STUDY DESIGNS

Molecular spandrels: tests of adaptation at the genetic level

Rowan D. H. Barrett and Hopi E. Hoekstra

Abstract | Although much progress has been made in identifying the genes (and, in rare cases, mutations) that contribute to phenotypic variation, less is known about the effects that these genes have on fitness. Nonetheless, genes are commonly labelled as 'adaptive' if an allele has been shown to affect a phenotype with known or suspected functional importance or if patterns of nucleotide variation at the locus are consistent with positive selection. In these cases, the 'adaptive' designation may be premature and may lead to incorrect conclusions about the relationships between gene function and fitness. Experiments to test targets and agents of natural selection within a genomic context are necessary for identifying the adaptive consequences of individual alleles.

Spandrel
An architectural feature that is necessary in the construction of domed cathedrals, but because of its aesthetic and purposeful design it can easily be confused as the featured element rather than a by-product of an engineering constraint.

Selection coefficient
The strength of selection as measured by the difference in fitness among genotypes.

Directional selection
Natural selection that favours one end of a distribution of a quantitative trait.

Stabilizing selection
Natural selection that favours intermediate values of a quantitative trait.

Department of Organismic and Evolutionary Biology, Department of Molecular and Cellular Biology and the Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts 02138, USA.
e-mails: rbarrett@fas.harvard.edu; hoekstra@oeb.harvard.edu
doi:10.1038/nrg3015
Corrected online
29 November 2011

"Evolutionary adaptation is a special and onerous concept that should not be used unnecessarily, and an effect should not be called a function unless it is clearly produced by design and not by chance."
George C. Williams, 1966.

An adaptive trait is one that increases organismal fitness in a particular environmental context and is hence the target of natural selection. This apparently simple concept has been the subject of extensive debate concerning the prevalence of adaptive traits, the mechanisms by which they arise and the levels at which selection operates. In the research programme known as 'adaptationism', which was famously criticized by Gould and Lewontin[1], organisms were partitioned into traits that are viewed *a priori* as being optimally designed by natural selection for their current function. The backlash against this movement gave rise to the famous metaphor of the spandrel, which highlighted our reliance on plausibility as a criterion for accepting adaptive stories while largely ignoring alternative possibilities, such as the trait having evolved by stochastic processes or one trait being changed indirectly by selection on a correlated trait.

The criticisms raised against adaptationism spurred the development of rigorous quantitative methods for measuring selection on phenotypic traits[2–4]; this in turn served as a catalyst for field researchers to estimate selection coefficients in natural populations. The accumulated efforts to date have provided crucial information about the form of selection (that is, directional selection, stabilizing selection or disruptive selection) and the strength of selection commonly found in the wild. They also demonstrate that adaptive hypotheses can be formally tested (reviewed in REFS 5–7). In the past decade, new genomic technologies have allowed researchers to localize footprints of selection in the genomes of a diverse array of organisms. This advance has led to many studies reporting signatures of selection in patterns of genetic variation; one conclusion of these studies is that much of the genome is influenced by selection (in contrast to the neutral theory; BOX 1).

As a result of field-based studies of phenotypic selection and laboratory-based studies of genome-wide variation, it is now common to see loci designated as 'adaptive' if they either affect phenotypic traits that are known or suspected to be the target of selection or if they show statistical signatures of historical selection. These two approaches represent the connection of genotype and phenotype, and genotype and fitness, respectively. It is important to recognize that neither approach functionally connects genotype, phenotype and fitness, nor does linking the two approaches in tandem (see discussion in REF. 8). We argue that connecting all three of these components is necessary to fully understand the adaptive consequences of an allele (FIG. 1).

Here, we define an adaptive allele as 'an allele that has functional effects on a phenotypic trait that produce an increase in fitness'. Although this definition is restricted to targets of directional selection, it promotes the connection of genotype, phenotype and fitness because it requires evidence of the direct effects of alternative alleles on an ecologically relevant phenotype, as well as the effects of these alleles on fitness through known functional mechanisms. Ideally, we

Disruptive selection
Natural selection that favours extreme values of a quantitative trait over intermediate values.

Neutral theory
This theory holds that standing genetic variation is predominantly neutral, whereas most new mutations are deleterious.

would also like to know the agents of selection acting on the focal trait and driving changes in allele frequencies. We suggest that experimental tests of selection on genes that underlie phenotypic traits are a powerful way to avoid being led astray by intuitively appealing — but potentially incomplete or incorrect — adaptive stories (BOX 2).

As we describe below, the development of more rigorous observational and experimental measures of selection and fitness, together with population genomics studies, are allowing a more accurate characterization of the adaptive process. We also propose an experimental genomics approach for rigorously testing adaptation at the genetic level to avoid repeating the same mistakes that Gould and Lewontin cautioned against over two decades ago, but this time at the molecular level.

Box 1 | **How much of genome divergence is adaptive?**

Since Kimura proposed the neutral theory of molecular evolution[95,96], there has been considerable debate about the proportion of the genome that is adaptive. According to the neutral theory, random mutations are much more likely to be either deleterious or neutral than they are to confer a fitness advantage. Although deleterious and neutral mutations may be most likely to arise, only neutral alleles are expected to contribute to persistent polymorphisms because deleterious alleles are efficiently removed from the population by purifying selection[97]. The neutral theory and its extension, the nearly neutral theory[98], have provided support for four decades of empirical data that showed, in most organisms, that abundant quantities of DNA appear to have evolved with little selective constraint (for example, REF. 99; reviewed in REF. 100). Since then, statistical inference has improved and the amount data has increased, as we describe below.

The selectionists. Recently, several comparative genomic studies have claimed evidence for widespread selection in the genomes of certain species, leading to suggestions that the days of the neutral theory are over[101,102]. For example, analyses of genetic variation in *Drosophila* species reveal that much of the non-coding regions of the genome are under purifying selection, and thus of functional importance[103], and amino acid divergence between species is often driven by (weak) positive selection[104–105]. Moreover, many changes are strong enough that the fate of neutral mutations is as dependent on their linkage to adaptive mutations as it is on genetic drift[102,106]. Even in organisms with small effective population sizes, a considerable fraction of the genome is under selective constraint[107]. For example, there are thousands of candidate regions under selection in humans[108]. Similar to the situation in *Drosophila* species, selection affects non-coding regions throughout the human genome as well as coding regions[109]; more than one-fifth of the loci that were investigated show evidence of natural selection[110]. Additional evidence comes from amino acids that are associated with reduced synonymous site variation[111,112], which is consistent with recurrent selective sweeps.

The neutralists. These studies are now under renewed attack from 'neutralists', who suggest that the theoretical basis of the genome scans described above is not well established and often produces high rates of false-positive and false-negative results[100]. Nei *et al.*[100] have criticized genome scans as largely consisting of a collection of single-locus analyses, which can result in misleading conclusions. A well-recognized problem is that certain demographic events can mimic the molecular patterns produced by selection[21,23,83,100]. When these deficiencies are corrected, the results often conform to the expectations of the neutral theory. For example, in a re-analysis of one of the genome scan studies discussed above, introduction of hierarchical structures based on five previously established geographic regions reduces the frequency of selection candidates from 23% of sites investigated[110] to no more than would be expected by chance[83], suggesting that several studies claiming evidence for selection may need to be re-evaluated[8,21]. However, the prevalence of false negatives also indicates that there may be considerable selection that genome scans are missing[101]. Most criticism is directed at genome scans, but others have extended the scepticism to additional tests of neutrality[113].

Although the amount of genomic data continues to increase, resolution of the debate may hinge on both the development of more sensitive models to estimate selection and demography jointly and the development of experimental studies to provide functional validation for the genomic extent of selection (for example, REF. 91).

Observational studies of genotypes in nature

Observational studies have long had an important role in the study of adaptation. Two common approaches for studying the genetics of adaptation are to search for associations between allele frequency and environmental variables and to detect patterns of genomic variation that are indicative of natural selection (for example, high levels of differentiation, reduced heterozygosity, skews in the site frequency spectrum and extended linkage disequilibrium (LD)[9]). Studying patterns of genetic diversity in nature can help to develop adaptive hypotheses, but it can often be difficult to distinguish between causal mechanisms of selection[10]. Two general approaches, known as bottom-up (or reverse) genetics and top-down (or forward) genetics have been employed to hunt for gene variants that co-vary with the environment. Recently, there have been efforts made to link these methods to provide more insight than either approach in isolation.

Bottom-up approaches. The bottom-up approach identifies putatively adaptive genes by searching for a molecular signature of selection in populations from different environments or populations that have experienced a change in conditions. For example, consistent allele frequency differences have been found in *Drosophila* species across spatial gradients[11,12] as well as across temporal samples that are associated with recent climate warming[13]. In these studies, genetic signatures of selection are manifested as exceptionally high levels of differentiation relative to neutral expectations (often measured by the fixation index (F_{ST})). The challenge of this approach, as well as of other tests of selection, is to identify significant departures from neutral expectations under realistic demographic models[14,15] and to eliminate single-locus false positives[16]. Although statistical methodologies have advanced greatly in recent years[17–20], there is considerable evidence that loci can show unique patterns of nucleotide variation owing to chance alone, incorrect models of demographic history or ascertainment bias[21]. (It is also possible for loci to be under selection without yielding statistically significant results in tests for selection[22,23].)

One limitation of bottom-up approaches is the possibility that an anonymous locus showing unusual 'outlier' behaviour is not under selection itself, but is instead in LD with the selected site (or sites). The extent of LD between the marker locus and the functionally relevant mutation (or mutations) can vary markedly across study systems or across the genome. Similarly, the size and position of the genomic regions that show differentiation is often unknown, at least for species without complete genome sequences. Perhaps the most challenging aspect of bottom-up approaches is that it may not be clear which phenotypic traits differ between samples

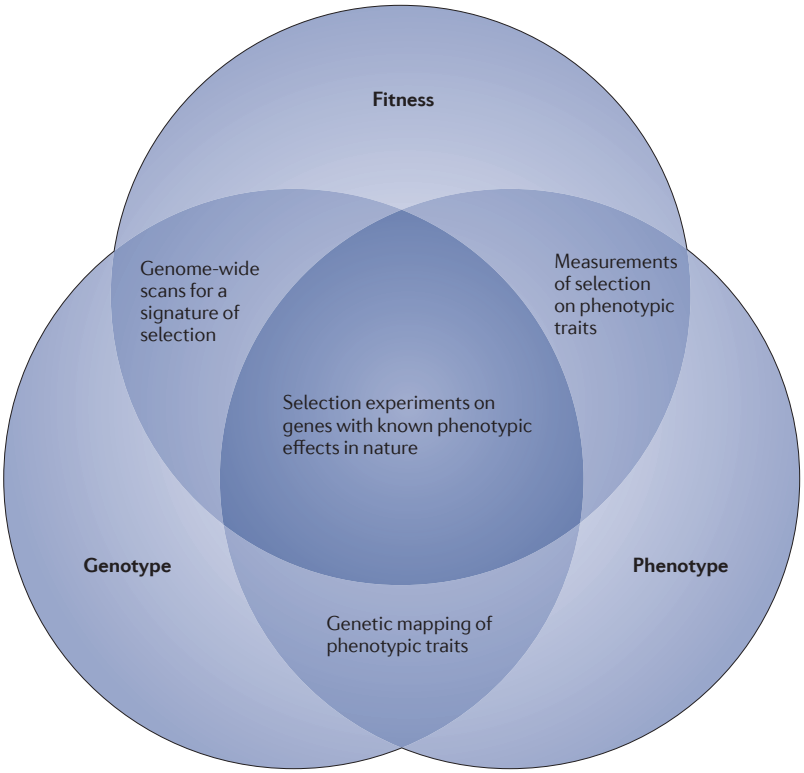


Figure 1 | **Connections between various approaches for studying the genetics of adaptation.** Ideally, an allele is designated as adaptive only when connections are made between its genotype, phenotype and fitness.

