Ellegren H, Sheldon BC. Genetic basis of fitness differences in wild populations. Nature 452: 169-175

Article *in* Nature · April 2008
DOI: 10.1038/nature06737 · Source: PubMed

2 authors, including:

Ben C Sheldon
University of Oxford
341 PUBLICATIONS **20,399** CITATIONS

SEE PROFILE

REVIEWS

Genetic basis of fitness differences in natural populations

Hans Ellegren[1] & Ben C. Sheldon[2]

Genomics profoundly influences current biology. One of many exciting consequences of this revolution is the potential for identifying and studying the genetic basis of those traits affecting fitness that are key to natural selection. Recent studies using a multitude of genomic approaches have established such genotype–phenotype relationships in natural populations, giving new insight into the genetic architecture of quantitative variation. In parallel, an emerging understanding of the quantitative genetics of fitness variation in the wild means that we are poised to see a synthesis of ecological and molecular approaches in evolutionary biology.

The darwinian evolutionary process can be summarized into three components: struggle for existence, variation in characters that influence success in that struggle, and transmission of that variation from parents to offspring. Understanding the causes of variation among individuals in their contribution to future generations—variation in fitness—and the way in which that variation is inherited—its genetic basis—thus lies at the heart of our understanding of evolution. In this review we discuss the genetic architecture of fitness traits in wild populations and how new genomic approaches to non-model organisms can pinpoint the genetic 'locus' of evolution. We argue that we are approaching a synthesis of population biology and genomics, with the potential greatly to advance our understanding of evolution in wild populations. We focus on animal populations, because these are easier to study in the wild; for related work on plants see ref. 1.

Fitness variation in the wild

For most biological problems, studying laboratory model organisms offers tremendous advantages in terms of control, replication and convenience. However, it is precisely those advantages that undermine the utility of these models for studying fitness variation (see Box 1), and its basis. In this case, field studies are of most relevance, because laboratory studies provide novel, stable, uniform, benign environments, where selection is unlikely to operate as it would in wild populations. In some cases (such as *Drosophila* populations used by biologists for experimental evolution studies[2]), laboratory populations have been maintained for long enough (hundreds of generations), and with sufficient competition, that adaptation to this novel environment has presumably occurred. Even in these cases, the relative invariance of the environment (or the arbitrariness of any imposed variation) suggests that they may be poor models of natural populations. On the other hand, laboratory models offer very clear advantages for testing the plausibility of evolutionary hypotheses, and for fine-scale dissection of their operation, and are often able to suggest novel hypotheses that can be further tested in field populations[3–5]. In addition to the environment influencing the relevance of fitness measures, the expression of genetic variance depends strongly on the environment in which it is measured. Comparisons across environments both in the laboratory and in the field indicate strong interactions between the expression of genetic variance and environmental axes[6]. Hence, studies of fitness, and the genetic factors influencing it, must be carried out in matching environments, ideally those to which organisms are adapted. This was appreciated by early ecological geneticists such as Ford and Dobzhansky; indeed, there are interesting parallels between their work and current interest in moving genomics into the field.

Measuring fitness in natural populations requires dedicated field effort, sometimes for decades, and the scale of these studies makes them particularly valuable for studying genetics of fitness in the wild. Studies over tens of generations result in long-term pedigrees of animals of known relatedness, inhabiting a wide range of environmental conditions. Currently, reliance on physical marking methods (banding or tagging), and the need for some observational information to frame inferences about parentage, has restricted such studies to vertebrates that undergo limited dispersal, particularly birds and mammals. This may change, and genetic-marker-based approaches may soon be used to reconstruct pedigrees in population types such as free-ranging invertebrates, or organisms with external fertilization. There are now numerous studies of selection on morphological, behavioural and life-history characters, replicated over many years within populations[7–9]. These studies confirm that natural selection often acts strongly and consistently within populations; the strength of selection is comparable to, or stronger than, that obtained from shorter-term studies of selection[10].

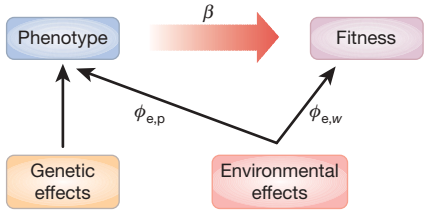
Quantitative genetics of fitness in the wild

The maturation of long-term population studies, and the opportunities they offer for measuring selection, has been accompanied by a surge of interest in applying quantitative genetic models, particularly the 'animal model' approach[11] to wild populations. This has confirmed the empirical generalization from laboratory studies that additive genetic variance varies with respect to trait type: characters under strong selection have low heritabilities, high additive genetic variance, and even higher environmental variance[12]. Estimates of the heritability of fitness itself have generally been unable to exclude zero in their range when using traditional measures of fitness; those that have taken account of demography, and which are measured at a per-generation scale (for example, Box 1) have found small, but significant, heritability of lifetime fitness[5,13]. Simultaneous estimation of environmental and genetic influences on characters, and their separation in individuals, which animal models allow, is vital to gain unbiased estimates of the force of selection on characters. Both

[1]Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden. [2]Edward Grey Institute, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Box 1 | Fitness variation and natural selection in the wild

