DOI: 10.1534/genetics.109.109991

Perspectives

Anecdotal, Historical and Critical Commentaries on Genetics

Darwin and Genetics

Brian Charlesworth[1] and Deborah Charlesworth

Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh EH9 3JT, United Kingdom

ABSTRACT

Darwin's theory of natural selection lacked an adequate account of inheritance, making it logically incomplete. We review the interaction between evolution and genetics, showing how, unlike Mendel, Darwin's lack of a model of the mechanism of inheritance left him unable to interpret his own data that showed Mendelian ratios, even though he shared with Mendel a more mathematical and probabilistic outlook than most biologists of his time. Darwin's own "pangenesis" model provided a mechanism for generating ample variability on which selection could act. It involved, however, the inheritance of characters acquired during an organism's life, which Darwin himself knew could not explain some evolutionary situations. Once the particulate basis of genetics was understood, it was seen to allow variation to be passed intact to new generations, and evolution could then be understood as a process of changes in the frequencies of stable variants. Evolutionary genetics subsequently developed as a central part of biology. Darwinian principles now play a greater role in biology than ever before, which we illustrate with some examples of studies of natural selection that use DNA sequence data and with some recent advances in answering questions first asked by Darwin.

The power of Selection, whether exercised by man or brought into play under nature through the struggle for existence and the consequent survival of the fittest, absolutely depends on the variability of organic beings. Without variability, nothing can be effected; slight individual differences, however, suffice for the work, and are probably the chief or sole means in the production of new species.

Charles Darwin (1868)

CHARLES Darwin was the first person to appreciate clearly that evolution depends on the existence of heritable variability within a species to generate the differences between ancestral and descendant populations. The development of Darwin's thoughts on the nature and causes of evolution is clearly documented in his "transmutation" notebooks of 1836–1838 (Barrett *et al.* 1987). Once he had decided that species originated by "descent with modification," Darwin quickly realized the need to find a mechanism for accomplishing the changes involved. In formulating the idea of natural selection, he was greatly influenced by the experience of breeders in artificially selecting populations of domestic animals and plants. Chapter 1 of *The Origin of Species* (Darwin 1859) is famously devoted to documenting the existence of variability in these populations and the effectiveness of artificial selection:

> The key is man's power of cumulative selection: nature gives successive variations; man adds them up in certain directions useful to himself (Darwin 1859, p. 30).

It was only a short step to applying this observation to selection in nature:

> Can it, then, be thought improbable, seeing that variations useful to man have undoubtedly occurred, that other variations useful in some way to each being in the great and complex battle of life, should sometimes occur in the course of thousands of generations? ... This preservation of favourable variations and the rejection of injurious variations, I call Natural Selection (Darwin 1859, pp. 81–82).

Most of the books and papers that Darwin published after *The Origin of Species* were devoted to describing how a vast range of biological phenomena—from the sexual systems of plants to human anatomy and behavior—could be interpreted in terms of evolution by natural selection or by the special form of natural selection represented by sexual selection. Surprisingly (at least from today's perspective), many biologists were, for a long time, far from convinced that natural selection was the predominant guiding force in evolution. This continued into the 1920s. In the Introduction to Volume 1 of his treatise on evolutionary genetics, Sewall Wright noted:

> Along with the universal acceptance by biologists of evolution as a fact, there came to be increasing dissatis-

[1]*Corresponding author:* Institute of Evolutionary Biology, Ashworth Laboratories, King's Buildings, West Mains Rd., Edinburgh EH9 3JT, United Kingdom. E-mail: brian.charlesworth@ed.ac.uk

faction, during the latter part of the nineteenth century, with natural selection as the master theory of causation (Wright 1968, pp.7–8).

SKEPTICISM ABOUT NATURAL SELECTION

Prominent early geneticists such as William Bateson, Hugo de Vries, and Richard Goldschmidt were notorious skeptics about natural selection and the evolutionary role of the small individual differences relied on by Darwin, emphasizing instead the role of mutations with large and manifold effects (Provine 1971). Many naturalists and paleontologists held what now seem to us to be semi-mystical theories, such as internal drives to improvement or perfection; many of them espoused Lamarckian views up until the 1930s (in France and in the Soviet Union and its satellites, Lamarckism persisted well into the 1960s). In his classic history of modern science, *The Edge of Objectivity*, Charles Coulston Gillispie quotes the leading historian of biology in 1929, Erik Nordenskiold, as stating that the proposition that natural selection "does not operate in the form imagined by Darwin must certainly be taken as proved" (Gillispie 1960, p. 320). The book *Evolution in the Light of Modern Knowledge*, a compendium of essays by 13 leading British biologists, published by Blackie and Son in 1925 to provide (according to the publisher's note) "an authoritative statement" about "the doctrine of evolution...after the general upheaval of fundamental theories in the past 20 years", has no index reference to natural selection. This contrasts with 3253 articles mentioning natural selection and evolution in 2008 in the Web of Science database. For a detailed discussion of anti-Darwinian evolutionary ideas, see Bowler (1983) and Gayon (1998).

Why was there such skepticism toward natural selection, and why have things changed so much? One reason was the lack during Darwin's lifetime of direct evidence for natural selection. This started to change in the late 19th and early 20th centuries through the work of Bumpus (1899) in the United States, and Weldon (1895, 1901) and his student Di Cesnola (1907) in Europe. These scientists initiated the field now known as ecological genetics, and we now have literally thousands of examples where field naturalists have demonstrated the operation of natural selection in the wild on both discrete polymorphisms and quantitative traits (Kingsolver *et al.* 2001; Bell 2008; Leimu and Fischer 2008).

