SHORT REVIEW

Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits

JR Stinchcombe[1] and HE Hoekstra[2]

[1]*Department of Ecology and Evolutionary Biology, Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario, Canada and [2]Department of Organismic and Evolutionary Biology and the Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA*

A central challenge in evolutionary biology is to identify genes underlying ecologically important traits and describe the fitness consequences of naturally occurring variation at these loci. To address this goal, several novel approaches have been developed, including 'population genomics,' where a large number of molecular markers are scored in individuals from different environments with the goal of identifying markers showing unusual patterns of variation, potentially due to selection at linked sites. Such approaches are appealing because of (1) the increasing ease of generating large numbers of genetic markers, (2) the ability to scan the genome without measuring phenotypes and (3) the simplicity of sampling individuals without knowledge of their breeding history. Although such approaches are inherently applicable to non-model systems, to date these studies have been limited in their ability to uncover functionally relevant genes. By contrast, quantitative genetics has a rich history, and more recently, quantitative trait locus (QTL) mapping has had some success in identifying genes underlying ecologically relevant variation even in novel systems. QTL mapping, however, requires (1) genetic markers that specifically differentiate parental forms, (2) a focus on a particular measurable phenotype and (3) controlled breeding and maintenance of large numbers of progeny. Here we present current advances and suggest future directions that take advantage of population genomics and quantitative genetic approaches – in both model and non-model systems. Specifically, we discuss advantages and limitations of each method and argue that a combination of the two provides a powerful approach to uncovering the molecular mechanisms responsible for adaptation.
Heredity (2008) **100,** 158–170; doi:10.1038/sj.hdy.6800937; published online 21 February 2007

Introduction

Understanding the genetic basis of ecologically important traits – traits that increase an organism's ability to survive and reproduce in natural environments – has been and continues to be a central goal for ecological and evolutionary genetics (Feder and Mitchell-Olds, 2003). Identifying the genes for ecologically relevant traits will allow a host of important genetic *and* ecological questions to be answered: how many genes influence ecologically important traits, and what are their relative effect sizes (Orr and Coyne, 1992; Orr, 1998)? Do these genes show evidence of non-neutral evolution at the sequence level (Stahl *et al.*, 1999; Tian *et al.*, 2002; Mauricio *et al.*, 2003)? What ecological and evolutionary forces lead to the maintenance of variation at these loci (Mitchell-Olds and Schmitt, 2006)? Do ecologically similar environments favor the same genes (Calboli *et al.*, 2003; Colosimo *et al.*, 2004, 2005; Protas *et al.*, 2006), or is it possible to achieve a similar phenotype

with different genetic mechanisms (Hoekstra and Nachman, 2003; Hoekstra *et al.*, 2006)? Answering these questions is not trivial, yet to begin to make progress on them, identifying the genes that influence ecologically important traits is a prerequisite. In addition, these questions must be answered in a number of organisms, including and extending beyond traditional model systems – representing diverse taxonomic groups, life histories and ecological roles – before a clear picture of the ecology and genetics of adaptation emerges. Here we review recent contributions of a relatively new approach, population genomics and an old-mainstay, quantitative genetics, to the challenge of finding genes that underlie ecologically important traits. We argue that combining these approaches provides a powerful and promising way to move from chromosomal regions to genes and even to mutations underlying adaptive phenotypic variation.

What is population genomics?

At its core, population genomics is simply population genetics writ large – that is, population genetic analyses of a large number of loci, distributed throughout the genome (Black *et al.*, 2001; Luikart *et al.*, 2003; Schlotterer, 2003). Population genomics can be narrowly defined as

separating locus-specific effects (recombination, selection, mutation and so on) that affect one or a few loci at a time from genome-wide demographic effects (genetic bottlenecks, founder events, inbreeding and so on). By utilizing a large number of loci spread throughout the genome, the effects of selection on a beneficial mutation and neutral variation at flanking sites (genetic hitchhiking; Maynard Smith and Haigh, 1974) can be compared to genome-wide demographic effects, which are not locus specific. As such, the population genomic approach can be described in four phases (Luikart et al., 2003): (1) sample many individuals, (2) genotype this large population sample for many independent loci, (3) identify statistical 'outlier' loci and (4) either estimate demographic parameters and statistics (e.g., F_{ST}, phylogeographic structure, evidence of past bottlenecks) in a large data set with outlier loci removed, or alternatively, study the outlier loci specifically in an attempt to infer potential selective mechanisms underlying them.

At its core, population genomics relies on two key factors. First, it requires genotyping of a large number of loci, whether through amplified fragment length polymorphism (AFLP)'s, microsatellites, single-nucleotide polymorphism (SNP)'s or sequences. The current explosion of molecular techniques and genomic tools available suggests that this is unlikely to be a rate-limiting step, even for non-model species. One key working assumption of population genomics approaches, particularly