Purifying selection
Natural selection that favours the current condition by removing deleterious alleles that arise in the population.

Nearly neutral theory
An extension of the neutral theory that suggests that polymorphism at functionally important sites is predominantly nearly, rather than completely, neutral.

Effective population sizes
The size of an ideal population of breeding individuals that would experience the same amount of genetic drift as the observed population.

Selective sweeps
The increase in frequency of an allele (and closely linked chromosomal segments) that is caused by positive selection for the allele. Sweeps initially reduce variation and subsequently lead to a local excess of rare alleles (and an excess of homozygosity) as new unique mutations accumulate.

and whether any are influenced by the implicated loci, even when the gene can be pinpointed[24–26]. The absence of knowledge about phenotype naturally limits the ecological investigation of putative selective agents. Thus, the greatest strength of this approach is also its greatest weakness — by ignoring phenotypic measurements, the study is not biased by a few visible (and measureable) traits, but once genomic regions are identified, it is challenging to study their function and relevance to fitness without phenotypic information.

Top-down approaches. A top-down approach avoids the problem of not knowing the phenotypic targets of selection by starting with traits that are known to vary between environments or to have functional importance and then working to identify the genes underlying those traits. Most often, this is accomplished using a large number of genetically heterogeneous individuals that are scored for phenotypic traits and genotyped at many markers to test for co-segregation of genotype and phenotype. For example, genotype–phenotype co-segregation could be assessed through linkage mapping by using controlled crosses between parents with extreme trait values or through association mapping by taking advantage of a naturally admixed population. The premise is that markers showing strong associations with phenotype are in LD with the causal locus[27]. These methods can also be used in conjunction with knowledge of candidate genes to aid in quickly identifying genes that are associated with a particular phenotype. Candidate genes can either be used *a priori* to select associated markers that can be used for mapping[28], or they can be used *post hoc* to determine causative loci within regions that are associated with phenotype[29–31]. Although this approach is a powerful way of determining the genetic basis of divergent phenotypes, it does not alone provide any measure of the selection that may be driving differences in trait values or an understanding of the fitness consequences of the divergent phenotypes.

Combining bottom-up and top-down approaches. Combining mapping approaches with bottom-up methods can provide this missing information about whether selection acts on the traits under investigation. For example, genome scans have been used to detect signatures of selection acting on loci that underlie divergent phenotypes in pea aphids, *Arabidopsis thaliana*, stick insects and lake whitefish[32–35]. Or, once a gene has been identified, population genetics approaches can be used to implicate selection acting on that locus, as has recently been done in the iconic example of the peppered moth (BOX 3). In another example, it has been shown that the pale pigmentation of deer mice living in the light-coloured Sand Hills of Nebraska in the United States is caused by a mutation in the *Agouti* locus, which has a molecular signature of recent natural selection[36] (FIG. 2, TABLE 1 and Supplementary information 1 (table)). Although intuitively appealing, there is no direct evidence that the cause of this historical selection is increased survival that is due to crypsis against predation. Even if natural selection did act on the *Agouti* mutation, it is still necessary to rule out the possibility that pleiotropic effects of the mutations were the direct target of selection.

Thus, the combination of functional effect and selection does not demonstrate that selection only acted on the specific trait in question or that the gene is necessarily adaptive. Although outstanding progress has been made in identifying genes that contribute to phenotypic variation as well as in documenting evidence of selection, more work is needed to determine precisely how selection pressures lead to changes in allele frequency through time.

Experimental evolution studies in the laboratory

Adaptation can be studied as a process or as an outcome. Experimental studies help to clarify the genetics of adaptation because they can rigorously document the mechanisms and targets of selection that drive changes in allele frequency (the process) in ways that are not possible when investigating historical signatures of selection (the outcome). Experimental evolution — which is often done with populations that have large effective sizes, short generation times and small, easily manipulated habitats — has been a powerful approach in examining the genetic details of adaptation. Moreover, many of these organisms, such as bacteria and viruses, have small genomes; in some cases, this makes replicated whole-genome sequencing a feasible approach for discovering causal

mutations. Perhaps most importantly, researchers have the advantage of being able to resurrect ancestors from frozen stocks to compare directly the fitness consequences of evolved differences between pre- and post-selection populations.

Linking genetic variants to fitness. Using an experimental approach, a number of studies have explicitly measured the differential performance of genetic variants in laboratory-based assays. For example, Rokyta *et al.*[37] carried out a selection experiment in which 20 replicate populations of an ssDNA virus were allowed to adapt independently to the same conditions. In each replicate, the first mutation to arise and spread was identified by whole-genome sequencing, and the fitness effect was measured as the growth rate of the virus. Similarly, Araya *et al.*[38] sequenced >93% of the genome of a laboratory-evolved strain of *Saccharomyces cerevisiae* and its ancestor to reveal a number of SNPs and copy number amplifications that are associated with laboratory adaptation. Importantly, these studies carefully identified the exact mutations that confer fitness gains, although the functional mechanisms by which the experimental populations adapted to the selection environment were not tested. One indirect approach that has been used to identify the potential phenotypic basis for adaptation is to infer effects from gene identity. For instance, a selection experiment in an RNA bacteriophage showed that a disproportionate number of substitutions measured across the genome occurred in a single gene; the effects of these substitutions were then inferred from previous studies that had characterized the phenotypic consequences of mutations at this locus[39].

Box 2 | **Avoiding spandrels: using genes to distinguish between targets of phenotypic selection**

Just as trait correlations can lead to incorrect conclusions about the phenotypic targets of selection[1–4], genetic architecture (for example pleiotropy, linkage and epistasis) can alter the overall selection coefficient acting on a particular allele by creating unexpected correlations between genetic variants and fitness. For example, an allele may have beneficial fitness effects on a particular trait of interest, but may nevertheless have a negative selection coefficient because of its deleterious pleiotropic effect on additional unmeasured traits[114]. Similarly, an advantageous allele may reside at a locus that is physically linked to another gene (or genes) relevant for fitness; depending on the direction and magnitude of the contribution of each locus to fitness, the positive effects of the focal allele may be swamped by the negative effects of neighbouring alleles[115]. Finally, putatively adaptive alleles, which have been characterized in one genetic background, may have different effects in another genetic background such that they no longer yield positive selection coefficients[92]. Although it is not yet clear how common these outcomes may be, several recent studies have demonstrated that when fitness is tested explicitly, our simple adaptive stories are, at best, often incomplete or, at worst, incorrect.

TYRP1* in soay sheep, *Ovis aries

Dark coat colour is associated with large size, which is heritable and positively correlated with fitness; however, the frequency of dark sheep has decreased over time. This unexpected microevolutionary trend is explained by genetic linkage between the tyrosinase-related protein 1 (*TYRP1*) pigmentation gene underlying the colour polymorphism and QTLs with antagonistic effects on size and fitness. As a consequence, homozygous dark sheep are large, but have reduced fitness relative to phenotypically indistinguishable dark heterozygotes and to homozygous light sheep[30].

Eda* in threespine sticklebacks, *Gasterosteus aculeatus

A derived allele at the *Ectodysplasin* (*Eda*) locus causing reduced armour plating has been fixed repeatedly after marine sticklebacks colonized freshwater from the sea (FIG. 2Aa–Ea). An experiment that introduced marine sticklebacks to freshwater ponds confirmed positive selection on the *Eda* allele following armour development, but also negative selection occurring before plates had completed development. This demonstrates not only that countervailing effects diminish the advantage of the allele over the whole lifespan but also that the evolution of armour is influenced by these effects[77].

FRI* in thale cress, *Arabidopsis thaliana

Although early flowering plants produce more fruits under spring conditions, and nonfunctional alleles of the *FRIGIDA* (*FRI*) gene are associated with early flowering, variation at *FRI* is not associated with fitness. Nonfunctional *FRI* alleles have negative pleiotropic effects on fitness by reducing inflorescence structures, which reduces the adaptive value of the locus and allows the maintenance of alternative life history strategies[75].

CHSD* in common morning glory, *Ipomea purpurea

Natural selection from pollinators favours white-flowered variants at the chalcone synthase D (*CHSD*) locus (the photograph illustrates colour polymorphisms in the *Ipnomea* genus). Despite this apparent advantage, the white colour allele (*a*) remains rare in natural populations. Field experiments demonstrate that the selection advantage from pollinators is opposed by pleiotropic effects of the *a* allele: this is manifested as reduced survival from germination to flowering that, along with inbreeding depression, more than compensates for the pollinator advantage[116].

Linkage disequilibrium
(LD). Nonrandom association of alleles at two or more loci. The pattern and extent of linkage disequilibrium in a genomic region is affected by mutation, recombination, genetic drift, natural selection and demographic history.

Fixation index
(F_{ST}). A measure of population subdivision, indicating the proportion of genetic diversity found between populations relative to the amount of genetic diversity found within populations.

Ascertainment bias
A consequence of collecting a nonrandom subsample with a systematic bias, so that results based on the subsample are not representative of the entire sample.

Pleiotropic
The effect of a gene on more than one phenotypic trait.

