (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461, 1243–1247
8 Anderson, J.B. *et al.* (2003) Mode of selection and experimental evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics* 163, 1287–1298
9 Orr, H.A. (2005) The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.* 6, 119–127
10 Turner, J.R. (1988) The evolution of mimicry: a solution to the problem of punctuated equilibrium. *Am. Nat.* 131, S42–S66
11 Orr, H.A. (1998) The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52, 935–949
12 Cain, A.J. *et al.* (1960) New data on the genetics of polymorphism in the snail *Cepaea Nemoralis* L. *Genetics* 45, 393–411
13 Sheppard, P.M. *et al.* (1985) Genetics and the evolution of Müllerian mimicry in *Heliconius* Butterflies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 433–610
14 Gross, J.B. *et al.* (2009) A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genet.* 5, e1000326
15 Colosimo, P.F. *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of *Ectodysplasin* alleles. *Science* 307, 1928–1933
16 Miller, C.T. *et al.* (2007) *Cis*-regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* 131, 1179–1189
17 Shapiro, M.D. *et al.* (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* 428, 717–723
18 Carbone, M.A. *et al.* (2005) Quantitative trait loci affecting the difference in pigmentation between *Drosophila yakuba* and *D. santomea*. *Genetics* 171, 211–225
19 Slate, J. *et al.* (2010) Genome mapping in intensively studied wild vertebrate populations. *Trends Genet.* 26, 275–284
20 Xu, S. (2003) Theoretical basis of the Beavis effect. *Genetics* 165, 2259–2268
21 Hohenlohe, P.A. *et al.* (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PloS Genet.* 6, e1000862
22 González, J. *et al.* (2010) Genome-wide patterns of adaptation to temperate environments associated with transposable elements in *Drosophila*. *PLoS Genet.* 6, e1000905
23 Turner, T.L. *et al.* (2008) Genomic analysis of adaptive differentiation in *Drosophila melanogaster*. *Genetics* 179, 455–473
24 Sucena, E. *et al.* (2003) Regulatory evolution of *shavenbaby/ovo* underlies multiple cases of morphological parallelism. *Nature* 424, 935–938
25 McKechnie, S.W. *et al.* (2010) A clinally varying promoter polymorphism associated with adaptive variation in wing size in *Drosophila*. *Mol. Ecol.* 19, 775–784
26 True, J.R. and Haag, E.S. (2001) Developmental system drift and flexibility in evolutionary trajectories. *Evol. Dev.* 3, 109–119
27 Fisher, R.A. (1930) *The Genetical Theory of Natural Selection*, (1st edn), Oxford University Press
28 McGregor, A.P. *et al.* (2007) Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* 448, 587–590
29 Orgogozo, V. *et al.* (2006) High-resolution quantitative trait locus mapping reveals sign epistasis controlling ovariole number between two *Drosophila* species. *Genetics* 173, 197–205
30 Knight, C.G. *et al.* (2001) Testing life-history pleiotropy in *Caenorhabditis elegans*. *Evolution* 55, 1795–1804
31 Ffrench-Constant, R.H. *et al.* (2004) The genetics and genomics of insecticide resistance. *Trends Genet.* 20, 163–170
32 Cadieu, E. *et al.* (2009) Coat variation in the domestic dog is governed by variants in three genes. *Science* 326, 150–153
33 Hoekstra, H.E. and Coyne, J.A. (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61, 995–1016
34 Steiner, C.C. *et al.* (2007) Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol.* 5, e219
35 Mundy, N.I. (2005) A window on the genetics of evolution: MC1R and plumage colouration in birds. *Proc. R. Soc. Lond. B Biol. Sci.* 272, 1633–1640
36 Protas, M.E. *et al.* (2006) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat. Genet.* 38, 107–111