important for studies using anonymous markers, is that the loci are independent. Even with markers of known locations in the genome, as the number of markers increases, a degree of auto-correlation will be introduced, potentially resulting in misleading inferences (see Hahn, 2006). Second, the population genomics approach requires a reliable means to detect outlier loci that may indicate regions that have been under selection – either to remove these loci to study genome-wide effects, or to identify such loci as the focus of study (Figure 1a). Because local adaptation and directional selection should have locus-specific effects of reducing genetic variability within populations and increasing differentiation between populations, loci that are outliers for these characteristics are strong candidate regions for involvement in adaptation. Determining whether an individual locus behaves as an outlier can be statistically evaluated with a battery of approaches, among them: testing whether F_{ST} is significantly different from either zero or neutral expectations (Lewontin and Krakauer, 1973; Beaumont and Nichols, 1996; Vitalis et al., 2001; Beaumont and Balding, 2004); the lnRV and lnRH statistics (natural log of the ratio of the variance and heterozygosity of alleles between two populations; (Schlotterer, 2002)); and the Ewens–Watterson test (Ewens, 1972; Watterson, 1978; Vigouroux et al., 2002); see Storz (2005) for a recent review of such tests. Importantly, the statistical significance of these estimates
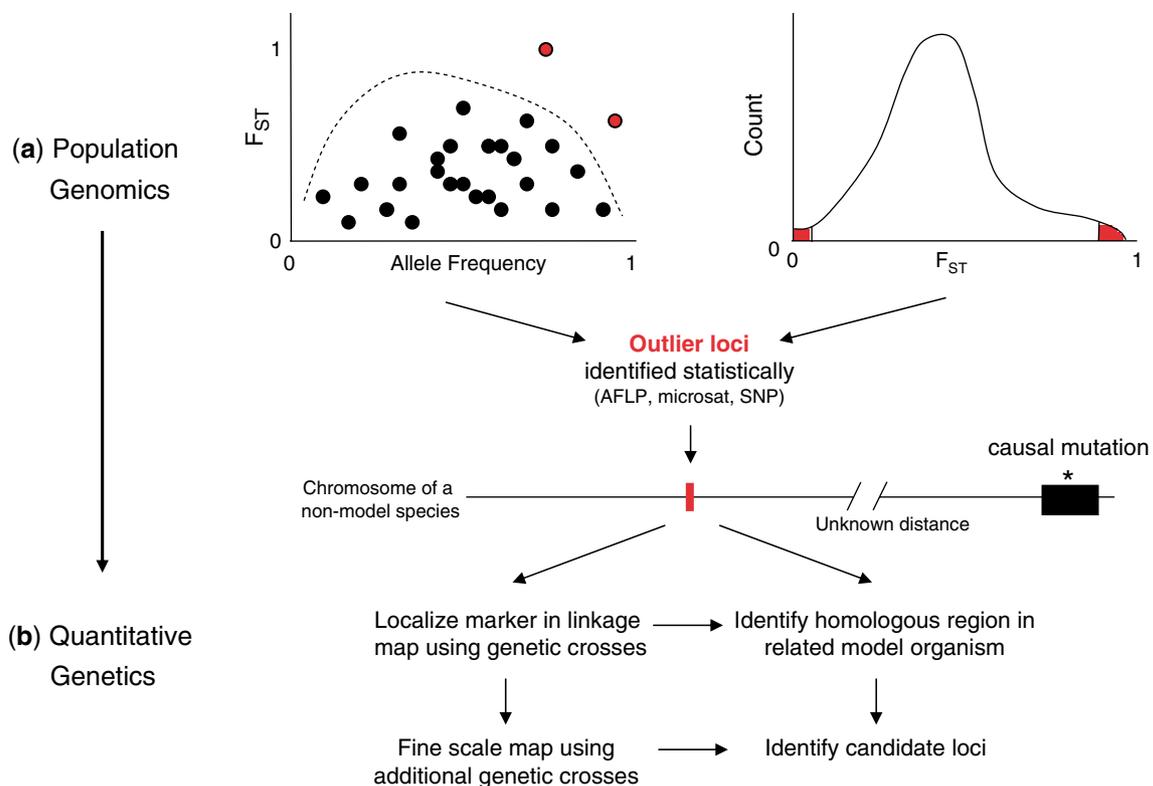

Figure 1 Conceptual model for the integration of population genomics and quantitative genetics. In (a), using population genomics approaches, outlier loci can be identified either by identifying loci with F_{ST} values that exceed confidence limits or intervals based on neutral coalescent simulations (dashed line, left panel) or are in the tails of the empirical, genome-wide distributions (filled portions of distribution, right panel). Once these outlier loci (shown in red), which are some unknown distance from the causal mutations, have been identified statistically, the next step of identifying the causal gene and mutation(s) can be pursued using genetic mapping techniques common to quantitative genetics (b). These mapping approaches can entail genetic crosses, or identification of homologous regions/candidate loci in related model organisms, or both.

can be determined from the genome-wide empirical distributions of the test statistics (Akey *et al.*, 2002), or by comparing observed statistics to a distribution generated by neutral coalescent simulations (Beaumont and Nichols, 1996; Beaumont and Balding, 2004), or neutral, non-equilibrium simulations using parameters estimated from the data (Thornton and Andolfatto, 2006).

Applications: estimating genome-wide effects
One consequence of gathering anonymous genome-wide polymorphism data is the potential it offers for investigators to separate locus- and region-specific selective effects from genome-wide effects such as demography. Several recent studies provide new methods to distinguish the effects of demography and selection in shaping genome-wide levels of polymorphism. These studies also caution that for candidate genes or loci linked with so-called outlier loci (*see below*), the challenge of distinguishing between purely demographic factors and the combined effects of demography and selection will be difficult (Przeworski *et al.*, 2005; Teshima *et al.*, 2006).

Because many demographic factors can affect patterns of nucleotide polymorphism in a way similar to the effects of selection, methods that can differentiate the effects of these forces are necessary before inferences can be made about their relative importance. Indeed, several recent studies have detected genome-wide departures from predictions of equilibrium neutral models in standard tests of selection (see Ford, 2002 for a review of such tests), presumably because of the effects of population genetic structure and demography (Andolfatto and Przeworski, 2000; Nordborg *et al.*, 2005; Schmid *et al.*, 2005, 2006). Although application in more traditional ecological settings is limited, three recent papers have used alternative approaches to distinguish between demographic and selective forces in shaping human polymorphism levels (Nielsen *et al.*, 2005; Stajich and Hahn, 2005; Williamson *et al.*, 2005), and some generalizations appear to be emerging. First, purely demographic factors can generate much of the observed variation in the amount and frequency of polymorphism in human populations. Based on this result, it seems likely that demography can have a large effect on genetic variability in many species that have similar ecological, demographic and genetic histories. Second, against this backdrop of demographic factors, it is still possible to detect loci that appear to have been under natural selection, either because patterns of variation at individual loci show a poor fit to a purely demographic model (Stajich and Hahn, 2005), or models incorporating selection provided a better fit to the data than demographic models parameterized with putatively neutral non-coding SNPs (Williamson *et al.*, 2005), or because individual regions of the genome show allele frequency distributions that differ from global, genome-wide allele frequency distributions (Nielsen *et al.*, 2005). Importantly, an ongoing challenge will be to distinguish whether patterns of variation at these loci truly show evidence of natural selection, or could as easily be explained by slightly more complicated (yet still realistic) demographic models.

In situations in which an ancestor-descendant relationship exists between different species or samples within a species (e.g., colonization of an island or novel habitat, domestication), it is possible to gain additional information by utilizing data from the ancestral population (Ometto *et al.*, 2005; Wright *et al.*, 2005; Yamasaki *et al.*, 2005). In the case of maize and its wild ancestor, teosinte, Wright *et al.* (2005) used a simulation approach to partition selective and demographic effects on polymorphism levels at 774 genes. By running coalescent simulations conditioned on the simulated data fitting multiple summaries of teosinte data, the authors were able to control for the shared history of demography, mutation, and recombination of the maize and teosinte lineages before domestication. Within this context, the severity of the bottleneck that accompanied domestication was estimated for each locus to arrive at a multilocus (genome-wide) estimate of the bottleneck severity. By comparing these models to other models that allowed a fraction of loci to show evidence of a more severe bottleneck that is indicative of artificial selection, Wright *et al.* (2005) estimated that approximately 2–4% of genes in the maize genome were targets of artificial selection. Importantly, these candidate loci were then aligned with published linkage and quantitative trait locus (QTL) maps, showing a significant clustering between candidate loci and QTL for morphological differences between teosinte and maize.