Filling the phenotype gap. One type of laboratory evolution study in which phenotypic information is often explicitly linked to fitness differences is studies that test for antibiotic resistance (reviewed in REF. 40). For example, there are over 900 unique mutations in 36 genes that confer resistance to several different antibiotics in *Mycobacterium tuberculosis*[41]. Other studies of adaptation in *Escherichia coli* fill the gap in phenotypic information by identifying both functional changes to specific proteins[42] and global changes in gene regulation[43,44]. For example, site-directed mutagenesis has been used to characterize the fitness effects of individual mutations detected by whole-genome sequencing[44] during adaptation to a glycerol-based growth medium. A study of the bacterium *Pseudomonas fluorescens* and its viral parasite ϕ2 also links genome-wide patterns of variation in both genomes to phenotypic adaptation and, in doing so, provides an empirical test of the Red Queen hypothesis that coevolution in interacting species drives molecular evolution[45]. Finally, experimental genomic approaches may also be applicable to sexually reproducing systems with larger genomes: Burke *et al.*[46] resequenced *Drosophila melanogaster* populations that had experienced >600 generations of laboratory selection for accelerated development. Several dozen genomic regions were identified that show a strong association with a 20% reduction in development time from egg to adult in selected populations. This represents one of the first genome-wide experimental evolution studies conducted in an animal.

Strengths and limitations of experimental evolution studies. A great strength of the laboratory studies described above is the ability of experimenters to limit variation in as many environmental variables as possible and thereby test the genetic response to one or a few selection pressures. This is one of the most powerful ways of investigating mechanisms and targets of selection, but it comes at a cost. In stripping the dimensions of an environment down to just a few variables, many of the important ecological interactions that drive adaptation in natural contexts are removed. The prevalence of genotype–environment[47] and genotype–genotype–environment[48] interactions in nature suggests that adaptive evolution is often context-dependent. Thus, laboratory environments may result in misleading inferences about which genes contribute to adaptation and the mechanisms through which they act. In addition, unanticipated selective effects caused by culturing organisms in highly artificial laboratory conditions can bias the interpretation of adaptive effects. For example, genetic responses to culturing practices in stocks of *Drosophila* species have had large effects on longevity: the result is that studies purporting to have found genes that increase survival (for example, *mth*, *Chico*, *Lin-12/Notch* repeat, *Indy* and superoxide dismutase (*Sod*)) may instead be genes that simply reverse these laboratory effects, rather than genes that are likely to contribute to adaptation in wild populations (reviewed in REF. 49). It is clear that laboratory evolution can provide a useful way of testing mechanistic hypotheses about the genetics of adaptation, but it is also important to consider the potential artefacts and confounding effects of these artificial systems before extending conclusions to more complex organisms (with larger genomes) or to natural contexts.

Experimental studies of adaptation in nature

Laboratory experiments can provide rigorous tests of the plausibility of adaptive hypotheses, but they cannot determine their importance in natural populations[50]. Of course, no field experiment can investigate all of the factors that are important for adaptation, but the hope is that by conducting empirical studies in nature, the single or few factors that are explicitly manipulated and tested will be subject to all of the interactions that occur under natural conditions. Pioneering studies of natural selection in wild populations demonstrated

Box 3 | **The rise, fall and resurgence of an adaptation icon**

The rise

One of the best-known examples of contemporary adaptive evolution is industrial melanism in English peppered moths, *Biston betularia*. The 'typical' peppered moth is white and speckled. The appearance of a melanic form, carbonaria, coincided with the peak of the Industrial Revolution. The first carbonaria form was recorded in Manchester in 1848, and in the ensuing 47 years climbed to 98% frequency. Tutt hypothesized that the increase in melanism resulted from bird predation favouring more cryptic carbonaria on sooty versus unpolluted resting sites on bark (reviewed in REF. 118). Haldane[119] calculated that to explain the rapid change in frequency, carbonaria would have to be 1.5 times as fit as the typical form. It took several decades before Tutt's hypothesis was empirically tested. In a series of field experiments, Kettlewell[120] showed that conspicuous moths were attacked more frequently than those that matched their background. Furthermore, as expected, carbonaria frequencies declined following the Clean Air Acts of 1956 and 1968 (REF. 121). The reciprocal nature of Kettlewell's results, conducted in polluted and in non-polluted woodlands, as well as those from >30 other experiments, in conjunction with strong correlations between carbonaria frequency and industrial pollution (reviewed in REF. 118), made the case highly persuasive.

The fall

Around the turn of the century, the reputation of this example declined as criticism about the design of predation experiments and ill-founded claims of fraud against Kettlewell raised uncertainty about the role of natural selection as a cause for changes in moth colour. In a review of Majerus's 1998 book *Melanism: Evolution in Action*[122], which outlines many of the defects of earlier experiments, Coyne concluded "for the time being we must discard Biston as a well-understood example of natural selection in action…"[123]. Although Majerus[122] had pointed out some weaknesses in Kettlewell's experimental design, he never doubted the importance of predation. Coyne[123], on the other hand, contested the strength of the data supporting predation as the definitive agent of selection. This case highlights the difficulty in identifying precise selective agents acting on phenotypic traits in natural populations.

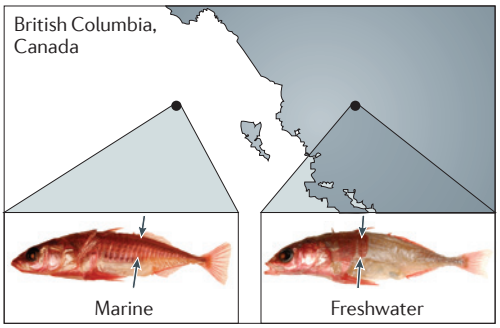
The resurgence

Two recent developments provide new evidence in support of Tutt's original adaptive hypothesis. First, Majerus conducted the largest field experiment (4,864 released moths) ever carried out on industrial melanism. This study shows that differential bird predation is sufficient to explain changes in the frequency of the morphs observed over a 6-year period (data recently re-analysed by L. M. Cook, B. S. Grant, I. J. Saccheri and J. Mallet, unpublished observations). Second, the genetic basis of the colour polymorphism has been identified: melanism has a single mutational origin with a signature of recent strong selection. Van't Hof and colleagues[124] genetically mapped the locus determining colour morph to a 200 kb region and showed a statistical association between a single SNP and colour in a geographically diverse sample, which was consistent with the presence of a single major-effect allele in all carbonaria forms tested. This molecular evidence supports nineteenth century records that show the spread of carbonaria morphs across Britain emanated from a single location in Manchester. Future research can now take advantage of these new genetic data to conduct selection experiments on carbonaria sequence variants, which could reveal pleiotropic effects of the gene that contributed to patterns of evolutionary change.