37 Rebeiz, M. *et al.* (2009) Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* 326, 1663–1667
38 Chan, Y.F. *et al.* (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* 327, 302–305
39 Flint, J. and Mackay, T.F.C. (2009) Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res.* 19, 723–733
40 Stern, D.L. and Orgogozo, V. (2008) The loci of evolution: how predictable is genetic evolution? *Evolution* 62, 2155–2177
41 Stern, D.L. and Orgogozo, V. (2009) Is genetic evolution predictable? *Science* 323, 746–751
42 Karasov, T. *et al.* (2010) Evidence that adaptation in *Drosophila* is not limited by mutation at single sites. *PLoS Genet.* 6, e1000924
43 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
44 McCracken, K.G. *et al.* (2009) Parallel evolution in the major haemoglobin genes of eight species of Andean waterfowl. *Mol. Ecol.* 18, 3992–4005
45 Hoekstra, H.E. *et al.* (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313, 101–104
46 Majerus, M.E. and Mundy, N.I. (2003) Mammalian melanism: natural selection in black and white. *Trends Genet.* 19, 585–588
47 Hubbard, J.K. *et al.* (2010) Vertebrate pigmentation: from underlying genes to adaptive function. *Trends Genet.* 26, 231–239
48 Nadeau, N.J. *et al.* (2007) Evolution of an avian pigmentation gene correlates with a measure of sexual selection. *Proc. R. Soc. Lond. B Biol. Sci.* 274, 1807–1813
49 Stinchcombe, J.R. and Hoekstra, H.E. (2007) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100, 158–170
50 Joron, M. *et al.* (2006) *Heliconius* wing patterns: an evo–devo model for understanding phenotypic diversity. *Heredity* 97, 157–167
51 Baxter, S.W. *et al.* (2010) Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in the *Heliconius melpomene* clade. *PLoS Genet.* 6, e1000794
52 Counterman, B.A. *et al.* (2010) Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*. *PLoS Genet.* 6, e1000796
53 Baxter, S.W. *et al.* (2008) Butterfly speciation and the distribution of gene effect sizes fixed during adaptation. *Heredity* 102, 57–65
54 Derome, N. *et al.* (2006) Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchill) ecotypes. *Mol. Ecol.* 15, 1239–1249
55 Derome, N. and Bernatchez, L. (2006) The transcriptomics of ecological convergence between 2 limnetic coregonine fishes (Salmonidae). *Mol. Biol. Evol.* 23, 2370–2378
56 Roelants, K. *et al.* (2010) Identical skin toxins by convergent molecular adaptation in frogs. *Curr. Biol.* 20, 125–130
57 Ng, C.S. *et al.* (2008) Genetic basis of sex-specific color pattern variation in *Drosophila malerkotliana*. *Genetics* 180, 421–429
58 Henning, F. *et al.* (2010) Genetic, comparative genomic, and expression analyses of the *Mc1r* locus in the polychromatic midas cichlid fish (Teleostei, *Cichlidae Amphilophus* sp.) species group. *J. Mol. Evol.* 70, 405–412
59 Cheviron, Z.A. *et al.* (2006) Sequence variation in the coding region of the melanocortin-1 receptor gene (*MC1R*) is not associated with plumage variation in the blue-crowned manakin (*Lepidothrix coronata*). *Proc. R. Soc. B Biol. Sci.* 273, 1613–1618
60 Mullen, L.M. and Hoekstra, H.E. (2008) Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution* 62, 1555–1570
61 Jeffery, W.R. (2009) Regressive evolution in *Astyanax* cavefish. *Annu. Rev. Genet.* 43, 25–47
62 Hermisson, J. and Pennings, P.S. (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169, 2335–2352
63 Tishkoff, S.A. *et al.* (2007) Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 39, 31–40
64 Prud'homme, B. *et al.* (2006) Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* 440, 1050–1053
65 Jackson, I.J. (1994) Molecular and developmental genetics of mouse coat color. *Annu. Rev. Genet.* 28, 189–217
66 Chipman, A.D. (2010) Parallel evolution of segmentation by co-option of ancestral gene regulatory networks. *BioEssays* 32, 60–70
68 McBride, C.S. (2007) Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 4996–5001
69 Protas, M. *et al.* (2007) Regressive evolution in the mexican cave tetra, *Astyanax mexicanus*. *Curr. Biol.* 17, 452–454
70 Werner, T. *et al.* (2010) Generation of a novel wing colour pattern by the Wingless morphogen. *Nature* 464, 1143–1148
71 Gompel, N. *et al.* (2005) Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* 433, 481–487
72 Moczek, A.P. and Rose, D.J. (2009) Differential recruitment of limb patterning genes during development and diversification of beetle horns. *Proc. Natl. Acad. Sci. U. S. A.* 106, 8992–8997