Applications: detecting outlier loci
Many applications of the population genomics approach have concentrated on attempts to detect outlier loci, either by screening a large number of anonymous loci or by comparing test statistics between candidate genes and a random sample of unlinked loci. There have been numerous applications of both approaches utilizing data from humans (Payseur *et al.*, 2002; Akey *et al.*, 2004; Hahn *et al.*, 2004; Rockman *et al.*, 2004, 2005; Storz *et al.*, 2004; Voight *et al.*, 2006), *Drosophila* (Harr *et al.*, 2002; Glinka *et al.*, 2003; Kauer *et al.*, 2003; Orengo and Aguade, 2004; Schofl and Schlotterer, 2004; Pool *et al.*, 2006), *Mus musculus* (Ihle *et al.*, 2006) and *Arabidopsis thaliana* (Cork and Purugganan, 2005). However, because in most of these cases neither the ecological context in which selection occurred nor the potential selective agent are known (but see Cork and Purugganan, 2005), here we focus on other recent applications.

Two clear cases in which the 'ecological' context and agent of selection are known are artificial selection/domestication and pesticide use. These cases provide a test for population genomics methods, at least in cases in which selection is strong and recent. To date, the population genomics approach has been used successfully to confirm loci that might have undergone a selective sweep in maize during domestication (Vigouroux *et al.*, 2002), genes for coat color and shortened limbs in dog breed formation (Pollinger *et al.*, 2005), chloroquine resistance in the malaria-inducing parasite *Plasmodium falciparum* (Wootton *et al.*, 2002) and warfarin resistance in rats (Kohn *et al.*, 2003). However, it is important to note that in all of these cases, strong artificial rather than natural selection is driving phenotypic divergence. In a 'proof of concept' paper, Anderson *et al.* (2005) compared F_{ST} for 10 non-synonymous mutations in four loci known to be involved in antimalarial drug resistance to F_{ST} for 10 synonymous

mutations in housekeeping genes or genes of unknown function. They found that not only was F_{ST} higher for non-synonymous mutations in drug resistance loci than for synonymous mutations at other loci, but that it was higher than neutral coalescent simulations that had been based on their putatively neutral loci, confirming that in this case loci subject to natural selection indeed exhibit higher F_{ST} relative to neutral loci.

In more traditional ecological settings, the population genomics approach has been applied in several cases in which species show clinal variation or ecotypic differentiation. Although not at a genomic scale, Storz and Dubach (2004) showed a clear example of detecting outlier loci: the albumin (*Alb*) locus in the deer mouse *Peromyscus maniculatus* showed significant altitudinal differentiation that exceeded neutral expectations based on 18 other allozyme markers, although the precise selective agent remains unclear. Studies that implicate an environmental gradient as the selective force producing differentiation are clearly strengthened by multiple tests (e.g., multiple altitudinal or latitudinal transects), and preferably using multiple statistical approaches (Campbell and Bernatchez, 2004; Storz *et al.*, 2004; Vasemagi and Primmer, 2005). However, identifying truly independent tests may prove to be a challenge because before population divergence, individual loci will share mutational environment and coalescent histories, potentially introducing some degree of correlation between populations.

To date, four studies, using anonymous genome-wide markers, have used multiple comparisons to test for consistent or repeatable outlier loci, using a variety of species (Table 1). For example, the common frog (*Rana temporaria*) exhibits altitudinal clines in a host of life history traits in Europe. Bonin *et al.* (2006) showed that approximately 2% of the AFLP loci they screened also exhibited elevated altitudinal differentiation; to guard against false positives, the authors only considered true outlier loci to be those that showed elevated differentiation in multiple tests. Regions in linkage with these AFLP loci would be strong candidates to contain genes under-

lying life history traits in this species that have been subject to altitudinally varying selection. Results from these studies (Table 1) suggest that <5–10% of loci screened show significantly elevated F_{ST} between differentiated ecotypes or populations, although the small number of examples available means generalizations are tentative.

Limitations of population genomics

Despite the appeal of these methods, especially for non-model organisms, they suffer from three glaring weaknesses from the standpoint of ecological and evolutionary functional genomics when applied in isolation. First, and perhaps most importantly, in cases where anonymous genetic markers are used to scan the genome, it is extremely likely that any anonymous locus showing 'outlier' behavior is not the causal locus itself, but is either physically linked or in linkage disequilibrium (LD) with the selected site(s). The extent of LD between the marker locus and the functionally relevant mutation can vary dramatically across the genome and also study systems, and will be affected by population history, mating system, recombination rate, the age of the selected allele, the strength of selection and many other factors (Nordborg and Tavare, 2002), making it difficult to localize the functionally relevant mutation. Similarly, the size and position of the genomic regions that show differentiation will be unknown, at least for species without detailed linkage maps (*see below*). On their own, most population genomic studies in natural populations have been limited to detecting a few statistical outlier loci, often in regions of unknown position in the genome. Therefore, the next and most important step of moving from anonymous marker to functional gene/mutation is unclear.

Second, population genomic studies are usually carried out in the absence of any information about phenotype. Thus, although genetic loci that show significant differentiation may be indicative of the effects of natural selection and local adaptation, in many cases it

Table 1 Examples of recent studies using the population genomics approach to identify 'outlier loci' involved in differentiation between habitats

Species	Marker type	Comparison	Percent of differentiated loci (# differentiated loci/ # loci screened)	Number of independent comparisons	References
Periwinkle snail (*Littorina saxatalis*)	AFLP	Between shell shape morphs that vary clinally with vertical elevation	4.9% (15/306)	Three bays	Wilding *et al.* (2001)
Atlantic salmon (*Salmo salar*)	Genomic and EST-derived microsatellites	Between habitats (fresh, brackish, and saltwater) and seas (Barents versus White Sea)	9.5% (9/95)	Two populations each for Barents versus White Sea; 2–4 populations for fresh, brackish, saltwater habitats	Vasemagi *et al.* (2005)
Whitefish (*Coregonus clupeaformis*)	AFLP	Between dwarf and normal ecotypes	1.4% (6/440)	Four lakes, each with sympatric pairs of ecotypes	Campbell and Bernatchez (2004)
Common frog (*Rana temporaria*)	AFLP	Low, intermediate, high altitude	2% (8/392)	Two altitudinal transects	Bonin *et al.* (2006)

Abbreviations: AFLP, amplified fragment length polymorphism; EST, expressed sequence tag.

is unclear which traits may differ between samples, and if any correspond to the differentiated loci. The absence of knowledge about the phenotype under selection limits both ecological investigation about the putative selective agents as well as any knowledge or future use of candidate genes (see below).