A mathematical model for the relationship between natural selection and evolution was developed in the 1960s and 1970s by Robertson, Li and particularly Price, encapsulated by the Price equation, which specifies the relationship between evolutionary change and selection, conditional on the transmission of variation[65]. The link between genetics and selection was further developed in multivariate form by Lande and co-workers[66,67], who proposed an intuitive multiple-regression-based approach for the quantification of selection in natural populations. The essence of this approach is that selection gradients are derived that describe associations between fitness and trait variation. Its appeal is that, in principle, in combination with information about genetic variances and covariances, inferences about the evolutionary forces acting on populations, and potential evolutionary trajectories, may be derived. A central challenge for studies in the wild is to understand the importance of environmental covariance between traits (Box 1 figure). In this hypothetical example, selection β is assessed by measuring the covariance ϕ between phenotype (subscript 'p') and fitness w, but the true relationship between the genetic component of a character and fitness (that is, the degree to which the trait will respond to selection) may be masked by environmentally induced covariances (where subscript 'e' refers to the environment) between the trait and fitness, which will be determined by the product of $\phi_{e,p}$ and $\phi_{e,w}$; when this covariance is positive, selection will be overestimated. To understand selection and fitness variation requires that fitness be measured. While fitness can be defined simply as the number of offspring contributed to future generations, either for individuals or for populations, this definition may not be reliable in non-stable populations, for which a measure of the intrinsic rate of increase at either population or individual levels that takes into account the timing during life that offspring are produced, is preferred. The appropriate definition of fitness may also depend on the context: theoretical approaches typically focus on population fitness measures, whereas empirically oriented studies tend to focus on individual fitness[68]. A recently developed fitness measure that combines individual and population aspects of fitness, designed to deal with unstable population dynamics and to take account of the continuous nature of evolution, has been termed 'de-lifing', because it assesses the effect on population growth of removing each individual's contribution to the population[69]. Application of this and other methods to take account of population demography may yield insights about the operation of selection and evolution in natural populations[70,71].



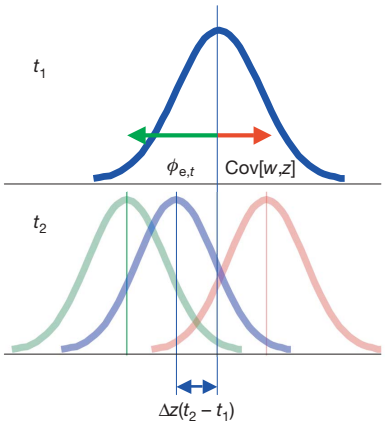
Box 1 Figure | Genetic and environmental effects on fitness.

Darwin and Fisher discussed cases of what we would now term environmental covariance between phenotype and fitness, where selection on a character is apparent, rather than real, because both are correlated with a third unmeasured character (Box 1). Such effects can be very important, with the result that selection is much weaker than would be apparent. For example, Kruuk and colleagues showed[7] in red deer stags that while the males with the largest antlers enjoyed highest reproductive success, most of this effect was due to correlated effects of the environment: the true force of selection on antler size was much weaker.

An emerging theme from recent studies of wild populations is the conditional nature of much of the expressed genetic variation, both with respect to the environment, but also to the life-stage of the focal individual[14], and its sex[5]. This has important consequences for our understanding of the evolutionary dynamics of characters in the wild. For example, Wilson and colleagues[6] showed that additive genetic

Box 2 | Evolutionary change while the environment changes

Whether populations are usually at evolutionary equilibria, evolving only in response to shifts in the adaptive landscape, or whether evolution is continually occurring, is fundamental to our understanding of the evolutionary process. Recent studies of wild populations suggest that observations of phenotypes alone offer little chance of resolving this question, because evolutionary responses may be hidden by environmental deterioration (for example, due to density-dependent competition for resources). Indeed, Fisher[70] and Price[71] showed that this can be expected to be the case when a change in gene frequencies itself causes a form of environmental deterioration, which will be true when fitness depends to a large extent on success in competition with conspecifics. If generally true, this implies that much evolutionary change, and differentiation in wild populations, may not be apparent at the level of the phenotype (Box 2 figure): detecting these differences becomes increasingly challenging using quantitative genetics alone, because the models and traits become ever more complex, and removed from raw phenotypes. Understanding the molecular genetic basis of the traits can make this problem far more tractable.



Box 2 Figure | Simplified representation of opposing effects of the environment and selection on the phenotypic evolution of a quantitative character. If a trait is under directional selection, the mean phenotype is expected to change as a function of the genetic variance and the strength of selection on that character (Cov[w, z]), as shown in the top panel; the red curve at time t_2 (bottom panel) indicates a hypothetical change in the distribution of phenotypes due to selection. However, a simultaneous change in the environment($\phi_{e,t}$ and the green curve) can act to mask any change due to selection. The resultant phenotype (blue curve), and the extent to which any evolutionary response over time $\Delta z(t_2 - t_1)$ is observed, depends on the balance between these two effects.

variance in birth weight in wild sheep, a trait under strong selection, was higher in years in which survival was high; conversely, selection was weakest in such environments. This had the effect of producing a negative covariance between the selection on this character, and the variance available for selection, a process that would act to constrain the evolutionary response to selection. A second recently emergent theme is that analysis at the level of phenotypes may give a very incomplete picture of the degree of genetic change occurring over time. Ecological geneticists have long been aware of the phenomenon of counter-gradient variation, where an ecological or environmental gradient acts to hide genetic differentiation along the same axis[15]. For example, selection may favour faster developmental rates at higher latitudes, but in ectotherms, lower ambient temperatures may reduce the effective rate of development[16]; 'common garden' experiments are used to separate such effects. Several recent long-term studies have demonstrated a form of temporal counter-gradient variation, where responses to selection on morphology (body mass in two bird species[17,18] and in wild sheep[19]) at the level of the breeding value were much larger than was apparent at the phenotypic level, because of simultaneous changes in the environment. Such a pattern of 'cryptic evolution' may be quite general (Box 2).

