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.

“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 organisinal 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 Lewontin1, 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 traits2–4; 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 REF S 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 S 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
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.

**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 heterozygosities, skews in the site frequency spectrum and extended linkage disequilibrium (LD)). 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. 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 as well as across temporal samples that are associated with recent climate warming. 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 (Fst)). 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 and to eliminate single-locus false positives. Although statistical methodologies have advanced greatly in recent years, 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. (It is also possible for loci to be under selection without yielding statistically significant results in tests for selection.)

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.

---

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

Since Kimura proposed the neutral theory of molecular evolution, 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. The neutral theory and its extension, the nearly neutral theory, 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, reviewed in REF. 100). Since then, statistical inference has improved and the amount of data has increased, as we describe below.

**The selectionists.** Several recent 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. 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, and amino acid divergence between species is often driven by (weak) positive selection. 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. Even in organisms with small effective population sizes, a considerable fraction of the genome is under selective constraint. For example, there are thousands of candidate regions under selection in humans. Similar to the situation in Drosophila species, selection affects non-coding regions throughout the human genome as well as coding regions, more than one-fifth of the loci that were investigated show evidence of natural selection. Additional evidence comes from amino acids that are associated with reduced synonymous site variation, 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. Nei et al. 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. 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 to no more than would be expected by chance, suggesting that several studies claiming evidence for selection may need to be re-evaluated. However, the prevalence of false negatives also indicates that there may be considerable selection that genome scans are missing. Most criticism is directed at genome scans, but others have extended the scepticism to additional tests of neutrality.

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).

---

**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.
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. 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, or they can be used post hoc to determine causative loci within regions that are associated with phenotype. 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. 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 (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 preation. 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. 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. 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.

**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, 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. 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. 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. 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.

**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. 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.

**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.

**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 *Ipomea* 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.

Image is reproduced, with permission, from REF. 117 © (2011) Macmillan Publishers Ltd. All rights reserved.

---

**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.

**Ascertaining 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.

**Pleiotropy**: 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 proteins42 and global changes in gene regulation43,44. For example, site-directed mutagenesis has been used to characterize the fitness effects of individual mutations detected by whole-genome sequencing44 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 evolution45. 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 interactions47 and genotype–genotype–environment 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.

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 in 1848, and in the ensuing seven 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, Kettlewell120 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 Majerus122 had pointed out some weaknesses in Kettlewell’s experimental design, he never doubted the importance of predation. Coyne122, 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 colleagues124 genetically mapped the locus determining colour morph to a 200kb 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.
QTL mapping of fitness. One approach that has been successfully used to investigate the genetic basis of adaptation in nature is to conduct reciprocal transplant experiments 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. For example, RILs have been used to investigate local adaptation in coastal and inland races of yellow monkeyflower (Mimulus guttatus) and in contrasting temperature and photoperiod environments in a rockcress plant (Boechera stricta) and in contrasting moisture environments in annual grasses Avena barbata and Hordeum spontaneum. Second-generation backcross lines were similarly exploited to investigate natural selection on viability and fertility traits associated with QTLs in A. thaliana.
### Table 1 | Examples of genes commonly described as adaptive