The third potential weakness of the approach is with the logical inference that loci showing patterns of high differentiation (or reduced variation) have been subject to selection, whereas loci that do not show these patterns have not. Existing evidence suggests that it is possible and even probable for some loci to show high levels of differentiation (or reduced within population variation) without having been targets of selection either owing to chance alone, or for instance, due to incorrect models of demographic history used in estimating parameters like F_{ST} (e.g., island versus stepping stone models; see Akey et al., 2004) or ascertainment bias (Thornton and Jensen, in press). Similarly, it is also possible for loci to be under selection without yielding statistically significant results in tests for selection (Gallavotti et al., 2004; McVean et al., 2005; Przeworski et al., 2005; Teshima et al., 2006). Simulation studies by Teshima et al. (2006) suggest that a sizable proportion of loci under selection will be missed in empirical genome-wide scans, especially if the loci selected had previously been neutral. In addition, requiring loci to show outlier behavior in independent population comparisons or transects, while helpful in guarding against false positives, implicitly assumes that the same loci will be fixed in response to similar environmental conditions (Bonin et al., 2006). Existing evidence demonstrates that this may not be the case even when both phenotypes and selective environments are very similar (Hoekstra and Nachman, 2003; Hoekstra et al., 2006), suggesting that this criterion will lead investigators to miss some loci involved with adaptation.

New contributions from quantitative genetics

Unlike population genomics, quantitative genetics is not a novel approach, but is instead rooted in a long history

(Galton, 1869, 1889). More recently, molecular tools have reinvigorated quantitative genetics through LD and QTL mapping. Like population genomics approaches, both LD and QTL mapping require the survey of a large number of genome-wide molecular markers (Figure 2). Specifically, LD mapping relies on surveys of genetic polymorphism data from a collection of samples (inbred lines, accessions, individuals and populations) to test for statistical associations between these genetic markers and particular phenotypes, again based on the premise that the marker(s) is in LD with the causal locus, or less likely, is in fact the causal mutation itself (Box 1; see Mackay, 2001; Clark, 2003; Mitchell-Olds and Schmitt, 2006). By contrast, in a QTL mapping approach, statistical analyses of genome-wide molecular markers and phenotypes measured in progeny of controlled crosses are used to identify chromosomal regions contributing to phenotypic differentiation (reviewed in Mackay, 2001; Erickson et al., 2004).

LD mapping and related methods (Box 1) offer the prospect of identifying genes for ecologically important traits. By utilizing naturally occurring variation sampled in wild populations that have accumulated hundreds to thousands of recombination events over time (compared to a few generations in laboratory crosses), LD mapping is expected to (1) necessitate more markers than traditional QTL studies to provide complete coverage of the genome, but (2) have substantially higher resolution for fine-scale mapping of genomic regions. This approach offers great potential, especially if candidate genes are available for association tests. However, one of the major hurdles facing LD mapping is the need to control for cryptic population structure or stratification, which can lead to false positives (see Pritchard et al., 2000a, b; Cardon and Palmer, 2003; Marchini et al., 2004; Yu et al., 2005). The LD mapping approach has successfully been applied in Drosophila and maize (Long et al., 1998; Thornsberry et al., 2001; Palsson and Gibson, 2004), and is starting to be applied in ecological settings. For example, Stinchcombe et al. (2004, 2005) showed that accessions of Arabidopsis thaliana with
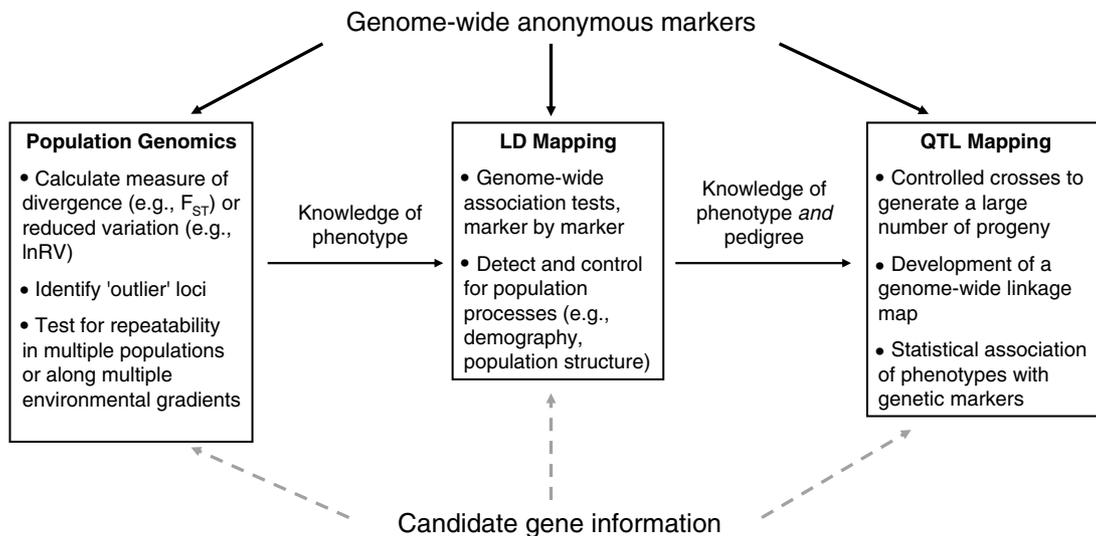

Figure 2 Schematic illustration of the relationships between population genomics, LD mapping and QTL mapping, emphasizing the different types of data required.

Box 1 Recent approaches for gene mapping in populations without a known cross or pedigree structure

LD mapping: A strategy to identify genes or genetic regions influencing a trait by comparing the phenotype of individuals with alternate alleles at a genetic marker which is presumed to be in LD with the causal loci. Phenotypes can either be the mean phenotype of a quantitative trait, or the frequency of occurrences for traits that are scored as presence/absence (e.g., cases or controls in medical studies). For many self-fertilizing plant species, inbred lines are used in lieu of individuals, provided there is little within-line genetic variation. For an example, see Palsson and Gibson (2004) and Hirschhorn and Daly (2005) for a review.
Candidate gene/association mapping: A variation on LD mapping, with the difference that associations are examined between phenotypes and alternate alleles at a candidate gene. For a review, see Long and Langley (1999) and for examples, see Thornsberry *et al.* (2001), Nachman *et al.* (2003) and Wilson *et al.* (2004).
Haplotype mapping: Another a variation on LD mapping, with the difference that haplotype blocks rather than individual genetic markers or candidate genes are utilized. For an example, see Olsen *et al.* (2004) and Aranzana *et al.* (2005).
Admixture-LD mapping: A strategy to identify genes or genetic regions influencing a trait in genetically admixed populations by testing for a non-random association between a phenotype and a genetic region that has ancestry predominantly from one of the parental populations. See Smith and O'Brien (2005) for a review, and Reich *et al.* (2005) for an example in human medical genetics.
Hitchhiking mapping: A mapping strategy to identify regions of the genome that have recently been under positive selection by detecting regions of reduced levels of genetic variation, due to the fact that fixation of beneficial mutation also reduces genetic variation at linked sites. In contrast to the approaches outlined above, hitchhiking mapping can be pursued without knowledge of the phenotype associated with the genetic region. For reviews, see Schlotterer (2003) and Storz (2005).

putatively functional *FRIGIDA* alleles exhibited significant latitudinal clines for flowering time and vernalization sensitivity, as would be predicted based on *FRIGIDA*'s role in the vernalization flowering time pathway (Simpson and Dean, 2002). In like fashion, Aranzana *et al.* (2005) showed that genome-wide association tests could successfully identify known flowering time and pathogen resistance genes in *Arabidopsis thaliana*, despite appreciable population structure. At present, most success stories in non-model organisms are limited to associations between a phenotype (inherited in a simple Mendelian manner) and one or a few candidate genes. For example, allelic variation at the melanocortin-1 receptor (*Mc1r*) was perfectly associated with coat color phenotype (melanic versus wild-type dorsal pelage) within populations (Nachman *et al.*, 2003) and with environmental variation (dark-colored lava versus light-colored granitic habitat) among populations (Hoekstra *et al.*, 2004); similar statistical associations were not observed at neutral mtDNA markers.