DARWIN'S DIFFICULTIES WITH INHERITANCE: MENDEL TO THE RESCUE

The other major factor, of course, was the fact that Darwin failed to arrive at an understanding of the mechanism of inheritance, despite realizing its importance and devoting a vast effort to assembling evidence in his *Variation of Animals and Plants Under Domestication* (Darwin 1868). Unfortunately, he was unaware of Mendel's work, despite its publication 2 years earlier (Mendel 1866). Mendel's work has now, of course, permanently revolutionized our understanding of heredity, and his tragic failure to obtain recognition in his lifetime is a well-known story. It is less well known that Mendel was well aware of the importance for evolution of understanding genetics:

> This seems to be the one correct way of finally reaching a solution to a question whose significance for the evolutionary history of organic forms cannot be underestimated (Mendel 1866, p. 2).

Sadly, even if Mendel had lived to see the rediscovery of his work, he probably would not have had the satisfaction of seeing it contribute to evolutionary understanding because, even after genetics had begun its rapid development in the early decades of the 20th century, evolutionary biologists initially failed to understand how to incorporate genetics into their work. We will outline these failures to achieve a synthesis later, but first consider Darwin's efforts to understand inheritance and how his approach fell short of Mendel's.

Mendel's ability to solve the most difficult problem in 19th century biology after the mechanism of evolution rests on his use of a then-unique approach: combining rigorous genetic experiments with quantitative, probabilistic predictions about their expected outcomes: in other words, using biological data to test a quantitative hypothesis. It is a triumph of productive theoretical reasoning that Mendel proposed his particulate inheritance hypothesis well before a proper understanding of the cellular basis of sexual reproduction was achieved by either animal or plant biologists (Farley 1982).

This achievement eluded Darwin, the other greatest mind in 19th century biology, although he came close to seeing the same phenomena as Mendel did and frequently looked at data in a quantitative manner (Howard 2009). Darwin repeatedly referred to the phenomenon of reversion to ancestral types in *Variation of Animals and Plants Under Domestication* (Darwin 1868). He also compiled examples of the transmission of traits down several generations of pedigrees and obtained help from the mathematical physicist Sir George Stokes to show that these cases are unlikely to be due to chance, one of the first examples of a test of statistical significance in biology (Darwin 1868, chap. 12).

Ironically, Darwin analyzed data from his own crossing experiments on distyly in Primula species (summarized in Darwin 1877, chap. 5), which gave what we can now see as clear evidence for Mendelian ratios (see also Bulmer 2003, p.112, and Howard 2009). In distylous species (Figure 1), the long-styled morphs (L) are now known to be homozygotes *ss* for the alleles at several loci in a supergene controlling style length, stamen position,

pollen and stigma placement, morphology, and incompatibility, whereas the short-styled morph (S) is heterozygous *Ss*. The only matings that invariably succeed are L × S and S × L (Darwin called these "legitimate pollinations"), and these give a 1:1 ratio of L and S plants (Table 1). It is occasionally possible to obtain seeds by self-fertilization, in which case L plants produce only L offspring (Table 1). Darwin stated:

> From the long-styled form, always fertilised with its own-form pollen, I raised in the first generation three long-styled plants, from their seed 53 long-styled grandchildren, from their seed 4 long-styled great-grandchildren, from their seed 20 long-styled great-great-grandchildren, and lastly, from their seed 8 long-styled and 2 short-styled great-great-great-grandchildren. . . .altogether 162 plants were raised, . . . 156 long-styled and 6 short-styled (Darwin 1877, pp. 228–229).

The few short-styled plants in the final generation were presumably contaminants (Darwin's experiments were remarkably free from them). Self-pollination of S plants should generate a 3:1 ratio, as Darwin found (see Table 1; none of the ratios deviates significantly from the expected ratio). He remarked:

> I raised at first from a short-styled [*P. sinensis*] plant fertilised with its own-form pollen one long-styled and seven short-styled illegitimate seedlings. . . . Dr. Hildebrand raised fourteen plants, of which eleven were short-styled and three long-styled (Darwin 1877, p. 216).

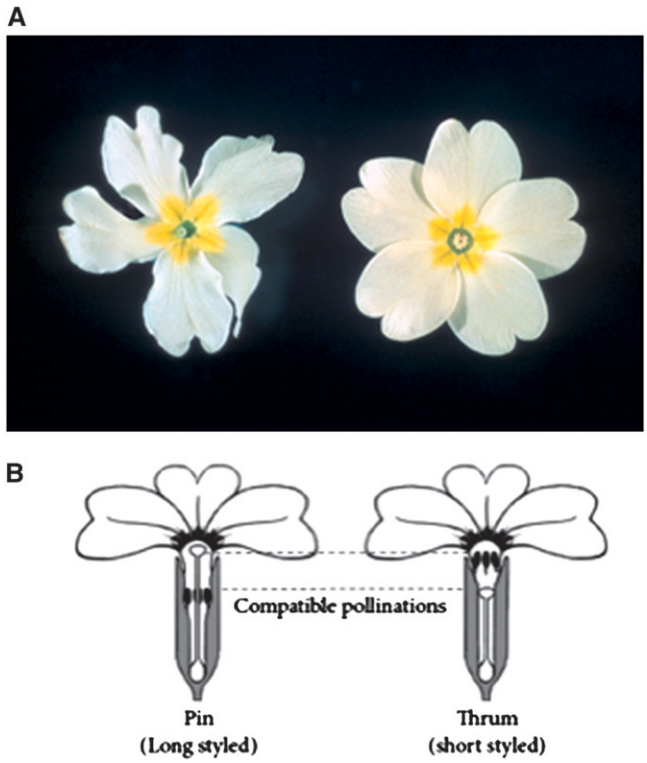

Figure 1.—Distyly in primroses. (A) Long-styled (pin) and short-styled (thrum) flowers of *Primula vulgaris*. (B) Vertically sectioned flowers, with the compatible pollinations indicated. (Pollen from high anthers is compatible with stigmas of long-styled plants, and pollen from low anthers is compatible with stigmas of short-styled plants, while the other two types of pollinations are incompatible.)