What is the 'locus' of evolution?

The application of quantitative genetics should, ideally, be done bearing in mind the assumptions of this method of analysis: for example, that very large numbers of loci of small effect underlie traits. It will not be until the union of ecological and molecular genetics that we have an understanding of how realistic these assumptions are. Before we discuss the potential for unravelling the molecular basis of fitness differences in natural populations using genomic approaches, we must consider in which type of DNA sequence we expect to find adaptive mutations affecting fitness. What is, as Hoekstra and Coyne[20] put it, ''the locus of evolution''? More than 30 years ago, King and Wilson[21], inspired by the similarity of human and chimpanzee proteins, suggested that it may be changes in regulation of gene expression, rather than changes in their structure, that mainly drive phenotypic evolution. Whether this is actually the case has been a contentious issue, perhaps more now (paradoxically, given the availability of sequence data) than ever before. One school of thought, whose advocates are mainly concerned with morphological evolution ('evo-devo'), holds that *cis*-regulatory elements are indeed the primary targets for genetic changes underlying new phenotypes[22]. Another school of thought maintains that changes in amino-acid-coding regions of genes, leading to new protein structures, are the key to functional evolution[20]. There is no reason to expect one or other of these mechanisms to the sole explanation, so the debate concerns their relative importance.

Work on model organisms rarely focuses on population data so there is little information available on naturally occurring variants that affect fitness that is relevant to this debate. Comparisons of closely related species in *Drosophila* and *Saccharomyces* show that regulatory mutations underlie several key fixed differences in phenotype[23–25]. However, there are also examples of structural changes in genes affecting phenotypes of central importance[26]. More generally, much recent work has demonstrated the role of positive selection in adaptive protein evolution across many different organisms[27]. As genomics is taken into natural populations, it seems prudent to expect that the loci of evolution will be both regulatory and structural in origin. We now review the resources and approaches (Fig. 1) that are currently available to find these loci.

Genomic resources

Linking genes or other DNA sequences to fitness ultimately requires genome sequence information. Until recently, obtaining such data has been a limiting step for progress on non-experimental models. With the first gigabase-pair-sized genomes targeted, it was foreseen that genomic research would move rapidly into post-genomic approaches focusing on functional aspects rather than on gathering more sequence data. However, while functional genomics has indeed flourished, DNA sequence data continues to accumulate exponentially. New high-throughput technologies have hugely decreased the unit time and cost of obtaining sequence data and open up immense possibilities for large-scale sequencing initiatives in ecologically important species. One of the first such examples is the sequencing of a large fraction of the coding part of genome of the Glanville fritillary butterfly *Melitaea cinxia*, a key model species in studies of metapopulation ecology[28]. Moreover, it is increasingly recognized that comparative genomics, where sequences from two or more species are aligned and compared, is a powerful tool for detecting regions that evolve under negative or positive selection, indicative of functionality. Adaptive evolution can be inferred from, for example, gene sequences showing sites of repeated non-synonymous substitutions in multiple species alignments, an increased rate of non-synonymous substitutions in divergence compared to diversity data, or a high frequency of derived alleles[29].

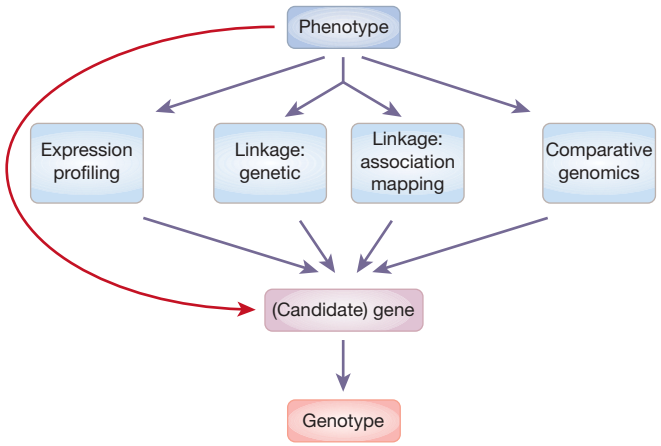


Figure 1 | Linking phenotypes to genotypes. Routes towards linking variation in phenotypes to variation in genotypes.

Genetic mapping in natural populations

Linkage analysis is the traditional way of identifying chromosomal regions containing trait loci in model organisms. It relies on following the inheritance of segregating traits in pedigrees and seeks to find co-inheritance of traits and genetic markers; if this can be established, trait loci are inferred to map in the vicinity of marker loci. As the power of linkage analysis increases with number of meioses that can be studied, linkage maps have so far mainly been constructed for species that can be bred in captivity, including fishes, insects and mammals. However, linkage maps from wild populations are starting to accumulate, including some of model species for ecological research.

Decades of work on human disease genetics and trait mapping in model organisms has revealed that finding the causal genetic basis of segregating phenotypes can be extremely demanding. In general, the lower the heritability and the more loci involved, the more difficult it is to dissect genotype–phenotype relationships[30]. Added complexity results from the fact that genetic variation at different loci may contribute to similar phenotypes in different populations, notably due to gene–environment interactions. Given this, will quantitative trait locus (QTL) mapping be a useful method for unravelling genetic architecture of traits of modest heritability that demonstrate continuous variation in natural populations[31]? The answer is likely to depend on the character of the study system. The use of inbred line crosses is the most powerful method of QTL analysis, because it maximizes linkage disequilibrium between markers and trait loci[32]. Accordingly, if organisms from natural populations can be bred and selected for divergent phenotypes under controlled conditions, mapping is quite feasible, although, as stated above, laboratory conditions do not replicate the full range of environments experienced by organisms in the wild. Another caveat is that the use of inbred lines fails to mirror epistasis based on standing genetic variation shown by organisms in outbred populations.