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ace1</td>
<td>Resistance to insecticide in mosquitoes (Culex pipiens)</td>
<td>Single amino acid substitution associated with loss of sensitivity to organophosphates and carbamates</td>
<td>Insecticide</td>
<td>Recent and rapid spread of resistance to organophosphates and carbamate insecticides; no resistance alleles detected in populations not exposed to insecticide (phenotypic (P) and genotypic (G))</td>
<td>Not calculated</td>
<td>No</td>
<td>128</td>
</tr>
<tr>
<td>CHSD</td>
<td>Flower colour in morning glory (Ipomoea purpurea)</td>
<td>Transposon insertion in the gene that codes for chalcone synthase, the first enzyme of the anthocyanin pathway</td>
<td>Flower colour under different pollination syndromes</td>
<td>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)</td>
<td>Selection coefficient against amino acid = 0.05 (G)</td>
<td>Yes (background genotype randomized)</td>
<td>116,129</td>
</tr>
<tr>
<td>ebony</td>
<td>Abdominal melanin in Drosophila melanogaster</td>
<td>Five CRMs alter expression of an enzyme in the biogenic amine synthesis pathway that affects pigmentation</td>
<td>Thermal regulation at different temperatures; differential absorption of heat from solar radiation</td>
<td>Correlation between elevations and pigmentation; patterns of nucleotide variation in the ebony region consistent with a partial selective sweep (P and G)</td>
<td>Not calculated</td>
<td>No</td>
<td>130,131</td>
</tr>
<tr>
<td>Eda</td>
<td>Lateral plate armour in threespine stickleback (Gasterosoleus aculeatus)</td>
<td>Haplotype associated with changes to ectodermal signalling pathway that affects development of dermal plates</td>
<td>Predation regime (insect versus fish); growth rate (trade-off between armour and growth)</td>
<td>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)</td>
<td>Selection coefficients = 0.50 against the low (L) allele in juveniles, 0.52 against the complete (C) allele in adults (G)</td>
<td>Yes (rare genetic variants that segregate at Eda in a homogenous genetic background, but closely linked genes likely to co-segregate)</td>
<td>29,31, 77,132, 133</td>
</tr>
<tr>
<td>FRI</td>
<td>Flowering time in Arabidopsis thaliana</td>
<td>Deletion polymorphism disrupting ORFs associated with early flowering times</td>
<td>Seasonal timing based on climatic conditions</td>
<td>Overwinter survival and seed production differences between genotypes with different vernalization times (P and G)</td>
<td>Selection gradient = -0.43 in autumn cohort, 0.66 in spring cohort (P)</td>
<td>Yes (inbred accession lines)</td>
<td>75,134</td>
</tr>
<tr>
<td>Hb</td>
<td>Aerobic metabolism in deer mice (Peromyscus maniculatus)</td>
<td>Combined effects of 12 amino acid substitutions alter oxygen affinity binding</td>
<td>Physiological performance at different altitudes</td>
<td>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)</td>
<td>Selection gradient = 0.52 (P)</td>
<td>No</td>
<td>135,136, 137</td>
</tr>
<tr>
<td>LCT</td>
<td>Ability to digest lactose in adult humans</td>
<td>Three CRMs associated with transcription level of the enzyme LP, which hydrolyses lactose into sugars</td>
<td>Lactase persistence in pastoralist populations</td>
<td>LD patterns consistent with a selective sweep (G)</td>
<td>Selection coefficient = 0.04–0.097 (G)</td>
<td>No</td>
<td>138</td>
</tr>
<tr>
<td>Mclr</td>
<td>Coat pigmentation in beach mice (Peromyscus polionotus)</td>
<td>Single amino acid substitution reduces receptor signalling, affecting melanogenesis</td>
<td>Visual predation against conspicuous phenotypes</td>
<td>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</td>
<td>Selection index = 0.50 (P — on plasticine mouse models); selection coefficient = 0.07–0.21 depending on body region (P)</td>
<td>No</td>
<td>28,139, 140,141</td>
</tr>
</tbody>
</table>
from the relative fitness of different allelic classes at a number of survivorship and fecundity QTLs in 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.53

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 and RILs 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.51 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 compromised 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).53 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 occurring variants at the Ectodysplasin (Eda) locus, which underlies differences in armour plates in three-spine sticklebacks (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