DARWIN'S DIFFICULTIES WITH BLENDING INHERITANCE AND HIS THEORY OF PANGENESIS

Darwin failed to understand the significance of these results because he had no model of particulate inheritance that could be applied to genetic data. Indeed, Darwin appears to have maintained a belief in the predominance of blending inheritance, as did nearly all of his contemporaries. As Fisher pointed out in chapter 1 of *The Genetical Theory of Natural Selection* (Fisher 1930), there are few explicit statements on this in Darwin's published works, although they appear in some of his unpublished notes and essays. In addition, chapter 15 of *Variation of Animals and Plants Under Domestication* (Darwin 1868) starts with a section *On Crossing as a Cause of Uniformity of Character*, which implicitly assumes that crossing leads to blending. It is unclear, however, to what extent he thought that an offspring was a product of the complete fusion of the genetic contributions of its parents (Bulmer 2003, chap. 4).

Blending inheritance leads to a difficulty that was forcefully pointed out by Fleeming Jenkin (Jenkin 1867), the professor of engineering at the University of Edinburgh (the building next to ours is, somewhat unfortunately perhaps, named after him). Under blending inheritance, variation decays rapidly because the genotypes of the offspring of a cross are all the same and are intermediate between those of the two parents. With random mating, the genetic variance of a quantitative trait then decays by a factor of one-half each generation (Fisher 1930, p. 4). Acceptance of blending inheritance clearly raises doubts about the ability of either natural or artificial selection to make permanent changes in a population. In the sixth edition of *The Origin of Species*, published in 1872, Darwin reacted to Jenkin as follows:

> Nevertheless, until reading an able and valuable article in the "North British Review" (1867), I did not appreciate how rarely single variations, whether slight or strongly-marked, could be perpetuated (Darwin 1859, pp. 111–112).

Since heritable variability is required for selection to be effective, and Darwin's survey of the results of artificial selection had convinced him that there is enough variation for it to be effective, Darwin sought a way of generating an abundance of such variation. This was provided by his theory of pangenesis, according to which variations experienced by the individual during its lifetime are transmitted to the germ cells by hypothetical "gemmules" (Darwin 1868, chap. 27). This is an hypothesis of the inheritance of acquired characters, which Darwin accepted as an experimentally established

TABLE 1

Darwin's results for progeny of long- and short-styled Primula crossed with the same morph

Primula species	Parental morph	No. of long-styled progeny	No. of short-styled progeny
P. vulgaris	Long	69	0
P. veris (cowslip)		156	6
P. sinensis		52	0
P. auricula	Short	25	75
P. veris (cowslip)		5	9
P. sinensis		1	24

fact (there is an extensive discussion on the transmission of mutilations in Darwin 1868, chap. 12).

However, Darwin was clearly not quite sure about this. For example, he mentioned that the circumcision of male infants has not led to a loss of the foreskin in the Jewish community (Darwin 1868, Vol. 1, p. 558). He also noted that there are some instances of evolution that cannot be explained by this hypothesis, notably the adaptive characteristics of the sterile castes of social insects:

> For no amount of exercise, or habit, or volition, in the utterly sterile members of a community could possibly have affected the structure or instincts of the fertile members, which alone leave descendants. I am surprised that no one has advanced this demonstrative case of neuter insects, against the well-known doctrine of Lamarck (Darwin 1859, p. 242).

Darwin's use of this natural case of sib selection to refute Lamarckism anticipates later uses of the same reasoning, which reached a peak of perfection in the Lederbergs' experiments on replica plating in *Escherichia coli* (Lederberg and Lederberg 1952).

EVOLUTION AS A PROCESS OF CHANGE IN THE FREQUENCIES OF MENDELIAN VARIANTS WITHIN POPULATIONS

Unlike Darwin, who regarded the inheritance of acquired characters largely as a source of variation on which selection could act, the 20th century advocates of Lamarckian inheritance viewed it as an alternative explanation of adaptive evolution. As was brilliantly laid out by Fisher in chapter 1 of *The Genetical Theory of Natural Selection*, and as is no doubt familiar to readers of Genetics, all the difficulties posed by blending disappear with Mendelian inheritance: variability within a population is conserved, not lost, when no evolutionary forces are acting, a genetic equivalent to Galileo's law of inertia. The inheritance of acquired characters is therefore not needed for the regeneration of genetic variability.

It is, of course, well known that our knowledge of the physical basis of genes and of their behavior now largely excludes Lamarckian inheritance. However, recent studies have uncovered some situations in which the DNA of certain genome regions is modified during the life of an individual, and these epigenetic marks with functions in developmental control and other processes can sometimes pass via meiotic divisions to descendant generations (*e.g.*, Cubas *et al.* 1999; Richards 2006; Namekawa *et al.* 2007; Heijmans *et al.* 2008; Sidorenko and Chandler 2008). If variants that arise in this way are stably transmitted, then they can be treated as Mendelian variants that can be exploited in evolution. If their inheritance is unstable, as is often the case, they cannot contribute significantly to evolution.

The breakthrough in understanding the nature of variation in quantitative traits (equivalent to Darwin's "slight differences, which may be called "individual differences"; see the epigraph to this article) came in the early years of genetics, starting with experiments with pure lines, whose individuals have virtually identical genotypes. These experiments showed that plentiful phenotypic variation exists among such individuals but is not transmissible to the offspring (Johannsen 1909; Wright 1920), leading to the rejection of Lamarckian inheritance by the genetics community. Furthermore, the variability of quantitative traits (which often show apparent blending in F_1 crosses between pure lines) increases in F_2 and later generations (Nilsson-Ehle 1909; East 1910), as expected with particulate Mendelian inheritance. Moreover, the factors responsible can be mapped to chromosomal regions and sometimes (with modern methods) to single genes or nucleotide variants (Flint and Mackay 2009). Even initially puzzling cases of very complex patterns of inheritance, such as beaded and truncate wing in Drosophila, were traced to factors linked to chromosomal genes, and the virtual universality of Mendelism was established by the early 1920s (Altenburg and Muller 1920). In contrast to the inheritance of acquired characters, mutations were found to be very rare, stable modifications of genes and to arise independently of whether or not they confer increased fitness in a given environment (Muller 1932).