Unlike LD mapping, QTL approaches require the breeding of a large number of progeny, but thereby skirt the complications associated with genetic structure in natural populations. The genetic architecture of one phenotype, bristle number in *Drosophila*, has perhaps been the most intensively studied in a QTL context (reviewed by Mackay, 1995, 1996), and after tireless work, genes underlying bristle variation have been identified (Lai *et al.*, 1994; Long *et al.*, 1995). Although the precise molecular mechanisms remain elusive and the ecological relevance of bristle number is unclear, the progress in identifying the genes underlying bristle number suggest that moving from QTL to gene can be daunting even in model systems. Moreover, available data from both *Drosophila melanogaster* and *Arabidopsis thaliana* suggests that considerable heterogeneity exists in the causal mutations for ecologically important traits, either because of different loci affecting traits in natural populations than in mapping crosses (McDonald and Long, 2004), or because of genotype × environment interactions lead to different loci being identified in field versus laboratory settings (Weinig *et al.*, 2002).

For these reasons, most QTL studies have been limited to describing the genetic architecture of traits, with little progress in reaching the level of genes and mutations (Flint *et al.*, 2005), especially in non-model systems.

Nonetheless, a small but growing number of exceptions exist (e.g., Johanson *et al.*, 2000; El-Assal *et al.*, 2001; Shapiro *et al.*, 2004; Colosimo *et al.*, 2005; Balasubramanian *et al.*, 2006; Protas *et al.*, 2006), suggesting that QTL mapping is a feasible method of identifying the genes for ecologically important traits. And, although time intensive, costly and challenging, QTL approaches arguably represent the most comprehensive way to identify genomic regions and ultimately genes contributing to adaptive variation, especially for multigenic traits (Price, 2006).

There are three major ways in which genetic mapping approaches can interface with population genomics approaches in natural populations. First, data from genetic mapping studies (such as QTL studies) can be applied to population genomics studies. By scoring genetic markers in controlled crosses or pedigrees, genetic linkage maps can be generated, allowing for the possibility of linking outlier loci detected using population genomics approaches to 'real' chromosomal positions in the genome – representing a first step in localizing the genes of interest (Figure 1b). Second, by providing a large number of anonymous markers for study, the data gathered for population genomics approaches can do 'double duty' and be used to test and control for population genetic structure in subsequent studies using an LD mapping approach. Finally, population genomics approaches can be used to fine-scale map within the large chromosomal regions identified by lab-based QTL studies.

Applications: linkage map development and QTL mapping

A prerequisite for QTL mapping is the development of a linkage map, which allows investigators to associate phenotypes with specific identifiable regions of genome. Although the development of a linkage map and QTL mapping are clearly distinct issues, and the development of linkage maps is no longer necessary in many model systems with complete genome sequences, generating linkage maps can remain a challenge in many novel systems. Recently much effort has been spent generating linkage maps in non-model species, using a variety of experimental approaches and a diversity of molecular markers, with great potential for identifying genes underlying ecologically relevant variation (Table 2).

Table 2 A sampling of non-model species for which robust linkage maps have been developed