Environment–phenotype association

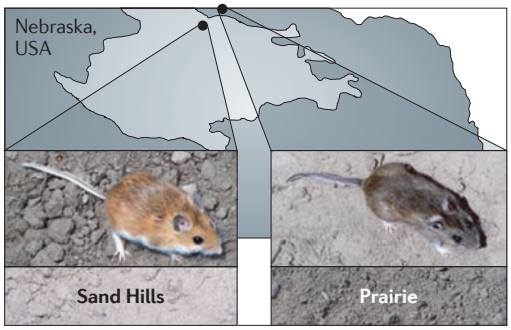
A Sticklebacks

Aa Armour plating in stickleback fish



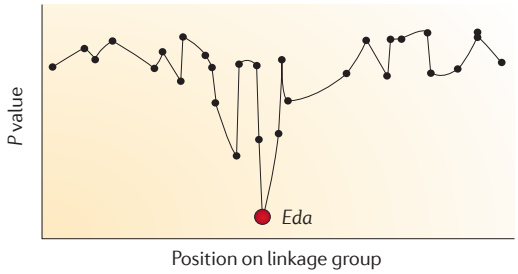
B Deer mice

Ba Coat colour in deer mice

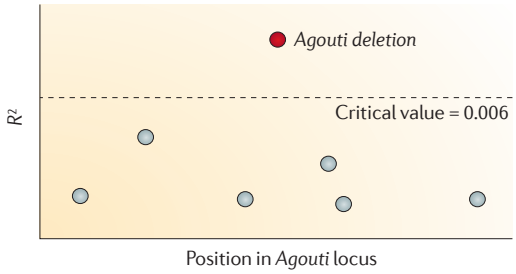


Identifying the gene or mutation

Ab LD mapping



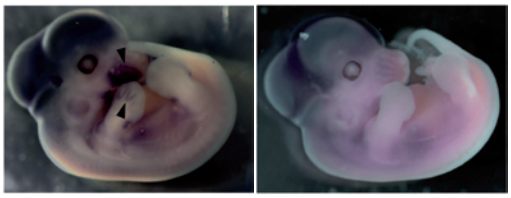
Bb Association mapping



Test of gene function

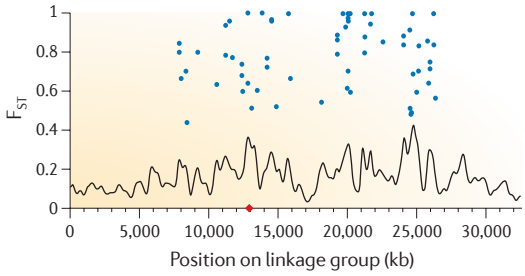
Ac Gain-of-function

Bc Loss-of-function

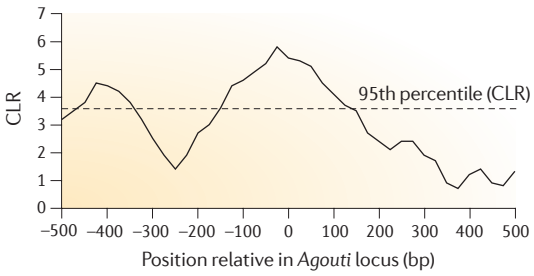


Signatures of selection

Ad Genetic differentiation

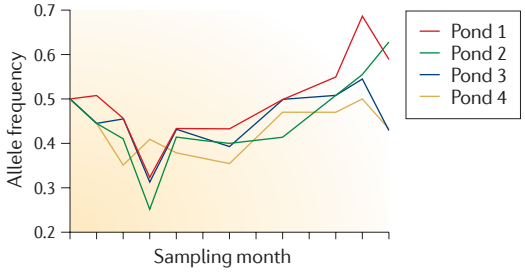


Bd Composite likelihood ratio test

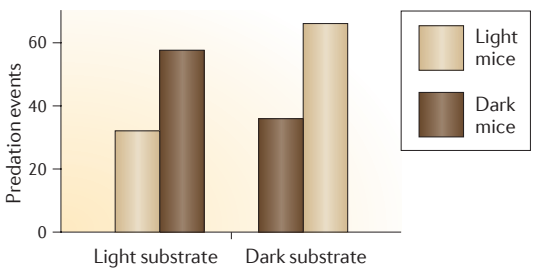


Experimental tests of selection

Ae Colonization experiment



Be Predation experiment



◀ Figure 2 | **Examples of using complementary ecological and molecular approaches to investigate the genetic basis of adaptation. Aa** | Marine threespine stickleback fish typically have 30 or more lateral armour plates (left fish, indicated by the arrows), whereas most freshwater stickleback have between zero and nine plates (right fish, indicated by the arrows). **Ab** | Linkage disequilibrium (LD) screening shows large differences in allele frequency between completely plated and low-plated fish at a microsatellite marker located in an intron of the *Ectodysplasin* (*Eda*) gene (see also TABLE 1 and Supplementary information 1 (table)). **Ac** | Introduction of an *Eda* cDNA construct stimulates lateral plate formation. The fish on the left is a control, low-plated morph; the fish on the right is a sibling from the same clutch after introduction of an *Eda* transgene. Six extra plates have formed in the transgenic fish (indicated by the arrow). **Ad** | The fixation index (F_{ST}) between marine and freshwater populations on the linkage group containing *Eda* (indicated by the red diamond). Blue circles represent SNPs with F_{ST} values at which population differentiation is significant at the level of $\alpha = 10^{-20}$. **Ae** | Concordant changes in the frequency of the low allele at *Eda* in natural populations of marine stickleback transplanted to four replicate freshwater ponds. **Ba** | Deer mice inhabiting the Nebraska Sand Hills, USA, which has light-coloured substrate, have a lighter dorsal coat colour than nearby ancestral populations inhabiting prairie habitats with darker soil. **Bb** | Correlation (R^2) between genotype at polymorphic sites within the *Agouti* locus and coat colour phenotype in an admixed population from the edge of the Sand Hills. **Bc** | *In situ* hybridizations for *Agouti* in deer mouse embryos. Wild-type embryos express *Agouti* (arrowheads point to expression in the whisker plate and limbs) and will develop into adults with a light-coloured belly and banded dorsal hairs; in an *Agouti* knockout (right), expression is not detected and resulting adult mice are uniformly pigmented with completely eumelanic hairs. **Bd** | Composite likelihood ratio (CLR) as a function of distance from candidate polymorphism within *Agouti*. Values above the line reject the neutral model, which is consistent with recent, strong selection at these sites. **Be** | Increased attack rates from owls on conspicuous mice relative to cryptic mice on both light and dark substrates in a predation experiment. Fish images in panel **Aa** are reproduced, with permission, from REF. 77 © (2008) American Academy for the Advancement of Sciences (AAAS). Panel **Ab** is reproduced and panel **Ac** is modified, with permission, from REF. 29 © (2005) AAAS. Panel **Ad** is modified from REF. 125. Panel **Ae** is modified, with permission, from REF. 77 © (2008) AAAS. Panel **Ba**, **Bb** and **Bd** are modified, with permission, from REF. 36 © (2009) AAAS. Panel **Bc** is reproduced from REF. 126. Panel **Be** is calculated from data in REF. 127.

the feasibility of linking variation in phenotypic traits with fluctuations in the environment, selection and evolution[51,52]. However, the most powerful test of cause-and-effect relationships is a planned experiment with manipulative treatments.

A few examples of this approach include controlled introductions of wild Trinidadian guppies to high and low mortality rate environments[53,54], *Anolis* lizards to lizard-free islands in the Caribbean[55,56], stick insects between host plants[57,58] and threespine sticklebacks to different competition and predation regimes in freshwater ponds[59,60] (FIG. 2Aa–Ea). All of these research programmes have used manipulations that mimic events frequently occurring in nature to make direct inferences about adaptation. However, these studies have largely been conducted at the phenotypic level without knowledge of the specific genes that influence the functional traits under selection (that is, colour patterns and life histories in guppies, limb morphology in lizards, colour patterns in stick insects and trophic morphology in sticklebacks) or how this genetic variation evolves in response to the experimental treatments. In this section, we consider recent studies that have used the same powerful experimental framework described above, but extend inferences about adaptation to the genetic level.

Eumelanic
Consisting of brown to black pigments (that is, eumelanin) found in vertebrate hair, feathers or scales.

Recombinant inbred lines
(RILs). A population of fully homozygous individuals obtained through the repeated self-fertilization of an F1 generation hybrid.

Backcross lines
A population of individuals obtained by crossing a hybrid with one of its parents or to an individual that is genetically similar to its parent in order to generate offspring with genetic identity closer to that of the parent.

QTL mapping of fitness. One approach that has been successfully used to investigate the genetic basis of adaptation in nature is to conduct reciprocal transplants of hybrid individuals between two environments and subsequently carry out mapping experiments to identify the genetic regions that are associated with fitness in each habitat. Such studies provide selection differentials for segregating traits and then use the covariation between phenotype and genotype to infer the fitness effects of alleles at these loci. Thus, these studies do not directly measure selection at the genetic level, but rather they identify QTLs that are correlated with phenotypic traits that have an impact on fitness in natural conditions (that is, they map rather than measure fitness at the genetic level). Nevertheless, these associations serve as an excellent starting point for future work to follow genetic changes at these loci explicitly.

The approach has so far been used almost exclusively in self-fertilizing plants, which are convenient for reciprocal transplant experiments. In particular, recombinant inbred lines (RILs) are ideal for field studies of selection because siblings within a line are genetically identical and can be replicated across environments[61]. For example, RILs have been used to investigate local adaptation in coastal and inland races of yellow monkey flower (*Mimulus guttatus*)[62,63], in contrasting temperature and photoperiod environments in a rockcress plant (*Boechera stricta*)[64] and in contrasting moisture environments in annual grasses *Avena barbata*[65] and *Hordeum spontaneum*[66]. Second-generation backcross lines were similarly exploited to investigate natural selection on viability and fertility traits associated with QTLs in *A. thaliana*[67,68].

A complementary approach for conducting selection experiments is to genotype individuals in natural populations in which individual fitness has already been measured. For example, existing information on relationships among individuals (such as pedigree information) permits statistical tests of the effects of candidate alleles in various genetic backgrounds or the effects of alternative alleles among family members[30,69]. A comprehensive review by Morjan and Rieseberg[6] retroactively applies the general conceptual approach of linking phenotypic selection with QTLs in a wide diversity of systems, including animals, bacteria, fungi and plants. Interpretation of these studies must consider several sources of bias in estimating the number of loci and their effects on traits under selection[70,71]. Nevertheless, a common finding appears to be that the QTLs associated with selection vary across environments and that the effects of selection are highly dependent on genetic background.

Directly measuring selection on genes. To establish fully that an allele is adaptive, it is important to demonstrate that known agents of natural selection on focal traits cause allele frequencies at the locus to change in the expected direction. This closes the genotype–phenotype–fitness loop and enables direct inferences about the adaptive importance of a gene (FIGS 1,2). For example, Weinig *et al.*[72] calculated selection coefficients from the relative fitness of different allelic classes at a number of survivorship and fecundity QTLs in

Table 1 | **Examples of genes commonly described as adaptive**

Gene	Phenotypic effect	Molecular change and functional effect	Putative agent of selection	Evidence for selection	Estimate of selection	Control for genetic background?	Example refs
Ace1	Resistance to insecticide in mosquitoes (*Culex pipiens*)	Single amino acid substitution associated with loss of sensitivity to organophosphates and carbamates	Insecticide	Frequency pattern in space (clines across treated and non-treated areas) and time (seasonal and year-to-year variation) are consistent with a migration–selection process (P and G). Fitness differences measured in the laboratory (P and G)	Selection coefficient for G119S mutation ranges from 0 to 0.4 in treated areas and –0.1 to –0.5 in non-treated areas, depending on the season (G)	Yes (field estimate: background genotype randomized and indirect selection caused by a nearby resistance locus (*Ester*) taken into account. Laboratory estimates: controlled for using repetitive backcrosses)	128, 150–153
CHSD	Flower colour in morning glory (*Ipomoea purpurea*)	Transposon insertion in the gene that codes for chalcone synthase, the first enzyme of the anthocyanin pathway	Flower colour under different pollination syndromes	Field experiments demonstrate a transmission advantage for white-flower allele (*a*) owing to increased self-fertilization and lack of pollen discounting, but also that pleiotropic effects reducing survival cause a net selective disadvantage (G)	Selection coefficient against amino acid = 0.05 (G)	Yes (background genotype randomized)	116,129
ebony	Abdominal melanism in *Drosophila melanogaster*	Five CRMs alter expression of an enzyme in the biogenic amine synthesis pathway that affects pigmentation	Thermal regulation at different temperatures; differential absorption of heat from solar radiation	Correlation between elevations and pigmentation; patterns of nucleotide variation in the *ebony* region consistent with a partial selective sweep (P and G)	Not calculated	No	130,131
Eda	Lateral plate armour in threespine stickleback (*Gasteroseus aculeatus*)	Haplotype associated with changes to ectodermal signalling pathway that affects development of dermal plates	Predation regime (insect versus fish); growth rate (trade-off between armour and growth)	Correlations between armour and environment type (P and G); significant changes in allele frequency following experimental introduction to freshwater and association between genotype and growth rate (P and G)	Selection coefficients = 0.50 against the low (*L*) allele in juveniles, 0.52 against the complete (*C*) allele in adults (G)	Yes (rare genetic variants that segregate at *Eda* in a homogenous genetic background, but closely linked genes likely to co-segregate)	29,31, 77,132, 133
FRI	Flowering time in *Arabidopsis thaliana*	Deletion polymorphism disrupting ORFs associated with early flowering times	Seasonal timing based on climatic conditions	Overwinter survival and seed production differences between genotypes with different vernalization times (P and G)	Selection gradients = –0.43 in autumn cohort, 0.66 in spring cohort (P)	Yes (inbred accession lines)	75,134
Hb	Aerobic metabolism in deer mice (*Peromyscus maniculatus*)	Combined effects of 12 amino acid substitutions alter oxygen affinity binding	Physiological performance at different altitudes	Differential survival of mice with different thermogenic capacities (P); nucleotide divergence and LD patterns between high- and low- altitude populations, indicative of natural selection (G)	Selection gradient = 0.52 (P)	No	135,136, 137
LCT	Ability to digest lactose in adult humans	Three CRMs associated with transcription level of the enzyme LPH, which hydrolyses lactose into sugars	Lactase persistence in pastoralist populations	LD patterns consistent with a selective sweep (G)	Selection coefficient = 0.04–0.097 (G)	No	138