For QTL mapping to become widely used for studies of outbred natural populations will require that large pedigrees, or extensive series of sibling-pairs, can be sampled and components of fitness measured in these individuals[33,34]. If this proves impractical, for instance because of long generation times or because fitness can be measured for few individuals in the pedigree, alternative approaches may be more promising, as discussed below. Moreover, it must be recognized that mapping chromosomal regions that co-segregate with traits of interest represents just the first step towards identification of causative genetic variants. Numerous genes usually reside within a targeted chromosomal region and the eventual identification of such variants may require positional cloning, refined mapping and nomination of candidate genes. Only a few studies of natural populations have yet gone all this way, notably the mapping[35] and subsequent identification of ectodysplasin[36] underlying armour plate patterning in different populations of threespine sticklebacks. Mutations at this locus have led freshwater forms of sticklebacks to evolve a loss of pelvic structure, possibly an adaptation to a change in the risk of predation.

An alternative method is to search for linkage between markers and trait loci by genome scans of population samples, rather than by pedigree analysis. Known as association mapping or linkage disequilibrium mapping, the approach relies on a statistical association between marker and trait loci that are in linkage disequilibrium, typically with much higher resolution than in conventional pedigree analysis. The efficiency of association mapping depends, among other things, on the number and distribution of markers used to scan the genome, and the extent of linkage disequilibrium. Extensive linkage disequilibrium facilitates detecting disequilibrium between trait loci and markers but comes with the price of reduced resolution; short-range linkage disequilibrium requires more markers but simplifies the subsequent process of identification of causative sequence variants. Data on linkage disequilibrium in natural populations are now accumulating and, not surprisingly, indicate that it varies among species and is at least in part dependent on population history[37,38]. It is important to note that population structure may result in artefactual associations[39].

If genetic markers spread across the genome are available, other approaches can be used for the identification of regions subject to adaptive evolution, using population samples. Selective sweep mapping[40] takes advantage of the fact that strong positive selection is expected to leave a local footprint in form of reduced genetic diversity around the selected locus; when a selected allele goes to fixation, linked neutral variants will also increase in frequency. However, it can be demanding to achieve sufficient statistical power in genome-wide scans, in part relating to dominance and whether selection has been acting on standing variation or a new mutation[41,42]. Finally, it has recently been argued that, by studying the distribution of inbreeding coefficients across the genome, regions that show apparent adaptive divergence among populations can be identified[43,44]. Often referred to as population genomics[45], this approach is particularly suited for identifying loci under disruptive selection, such as in recently derived species or in diverging populations. It has been applied to a variety of species in which genomic regions with a high distribution of inbreeding coefficients, suggestive of strong assortative mating, or even reproductive isolation, contrast with a genomic background of almost free gene flow[46]. Local adaptation often occurs along environmental gradients associated with continuously varying phenotypic characters, so it is sensible to search for outlier loci in consecutive pairwise comparisons along such gradients. For example, in the common frog (*Rana temporaria*), a number of candidate loci for adaptation to altitude have been found, identified through showing a significantly higher degree of differentiation than under neutral expectation[47].

In combination, genome scans are important in confirming that QTLs identified in contemporary populations have played a part in adaptive phenotypic differentiation, driven by directional selection. If signs of selection as revealed by scans for reduced within-population variability or increased between-population divergence coincide with the chromosomal location of QTLs, this highlights the significance of genes within these regions in adaptive evolution[48].

Candidate gene approaches

A major advantage of studies of the genetic basis of fitness differences in natural populations is that candidates for trait loci can be nominated on the basis of knowledge of similar phenotypes in model species, circumventing the tedious process of unprejudiced genome-wide approaches. A successful example of this approach is the analysis of adaptive variation in vertebrate pelage and plumage traits. Coat-colour mutations in laboratory mice provide a detailed catalogue of proteins involved in pigmentation with associated information on how these proteins interact with each other. The melanin-based pigmentation pathway leading to melanocyte development and migration is highly conserved in vertebrates, so the wealth of information provided by mouse genetics is broadly applicable. The melanocortin-1 receptor (Mc1r), a G-protein-coupled receptor that induces cAMP production needed in the synthesis of eumelanin, is known to be a major determinant of pigmentation phenotypes[49]. A candidate gene approach focusing on allelic variation in *Mc1r* has successfully linked this gene to colour morphs of adaptive significance in several natural populations, including pigmentation variation relating to crypsis in pocket mice[50], little striped whiptail and lesser earless lizards[51], and to mate choice in arctic skuas and lesser snow geese[52]. Other genes involved in this pathway, like *Tyrp1*, *Agouti*[53] and *Kitlg*[54], have also been shown to be associated with pigmentation polymorphisms in natural populations. In the case of *Mc1r*, there is compelling evidence that the causative genetic background to phenotypic variation is coding sequence polymorphism. As such, fitness-related coat-colour polymorphism mediated by *Mc1r* exemplifies a mendelian trait governed by variation in gene structure rather than gene expression.