Common name	Scientific name	Experimental design	Primary markers	References
Plants				
Tomato	*Lycopersicon esculentum* and *L. pennellii*	Backcross and F2	RFLP	Bernatzky and Tanksley (1986)
Monkey flowers	*Mimulus guttatus* and *M. platycalx*	Backcross	RAPD	Lin and Ritland (1996)
Monkey flowers	*M. cardinalis* and *M. lewisii*	F2	RAPD	Bradshaw *et al*. (1998)
Monkey flowers	*M. guttatus* and *M. nasutus*	F2	AFLP	Fishman *et al*. (2001)
Shepherd's purse	*Capsella bursa-pastoris*	F2	RAPD	Linde *et al*. (2001)
Poplar	*Populus deltoides, P. nigra* and *P. trichocarpa*	Linkage analysis with two-way pseudo test cross design	AFLP	Cervera *et al*. (2001)
Western and Sierra columbines	*Aquilegia formosa* and *A. pubescens*	F2	AFLP	Hodges *et al*. (2002)
Sunflowers	*Helianthus annuus* and *H. petiolaris*	Backcross 2	Microsatellites	Rieseberg *et al*. (2003)
Lyrate rockcress	*Arabidopsis lyrata*	F2	CAPS markers	Kuittinen *et al*. (2004)
Pink shepherd's purse	*Capsella rubella* and *C. grandiflora*	F2	RFLPs, ESTs	Boivin *et al*. (2004)
Louisiana Iris	*Iris fulva* and *I. brevicaulis*	Backcross (recipcrocal)	IRRE retrotransposon display markers	Bouck *et al*. (2005)
Birch trees	*Betula pendula*	Full and half-sib families	AFLP	Pekkinen *et al*. (2005)
Bladder campion	*Silene vulgaris*	F2	AFLP	Bratteler *et al*. (2006)
Alpine pennycress	*Thlaspi caerulescens*	F3	AFLP	Assuncao *et al*. (2006)
Linanthus	*Leptosiphon bicolor* and *L. jepsonii*	F2, and reciprocal backcrosses	AFLP	Goodwillie *et al*. (2006)
Vertebrates				
Fire-bellied toads	*Bombina bombina* and *B. variegata*	F2	SSCP	Nürnberger *et al*. (2003)
Tiger salamander	*Ambystoma tigrinum* and *A. mexicanum*	Backcross	AFLP	Voss and Shaffer (1997)
Tiger salamander	*Ambystoma tigrinum* and *A. mexicanum*	Backcross	PCR-based ESTs	Smith *et al*. (2005)
Tilapia	*Oreochromis niloticus*	Haploid embryos	AFLPs, microsatellites	Kocher *et al*. (1998)
Tilapia	*Oreochromis niloticus* and *O. aureus*	F2	Microsatellites, SNP	Lee *et al*. (2005)
Three-spined stickleback	*Gasterosteus aculeatus*	F2	Microsatellites	Peichel *et al*. (2001)
Atlantic salmon	*Salmo salar*	Half and full-sib family	AFLP	Moen *et al*. (2004)
Brown trout	*Salmo trutta*	Backcross	Microsatellites	Gharbi *et al*. (2006)
Mexican tetra (cavefish)	*Astyanax mexicanus*	Backcross	Microsatellites	Protas *et al*. (2006)
Turkey	*Meleagris gallopavo*	F2	Microsatellites, SNP	Reed *et al*. (2005)
Great reed warbler	*Acrocephalus arundinaceus*	Pedigree	Microsatellites	Hansson *et al*. (2005)
Passerine birds	—	Homology mapping	Microsatellites	Dawson *et al*. (2006)
Tammar wallaby	*Macropus eugenii*	Backcross	Microsatellites	Zenger *et al*. (2002)
Opossum	*Monodelphis domestica*	Backcross	Microsatellites	Samollow *et al*. (2004)
Hamadryas baboon	*Papio hamadryas*	Pedigree	Microsatellites	Rogers *et al*. (2000)
Red deer	*Cervus elaphus*	Pedigree	Microsatellites	Slate *et al*. (2002)
Soay sheep	*Ovis aries*	Pedigree	Microsatellites	Beraldi *et al*. (2006)
Oldfield mice	*Peromyscus polionotus*	F2	Microsatellites, SNP	Steiner *et al*. (in review)
Insects				
Mosquito	*Anopheles gambiae*	Backcross	Microsatellites	Zheng *et al*. (1996)
Apple maggot fly	*Rhagoletis pomonella*	Single pair crosses	RFLP/SSCP	Roethele *et al*. (1997)
Ground crickets	*Allonemobius fasciatus* and *A. socius*	Backcross (bidirectional)	RAPD	Chu and Howard (1998)
Colorado potato beetle	*Leptinotarsa decemlineata*	Backcross	AFLP	Hawthorne (2001)
Pea aphids	*Acyrthosiphon pisum*	F2	AFLP	Hawthorne and Via (2001)
Hawaiian crickets	*Laupala paranigra* and *L. kohalensis*	F2	AFLP	Parsons and Shaw (2002)
Cactus fruit fly	*Drosophila mojavensis*	Backcross	Microsatellites	Staten *et al*. (2004)
Sulfur butterfly	*Colias eurytheme* and *C. philodice*	Backcross	AFLP	Wang and Porter (2004)
Corn borer	*Ostrinia nubilalis*	Backcross	AFLP	Dopman *et al*. (2004)
Red flour beetle	*Tribolium castaneum*	F2	AFLP	Zhong *et al*. (2004)
Heliconius butterfly	*Heliconius melpomene*	F2	AFLP	Jiggins *et al*. (2005)
Heliconius butterfly	*Heliconius erato*	Backcross	AFLP	Tobler *et al*. (2005)
Other invertebrates				
Pacific oysters	*Crassostrea gigas*	Backcross	Microsatellites	Hubert and Hedgecock (2004)
Nematode	*Pristionchus pacificus*	F2	SSCP	Srinivasan *et al*. (2002)

Abbreviations: AFLP, amplified fragment length polymorphism; CAPS, cleaved amplified polymorphic sequences; IRRE, iris retroelement; RAPD, random amplification of polymorphic DNA; RFLP, restriction fragment length polymorphism; SNP, single-nucleotide polymorphism; SSCP, single strand conformation polymorphism.
Where more than one species name is provided, the linkage map was generated from interspecific crosses or backcrosses. The experimental design and predominate molecular marker used in the linkage map are given. For many of these species, newer linkage maps have been created using additional markers, or genome-sequencing projects are currently underway.

Species that can be maintained in captivity, bred in the lab, and have relatively large brood sizes are often ideal for generating linkage maps using traditional crosses (e.g., butterflies (*Heliconius*, *Bicyclus*), sticklebacks (*Gasterous*), deermice (*Peromyscus*), monkeyflowers (*Mimulus*) and columbines (*Aqueligia*)). In other cases, linkage maps can be generated by following large pedigrees in natural populations (e.g., red deer (*Cervus elaphus*), soay sheep (*Ovis aries*), great reed warblers (*Acrocephalus arundinaceus*)); such long-term studies are time intensive and are only applicable to species that can be easily followed over time.

It is clear that many systems of ecological interest are not easily manipulated in the laboratory (i.e., genetic crosses are not feasible or generation times are prohibitively long). In many cases, ecological systems can take advantage of either closely related genetic model systems with genetic linkage maps or even complete genome sequences (e.g., Dawson *et al.*, 2006; Windsor *et al.*, 2006). For example, a recent study generated a predicted linkage map for passerine birds by taking advantage of the sequence similarity of available microsatellites and the draft chicken genome sequence (Dawson *et al.*, 2006), and then evaluated the accuracy of the predicted linkage map by comparing it to a previously published map for the great reed warbler (*Acrocephalus arundinaceus*). Despite the fact that chickens and warblers are diverged by millions of years, 24 microsatellite markers were conserved between the linkage maps, and synteny was maintained across genomes, highlighting the utility of the chicken genome for generating genomic resources for other avian species. Similar levels of conserved linkage have been reported between model organisms and non-model relatives, including *Drosophila* and the apple maggot fly (*Rhagoletis*; Roethele *et al.*, 2001), *Mus* and deer mice (*Peromyscus*; Steiner *et al.*, in review), and zebrafish and salamanders (Voss *et al.*, 2001). The availability of linkage maps for non-model species can be extremely useful for two primary reasons: (1) evenly spaced markers representing even coverage of the genome can be chosen for use in population genomic scans of the genome, or (2) alternatively, once regions of interest are identified, homologous regions in a closely related species (either a model with a complete genome sequence or one more amenable to controlled crosses and breeding) can be used to either design additional markers for fine-scale mapping or to search for candidate loci.

The benefits of combining the population genomics approach with traditional linkage maps can be seen in two studies that focused on closely related plant species (maize and teosinte: Vigouroux *et al.*, 2002; pedunculate and sessile oak: Scotti-Saintagne *et al.*, 2004). Both Vigouroux *et al.* (2002) and Scotti-Saintagne *et al.* (2004) detected loci that behaved as outliers in comparisons of population samples between closely related species. Because linkage maps have been made from experimental crosses, it is possible to determine (1) the genomic position in which these outliers occur, and (2) in some cases, test if loci showing elevated differentiation are also the loci closest to QTL for traits that are differentiated between the species. In the maize example, two of the outlier loci were located near known QTL for ear structure and endosperm weight, two traits that differ dramatically between maize and teosinte and could have

been past targets of artificial selection (Vigouroux *et al.*, 2002). In fact, because even the largest QTL mapping populations are limited by the number of recombination events, population genomic approaches may be useful in this context to fine-scale map genes.