By the 1920s, it was clear that (contrary to the beliefs of many early geneticists, who emphasized the large effects of dramatic mutations and ignored the evidence for the Mendelian basis of quantitative trait variation), Darwin-

ian evolution by natural selection is, in fact, enabled by Mendelian inheritance: mutations in genes provide the source of new, stable variants on which selection can act. This set the stage for understanding that evolution is fundamentally a process of change in the frequencies of Mendelian variants within populations and species, leading to the development of classical population and quantitative genetics. The fascinating struggle to reach this understanding is ably described by Provine (1971).

HOW RELEVANT IS NATURAL SELECTION TO MODERN BIOLOGY?

The chief post-Darwin component of major importance in modern evolutionary thinking is the idea of genetic drift and, specifically, the possibility that a significant portion of variability and evolution of DNA sequence variants is driven by random fluctuations in the frequencies of variants with little or no effects on fitness (Kimura 1983). Darwin himself had the idea of selective neutrality:

> Variations neither useful nor injurious would not be affected by natural selection, and would be left a fluctuating element, as perhaps we see in the species called polymorphic (Darwin 1859, p. 81).

In a surprising turn of events, the concept of selective neutrality has become a cornerstone of modern tests for natural selection, by providing a null hypothesis that can be tested against data on sequence variation and evolution. Evolutionary biology is now mature enough to repay its debt to genetics and indeed is now (together with genetic and molecular genetic approaches) central to work initiated with largely functional genetic motivations, including genome sequencing.

Given some genetic variation in a phenotype of interest, ecological genetic approaches can relate fitnesses to the differences between individuals within a single natural population, sometimes using data on undisturbed individuals (Bell 2008). With more disturbance to the organisms, between-population differences can also be tested for their selective importance by using methods such as reciprocal transplant experiments. Changes in genotype frequencies can be followed over time in such experiments or after perturbing alleles from their natural frequencies. These approaches have firmly documented the action of selection, sometimes on unexpected characters such as the inversion polymorphisms of Drosophila (Wright and Dobzhansky 1946). However, this approach may miss many instances of selection, because even the largest and most sensitive experiments, such as those involving competition between strains of yeast or bacteria, cannot detect selective differences $<\sim 0.1\%$ in magnitude (Dykhuizen 1990).

At the other extreme of the evolutionary timescale, the comparative approach can be used to relate differences in ecological conditions experienced by different evolving lineages to differences in the outcome of evolution by natural selection (Harvey and Pagel 1991). Darwin was the first biologist to explicitly use the comparative approach for this purpose. This approach is now highly statistical (Felsenstein 2004) and often uses sequence-based phylogenies, which have the advantage of being much less susceptible to the action of natural selection in causing variation in the rate and direction of character change than the morphological traits formerly used in phylogenetic analysis. Even without modern methods, Darwin used the comparative method to good effect in his work on plant mating system evolution, for example, in his review of the literature to show that inbreeding plants have smaller flowers and are generally less attractive to pollinators compared with outcrossing ones (Darwin 1876), a finding that has held up in more comprehensive modern studies and that tells us that attracting pollinators consumes resources (*e.g.*, Ornduff 1969). The comparative approach is, however, incapable of providing estimates of the intensity of selection involved in causing the changes observed.

Modern DNA sequencing technology provides population geneticists with the ability to study the extent to which selection acts on variants across the genome, as opposed to mutation and random genetic drift. After several decades of using the ecological genetic and comparative approaches to detect selection in nature on visible or physiological traits, biologists can now test for the selective effects of specific genetic differences between individuals without needing to know their phenotypic effects. For these tests, neutrality provides an essential null hypothesis. With our newly acquired ability to apply statistical population genetics methods to the analysis of patterns of within-species variation and between-species divergence in large, genomewide data sets, extremely weak pressures of selection, well below the resolution of experimental methods, can be detected and measured. Many of the approaches currently being used are closely based on the classical work of Fisher, Kimura, and Wright on the behavior of variants subject to mutation, selection, and genetic drift, which are summarized in Kimura's (1983) book, *The Neutral Theory of Molecular Evolution.* These methods are often extremely computationally intensive, especially when complications like recent changes in population size are taken into account.

With the increasing availability of large data sets on DNA sequence variation across the genomes of humans and *Drosophila melanogaster*, we are getting close to answering questions such as: What is the distribution of selection coefficients for newly arising deleterious amino acid mutations? What fraction of amino acid variants distinguishing related species are fixed by natural selection, as opposed to genetic drift acting on neutral or slightly deleterious variants? To what extent

are variants at synonymous coding sites and noncoding sites subject to selection, and how strong is this selection?

The results are sometimes quite startling. It has been fairly conclusively established, for example, that a typical human being is heterozygous for several hundred amino acid mutations, most of which have only very small effects on fitness (of the order of 10^{-3}) (Boyko *et al.* 2008), that ~50% of amino acid variants distinguishing related Drosophila species have been fixed by selection (Sella *et al.* 2009), and that more noncoding sites than coding sites in both humans and Drosophila can mutate to selectively deleterious alternatives that are rapidly removed by selection (Encode Project Consortium 2007; Haag-Liautard *et al.* 2007).

In addition to these direct tests of selection on variants, we can also use information on neutral or nearly neutral variants that are not themselves under selection to make inferences about selection at linked sites in the genome. One example is the detection of selective sweeps caused by the recent spread of selectively favorable mutations. The spread of an advantageous allele can quickly lead to very low variability in the gene affected, and closely linked regions may also have reduced diversity as a result of hitchhiking through the population of the segment of the chromosome that contained the original beneficial mutation (Maynard Smith and Haigh 1974). These effects on linked neutral or nearly neutral variants can be used in statistical tests for the action of natural selection. This has enabled geneticists to detect and estimate the strength of selection acting on genes such as drug resistance genes in the human malaria parasite by using the variability of microsatellite markers (*e.g.*, Nash *et al.* 2005) to detect numerous examples of recent adaptations in human populations from their effects on patterns of variation at linked SNPs (*e.g.*, Currat *et al.* 2002; Sabeti *et al.* 2002; Williamson *et al.* 2007; Akey 2009) and to search for genes involved in responses to artificial selection (Walsh 2008).