Table 1 (cont.) | **Examples of genes commonly described as adaptive**

Gene	Phenotypic effect	Molecular change and functional effect	Putative agent of selection	Evidence for selection	Estimate of selection	Control for for genetic background?	Example refs
Mc1r	Coat pigmentation in beach mice (*Peromyscus polionotus*)	Single amino acid substitution reduces receptor signalling, affecting melanogenesis	Visual predation against conspicuous phenotypes	Correlations between pigmentation and substrate colour (P); higher frequency of attacks on conspicuously coloured clay model mice (P); clinal models of selection based on phenotype distributions	Selection index = 0.50 (P — on plasticine mouse models); selection coefficient = 0.07–0.21 depending on body region (P)	No	28,139, 140,141
MCPH1	Brain size in humans	Loss-of-function mutations alter the proliferation and differentiation of neuroblasts during neurogenesis, which can affect cranial capacity	Cognitive ability	Nucleotide diversity and LD patterns consistent with a recent selective sweep (G)	Not calculated	No	141,142
Scn4a	Resistance to tetrodotoxin (TTX) in garter snakes	Amino acid substitution associated with binding of TTX to *Scn4a*, affecting muscle contraction after ingestion	Defensive toxin in prey (rough-skinned newts)	Correlations between snake TTX resistance and levels of newt predation and newt TTX level (P)	Not calculated	No	143,144
PGI	Flight performance in *Colias* butterflies	Single amino acid substitution associated with glycolysis of nectar used to fuel flight	Optimal sugar metabolism depending on environmental temperature	Nucleotide diversity and LD patterns consistent with a recent selective sweep (G); differences in survival, male mating success and female fecundity among common genotypes (G)	Selection coefficient against genotype with lowest relative fitness = 0.82 (males), 0.83 (females) (G)	No	145,146

An expanded version of this table is available in Supplementary information 1 (table). A pollination syndrome is a suite of flower traits that have evolved in response to natural selection imposed by different pollen vectors, which can be abiotic or biotic. Pollen discounting describes the situation in which the transmission of genes through pollen is reduced below a random expectation. *Ace1*, acetylcholinesterase type 1; *CHSD*, chalcone synthase D; CRM, *cis*-regulatory mutation; *Eda*, *Ectodysplasin*; *FRI*, *FRIGIDA*; *Hb*, Haemoglobin; *LCT*, lactase; LD, linkage disequilibrium; LPH, lactase-phlorizin hydrolase; *Mc1r*, melanocortin 1 receptor; *MCPH1*, microcephalin 1; *PGI*, phosphoglucose isomerase; *Scn4a*, sodium channel protein type 4 subunit (also known as $Na_v1.4$).

Likelihood ratio
The ratio of how many times more likely the observed data are under one model than the other, computed for all pairwise models.

Identical-by-descent lines
A population of individuals sharing identical copies of the same ancestral allele.

Near-isogenic lines
A population of individuals that differs from its parent in only one genomic location, typically a QTL of interest. Using markers that are diagnostic for that QTL, backcrosses are made to the parent until the entire genome of the line is exactly like the parent except in the region around the marker locus.

A. thaliana that had been grown under different seasonal conditions. Similarly, a study of natural selection for salt tolerance in wild sunflower hybrids explicitly tested how specific QTLs have an impact on fitness through increased calcium uptake coupled with greater exclusion of sodium and related mineral ions[73].

In both of these studies, selection is calculated using the nearest molecular marker for each QTL (that is, the marker that is closest to the likelihood ratio peak) as a surrogate for the gene under selection. The limited number of recombination events occurring in these backcross lines[73] and RILs[72] prevents identification of a causal gene (or genes), and thus it is possible that the aggregate selective effects of closely linked genes appear as a single locus with a larger selection coefficient[74]. Selection experiments that are designed to measure fitness differences between genotypes at candidate loci obviate these problems, but they risk missing other important genes that affect selection on the focal trait. A compromise solution was taken in an experiment that focused on fitness effects of the candidate flowering time gene *FRIGIDA* (*FRI*), but also tested for the influence of genetic background at the interacting gene *FLOWERING LOCUS C* (*FLC*)[75]. In concordance with other studies, the direction of selection on *FRI* was found to vary owing to genotype–genotype interactions (allelic variation at *FLC*) and genotype–environment interactions (season of germination).

Controlling for genetic background. The prevalence of genotype–genotype interactions highlights the importance of controlling for genetic background when attempting to make inferences about the fitness consequences of a particular allele. This can be accomplished by using populations in which the alleles of interest are segregating within either a randomized genetic background (for example, in the F2 generation, in more advanced generation intercrosses, in RILs and in identical-by-descent lines) or a homogenous genetic background (for example, in near-isogenic lines (NILs)) (TABLE 2, reviewed in REF. 76). Ideally, natural variation rather than artificial crosses could be used to test the effects of putatively adaptive alleles. This may be possible if the alleles of interest in a locally adapted population were selected from standing genetic variation that is still segregating in the ancestral population. In this case, the fitness effects of alternative alleles can be tested in the genetic background of the ancestral population. For example, researchers measured selection on naturally

Table 2 | **Advantages and disadvantages of using different types of genetic variants in selection experiments**

Genetic variant	Advantage	Disadvantage	Example reference
Random sample from natural populations in derived habitat or distinct habitats	Represents natural variation available to selection in the wild. High genetic diversity	Difficult to isolate selection on specific traits or genes	30
Phenotypic extremes	Represents an extreme subset of natural variation available to selection in the wild. Increases fitness differences between variants and thus the likelihood of detecting selection	Organisms carry alternate alleles at most loci contributing to trait, making it difficult to isolate selection on different genes contributing to a polygenic trait	145
Rare genetic variants in ancestral or alternate environment type	Represents an extreme subset of natural variation available to selection in the wild. Controls for genetic background to allow isolation of selection on gene of interest	Depending on levels of linkage disequilibrium, may be difficult to distinguish between effects of focal locus and closely linked genes	77
F2 or backcross lines	Contain alleles segregating from both parental types, so all traits can be analysed at the same time. Can detect heterosis and dominance effects	Power to isolate selection on individual loci highly dependent on sample size and genetic architecture of traits. Limited genetic diversity	73
Recombinant inbred lines (RILs)	Simple segregation patterns facilitate detection of epistasis and genotype–environment interactions. Populations are almost completely homozygous, making it possible to replicate genotypes among several environmental treatments. Only one individual needs to be genotyped per population	Fairly course genetic mapping. Limited genetic diversity	63
Near-isogenic lines (NILs)	Contains only small segments of donor genome, which removes confounding effects of other loci segregating in the population. Populations are almost completely homozygous, so it is possible to replicate genotypes among several environmental treatments. Should facilitate studies of multigenerational selection (by tracking changes in allele frequency at single genes). Only one individual needs to be genotyped per population. Increased power to detect small-effect loci	Size of introgression depends on number of backcross generations, making fine mapping time-consuming. Limited genetic diversity	147
Transgenic lines	Contains only a single gene from a donor genome, which removes confounding effects of other loci segregating in the population. Should facilitate studies of multigenerational selection (by tracking changes in allele frequency at single genes)	Limited genetic diversity. Challenging to obtain within-line replicates	148
Knockouts	Exact targeting of selective consequences of loss-of-function of a single locus	Limited to consequences of mutations conferring fitness effects resulting from complete loss-of-function	149

Heterosis describes a greater fitness in hybrid individuals than in either of the parental types.

occurring variants at the *Ectodysplasin* (*Eda*) locus, which underlies differences in armour plates in threespine sticklebacks[77] (FIG. 2Aa–Ea). A derived low allele causing reduced plate number has been repeatedly fixed after marine stickleback colonized freshwater from the sea, in which the ancestral (complete) allele predominates[29]. They introduced rare marine sticklebacks carrying both alleles to freshwater ponds and tracked allele frequencies over a generation. This approach controlled for background genetic variation because heterozygous marine fish were genetically ancestral with the exception of their *Eda* genotype and tightly linked loci.

The rise of experimental genomics

Recent sequencing and computational advances now allow genome-wide analyses to be conducted in a wide range of organisms[78,79]. The importance of these technological advances for population and ecological genomics has been well recognized[80–82]. For example, wider and deeper genomic coverage can help to distinguish between patterns that result from locus-specific selective effects versus genome-wide effects, such as demography[21,83], and provide greater power to identify the types of mutations that are involved in adaptation[84–86]. Perhaps less appreciated is the role that next-generation sequencing can play in experimental studies of selection. In particular, carefully designed experiments can be used to test hypotheses about mechanisms of selection, which can help to determine how and why allele frequencies change. For example, the stickleback experiment discussed above was designed to test whether the *Eda* low allele confers a selective advantage by increasing growth rate (owing to the reduced burden of producing armour plates). The results partly confirmed the hypothesized mechanism but also generated a surprising pattern of selection when it was least expected: before plates had completed development[77]. This implies that *Eda* or a closely linked gene has effects on other traits that are under selection, and it underscores how field experiments can lead to the discovery of unanticipated fitness

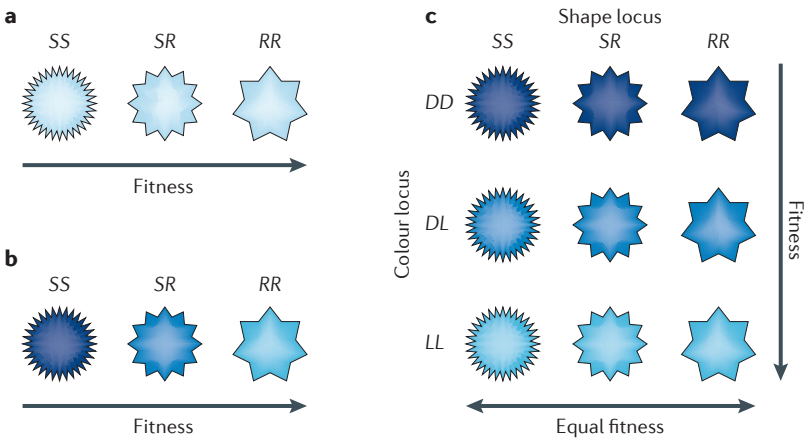


Figure 3 | **How understanding genetic architecture can inform us about selection. a** | Experiments show that an allele (*R*) affecting flower shape is associated with increased fitness. **b** | It is subsequently shown that this allele is also associated with flower colour, making it unclear whether the fitness effects of the allele are due to its role in controlling flower shape or flower colour. If the correlation between the allele and flower colour is due to linkage, then fine-scale mapping or near-isogenic lines could be used to show that the shape locus segregates independently of the colour locus, thereby revealing the true target of selection. **c** | If the correlation between the shape locus and the colour locus is because the same allele causing shape differences directly affects colour through pleiotropy, then these approaches will not be successful, and fitness associations will be the same as in part **b**. Elucidating the confounding patterns caused by complex genetic architecture can enable better predictions of the fitness consequences of individual genes.

effects. However, the unexpected patterns of selection also reveal the limitations of assaying fitness associated with allelic variation at only a single gene. Below, we discuss some of the ways that genome-wide selection experiments can improve understanding of the genetics of adaptation.