Utilizing knowledge of candidate genes

One appeal of both population genomics and quantitative genetic approaches is that anonymous markers can easily be generated in non-model species and then scored in a large number of individuals without any *a priori* knowledge of the genetic or developmental mechanisms responsible for ecological differentiation. However, the use of candidate genes, although not necessary, can certainly aid in moving from the identification of a genomic region (a QTL) to a single gene or even a nucleotide mutation (a QTN). The vast majority of successes in identifying genes responsible for adaptive phenotypic variation arguably have involved either candidate loci in the initial genomic scan or the identification of candidate loci within a genomic region of interest. For example, population genomic approaches need not be restricted to completely anonymous markers (e.g., AFLPs or microsatellites), and instead can include markers in candidate loci themselves or a subset of loci chosen to include possible candidate genes (e.g., markers based on expressed sequence tags developed in an appropriate tissue type or from microarray experiments). Similarly, association studies in natural populations can include candidate loci; for example, Olsen *et al.* (2004) used this approach to assess how allelic variation at the photoperiod receptor gene *CRY2* contributes to variation in flower timing in 95 wild accessions of *Arabidopsis*.

Even in large genetic crosses, candidate genes have played a major role in the success stories of linking adaptive phenotypic variation to genes. For example, in three-spine sticklebacks (*Gasterosteus aculeatus*), a QTL approach identified a 10 Mb region containing a large effect region contributing to adaptive variation in pelvic morphology between oceanic and lake populations (Shapiro *et al.*, 2004). A candidate gene, the *Pitx1* gene, was identified in this region based on its knockout phenotype in laboratory mice, which affects pelvic morphology. When interrogated in sticklebacks, *Pitx1* expression differences were associated with a reduced pelvis in lake populations, although the precise molecular change is yet to be identified. Additional phenotypes, like pigmentation variation, have been well explored in vertebrate systems, in part because the wealth of genetic and developmental information on pigmentation provides an extensive list of well-characterized candidate loci (Hoekstra, 2006). First, mutations in the tyrosine-related protein 1 (*Tyrp-1*) gene have been mapped in a pedigreed population of Soay sheep (*Ovis aries*), and are associated with a naturally segregating light/dark coat color polymorphism (Gratten *et al.*, 2007). Second, genetic crosses and exploration of candidate genes in Mexican tetra (*Astyanax mexicanus*) led to the discovery that multiple independent deletions in the ocular and cutaneous albinism-2 (*Oca2*) gene were responsible for parallel loss of pigmentation in cave-dwelling tetra populations (Protas *et al.*, 2006). Finally, a QTL study of adaptive color pattern in beach mice (*Peromyscus polionotus*) identified several regions of major

effect (Steiner *et al.*, in review), one of which contained the candidate gene, *Mc1r*. A single amino-acid change in the *Mc1r* coding region is associated with between 10 and 36% of the variation in several adaptive pigment traits and the functional effects of this amino-acid change was verified in pharmacological assays (Hoekstra *et al.*, 2006).

Future directions: combining data from laboratory crosses and natural populations

Both population genomics and quantitative genetic approaches have limitations, especially in non-model systems, which often lack complete genome sequences. Although population genomic studies have been largely successful in generating large-scale genomic data for comparisons between populations, disentangling the effects of demography and sifting through false positives have been major challenges. Beyond the statistical challenges, the next step of moving from anonymous markers to known genetic regions and eventually to genes is perhaps even more daunting. Whereas QTL studies have successfully identified chromosomal regions contributing to phenotypic variation for those species which are amenable to genetic crossing experiments, narrowing these regions to genes, especially for traits with limited candidate loci requires enormous sample sizes and a plethora of genetic markers to detect rare recombination events (Flint *et al.*, 2005; Slate, 2005). Because of the limitations of each respective method, combining these approaches has the potential to be extremely powerful for identifying genes responsible for ecologically relevant variation.

Here we provide a powerful example of how combining multiple approaches can yield more insight than a single method applied in isolation. Rogers and Bernatchez (2005) combined population genomics scans of the genome for outlier loci with QTL mapping to examine the genetic basis of growth rate differences between dwarf (limnetic) and normal (benthic) ecotypes of whitefish (*Coregonus clupeaformis*). By constructing a linkage map and performing QTL mapping using AFLP loci that had previously been used in population genomics scans (Campbell and Bernatchez, 2004), they were able to determine whether the loci closest to growth rate QTL were the same as loci showing elevated differentiation in genome-wide scans of natural populations. They found that eight loci closest to QTL for growth rate showed F_{ST} values outside the empirically determined 95% confidence limits estimated from 440 AFLP loci, suggesting that differentiation at these loci was due to selection on nearby growth rate loci. Moreover, because benthic and limnetic fish were sampled from four lakes, the authors were able to show that one AFLP locus corresponding to a growth rate QTL exhibited significantly higher levels of genetic differentiation between ecotypes than expected by neutrality in three of the four lakes, suggesting genetic parallelism in how growth rate differences have evolved in lakefish. By combining QTL mapping, population genomics and surveys of multiple populations, this study illustrates the potential utility of combining approaches to (1) link markers identified in population genomics scans to phenotype and (2) test for parallel evolution using comparative genomic scans. However, it is important to note that additional work in both natural and lab-based populations will be needed to narrow these genomic regions to genes and mutations.

Conclusions

Population genomics provides an alluring first glimpse into the genome of previously unexplored organisms. In isolation, this approach can provide estimates of the proportion of the genome that are inconsistent with simple patterns of neutrality and hints of the possibility of parallel evolution, but it is thus far limited in its ability to point us directly to genes underlying adaptive phenotypic variation. The recent explosion of genome-wide linkage maps in novel systems highlights the ease by which large-scale genomic markers can be generated, and represents a clear way in which population genomic data can be linked to genome/chromosomal position, bringing us one step closer to the adaptive alleles themselves. It is clear from recent studies that combining data from natural populations (e.g., population genomics approaches or LD mapping) with information from lab-based experiments (e.g., linkage maps and QTL) provides a powerful approach for identifying the genes responsible for adaptive phenotypes (e.g., Colosimo *et al.*, 2005).

Importantly, the identification of genes underlying ecologically relevant traits does not represent a scientific end point, but rather the beginning of a new set of questions! Are adaptations to similar environments due to the same genes or mutations either within or between species? Do adaptive alleles emerge from standing genetic variation or as new mutations? How does the strength of selection affect the genetic architecture of adaptive traits? How do demographic and stochastic factors affect the ability of organisms to adapt to changing environments? Although the tools for non-model systems will by definition lag behind model systems, the ecological and evolutionary questions that can be answered in a diversity of novel systems will often be unique. These questions and others can be more directly addressed once ecologically relevant genes are in hand for a diversity of systems and will together provide important insight into both the ecology and evolution of adaptation.

Acknowledgements

References

Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA *et al.* (2004). Population history and natural selection shape patterns of genetic variation in 132 genes. *PLoS Biol* **2**: 1591–1599.