73 Saenko, S.V. *et al.* (2008) Conserved developmental processes and the formation of evolutionary novelties: examples from butterfly wings. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 363, 1549–1556
74 Terai, Y. *et al.* (2003) The complexity of alternative splicing of *hagoromo* mRNAs is increased in an explosively speciated lineage in East African cichlids. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12798–12803
75 Baker, R. *et al.* (2009) Genomic analysis of a sexually-selected character: EST sequencing and microarray analysis of eye-antennal imaginal discs in the stalk-eyed fly *Teleopsis dalmanni* (Diopsidae). *BMC Genomics* 10, 361
76 Offen, N. *et al.* (2009) Identification of novel genes involved in the development of the sword and gonopodium in swordtail fish. *Dev. Dyn.* 238, 1674–1687
77 Metzker, M.L. (2010) Sequencing technologies – the next generation. *Nat. Rev. Genet.* 11, 31–46
78 Baird, N.A. *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3, e3376
79 Summerer, D. (2009) Enabling technologies of genomic-scale sequence enrichment for targeted high-throughput sequencing. *Genomics* 94, 363–368
80 Wilfert, L. and Jiggins, F.M. (2010) Disease association mapping in *Drosophila* can be replicated in the wild. *Biol. Lett.* DOI: 10.1098/rsbl.2010.0329
81 Luikart, G. *et al.* (2003) The power and promise of population genomics: from genotyping to genome typing. *Nat. Rev. Genet.* 4, 981–994
82 Schlötterer, C. (2003) Hitchhiking mapping – functional genomics from the population genetics perspective. *Trends Genet.* 19, 32–38
83 Haddrill, P.R. *et al.* (2005) Multilocus patterns of nucleotide variability and the demographic and selection history of *Drosophila melanogaster* populations. *Genome Res.* 15, 790–799
84 Pritchard, J.K. *et al.* (2010) The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* 20, R208–R215
85 Macdonald, S.J. and Long, A.D. (2004) A potential regulatory polymorphism upstream of *hairy* is not associated with bristle number variation in wild-caught *Drosophila*. *Genetics* 167, 2127–2131
86 Teshima, K.M. *et al.* (2006) How reliable are empirical genomic scans for selective sweeps? *Genome Res.* 16, 702–712
87 Uy, J.A.C. *et al.* (2009) Difference in plumage color used in species recognition between incipient species is linked to a single amino acid substitution in the melanocortin-1 receptor. *Am. Nat.* 174, 244–254
88 Norton, H.L. *et al.* (2007) Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Mol. Biol. Evol.* 24, 710–722
89 Coyle, S.M. *et al.* (2007) Parallel evolution of *Pitx1* underlies pelvic reduction in Scottish threespine stickleback (*Gasterosteus aculeatus*). *J. Hered.* 98, 581–586
90 Shapiro, M.D. *et al.* (2006) Parallel genetic origins of pelvic reduction in vertebrates. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13753–13758
91 Albertson, R.C. *et al.* (2010) Molecular pedomorphism underlies craniofacial skeletal evolution in Antarctic notothenioid fishes. *BMC Evol. Biol.* 10, 4
92 Odeen, A. and Hastad, O. (2003) Complex distribution of avian color vision systems revealed by sequencing the *SWS1* opsin from total DNA. *Mol. Biol. Evol.* 20, 855–861
93 Hunt, D.M. *et al.* (1998) Molecular evolution of trichromacy in primates. *Vision Res.* 38, 3299–3306

94 Yokoyama, R. and Yokoyama, S. (1990) Convergent evolution of the red- and green-like visual pigment genes in fish, *Astyanax fasciatus*, and human. *Proc. Natl. Acad. Sci. U. S. A.* 87, 9315–9318

95 Zhang, J. (2006) Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nat. Genet.* 38, 819–823

96 Wheat, C.W. (2010) Phosphoglucose isomerase (Pgi) performance and fitness effects among arthropods and its potential role as an adaptive marker in conservation genetics. *Conserv. Genet.* 11, 387–397

97 Oakeshott, J.G. *et al.* (1982) Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. *Evolution* 36, 86–96

98 Sawyer, L.A. *et al.* (2006) The *period* gene Thr-Gly polymorphism in Australian and African *Drosophila melanogaster* populations: implications for selection. *Genetics* 174, 465–480

99 Paaby, A.B. *et al.* (2010) Identification of a candidate adaptive polymorphism for *Drosophila* life history by parallel independent clines on two continents. *Mol. Ecol.* 19, 760–774