Conversely, high variability in a region can betray the action of natural selection acting in such a way that different alleles are maintained as polymorphisms for a long period by balancing selection; this divides the population into two or more compartments, between which neutral differences can accumulate at sites that are closely linked to the targets of selection, where recombination is ineffective at preventing differentiation between the compartments (Hudson 1990; Nordborg 1997). A well-known example is the human MHC region, in which not only are there many polymorphic amino acids in exon 2, which encodes most of the peptide-binding residues of the human mature MHC proteins, but after there are extraordinarily numerous polymorphic variants of synonymous and intron sites, compared with other loci in the same populations (Raymond *et al.* 2005). Similarly, frequency-dependent selection has clearly maintained sequence polymorphism for long evolutionary times at plant and fungal self-incompatibility loci, whose sequences are highly diverged (*e.g.*, Vekemans and Slatkin 1994; Richman *et al.* 1996; May *et al.* 1999).

Not only can selection within single populations be studied by molecular evolutionary approaches, but between-population differences due to local adaptation can also be revealed by increased divergence at sites linked to the targets of selection (Beaumont and Balding 2004). Indeed, scans of human and other species' genomes for sequences that are more differentiated than most genes are a major way of discovering candidates for genes that are currently under selection (Akey 2009).

MODERN PROGRESS TOWARD ANSWERING SOME OLD QUESTIONS: THE USE OF NATURAL GENETIC MARKERS

Another way in which modern evolutionary studies have contributed to genetics, as opposed to genetics contributing to evolutionary biology, is that an interest in quantifying the extent of genetic variation [initially motivated by a debate about whether variation within species is largely composed of recent mutations or includes a considerable proportion of variants maintained by balancing selection (Dobzhansky 1955; Lewontin 1974)] ultimately led to the discovery of vast numbers of DNA sequence variants that can be used as genetic markers for mapping (although these data did not in themselves settle the debate about whether selection maintains variation). The existence of abundant markers was predicted long ago:

> It would accordingly be desirable, in the case of man, to make an extensive and thorough-going search for as many factors as possible that could be used. . .as identifiers. They should, preferably, involve character differences that are (1) of common occurrence, (2) identifiable with certainty, (3) heritable in a simple Mendelian fashion. It seems reasonable to suppose that in a species so heterozygous there must really be innumerable such factors present. . . . It does seem clear that in the more tractable organisms, such as the domesticated and laboratory races of animals and plants, character analysis by means of linkage studies with identifying factors will come into more general use (Altenburg and Muller 1920).

In some species, naturally occurring markers can now be obtained so densely that new approaches are needed for genetic mapping because there is a very low chance of a crossover event between the closest markers (*e.g.*, Churchill *et al.* 2004; Van Os *et al.* 2005; Flibotte *et al.* 2009). The possibility of obtaining large numbers of genetic markers has produced renewed progress in mapping genes affecting quantitative characters, and new approaches are being developed for such studies, including association mapping that makes use of the population genetics concept of linkage disequilibrium

(associations between the allelic states of different loci or sites in a sequence; see Slatkin 2008 for an overview). The study of the population genetics of multilocus systems once appeared to be an esoteric field, remote from empirical data, which contributed to the reputation of theoretical population genetics for dryness and irrelevance to biology. Nevertheless, very important principles were developed that are now widely used by other geneticists, including ways to measure linkage disequilibrium [also now used to estimate recombination rates in genomes by using samples of sequences from populations (Myers *et al.* 2005)] and the concept that selection acting on a given sequence variant or allele has effects on closely linked variants (see above).

The kinds of approaches just mentioned are no longer restricted to humans and the genetics model organisms of most interest for functional molecular genetic work. One well-established use of markers is to infer the mating systems of populations in the wild (Ritland 1990). Darwin anticipated this when he used phenotypic differences, including flower colors, that he clearly assumed to be inherited, to infer the parentage of seeds:

> Altogether 233 plants were raised, of which 155 were mongrelised in the plainest manner, and of the remaining 78 not half were absolutely pure. I repeated the experiment by planting near together two varieties of cabbage with purple-green and white-green lacinated leaves; and of the 325 seedlings raised from the purple-green variety, 165 had white-green and 160 purple-green leaves. Of the 466 seedlings raised from the white-green variety, 220 had purple-green and 246 white-green leaves. These cases show how largely pollen from a neighbouring variety of the cabbage effaces the action of the plant's own pollen (Darwin 1876, p. 393).

It is now becoming possible to conduct fine-scale genetic mapping studies in nonmodel species, including those of applied interest, such as domesticated animals and plants and their pathogens, where QTL mapping is being aided by the abundant supply of new markers. Genetic mapping gives promise of testing hypotheses such as the close linkage of genes involved in heterostyly in Primula and other plant species (Li *et al.* 2007; Labonne *et al.* 2009) and mimicry in butterflies (Baxter *et al.* 2008), examples of problems that interested Darwin. Gene mapping is also important in modern work on the genetics of speciation, which is at last identifying genes involved in reproductive isolation between closely related species and is uncovering evidence for the Dobzhanky–Muller hypothesis that natural selection is important in causing genetic differences between populations that lower the survival or fertility of F_1 or F_2 hybrids, as a result of deleterious epistatic interactions between alleles derived from the two populations (*e.g.*, Barbash *et al.* 2003; Presgraves *et al.* 2003). As is well known, Darwin himself found the evolution of reproductive isolation puzzling:

> The importance of the fact that hybrids are very generally sterile, has, I think, been much underrated by some late writers. On the theory of natural selection the case is especially important as the sterility of hybrids could not possibly be of any advantage to them, and therefore could not have been acquired by the continued preservation of successive profitable degrees of sterility (Darwin 1859, p. 245).