Unravelling the genetic changes involved in fitness effects. By definition, a genome-wide approach does not consider the effects of genes in isolation, but rather considers them within their surrounding genetic milieu. This permits consideration of emergent fitness effects that are not predicted from experiments focusing on a single or a few candidate genes, but are essential for understanding the adaptive importance and future evolutionary trajectory of particular alleles. Increased marker density can lead to new insights by unravelling genetic correlations that may be driving unexpected selection patterns in experiments; for example, high-density markers could distinguish between the effects of closely linked loci versus pleiotropic effects of single alleles[72,87] (FIG. 3). Moreover, because it is becoming easier to position markers reliably in a genome, independent estimates of selection can be obtained for loci in increasingly narrow genomic windows[43,46,88]. If an allele shows unexpected fitness effects based on *a priori* knowledge of function, then demonstrating that linked sites do not show the same patterns of selection can help to strengthen the case for pleiotropy. Unfortunately, no amount of sequencing can help to determine the pleiotropic effects of an allele if the relevant phenotypic trait is not, or cannot, be measured. This highlights the importance of phenomics in understanding genotype–phenotype–fitness maps; as our genomic resources increase, we need to make sure that our phenotypic knowledge keeps pace[89].

Detecting weakly selected loci. An additional advantage of tracking genome-wide patterns in selection experiments is that it can provide an unbiased way of detecting loci that are under selection. Genome scans of historical selection detect loci under selection by identifying statistical outliers that are exceptional in their degree of genetic differentiation[90,91]. By contrast, an experimental genomic approach detects loci under selection by identifying genes that show a consistent response across experimental replicates. In this latter case, consistency in the direction of allele-frequency change across replicates, as well as the magnitude of the change itself, provides information that can be used to obtain refined estimates of selection. Furthermore, genetic drift can be directly modelled in experiments, because the numbers of individuals pre- and post-selection are known, as well as the pre- and post-selection allele frequencies. Thus, loci showing changes that are greater than would be expected under drift are likely to be under selection, and the power to distinguish between drift and selection increases when replication is high and experimental controls are included.

Measuring genome-wide fitness effects. Experimental genomics can help not only to understand the adaptive importance of specific genes but can also help to address questions about the broader effects of natural selection on the genome as a whole (for example, REF. 91). For instance, does rapid adaptation to a new environment typically use a small subset or large swathes of the genome? How does selection on these adaptive loci translate to genome-wide patterns of polymorphism at linked sites? It is well appreciated that different combinations of genes can lead to the same phenotype[92,93], and the few studies that have investigated interacting QTLs have found strong evidence for epistatic fitness effects[72,73,75,92], which is not surprising given the high prevalence of epistasis observed in genotype–phenotype relationships (reviewed in REF. 92). Expanding these pioneering studies from a handful of QTLs and genes to genome-wide analyses of selection in action will reveal how genetic architecture influences adaptation and vice versa.

We expect whole-genome approaches to become routine in future field studies of selection, as should rigorous standards of evidence for adaptation that account for the complex relationships between genetic architecture, organismal performance and fitness.

Conclusions

In this Review, we have discussed a diversity of approaches that have contributed to the tremendous progress in our understanding of the genetics of adaptation. It is increasingly possible to identify the genes and genetic pathways that underlie fitness-related traits and detect molecular evidence that they have been influenced by natural selection, even in non-model species and in relevant ecological contexts. This

has led to a proliferation of studies that investigate the genetics of adaptation by identifying genes that underlie fitness-related traits or show molecular signatures of selection.

Although frequent use of the qualifier 'putatively' in the results section of manuscripts suggests that many authors recognize that these approaches alone are often insufficient for showing that a particular allele is adaptive, this has not prevented strong claims about the adaptive importance of genes in the subsequent discussion sections (critiqued in REF. 8). There has been deserved attention brought to testing the functional effects of mutations on proteins and genes on phenotype[76,81,94]; we agree and further extend this call for rigor to tests of ecological function and fitness.

In highlighting the difficulty of constructing definitive arguments in favour of the adaptive nature of specific loci, our aim is to temper adaptationist story-telling, but not to dampen excitement about the applicability of new genomic and statistical techniques for studies of adaptation. Clearly, we are making great progress on both fronts. Instead, we wish to encourage the use of experiments, when possible, that determine the fitness consequences of genome-wide variation in a relevant ecological context. These studies will lead to a greater understanding of the ultimate and proximate mechanisms that drive adaptation and permit investigation of difficult conceptual issues (for example, the interacting and often confounded patterns of selection caused by pleiotropy, linkage and epistasis). Key to this research programme is the integration of different methods and knowledge from ecology, evolution and genomics that target several functional and biological levels (for example, mutations, genes, phenotypes, individuals and populations). This interdisciplinary approach will be invaluable for detecting (and avoiding) molecular spandrels, thereby allowing more precise tests of predictions from basic evolutionary theory.

1. Gould, S. & Lewontin, R. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B* **205**, 581–598 (1979).
This is a much-cited paper that criticized evolutionary biologists for adaptationist 'story-telling' in the absence of rigorous tests of alternative possibilities.
2. Arnold, S. Morphology, performance and fitness. *Am. Zool.* **23**, 347–361 (1983).
3. Lande, R. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**, 402–416 (1979).
4. Lande, R. & Arnold, S. J. The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226 (1983).
5. Kingsolver, J. G. *et al.* The strength of phenotypic selection in natural populations. *Am. Nat.* **157**, 245–261 (2001).
6. Morjan, C. L. & Rieseberg, L. H. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Mol. Ecol.* **13**, 1341–1356 (2004).
This is an innovative literature review that combines estimates of selection on phenotypic traits with the trait variance explained by underlying QTLs to obtain a rough measure of the selection acting on individual loci.
7. Siepielski, A., Dibattista, J. & Carlson, S. It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecol. Lett.* **12**, 1261–1276 (2009).
8. Nielsen, R. Adaptationism — 30 years after Gould and Lewontin. *Evolution* **63**, 2487–2490 (2009).
9. Thornton, K. R., Jensen, J. D., Becquet, C. & Andolfatto, P. Progress and prospects in mapping recent selection in the genome. *Heredity* **98**, 340–348 (2007).
10. Reznick, D. N. & Ghalambor, C. K. Selection in nature: experimental manipulations of natural populations. *Integr. Comp. Biol.* **45**, 456–462 (2005).
11. Tauber, E. *et al.* Natural selection favors a newly derived timeless allele in *Drosophila melanogaster*. *Science* **316**, 1895–1898 (2007).
12. Sezgin, E. Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in *Drosophila melanogaster*. *Genetics* **168**, 923–931 (2004).
13. Umina, P., Weeks, A., Kearney, M., McKechnie, S. & Hoffmann, A. A. A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. *Science* **308**, 691–693 (2005).
14. Beaumont, M. A. Adaptation and speciation: what can Fst tell us? *Trends Ecol. Evol.* **20**, 435–440 (2005).
15. Beaumont, M. A. & Balding, D. J. Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* **13**, 969–980 (2004).
16. Vasemagi, A. The adaptive hypothesis of clinal variation revisited: single-locus clines as a result of spatially restricted gene flow. *Genetics* **173**, 2411–2414 (2006).
17. Jensen, J. & Bachtrog, D. Characterizing recurrent positive selection at fast evolving genes in *Drosophila miranda* and *Drosophila pseudoobscura*. *Genom. Biol. Evol.* **2**, 371–378 (2010).
18. Andres, A. *et al.* Targets of balancing selection in the human genome. *Mol. Biol. Evol.* **26**, 2755–2764 (2009).
19. Nielsen, R. *et al.* Darwinian and demographic forces affecting human protein coding genes. *Genome Res.* **19**, 838–849 (2009).
20. Bazin, E., Dawson, K. J. & Beaumont, M. A. Likelihood-free inference of population structure and local adaptation in a Bayesian hierarchical model. *Genetics* **185**, 587–602 (2011).
21. Thornton, K. R. & Jensen, J. D. Controlling the false-positive rate in multilocus genome scans for selection. *Genetics* **175**, 737–750 (2007).
22. Przeworski, M., Coop, G. & Wall, J. D. The signature of positive selection on standing genetic variation. *Evolution* **59**, 2312–2323 (2005).
23. Teshima, K. M., Coop, G. & Przeworski, M. How reliable are empirical genomic scans for selective sweeps? *Genome Res.* **16**, 702–712 (2006).
24. Baird, N. A. *et al.* Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3**, e3376 (2008).
25. Andolfatto, P. *et al.* Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Res.* **21**, 610–617 (2011).
26. Xie, W. *et al.* Parent-independent genotyping for constructing an ultra high-density linkage map based on population sequencing. *Proc. Nat. Acad. Sci. USA* **107**, 10578–10583 (2010).
27. Mackay, T. F. C., Stone, E. A. & Ayroles, J. F. The genetics of quantitative traits: challenges and prospects. *Nature Rev. Genet.* **10**, 565–577 (2009).
28. Hoekstra, H. E., Hirschmann, R. J., Bundey, R. A., Insel, P. A. & Crossland, J. P. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**, 101–104 (2006).
29. Colosimo, P. F. *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).
30. Gratten, J. *et al.* A localized negative genetic correlation constrains microevolution of coat color in wild sheep. *Science* **319**, 318–320 (2008).
This study demonstrates the difficulty of predicting the microevolutionary consequences of selection in the absence of knowledge about the genetic basis of fitness variation.
31. Steiner, C. C., Weber, J. N. & Hoekstra, H. E. Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol.* **5**, e219 (2007).
32. Rogers, S. M. & Bernatchez, L. The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae) species pairs. *Mol. Biol. Evol.* **24**, 1423–1438 (2007).
33. Nosil, P., Egan, S. P. & Funk, D. J. Heterogeneous genomic differentiation between walking-stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. *Evolution* **62**, 316–336 (2008).
34. Via, S. J. W. The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Mol. Ecol.* **17**, 4334–4345 (2008).
35. Todesco, M. *et al.* Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature* **465**, 632–636 (2010).
36. Linnen, C. R., Kingsley, E. P., Jensen, J. D. & Hoekstra, H. E. On the origin and spread of an adaptive allele in deer mice. *Science* **325**, 1095–1098 (2009).
37. Rokyta, D., Joyce, P., Caudle, S. & Wichman, H. An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. *Nature Genet.* **37**, 441–444 (2005).
38. Araya, C. L., Payen, C., Dunham, M. J. & Fields, S. Whole-genome sequencing of a laboratory-evolved yeast strain. *BMC Genom.* **11**, 88 (2010).
39. Betancourt, A. J. Genome-wide patterns of substitution in adaptively evolving populations of the RNA bacteriophage MS2. *Genetics* **181**, 1535–1544 (2009).
40. Maclean, R. C., Hall, A. R., Perron, G. G. & Buckling, A. The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nature Rev. Genet.* **11**, 405–414 (2010).
41. Sandgren, A. *et al.* Tuberculosis drug resistance mutation database. *PLoS Med.* **6**, e1000002 (2009).
42. Stanek, M., Cooper, T. & Lenski, R. E. Identification and dynamics of a beneficial mutation in a long-term evolution experiment with *Escherichia coli*. *BMC Evol. Biol.* **9**, 302 (2009).
43. Barrick, J. E. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).
This was one of the first studies to investigate the genome-wide consequences of adaptation in experimentally evolving populations.
44. Herring, C. *et al.* Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nature Genet.* **38**, 1406–1412 (2006).
45. Paterson, S. *et al.* Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275–278 (2010).
46. Burke, M. K. *et al.* Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* **467**, 587–590 (2010).
47. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* (Addison Wesley Longman, Harlow, UK, 1996).
48. Thompson, J. N. *The Geographic Mosaic of Coevolution*. (Univ. Chicago Press, Chicago, 2005).
49. Spencer, C. C. & Promislow, D. E. L. Genes, culture, and aging flies—what the lab can and cannot tell us about natural genetic variation for senescence. *Sci. Aging Knowl. Environ.* **2002**, pe6 (2002).