An excellent example of how the power of the candidate gene approach can be strengthened by parallel genetic mapping is the case of albinism in Mexican tetra fish (*Astyanax fasciatus/mexicanus*). Several cave populations of these fish have independently evolved reduced pigmentation and regressed eyes compared with surface or river-dwelling sister populations. By constructing a microsatellite-based genetic map, Protas *et al.*[55] obtained a very strong signal of linkage between albinism and a single chromosomal region. When testing the segregation of a number of candidate genes known to have profound effect on pigmentation in model species, the gene ocular and cutaneous albinism-2 (*Oca2*) was found to map to the same region and thus co-segregate with albinism. Subsequent sequence analysis showed that amino-acid substitutions and coding sequence deletions occur in *Oca2* of albinistic cave fish. This study is of particular importance because, as is now commonplace in studies of, for example, human disease mutations, the identification of a causative locus was supported by functional studies in a mouse model. Although technologically demanding, such dual strategies (genetic and functional) will be very important in future studies of natural populations.

Another category of genes that are good candidates for those influencing fitness in natural populations is those encoding proteins involved with glucose metabolism ('energy production'). For example, the enzyme phosphoglucose isomerase (Pgi) catalyses the conversion of glucose-6-phosphate to fructose-6-phosphate in the first phase of glycolysis. *Pgi* polymorphism has major fitness effects in a diverse range of species[56]. In the Glanville fritillary butterfly, allelic variation at *Pgi* correlates with flight metabolic rate and is related to dispersal rate and the ability to establish new and isolated populations[57]. As such, this system constitutes one of the best examples of how genetic variation at a fitness-related locus affects metapopulation dynamics and population growth. The effect of *Pgi* polymorphism was found to be dependent on the size and isolation of populations; hence the genetic effect was strongly dependent on the environmental context, underlining the importance of combining ecological and genetic approaches. A caveat is that it remains to be formally demonstrated that *Pgi* is the causally responsible locus; in theory, a closely linked locus may be involved.

Transcriptome analysis

Improvement in array technology has had a significant impact on large-scale studies of gene expression—transcriptome profiling—with increasing application to natural populations[58]. A case study of the genetic background to adaptive variation in beak morphology among Darwin's finches is described in Box 3. A microarray experiment uses representative probes corresponding to known genes from a genome, or of most of the sequence from a genomic region (tiling paths or tiling arrays), to which tissue messenger RNA is hybridized and quantified on the basis of hybridization intensity. The experimental rationale is usually straightforward: hybridization data from two or more groups of samples (treatments) are compared to seek evidence of differentially expressed genes. The access to

Box 3 | The genetic basis of beak morphology in Darwin's finches

Adaptive radiation among Darwin's finches has resulted in more than ten different finch species exploiting a variety of ecological niches on the Galápagos islands, reflected in distinct beak morphology. Species feeding on cactus flowers have long, pointed beaks, species eating seeds on the ground have deep, wide beaks, while insect eaters have slender, pointed beaks. Differences in external beak morphology are consistent with corresponding differences in craniofacial skeletons and it has been shown experimentally in other bird species that the cellular origin of beak development is in the neural crest-derived mesenchyme[72]. Studying the expression pattern of growth factors implicated in avian craniofacial development, Abzhanov *et al.*[73] found a strong correlation between beak morphology and the level of expression of bone morphogenetic protein 4 (*Bmp4*) in the mesenchyme of the upper beak. This suggests that regulation of *Bmp4* expression is a key variable determining quantitative variation in beak morphology of Darwin's finches, although it cannot distinguish between a role of *cis*-regulatory elements or of variation in the induction/transduction of upstream factors. Subsequently, the same group used microarray hybridization to identify differentially expressed genes among finches with different beak morphology[74]. This led to the identification of calmodulin (CaM), a key component of a Ca^{2+}-dependent signal transduction pathway essential for the control of bone differentiation and growth, being expressed at much higher levels in cactus-feeding finches than in other species. These studies not only provide new insight into this textbook example of evolution by natural selection. They also show both the power and feasibility of combining several genomic approaches to non-model species, in this case including the construction of a species-specific expression array. In addition, they constitute an unusual example of experimental confirmation of the functional significance of the proposed genetic mechanism by experimental manipulation in a model system: misexpression of *Bmp4* and *CaM* in developing chicken embryos has an effect on beak morphology similar to that seen in finches[74].

species-specific arrays, typically limited to organisms in focus for genome projects, has somewhat impeded widespread use of microarrays. However, cross-species application of arrays developed for related species is a useful alternative.

Transcriptome profiling has in several cases shown distinct differences in gene expression between divergent natural populations, particularly for fish species[58]. For example, differences in the expression of genes involved with swimming activity were found between the dwarf and lake ecotypes of whitefish, consistent with differences in feeding behaviour and niche exploitation[59]. However, just as the relative importance of natural selection and genetic drift has been a long-standing issue in molecular evolution, there is ongoing debate concerning how much variation in gene expression results from neutral processes and selection, respectively. Similarly, the relative influence of directional selection and stabilizing selection is debated[60]. Many microarray studies use pooled rather than individual samples for analysis, but there is increasing evidence for significant inter-individual variation in patterns of gene expression, perhaps for the majority of genes[58]. Moreover, there is an emerging understanding that much of the expression variation seen under controlled environmental conditions is heritable[61]. Hence, in principle, there is raw material for evolution by natural selection. Also, as genes are typically part of networks, it is easy to see that many, if not most, of the differences in gene expression seen between two phenotypes do not necessarily represent genetic differences underlying phenotypic difference. As an illustrative example, thousands of transcripts can be found differentially expressed between males and females of many species[62], yet the great majority of these genes are not directly involved with sex determination.