However, the title *The Origin of Species* did not refer to this central puzzle concerning speciation, but rather to the evolution of adaptations and character differences; before the rise of genetics, it would have been virtually impossible for a correct interpretation of reproductive isolation to have been developed.

Another long-debated topic for which genetic marker availability should help our understanding is the question of the genetic basis of inbreeding depression and of heterosis. Although the deleterious effects of inbreeding were known to some earlier biologists, Darwin was the first to study the phenomenon thoroughly, because he realized that it provides an explanation for the existence of the elaborate adaptations of plants to avoid inbreeding. Darwin's book *The Effects of Cross and Self Fertilization in the Vegetable Kingdom* described his own experiments comparing progeny produced by self- and cross-fertilization in 57 plant species, and his summary of the main results anticipated future work that allowed us to measure inbreeding (in modern terms, inbreeding coefficients):

> That certain plants, for instance, *Viola tricolor, Digitalis purpurea, Sarothamnus scoparius, Cyclamen persicum*, etc., which have been naturally cross-fertilised for many or all previous generations, should suffer to an extreme degree from a single act of self-fertilisation is a most surprising fact. Nothing of the kind has been observed in our domestic animals; but then we must remember that the closest possible interbreeding with such animals, that is, between brothers and sisters, cannot be considered as nearly so close a union as that between the pollen and ovules of the same flower. Whether the evil from self-fertilisation goes on increasing during successive generations is not as yet known; but we may infer from my experiments that the increase if any is far from rapid. After plants have been propagated by self-fertilisation for several generations, a single cross with a fresh stock restores their pristine vigour; and we have a strictly analogous result with our domestic animals (Darwin 1876, p. 438).

As pointed out by Fisher in his *Design of Experiments* (Fisher 1935, chap. 3), Darwin used paired contrasts of the performance of an inbred and an outbred plant grown in the same pot, a method that is widely used in modern biological statistics. Darwin's insight into the utility of this approach was spoiled by the reanalysis of his data conducted by his cousin, Francis Galton, a supposedly more expert statistician.

Just as with his theory of sexual selection to explain male/female dimorphism (Darwin 1871), which was largely neglected until the 1970s, the individual selective advantage to outcrossing arising from inbreeding

depression postulated by Darwin was rejected by leading 20th century thinkers on plant evolution, such as C. D. Darlington and G. L. Stebbins, in favor of group selection hypotheses of advantages of increased variability to the population or species. The role of inbreeding depression in the evolution of mating systems is, however, now well established (Barrett 2002).

Although Darwin was unable to provide a satisfactory interpretation of his observations, inbreeding depression is now well known to be a genetic phenomenon, and hybrid vigor (heterosis) is widely used in agriculture. It is also well known that the genetic basis of these phenomena is difficult to ascertain and that this may impede efforts to make the best use of heterosis. There is no doubt that rare, deleterious mutations play an important role (Charlesworth and Charlesworth 1999): inbreeding, by producing homozygotes for such mutations, reduces survival and fertility because a large proportion of deleterious mutations are recessive, or partially so, and cause only slight harm when heterozygous, as was first clearly proposed by D. F. Jones (Jones 1917). Heterosis is also explicable on this basis because different inbred strains will be homozygous for different deleterious mutations, and different populations of a species in nature will differ similarly at some proportion of their genes, particularly if the populations are highly isolated (Ingvarsson *et al.* 2000; Escobar *et al.* 2008). It is still much less clear whether loci with overdominant alleles (alleles showing heterozygote advantage) also contribute any major part of inbreeding depression or heterosis, although it is intuitively easy to understand that, if such loci are common, these effects would be produced. Identification of the genetic factors involved in inbreeding depression or heterosis by the fine-scale mapping methods referred to above should help to answer these questions.

THE FUTURE IS BRIGHT!

The examples that we have outlined here show the value of the ongoing interaction between genetics and the study of evolution. From being a major headache for early supporters of evolution, genetics paved the way for models of evolution based on the known properties of inheritance, so that the constraints experienced by genes and genomes in evolution were correctly incorporated into quantitative models, and new possibilities, unknown to Darwin, were discovered.

Evolutionary genetics is inherently interdisciplinary, fruitfully combining models (often mathematical and often stochastic, given the nature of genetics) with empirical data. This intellectual tradition, now ~100 years old, deserves celebration along with Darwin's anniversaries. We hope that we have shown that evolution is more central to modern biological research than ever before and that this productive collaboration with genetics can be predicted to yield many further pure and applied scientific riches in the next hundred years. For this to happen, the need for a broad-enough education must be met. Biologists and doctors will need to understand genetics, and even some population genetics concepts, at least enough to collaborate with people with expertise in relevant quantitative methods. Mathematical ideas need to be demystified, as far as possible, so that biologists using phylogenetic and genetic marker or diversity analyses know what lies behind the computer programs that they use, an understanding without which the numbers that come out may lead to wrong conclusions. We need to regain a respect for the usefulness of statistics throughout biology and use it to test our ideas, as Darwin started to do. The same applies to theoretical modeling directed toward testable hypotheses, of which the idea of natural selection is still an excellent example, even though it has been extended to a far wider realm of biology than Darwin initially proposed and has given us many valuable tools at the interface between genetics and evolution. Darwin himself was interested in the functioning of organisms, not just in their morphology and relationships and the history of life, and he would surely have been delighted to see where his ideas have so far led us and how they have continued to be central within biology. In Dobzhansky's famous words:

> Nothing in biology makes sense except in the light of evolution (Dobzhansky 1973).

We thank Adam Wilkins and two reviewers for their helpful suggestions for improving the manuscript.

LITERATURE CITED

Akey, J. M., 2009 Constructing genomic maps of positive selection in humans: Where do we go from here? Genome Res. **19:** 711–722.

Altenburg, E., and H. J. Muller, 1920 The genetic basis of truncate wing: an inconstant and modifiable character in Drosophila. Genetics **5:** 1–59.