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MCPH1</td>
<td>Brain size in humans</td>
<td>Loss-of-function mutations alter the proliferation and differentiation of neuroblasts during neurogenesis, which can affect cranial capacity</td>
<td>Cognitive ability</td>
<td>Nucleotide diversity and LD patterns consistent with a recent selective sweep (G)</td>
<td>Not calculated</td>
<td>No</td>
<td>141,142</td>
</tr>
<tr>
<td>Scn4a</td>
<td>Resistance to tetrodotoxin (TTX) in garter snakes</td>
<td>Amino acid substitution associated with binding of TTX to Scn4a, affecting muscle contraction after ingestion</td>
<td>Defensive toxin in prey (rough-skinned newts)</td>
<td>Correlations between snake TTX resistance and levels of newt predation and newt TTX level (P)</td>
<td>Not calculated</td>
<td>No</td>
<td>143,144</td>
</tr>
<tr>
<td>PGI</td>
<td>Flight performance in Colias butterflies</td>
<td>Single amino acid substitution associated with glycolysis of nectar used to fuel flight</td>
<td>Optimal sugar metabolism depending on environmental temperature</td>
<td>Nucleotide diversity and LD patterns consistent with a recent selective sweep (G); differences in survival, mating success and female fecundity among common genotypes (G)</td>
<td>Selection coefficient against genotype with lowest relative fitness = 0.82 (males), 0.83 (females) (G)</td>
<td>No</td>
<td>145,146</td>
</tr>
</tbody>
</table>
Table 2 | Advantages and disadvantages of using different types of genetic variants in selection experiments

<table>
<thead>
<tr>
<th>Genetic variant</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Example reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sample from natural populations in derived habitat or distinct habitats</td>
<td>Represents natural variation available to selection in the wild. High genetic diversity</td>
<td>Difficult to isolate selection on specific traits or genes</td>
<td>30</td>
</tr>
<tr>
<td>Phenotypic extremes</td>
<td>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</td>
<td>Organisms carry alternate alleles at most loci contributing to trait, making it difficult to isolate selection on different genes contributing to a polygenic trait</td>
<td>145</td>
</tr>
<tr>
<td>Rare genetic variants in ancestral or alternate environment type</td>
<td>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</td>
<td>Depending on levels of linkage disequilibrium, may be difficult to distinguish between effects of focal locus and closely linked genes</td>
<td>77</td>
</tr>
<tr>
<td>F2 or backcross lines</td>
<td>Contain alleles segregating from both parental types, so all traits can be analysed at the same time. Can detect heterosis and dominance effects</td>
<td>Power to isolate selection on individual loci highly dependent on sample size and genetic architecture of traits. Limited genetic diversity</td>
<td>73</td>
</tr>
<tr>
<td>Recombinant inbred lines (RILs)</td>
<td>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</td>
<td>Fairly course genetic mapping. Limited genetic diversity</td>
<td>63</td>
</tr>
<tr>
<td>Near-isogenic lines (NILs)</td>
<td>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</td>
<td>Size of introgression depends on number of backcross generations, making fine mapping time-consuming. Limited genetic diversity</td>
<td>147</td>
</tr>
<tr>
<td>Transgenic lines</td>
<td>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)</td>
<td>Limited genetic diversity. Challenging to obtain within-line replicates</td>
<td>148</td>
</tr>
<tr>
<td>Knockouts</td>
<td>Exact targeting of selective consequences of loss-of-function of a single locus</td>
<td>Limited to consequences of mutations conferring fitness effects resulting from complete loss-of-function</td>
<td>149</td>
</tr>
</tbody>
</table>

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

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\(^{36-39}\). The importance of these technological advances for population and ecological genomics has been well recognized\(^{40-42}\). 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\(^{43,44}\), and provide greater power to identify the types of mutations that are involved in adaptation\(^{45-46}\). 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\(^{27}\). 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 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
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, \(\text{REF} \ 91\)). 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\(^{92}\), and the few studies that have investigated interacting QTLs have found strong evidence for epistatic fitness effects\(^{72,73,75}\), which is not surprising given the high prevalence of epistasis observed in genotype–phenotype relationships (reviewed in \(\text{REF} \ 92\)). 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.


18. This study demonstrates the difficulty of predicting the microevolutionary consequences of selection in the absence of knowledge about the genetic basis of fitness variation.


29. Boag, T. & Grant, P. 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.