50. Endler, J. A. *Natural Selection in the Wild* (Princeton Univ. Press, Princeton, 1986).
51. Boag, T. & Grant, P. R. Intense natural selection in a population of Darwin's finches (Geospizinae) in the Galapagos. *Science* **214**, 82–85 (1981).
This important study addresses the challenge raised by Gould & Lewontin's 'spandrel' paper (reference 1) by showing that natural selection can be measured in action in the wild. This work served as a catalyst for other field researchers to estimate selection coefficients in natural populations.
52. Price, T., Grant, P. R. & Gibbs, H. L. Recurrent patterns of natural selection in a population of Darwin's finches. *Nature* **309**, 787–789 (1984).
53. Reznick, D. N., Shaw, F. H., Rodd, F. H. & Shaw, R. G. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* **275**, 1934–1937 (1997).
54. Reznick, D. N. & Bryga, H. Life-history evolution in guppies (*Poecilia reticulata*).1. Phenotypic and genetic changes in an introduction experiment. *Evolution* **41**, 1370–1385 (1987).
55. Losos, J. B., Warheit, K. I. & Schoener, T. W. Adaptive differentation following experimental island colonization in *Anolis* lizards. *Nature* **387**, 70–73 (1997).
56. Schoener, T. W., Spiller, D. A. & Losos, J. B. Predators increase the risk of catastrophic exinction of prey populations. *Nature* **412**, 183–186 (2001).
57. Nosil, P., Crespi, B. & Sandoval, C. Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature* **417**, 440–443 (2002).
58. Nosil, P. & Crespi, B. Experimental evidence that predation promotes divergence in adaptive radiation. *Proc. Nat. Acad. Sci. USA* **103**, 9090–9095 (2006).
59. Schluter, D. Experimental evidence that competition promotes divergence in adaptive radiation. *Science* **266**, 798–801 (1994).
60. Rundle, H. D., Vamosi, S. M. & Schluter, D. Experimental test of predation's effect on divergent selection during character displacement in sticklebacks. *Proc. Nat. Acad. Sci. USA* **100**, 14943–14948 (2003).
61. Mackay, T. F. C. The genetic architecture of quantitative traits. *Ann. Rev. Genet.* **35**, 303–339 (2001).
62. Hall, M., Lowry, D. & Willis, J. Is local adaptation in *Mimulus guttatus* caused by trade-offs at individual loci? *Mol. Ecol.* **19**, 2739–2753 (2010).
This paper discusses a detailed ecological transplant experiment that demonstrated that local adaptation is largely controlled by non-overlapping loci.
63. Lowry, D., Hall, M., Salt, D. & Willis, J. Genetic and physiological basis of adaptive salt tolerance divergence between coastal and inland *Mimulus guttatus*. *New Phyt.* **183**, 776–788 (2009).
64. Anderson, J., Lee, C. & Mitchell-Olds, T. Life history QTLs and natural selection on flowering time in *Boechera stricta*, a perennial relative of *Arabidopsis*. *Evolution* **65**, 771–787 (2010).
65. Gardner, K. M. & Latta, R. G. Identifying loci under selection across contrasting environments in *Avena barbata* using quantitative trait locus mapping. *Mol. Ecol.* **15**, 1321–1333 (2006).
66. Verhoeven, K. J. F., Poorter, H., Nevo, E. & Biere, A. Habitat-specific natural selection at a flowering-time QTL is a main driver of local adaptation in two wild barley populations. *Mol. Ecol.* **17**, 3416–3424 (2008).
67. Ungerer, M., Linder, C. & Rieseberg, L. Effects of genetic background on response to selection in experimental populations of *Arabidopsis thaliana*. *Genetics* **163**, 277–286 (2003).
This was one of the first experimental studies to measure natural selection on phenotypic traits that are connected to specific QTLs.
68. Ungerer, M. & Rieseberg, L. Genetic architecture of a selection response in *Arabidopsis thaliana*. *Evolution* **57**, 2531–2539 (2003).
69. Beraldi, D. *et al.* Mapping quantitative trait loci underlying fitness-related traits in a free-living sheep population. *Evolution* **61**, 1403–1416 (2007).
70. Beavis, W. D. in *Molecular Dissection of Complex Traits* (ed. Paterson, A. H.) 145–162 (CRC Press, New York, 1998).
71. Otto, S. P. & Jones, C. D. Detecting the undetected: estimating the total number of loci underlying a quantitative trait. *Genetics* **156**, 2093–2107 (2000).
72. Weinig, C. *et al.* Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*. *Genetics* **165**, 321–329 (2003).
73. Lexer, C., Welch, M. E., Durphy, J. L. & Rieseberg, L. H. Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. *Mol. Ecol.* **12**, 1225–1235 (2003).
74. Barton, N. H. & Keightley, P. D. Understanding quantitative genetic variation. *Nature Rev. Genet.* **3**, 11–21 (2002).
75. Korves, T. *et al.* Fitness effects associated with the major flowering time gene *FRIGIDA* in *Arabidopsis thaliana* in the field. *Am. Nat.* **169**, 141–157 (2007).
76. Dalziel, A., Rogers, S. M. & Schulte, P. Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. *Mol. Ecol.* **18**, 4997–5017 (2009).
77. Barrett, R., Rogers, S. M. & Schluter, D. Natural selection on a major armor gene in threespine stickleback. *Science* **322**, 255–257 (2008).
78. Hudson, M. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Res.* **8**, 3–17 (2008).
79. Metzker, M. L. Sequencing technologies: the next generation. *Nature Rev. Genet.* **11**, 31–46 (2009).
80. Stapley, J. *et al.* Adaptation genomics: the next generation. *Trends Ecol. Evol.* **25**, 705–712 (2010).
81. Storz, J. & Wheat, C. Integrating evolutionary and functional approaches to infer adaptation at specific loci. *Evolution* **64**, 2489–2509 (2010).
This excellent perspective piece called for a synthesis of insights from indirect population genetic analyses and direct experiments on molecular function in making inferences about adaptation at the genetic level.
82. Ungerer, M., Johnson, L. & Herman, M. Ecological genomics: understanding gene and genome function in the natural environment. *Heredity* **100**, 178–183 (2008).
83. Excoffier, L., Hofer, T. & Foll, M. Detecting loci under selection in a hierarchically structured population. *Heredity* **103**, 285–298 (2009).
84. Ferguson, L. *et al.* Characterization of a hotspot for mimicry: assembly of a butterfly wing transcriptome to genomic sequence at the *HmYb/Sb* locus. *Mol. Ecol.* **19**, 240–254 (2010).
85. Renaut, S., Nolte, A. & Bernatchez, L. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* sp. Salmonidae). *Mol. Ecol.* **19**, 115–131 (2010).
86. Korbel, J. *et al.* The current excitement about copy-number variation: how it relates to gene duplications and protein families. *Curr. Op. Struct. Biol.* **18**, 366–374 (2008).
87. Christians, J. K. & Senger, L. K. Fine mapping dissects pleiotropic growth quantitative trait locus into linked loci. *Mamm. Genom.* **18**, 240–245 (2007).
88. Lawniczak, M. K. N., Emrich, S. J. & Holloway, A. K. Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. *Science* **330**, 512–514 (2010).
89. Houle, D., Govindaraju, D. R. & Omholt, S. Phenomics: the next challenge. *Nature Rev. Genet.* **11**, 855–866 (2010).
90. Nosil, P., Funk, D. & Ortiz-Barrientos, D. Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **18**, 375–402 (2009).
91. Michel, A. *et al.* Widespread genomic divergence during sympatric speciation. *Proc. Nat. Acad. Sci.* USA **107**, 9724–9729 (2010).
This novel study used a combination of laboratory-based selection experiments, genetic mapping, LD analyses and field surveys to show genomic patterns of divergent natural selection.
92. Phillips, P. C. Epistasis: the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Rev. Genet.* **9**, 855–867 (2008).
93. Cadieu, E. *et al.* Coat variation in the domestic dog is governed by variants in three genes. *Science* **326**, 150–153 (2009).
94. Dean, A. M. & Thornton, J. W. Mechanistic approaches to the study of evolution: the functional synthesis. *Nature Rev. Genet.* **8**, 675–688 (2007).
95. Kimura, M. Evolutionary rate at the molecular level. *Nature* **217**, 624–636 (1968).
96. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, UK, 1983).
97. Lynch, M. *The Origins of Genomic Architecture* (Sunderland, Massachusetts, 2007).
98. Ohta, T. Slightly deleterious mutant substitutions in evolution. *Nature* **246**, 96–98 (1973).
99. Lewontin, R. C. & Hubby, J. L. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* **54**, 595–609 (1966).
100. Nei, M., Suzuki, Y. & Nozawa, M. The neutral theory of molecular evolution in the genomic era. *Ann. Rev. Genom. Hum. Genet.* **11**, 265–289 (2010).
This recent review provides a comprehensive critique of the methods used to detect selection in genome scans.
101. Hermisson, J. Who believes in whole-genome scans for selection? *Heredity* **103**, 283–284 (2009).
102. Sella, G., Petrov, D., Przeworski, M. & Andolfatto, P. Pervasive natural selection in the Drosophila genome? *PLoS Genet.* **5**, e1000495 (2009).
103. Halligan, D. & Keightley, P. Ubiquitous selective constraints in the *Drosophila* genome revealed by a genome-wide interspecies comparison. *Genom. Res.* **16**, 875–884 (2006).
104. Shapiro, J. A. *et al.* Adaptive genic evolution in the *Drosophila* genomes. *Proc. Nat. Acad. Sci. USA* **104**, 2271–2276 (2007).
105. Sawyer, S. A., Parsch, J., Zhang, Z. & Hartl, D. L. Prevalence of positive selection among nearly neutral amino acid replacements in *Drosophila*. *Proc. Nat. Acad. Sci.* USA **104**, 6504–6510 (2007).
106. Macpherson, J., Sella, G., Davis, J. & Petrov, D. Genome-wide spatial correspondence between nonsynonymous *divergence* and neutral polymorphism reveals extensive adaptation in *Drosophila*. *Genetics* **177**, 2083–2099 (2007).
107. Asthana, S. *et al.* Widely distributed noncoding purifying selection in the human genome. *Proc. Nat. Acad. Sci. USA* **104**, 12410–12415 (2007).
108. Akey, J. M. Constructing genomic maps of positive selection in humans: where do we go from here? *Genom. Res.* **19**, 711–722 (2009).
109. Wright, S. & Andolfatto, P. The impact of natural selection on the genome: emerging patterns in *Drosophila* and *Arabidopsis*. *Ann. Rev. Ecol. Syst.* **39**, 193–213 (2008).
110. Foll, M. & Gaggiotti, O. E. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977–993 (2008).
111. Andolfatto, P. Hitchhiking effects of recurrent beneficial amino acid substitutions in the *Drosophila melanogaster* genome. *Genom. Res.* **17**, 1755–1762 (2007).
112. Sattath, S., Elyashiv, E., Kolodny, O., Rinott, Y. & Sella, G. Pervasive adaptive protein evolution apparent in diversity patterns around amino acid substitutions in *Drosophila simulans*. *PLoS Genet.* **7**, e1001302 (2011).
113. Hughes, A. L. Looking for Darwin in all the wrong places: the misguided quest for positive selection at the nucleotide sequence level. *Heredity* **99**, 364–373 (2007).
114. Otto, S. P. Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proc. R. Soc. Lond. B.* **271**, 705–714 (2004).
115. McVean, G. A. T. & Charlesworth, B. The effects of Hill–Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics* **155**, 929–944 (2000).
116. Coberly, L. & Rausher, M. Pleiotropic effects of an allele producing white flowers in *Ipomoea purpurea*. *Evolution* **62**, 1076–1085 (2008).
117. Baucom, R. S., Chang, S. M., Kniskern, J. M., Rausher, M. D. & Stinchcombe, J. R. Morning glory as a powerful model in ecological genomics: tracing adaptation through both natural and artificial selection. *Heredity* 30 Mar 2011 (doi:10.1038/hdy.2011.25).
118. Cook, L. The rise and fall of the carbonaria form of the peppered moth. *Q. Rev. Biol.* **78**, 399–417 (2003).
119. Haldane, J. B. S. A mathmatical theory of natural and artificial selection, part 1. *Trans. Camb. Phil. Soc.* **23**, 19–41 (1924).
120. Kettlewell, H. Further selection experiments on *Lepidoptera*. *Heredity* **10**, 287–301 (1956).
121. Grant, B. S., Cook, A. D., Clarke, C. A. & Owen, D. F. Geographic and temporal variation in the incidence of melanism in peppered moth populations in America and Britain. *J. Hered.* **89**, 465–471 (1998).
122. Majerus, M. E. N. *Melanism: Evolution in Action* (Oxford Univ. Press, Oxford, 1998).
123. Coyne, J. A. Not black and white. *Nature* **396**, 35–36 (1998).