With new, ultrahigh-throughput sequencing technology, the way gene expression is studied is likely to change soon. Rather than relying on measuring gene transcript abundance from hybridization intensity to microarrays, the number of times transcripts are called in deep-coverage sequencing affords direct quantification of expression levels. An attractive feature of this approach for studies of non-model organisms is that there is no need for the construction of species-specific microarrays.

Future directions

As outlined above, the past few years have seen remarkable growth in the range and scope of applications of genetic tools to evolutionary problems. Many conceptual, technological and analytical innovations, driven by work on the major genomic models, have hugely increased the applicability of these methods to non-model organisms. As sequence and genome structure data continue to accumulate, the possibilities for using these data for comparative purposes, and for leaping between species, will become ever richer. Given the rate of development, predicting where the limits of technology will lie in five years' time is very difficult.

Coupled with the huge increase in the applicability of genomic resources has been an improved understanding of the operation of natural selection from a quantitative genetic perspective in wild populations. This has been driven by the increased availability of high quality, long-term data sets from natural population studies, as well as the application of analytical techniques from other fields. These new studies have, in some cases, studied selection, phenotypic changes, and inheritance over tens of generations, providing richly detailed insights into the importance of spatial and temporal variation and their interaction with selection and inheritance.

The merging of these two research traditions offers the possibility for important insights into major problems in the evolutionary biology of wild populations; we outline three candidates here. First, the ability to identify specific genetic loci influencing phenotypes will enable a much more precise understanding of what constitutes a trait that can be the target of selection. For example, there is considerable interest at present in the evolutionary ecology of plasticity, partly owing to its relevance to adaptation to human-induced environmental change, and a reaction norm approach is often taken, where the expression of phenotypes across environment is analysed as a trait[63]. Knowledge of the molecular genetic basis of characters, coupled with expression studies, will enable the determination of whether phenotypes in different environments are really the same traits, and also potentially of the loci ('plasticity genes') which control the expression of these phenotypes across environments. In general, field biologists have tended to ignore the problem of the relationship between traits and underlying genetic causes (the 'phenotypic gambit'), but there is no longer any excuse to do so.

Second, with an increased focus on the loci underlying traits, several related issues are brought closer to solution. For example, recent interest in the role of sexually antagonistic genetic variation is based on classical quantitative genetics[4,5] or laboratory-based *Drosophila* breeding experiments[3]. A genomic perspective allows the possibility—for example through whole-genome scans or the analysis of sex-by-linkage-disequilibrium interactions—of identifying and characterizing any sex-antagonistic genes, and testing evolutionary theories about their genomic location.

Genomic approaches can also be used to determine the molecular genetic basis of genetic correlations, whether resulting from linkage disequilibrium or pleiotropy, and hence are of considerable importance in determining their role in directing evolutionary trajectories. A much-debated question in evolutionary biology concerns the stability of the genetic variance–covariance matrix. While assessing the temporal component of this problem may still be beyond reach, it might be approached at least from a spatial or environmental perspective. The genetic structure of traits is also of great importance for debates about the forces driving sexual selection, and the evolution of costly mate choice. The application of genomics has the potential to resolve questions such as the importance of additive and non-additive genetic variance in sexual display characters, and hence to

determine the extent to which female choice predicated solely on characteristics that can be transmitted to offspring is plausible[64].

Third, a general problem concerns the extent to which environmental and temporal variation act to maintain genetic variation for characters under fluctuating selection. There are many candidate processes that can maintain genetic variation in natural populations[29]; some of these processes can be studied by linking genes and fitness variation in ecological studies. For example, testing for the presence and importance of QTL and gene–environment interactions will shed light on this problem from a within-population perspective, whereas comparisons of selection on allelic variants in populations in different selective environments can place this question in a spatial context. Extending this question to a temporal framework offers the possibility of addressing (for example by applying a re-sequencing approach involving archived specimens or samples, and contemporary populations) the extent to which populations are continually undergoing adaptive evolution, or are relatively static.

Challenge and conclusion

A challenge must be accepted for this synthesis to occur. Each field has been based around the choice of different attributes for model organisms. It is inevitable that our understanding of evolutionary genomics is best developed in those organisms that have longest served as models for laboratory genetics (*Drosophila*, rodents), or that have most relevance to biomedical science (for example, primates). The most effective choice for studies of fitness in wild populations is diurnally active vertebrates that live at high densities (ungulate mammals, passerine birds, squamate reptiles). This has affected the kinds of problems that are tackled; more overlap between these questions should develop as more genomic information becomes available. For example, bird genomes appear to show high degrees of conservation in terms of organization, suggesting that with the impending addition of the zebra finch to the completed genome sequences of birds, much more work on wild birds that is genomic in focus will become possible. Equally, the increased ease of obtaining genomic information may reduce the need for observations of matings or parentage that motivates the choice of ecological models. Nevertheless, it would be naive to suggest that purely field-based models are soon likely to offer the opportunities that laboratory models do: there is a continuum between experimental model and ecological realism, and the most productive ground may lie with species that occupy an intermediate position on this continuum, such as sticklebacks or mice. Some field biologists may need to reassess their choice of model species, or be prepared to wait some time before they can apply the full range of genomic approaches. Equally, some laboratory-based scientists may need to consider whether placing their studies within an ecologically relevant context may offer, in the long run, more insight into contemporary evolution.