Barbash, D. A., D. F. Siino, A. M. Tarone and J. Roote, 2003 A rapidly evolving MYB-related protein causes species isolation in Drosophila. Proc. Natl. Acad. Sci. USA **6:** 5302–5307.

Barrett, P. H., P. J. Gautrey, S. Herber, D. Kohn and S. Smith, 1987 *Charles Darwin's Notebooks, 1836–1844.* Cornell University Press, Ithaca, NY.

Barrett, S. C. H., 2002 The evolution of plant sexual diversity. Nat. Rev. Genet. **3:** 274–284.

Baxter, S., R. Papa, N. Chamberlain, S. Humphray, M. Joron *et al.*, 2008 Convergent evolution in the genetic basis of Muellerian mimicry in Heliconius butterflies. Genetics **180:** 1567–1577.

Beaumont, M. A., and D. J. Balding, 2004 Identifying adaptive genetic divergence among populations from genome scans. Mol. Ecol. **13:** 969–980.

Bell, G., 2008 *Selection*, Ed. 2. Oxford University Press, Oxford.

Bowler, P. J., 1983 *Eclipse of Darwinism: Anti-Darwinian Evolution Theories in the Decade Around 1900.* Johns Hopkins University Press, Baltimore.

Boyko, A., S. H. Williamson, A. R. Indap, J. D. Degenhardt, R. D. Hernandez, *et al.*, 2008 Assessing the evolutionary impact of amino-acid mutations in the human genome. PLOS Genet. **5:** e1000083.

Bulmer, M. G., 2003 *Francis Galton.* Johns Hopkins University Press, Baltimore.

Bumpus, H. C., 1899 The elimination of the unfit as illustrated by the introduced sparrow, *Passer domesticus*. Biol. Lectures Woods Hole Marine Biol. Station **6:** 209–226.

Charlesworth, B., and D. Charlesworth, 1999 The genetic basis of inbreeding depression. Genet. Res. **74:** 329–340.

Churchill, G. A., D. C. Airey, H. Allayee, J. M. Angel, A. D. Attie *et al.*, 2004 The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat. Genet. **36:** 1133–1137.

Cubas, P., C. Vincent and E. S. Coen, 1999 An epigenetic mutation responsible for natural variation in floral asymmetry. Nature **401:** 157–161.

Currat, M., G. Trabuchet, D. Rees, P. Perrin, R. M. Harding *et al.*, 2002 Molecular analysis of the beta-globin gene cluster in the Niokholo Mandenka population reveals a recent origin of the beta(S) Senegal mutation. Am. J. Hum. Genet. **70:** 207–223.

Darwin, C. R., 1859 *The Origin of Species*. John Murray, London.

Darwin, C. R., 1868 *Variation of Animals and Plants Under Domestication*. John Murray, London.

Darwin, C. R., 1871 *The Descent of Man and Selection in Relation to Sex*. John Murray, London.

Darwin, C. R., 1876 *The Effects of Cross and Self Fertilization in the Vegetable Kingdom*. John Murray, London.

Darwin, C. R., 1877 *The Different Forms of Flowers on Plants of the Same Species*. John Murray, London.

Di Cesnola, A. P., 1907 A first study of natural selection in *Helix arbustorum* (Helicogena). Biometrika **5:** 387–399.

Dobzhansky, T., 1955 A review of some fundamental concepts and problems of population genetics. Cold Spring Harbor Symp. Quant. Biol. **20:** 1–15.

Dobzhansky, T., 1973 Nothing in biology makes sense except in the light of evolution. Am. Biol. Teach. **35:** 125–129.

Dykhuizen, D. E., 1990 Experimental studies of natural selection in bacteria. Annu. Rev. Ecol. Evol. Syst. **21:** 373–398.

East, E. M., 1910 A Mendelian interpretation of variation that is apparently continuous. Am. Nat. **44:** 65–82.

Encode Project Consortium, 2007 Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature **447:** 799–816.

Escobar, J., A. Nicot and P. David, 2008 The different sources of variation in inbreeding depression, heterosis and outbreeding depression in a metapopulation of *Physa acuta*. Genetics **180:** 1593–1608.

Farley, J., 1982 *Gametes and Spores*. Johns Hopkins Press, Baltimore.

Felsenstein, J., 2004 *Inferring Phylogenies*. Sinauer, Sunderland, MA.

Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Fisher, R. A., 1935 *The Design of Experiments*. Oliver & Boyd, Edinburgh.

Flibotte, S., M. Edgley, J. Maydan, J. Taylor, R. Zapf *et al.*, 2009 Rapid high resolution single nucleotide polymorphism-comparative genome hybridization mapping in *Caenorhabditis elegans*. Genetics **181:** 33–37.

Flint, J., and T. F. C. Mackay, 2009 Genetic architecture of quantitative traits in mice, flies, and humans. Genome Res. **19:** 723–733.

Gayon, J., 1998 *Darwinism's Struggle for Survival: Heredity and the Hypothesis of Natural Selection*. Cambridge University Press, Cambridge, UK.

Gillispie, C. C., 1960 *The Edge of Objectivity*. Princeton University Press, Princeton, NJ.

Haag-Liautard, C., M. Dorris, X. Maside, S. Macaskill, D. L. Halligan *et al.*, 2007 Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. Nature **445:** 82–85.

Harvey, P. H., and M. Pagel, 1991 *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.

Heijmans, B. T., E. W. Tobi, A. D. Stein, H. Putter, G. J. Blauw *et al.*, 2008 Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl. Acad. Sci. USA **6:** 17046–17049.

Howard, J., 2009 Why didn't Darwin discover Mendel's laws? J. Biol. **8:** 15.

Hudson, R. R., 1990 Gene genealogies and the coalescent process, pp. 1–44, in *Oxford Surveys in Evolutionary Biology*, Vol. 7, edited by Douglas Futuyma and Janis Antonovics. Oxford University Press, New York.

Ingvarsson, P. K., M. C. Whitlock and T. Hatfield, 2000 Local drift load and the heterosis of interconnected populations. Heredity **84:** 452–457.