124. van't Hof, A., Edmonds, N., Dalíková, M., Marec, F. & Saccheri, I. Industrial melanism in british peppered moths has a singular and recent mutational origin. *Science* **332**, 958–960 (2011).
125. Hohenlohe, P. A. *et al.* Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* **6**, e1000862 (2010).
126. Kingsley, E. P., Manceau, M., Wiley, C. D. & Hoekstra, H. E. Melanism in *Peromyscus* is caused by independent mutations in *Agouti*. *PLoS ONE* **4**, e6435 (2009).
127. Dice, L. R. Effectiveness of selection by owls of deer-mice (*Peromyscus maniculatus*) which contrast in color with their background. *Cont. Lab. Vert. Biol. Univ. Mich.* **34**, 1–20 (1947).
128. Labbe, P. *et al.* Independent duplications of the *Acetylcholinesterase* gene conferring insecticide resistance in the mosquito *Culex pipiens*. *Mol. Biol. Evol.* **24**, 1056–1067 (2007).
129. Habu, Y., Hisatomi, Y. & Iida, S. Molecular characterization of the mutableflaked allele for flower variegation in the common morning glory. *Plant J.* **16**, 371–376 (1998).
130. Rebeiz, M., Pool, J., Kassner, V., Aquadro, C. & Carroll, S. Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* **326**, 1663–1667 (2009).
131. Pool, J. E. & Aquadro, C. F. The genetic basis of adaptive pigmentation variation in *Drosophila melanogaster*. *Mol. Ecol.* **16**, 2844–2851 (2007).
132. Marchinko, K. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. *Evolution* **63**, 127–138 (2009).
133. Reimchen, T. E. & Nosil, P. Lateral plate asymmetry, diet and parasitism in threespine stickleback. *J. Evol. Biol.* **14**, 632–645 (2001).
134. Le Corre, V., *Roux.*, F. & Reboud, X. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time. *Mol. Biol. Evol.* **19**, 1261–1271 (2002).
135. Storz, J. & Kelly, J. Effects of spatially varying selection on nucleotide diversity and linkage disequilibrium: insights from deer mouse globin genes. *Genetics* **180**, 367–379 (2008).
136. Storz, J. F., Runck, A. M., Moriyama, H., Weber, R. E. & Fago, A. Genetic differences in hemoglobin function between highland and lowland deer mice. *J. Exp. Biol.* **213**, 2565–2574 (2010).
137. Hayes, J. P. & O'Connor, C. S. Natural selection on thermogenic capacity of high-altitude deer mice. *Evolution* **53**, 1280–1287 (1999).
138. Tishkoff, S. A. *et al.* Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genet.* **39**, 31–40 (2007).
139. Mullen, L. M. & Hoekstra, H. E. Natural selection along an environmental gradient: A classic cline in mouse pigmentation. *Evolution* **62**, 1555–1569 (2008).
140. Vignieri, S. N., Larson, J. N. & Hoekstra, H. E. The selective advantage of cryptic coloration in mice. *Evolution* **64**, 2153–2158 (2010).
141. Evans, P. *et al.* Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans. *Science* **309**, 1717–1720 (2005).
142. Jackson, A. P. *et al.* Identification of microcephalin, a protein implicated in determining the size of the human brain. *Am. J. Hum. Genet.* **71**, 136–142 (2002).
143. Geffeney, S., Brodie, E. D Jr. Ruben, P. C. & Brodie, E. D. Mechanisms of adaptation in a predator-prey arms race: TTX-resistant sodium channels. *Science* **297**, 1336–1339 (2002).
144. Brodie, E. D. & Brodie, E. D Jr. Evolutionary response of predators to dangerous prey: reduction of toxicity of newts and resistance of garter snakes in island populations. *Evolution* **45**, 221–224 (1991).
145. Wheat, C. W., Watt, W. B., Pollock, D. D. & Schulte, P. M. From DNA to fitness differences: sequences and structures of adaptive variants of colias phosphoglucose isomerase (*PGI*). *Mol. Biol. Evol.* **23**, 499–512 (2006).
146. Watt, W. B. in *Butterflies: Ecology and Evolution Taking Flight* (eds. Boggs, C. L., Watt, W. B. & Ehrlich, P. R.) 319–352 (Univ. Chicago Press, Chicago, 2003).
147. Bradshaw, H. D. & Schemske, D. W. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* **426**, 176–178 (2003).
148. Jackson, M., Stinchcombe, J., Korves, T. & Schmitt, J. Costs and benefits of cold tolerance in transgenic *Arabidopsis thaliana*. *Mol. Ecol.* **13**, 3609–3615 (2004).
149. Bell, G. Experimental genomics of fitness in yeast. *Proc. R. Soc. Lond. B.* **277**, 1459–1467 (2010).
150. Lenormand, T., Guillemaud, T., Bourguet, D. & Raymond, M. Evaluating gene flow using selected markers: a case study. *Genetics* **149**, 1383–1392 (2008).
151. Lenormand, T., Bourguet, D., Guillemaud, T. & Raymond, M. Tracking the evolution of insecticide resistance in the mosquito *Culex pipiens*. *Nature* **400**, 861–864 (1999).
152. Labbé, P. *et al.* Forty years of erratic insecticide resistance evolution in the mosquito *Culex pipiens*. *PLoS Genet.* **3**, 2190–2199 (2007).
153. Duron, O. *et al.* High *Wolbachia* density correlates with cost of infection for insecticide resistant *Culex pipiens* mosquitoes. *Evolution* **60**, 303–314 (2006).

Acknowledgements
We thank P. Andolfatto, A. Berry, J. Jensen, J. Losos, D. Lowry, R. Nielsen, P. Nosil, S. Otto, D. Schluter, J. Slate and M. Szulkin for useful discussion and comments on the manuscript. We thank W. Cresko and P. Hohenlohe for providing a modified figure and L. Cook, J. Mallet and I. Saccheri for providing unpublished data on *Biston betularia*. We are supported by National Sciences and Research Council of Canada (NSERC) Howard Alper and John Templeton Foundational Questions in Evolutionary Biology (FQEB) fellowships (R.D.H.B.) and a US National Science Foundation (NSF) grant DEB-0919190 (H.E.H.).

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
Rowan D. H. Barrett's homepage:
http://www.people.fas.harvard.edu/~rbarrett
Hopi E. Hoekstra's homepage:
http://www.oeb.harvard.edu/faculty/hoekstra
***Nature Reviews Genetics* series on Study Designs:**
http://www.nature.com/nrg/series/studydesigns/index.html

SUPPLEMENTARY INFORMATION
See online article: S1 **(table)**

ALL LINKS ARE ACTIVE IN THE ONLINE PDF