1. Mitchell-Olds, T. & Schmitt, J. Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* **441,** 947–952 (2006).
2. Rice, W. R. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature* **381,** 232–234 (1996).
3. Chippindale, A. K., Gibson, J. R. & Rice, W. R. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proc. Natl Acad. Sci. USA* **98,** 1671–1675 (2001).
4. Fedorka, K. M. & Mousseau, T. A. Female mating bias results in conflicting sex-specific offspring fitness. *Nature* **429,** 65–67 (2004).
5. Foerster, K. *et al.* Sexually antagonistic genetic variation for fitness in red deer. *Nature* **447,** 1107–1110 (2007).
6. Wilson, A. J. *et al.* Environmental coupling of selection and heritability limits evolution. *PLoS Biol.* **4,** e216 (2006).
7. Kruuk, L. E. B. *et al.* Antler size in red deer: heritability and selection but no evolution. *Evol. Int. J. Org. Evol.* **56,** 1683–1695 (2002).
8. Grant, P. R. & Grant, B. R. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296,** 707–711 (2002).
9. Garant, D., Kruuk, L. E. B., McCleery, R. H. & Sheldon, B. C. Evolution in a changing environment: a case study with great tit fledging mass. *Am. Nat.* **164,** 115–129 (2004).
10. Knapczyk, F. N. & Conner, J. K. Estimates of the average strength of selection are not inflated by sampling error or publication bias. *Am. Nat.* **170,** 501–508 (2007).
11. Kruuk, L. E. B. Estimating genetic parameters in natural populations using the "animal model". *Phil. Trans. R. Soc. Lond. B* **359,** 873–890 (2004).
12. Kruuk, L. E. B. *et al.* Heritability of fitness in a wild mammal population. *Proc. Natl Acad. Sci. USA* **97,** 698–703 (2000).
13. Pelletier, F., Clutton-Brock, T. H., Pemberton, J. M., Tuljapurkar, S. & Coulson, T. The evolutionary demography of ecological change: linking trait variation and population growth. *Science* **315,** 1571–1574 (2007).
14. Charmantier, A., Perrins, C., McCleery, R. H. & Sheldon, B. C. Quantitative genetics of age at reproduction in wild swans: support for antagonistic pleiotropy models of senescence. *Proc. Natl Acad. Sci. USA* **103,** 6587–6592 (2006).
15. Conover, D. O. & Schultz, E. T. Phenotypic similarity and the evolutionary significance of counter-gradient variation. *Trends Ecol. Evol.* **10,** 248–252 (1995).
16. Laugen, A. T. *et al.* Latitudinal countergradient variation in the common frog (*Rana temporaria*) development rates—evidence for local adaptation. *J. Evol. Biol.* **16,** 996–1005 (2003).
17. Merila, J., Kruuk, L. E. B. & Sheldon, B. C. Cryptic evolution in a wild bird population. *Nature* **412,** 76–79 (2001).
18. Garant, D., Kruuk, L. E. B., Wilkin, T. A., McCleery, R. H. & Sheldon, B. C. Evolution driven by differential dispersal within a wild bird population. *Nature* **433,** 60–65 (2005).
19. Wilson, A. J. *et al.* Quantitative genetics of growth and cryptic evolution of body size in an island population. *Evol. Ecol.* **21,** 337–356 (2007).
20. Hoekstra, H. E. & Coyne, J. A. The locus of evolution: evo devo and the genetics of adaptation. *Evol. Int. J. Org. Evol.* **61,** 995–1016 (2007).
21. King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188,** 107–116 (1975).
22. Carroll, S. B. *Endless Forms Most Beautiful: the New Science of Evo-Devo* (W. W. Norton & Co., New York, 2005).
23. Prud'homme, B. *et al.* Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. *Nature* **440,** 1050–1053 (2006).
24. Borneman, A. R. *et al.* Divergence of transcription factor binding sites across related yeast species. *Science* **317,** 815–819 (2007).
25. McGregor, A. P. *et al.* Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* **448,** 587–590 (2007).
26. ffrench-Constant, R. H., Rocheleau, T. A., Steichen, J. C. & Chalmers, A. E. A point mutation in a *Drosophila* GABA receptor confers insecticide resistance. *Nature* **363,** 448–451 (1993).
27. Bustamante, C. D. *et al.* Natural selection on protein-coding genes in the human genome. *Nature* **437,** 1153–1157 (2005).
28. Vera, J. C. *et al.* Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Mol. Ecol.* advance online publication, doi:10.1111/j.1365-294x.2008.03666.x (5 February 2008).
29. Mitchell-Olds, T., Willis, J. H. & Goldstein, D. B. Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Rev. Genet.* **8,** 845–856 (2007).
30. Williams, J. T. & Blangero, J. Power of variance component linkage analysis to detect quantitative trait loci. *Ann. Hum. Genet.* **63,** 545–563 (1999).
31. Barton, N. H. & Keightly, P. D. Understanding quantitative genetic variation. *Nature Rev. Genet.* **3,** 11–21 (2002).
32. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* (Sinauer Associates, Sunderland, Massachusetts, 1998).
33. Slate, J. *et al.* A genome scan for quantitative trait loci in a wild populations of red deer (*Cervus elaphus*). *Genetics* **162,** 1863–1873 (2002).
34. Beraldi, D. *et al.* Mapping quantitative trait loci underlying fitness-related traits in a free-living sheep population. *Evol. Int. J. Org. Evol.* **61,** 1403–1416 (2007).
35. Colosimo, P. F. *et al.* The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol.* **2,** e109 (2004).
36. Colosimo, P. F. *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307,** 1928–1933 (2005).
37. Backström, N., Qvarnstrom, A., Gustafsson, L. & Ellegren, H. Levels of linkage disequilibrium in a wild bird population. *Biol. Lett.* **2,** 435–438 (2006).
38. Slate, J. & Pemberton, J. M. Admixture and patterns of linkage disequilibrium in a free-living vertebrate population. *J. Evol. Biol.* **20,** 1415–1427 (2007).
39. Yu, J. *et al.* A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genet.* **38,** 203–208 (2006).
40. Williamson, S. *et al.* Localizing recent adaptive evolution in the human genome. *PloS Genet.* **3,** e90 (2007).
41. Pennings, P. S. & Hermisson, J. Soft sweeps. III: The signature of positive selection from recurrent mutation. *PloS Genet.* **2,** e186 (2006).
42. Teshima, K. M., Coop, G. & Preworski, M. How reliable are empirical genomic scans for selective sweeps? *Genome Res.* **16,** 702–712 (2006).
43. Beaumont, M. A. Adaptation and speciation: what can F_{ST} tell us? *Trends Ecol. Evol.* **20,** 435–440 (2005).
44. Storz, J. F. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol. Ecol.* **14,** 671–688 (2005).
45. Luikart, G., England, P. R., Tallman, D., Jordan, S. & Taberlet, P. The power and promise of population genomics: from genotyping to genome typing. *Nature Rev. Genet.* **4,** 981–993 (2003).
46. Campbell, D. & Bernatchez, L. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Mol. Biol. Evol.* **21,** 945–956 (2004).