Jenkin, F., 1867 Origin of Species. North British Rev. **46:** 277–318.

Johannsen, W., 1909 Elemente der exakten Erblichkeitslehre. Fischer, Jena, Germany.

Jones, D. F., 1917 Dominance of linked factors as a means of accounting for heterosis. Genetics **2:** 466–479.

Kimura, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri *et al.*, 2001 The strength of phenotypic selection in natural populations. Am. Nat. **157:** 245–261.

Labonne, J., A. Goultiaeva and J. Shore, 2009 High-resolution mapping of the S-locus in Turnera leads to the discovery of three genes tightly associated with the S-alleles. Mol. Genet. Genomics **281:** 673–685.

Lederberg, J., and E. M. Lederberg, 1952 Replica plating and indirect selection of bacterial mutants. J. Bacteriol. **63:** 399–406.

Leimu, R., and M. Fischer, 2008 A meta-analysis of local adaptation in plants. PLoS ONE **3:** e4010.

Lewontin, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

Li, J., M. Webster, M. Furuya and P. Gilmartin, 2007 Identification and characterization of pin and thrum alleles of two genes that cosegregate with the Primula *S* locus. Plant J. **51:** 18–31.

May, G., F. Shaw, H. Badrane and X. Vekemans, 1999 The signature of balancing selection: fungal mating compatibility gene evolution. Proc. Natl. Acad. Sci. USA **96:** 9172–9177.

Maynard Smith, J., and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. Genet. Res. **23:** 23–35.

Mendel, G., 1866 Versuche über Pflanzenhybriden. Ver. Naturforsch. Ver. Brünn **4:** 3–47.

Muller, H. J., 1932 Further studies on the nature and causes of gene mutations. Proceedings of the 6th International Congress on Genetics, Vol. 1, 213–255.

Myers, S., L. Bottolo, C. Freeman, G. McVean and P. Donnelly, 2005 A fine-scale map of recombination rates and hotspots across the human genome. Science **310:** 312–324.

Namekawa, S. H., J. L. VandeBerg, J. R. McCarrey and J. T. Lee, 2007 Sex chromosome silencing in the marsupial male germ line. Proc. Natl. Acad. Sci. USA **104:** 9730–9735.

Nash, D., S. Nair, M. Mayxay, P. N. Newton, J.-P. Guthmann *et al.*, 2005 Selection strength and hitchhiking around two antimalarial resistance genes. Proc. Biol. Sci. **272:** 1153–1161.

Nilsson-Ehle, H., 1909 Kreuzungsuntersuchungen an Hafer und Weizen. Lunds Univ. Aarskrift NS **5:** 1–12.

Nordborg, M., 1997 Structured coalescent processes on different time scales. Genetics **146:** 1501–1514.

Ornduff, R., 1969 Reproductive biology in relation to systematics. Taxon **18:** 121–133.

Presgraves, D. C., L. Balagopalan, S. M. Abmayr and H. A. Orr, 2003 Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature **423:** 715–719.

Provine, W. B., 1971 *The Origins of Theoretical Population Genetics*. University of Chicago Press, Chicago.

Raymond, C. K., A. Kas, M. Paddock, R. Qiu, Y. Zhou *et al.*, 2005 Ancient haplotypes of the HLA class II region. Genome Res. **15:** 1250–1257.

Richards, E. J., 2006 Inherited epigenetic variation: revisiting soft inheritance. Nat. Rev. Genet. **7:** 395–401.

Richman, A. D., M. K. Uyenoyama and J. R. Kohn, 1996 Allelic diversity and gene genealogy at the self-incompatibility locus in the Solanaceae. Science **273:** 1212–1216.

Ritland, K., 1990 A series of FORTRAN computer programs for estimating plant mating systems. J. Hered. **81:** 235–237.

Sabeti, P. C., D. E. Reich, J. M. Higgins, H. Z. Levine, D. J. Richter *et al.*, 2002 Detecting recent positive selection in the human genome from haplotype structure. Nature **419:** 832–837.

Sella, G., D. A. Petrov, M. Przeworski and P. Andolfatto, 2009 Pervasive natural selection in the Drosophila genome? PLOS Genet. **6:** e1000495.

Sidorenko, L., and V. Chandler, 2008 RNA-dependent RNA polymerase is required for enhancer-mediated transcriptional silencing

associated with paramutation at the maize *p1* gene. Genetics **180:** 1983–1993.

Slatkin, M., 2008 Linkage disequilibrium: understanding the evolutionary past and mapping the medical future. Nat. Rev. Genet. **9:** 477.

Van Os, H., P. Stam, R. G. Visser and H. J. Van Eck, 2005 RECORD: a novel method for ordering loci on a genetic linkage map. Theor. Appl. Genet. **112:** 30–40.

Vekemans, X., and M. Slatkin, 1994 Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics **137:** 1157–1165.

Walsh, B., 2008 Using molecular markers for detecting domestication, improvement, and adaptation genes. Euphytica **161:** 1–17.

Weldon, W. F. R., 1895 Attempt to measure the death-rate due to the selective destruction of *Carcinus moenas* with respect to a particular dimension. Proc. R. Soc. Ser B **58:** 360–379.

Weldon, W. F. R., 1901 A first study of natural selection in *Clausilia laminata* (Montagu). Biometrika **1:** 109–124.

Williamson, S., M. J. Hubisz, A. G. Clark, B. Payseur, C. Bustamante *et al.*, 2007 Localizing recent adaptive evolution in the human genome. PLoS Genet, **3:** 901–915.

Wright, S., 1920 The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. Proc. Natl. Acad. Sci. USA **6:** 320–333.

Wright, S., 1968 *Evolution and the Genetics of Populations*: *Genetic and Biometric Foundations*, Vol. 1. University of Chicago Press, Chicago.

Wright, S., and T. Dobzhansky, 1946 Genetics of natural populations XII. Experimental reproduction of some of the changes caused by natural selection in certain populations of *Drosophila pseudoobscura*. Genetics **31:** 125–156.