47. Bonin, A. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Mol. Ecol.* **23,** 773–783 (2006).
48. Rogers, S. M. & Bernatchez, L. Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeaformis*). *Mol. Ecol.* **14,** 351–361 (2005).
49. Lin, J. Y. & Fisher, D. E. Melanocyte biology and skin pigmentation. *Nature* **445,** 843–850 (2007).
50. 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).
51. Rosenblum, E. B., Hoekstra, H. E. & Nachman, M. W. Adaptive reptile color variation and the evolution of the *MC1R* gene. *Evolution Int. J. Org. Evolution* **58,** 1794–1808 (2004).
52. Mundy, N. I. *et al.* Conserved genetics basis of a quantitative plumage trait involved in mate choice. *Science* **303,** 1870–1873 (2004).
53. Steiner, C. C., Weber, J. N. & Hoekstra, H. E. Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol.* **5,** e239 (2007).
54. Miller, C. T. *et al. cis*-regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* **131,** 1179–1189 (2007).
55. Protas, M. E. *et al.* Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genet.* **38,** 107–111 (2006).
56. Dahlhoff, E. P. & Rank, N. E. Functional and physiological consequences of genetic variation at phosphoglucose isomerase: Heat shock protein expression is related to enzyme genotype in a montane beetle. *Proc. Natl Acad. Sci. USA* **97,** 10056–10061 (2000).
57. Haag, C. R., Saastamoinen, M., Marden, J. H. & Hanski, I. A candidate locus for variation in dispersal rate in a butterfly metapopulation. *Proc. Biol. Sci.* **272,** 2449–2456 (2005).
58. Whitehead, A. & Crawford, D. L. Variation within and among species in gene expression: raw material for evolution. *Mol. Ecol.* **15,** 1197–1211 (2006).
59. Derome, N. & Bernatchez, L. The transcriptomics of ecological convergence between two limnetic coregonine fishes (Salmonidae). *Mol. Biol. Evol.* **23,** 2370–2378 (2006).
60. Gilad, Y., Oshlack, A. & Rifkin, S. A. Natural selection on gene expression. *Trends Genet.* **22,** 456–461 (2006).
61. Gibson, G. & Weir, B. The quantitative genetics of transcription. *Trends Genet.* **21,** 616–623 (2005).
62. Ellegren, H. & Parsch, J. The evolution of sex-biased genes and sex-biased gene expression. *Nature Rev. Genet.* **8,** 689–698 (2007).
63. Nussey, D. H., Postma, E., Gienapp, P. & Visser, M. E. Selection on heritable phenotypic plasticity in a wild bird population. *Science* **310,** 304–306 (2005).
64. Charmantier, A. & Sheldon, B. C. Testing genetic models of mate choice evolution in the wild. *Trends Ecol. Evol.* **21,** 417–419 (2006).
65. Frank, S. A. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175,** 373–388 (1995).
66. Lande, R. A quantitative genetic theory of life history evolution. *Ecology* **63,** 607–615 (1982).
67. Lande, R. & Arnold, S. J. The measurement of selection on correlated characters. *Evolution Int. J. Org. Evolution* **37,** 1210–1226 (1983).
68. Metcalf, C. J. E. & Pavard, S. Why evolutionary biologists should be demographers. *Trends Ecol. Evol.* **22,** 205–212 (2007).
69. Coulson, T. *et al.* Estimating individual contributions to population growth: evolutionary fitness in ecological time. *Proc. R. Soc. Lond. Ser. B* **273,** 547–555 (2006).
70. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, 1930).
71. Price, G. R. Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.* **36,** 129–140 (1972).
72. Schneider, R. A. & Helms, J. A. The cellular and molecular origins of beak morphology. *Science* **299,** 565–568 (2003).
73. Abzhanov, A. *et al. Bmp4* and morphological variation of beaks in Darwin's finches. *Science* **305,** 1462–1465 (2004).
74. Abzhanov, A. *et al.* The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* **442,** 563–567 (2006).

Acknowledgements This work was supported by grants from the Swedish Research Council (H.E.) and by a Royal Society University Research Fellowship and an Erskine Fellowship (B.C.S).

Author Contributions H.E. and B.C.S. wrote the paper together.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.E. (hans.ellegren@ebc.uu.se) or B.C.S. (ben.sheldon@zoo.ox.ac.uk).