Akey JM, Zhang G, Zhang K, Jin L, Shriver MD (2002). Interrogating a high-density SNP map for signatures of natural selection. *Genome Res* **12**: 1805–1814.

Anderson TJC, Nair S, Sudimack D, Williams JT, Mayxay M, Newton PN *et al.* (2005). Geographical distribution of selected and putatively neutral SNPs in Southeast Asian malaria parasites. *Mol Biol Evol* **22**: 2362–2374.

Andolfatto P, Przeworski M (2000). A genome-wide departure from the standard neutral model in natural populations of *Drosophila*. *Genetics* **156**: 257–268.

Aranzana MJ, Kim JAS, Zhao K, Bakker E, Horton M, Jakob K *et al.* (2005). Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. *PLoS Genet* **1**: 531–539.

Assuncao AGL, Pieper B, Vromans J, Lindhout P, Aarts MGM, Schat H (2006). Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. *New Phytol* **170**: 21–32.

Balasubramanian S, Sureshkumar S, Agrawal M, Michael TP, Wessinger C, Maloof JN *et al.* (2006). The *PHYTOCHROME C* photoreceptor gene mediates natural variation in flowering and growth responses of *Arabidopsis thaliana*. *Nat Genet* **38**: 711–715.

Beaumont MA, Balding DJ (2004). Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol* **13**: 969–980.

Beaumont MA, Nichols RA (1996). Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc Lond B Biol Sci* **263**: 1619–1626.

Beraldi D, McRae AF, Gratten J, Slate J, Visscher PM, Pemberton JM (2006). Development of a linkage map and mapping of phenotypic polymorphisms in a free-living population of Soay Sheep (*Ovis aries*). *Genetics* **173**: 1521–1537.

Bernatzky R, Tanksley SD (1986). Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. *Genetics* **112**: 887–898.

Black WC, Baer CF, Antolin MF, DuTeau NM (2001). Population genomics: genome-wide sampling of insect populations. *Annu Rev Entomol* **46**: 441–469.

Boivin K, Acarkan A, Mbulu RS, Clarenz O, Schmidt R (2004). The *Arabidopsis* genome sequence as a tool for genome analysis in Brassicaceae. A comparison of the *Arabidopsis* and *Capsella rubella* genomes. *Plant Physiol* **135**: 735–744.

Bonin A, Taberlet P, Miaud C, Pompanon F (2006). Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Mol Biol Evol* **23**: 773–783.

Bouck A, Peeler R, Arnold ML, Wessler SR (2005). Genetic mapping of species boundaries in Louisiana Irises using IRRE retrotransposon display markers. *Genetics* **171**: 1289–1303.

Bradshaw HD, Otto KG, Frewen BE, McKay JK, Schemske DW (1998). Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics* **149**: 367–382.

Bratteler M, Lexer C, Widmer A (2006). A genetic linkage map of *Silene vulgaris* based on AFLP markers. *Genome* **49**: 320–327.

Calboli FCF, Kennington WJ, Partridge L (2003). QTL mapping reveals a striking coincidence in the positions of genomic regions associated with adaptive variation in body size in parallel clines of *Drosophila melanogaster* on different continents. *Evolution* **57**: 2653–2658.

Campbell D, Bernatchez L (2004). 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.

Cardon LR, Palmer LJ (2003). Population stratification and spurious allelic association. *Lancet* **361**: 598–604.

Cervera M-T, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V *et al.* (2001). Dense genetic linkage maps of three *Populus* species (*Populus deltoides, P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. *Genetics* **158**: 787–809.

Chu J, Howard DJ (1998). Genetic linkage maps of the ground crickets *Allonemobius fasciatus* and *Allonemobius socius* using RAPD and allozyme markers. *Genome* **41**: 841–847.

Clark AG (2003). Finding genes underlying risk of complex disease by linkage disequilibrium mapping. *Curr Opin Genet Dev* **13**: 296–302.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal Jr G, Dickson M, Grimwood J *et al.* (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**: 1928–1933.

Colosimo PF, Peichel CL, Nereng K, Blackman BK, Shapiro MD, Schluter D *et al.* (2004). The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol* **2**: 635–641.

Cork JM, Purugganan MD (2005). High-diversity genes in the *Arabidopsis* genome. *Genetics* **170**: 1897–1911.

Dawson DA, Burke T, Hansson B, Pandhal J, Hale MC, Hinten GN *et al.* (2006). A predicted microsatellite map of the passerine genome based on chicken-passerine sequence similarity. *Mol Ecol* **15**: 1299–1320.

Dopman EB, Bogdanowicz SM, Harrison RG (2004). Genetic mapping of sexual isolation between E and Z pheromone strains of the european corn borer (*Ostrinia nubilalis*). *Genetics* **167**: 301–309.

El-Assal SED, Alonso-Blanco C, Peeters AJM, Raz V, Koornneef M (2001). A QTL for flowering time in *Arabidopsis* reveals a novel allele of *CRY2*. *Nat Genet* **29**: 435–440.

Erickson DL, Fenster CB, Stenoien HK, Price D (2004). Quantitative trait locus analyses and the study of evolutionary process. *Mol Ecol* **13**: 2505–2522.

Ewens WJ (1972). The sampling theory of selectively neutral alleles. *Theoret Popul Biol* **3**: 87–112.

Feder ME, Mitchell-Olds T (2003). Evolutionary and ecological functional genomics. *Nat Rev Genet* **4**: 649–655.

Fishman L, Kelly AJ, Morgan E, Willis JH (2001). A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. *Genetics* **159**: 1701–1716.

Flint J, Valdar W, Shifman S, Mott R (2005). Strategies for mapping and cloning quantitative trait genes in rodents. *Nat Rev Genet* **6**: 271–286.

Ford MJ (2002). Applications of selective neutrality tests to molecular ecology. *Mol Ecol* **11**: 1245–1262.

Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter M, Doebley JF *et al.* (2004). The role of *barren stalk1* in the architecture of maize. *Nature* **432**: 630–635.

Galton F (1869). *Hereditary Genius*. Reprinted 1962, Meridian Books: NY.

Galton F (1889). *Natural Inheritance*. Macmillan: London.

Gharbi K, Gautier A, Danzmann RG, Gharbi S, Sakamoto T, Hoyheim B *et al.* (2006). A linkage map for brown trout (*Salmo trutta*): chromosome homologies and comparative genome organization with other salmonid fish. *Genetics* **172**: 2405–2419.

Glinka S, Ometto L, Mousset S, Stephan W, De Lorenzo D (2003). Demography and natural selection have shaped genetic variation in *Drosophila melanogaster*: a multi-locus approach. *Genetics* **165**: 1269–1278.

Goodwillie C, Ritland C, Ritland K (2006). The genetic basis of floral traits associated with mating system evolution in *Leptosiphon* (Polemoniaceae): an analysis of quantitative trait loci. *Evolution* **60**: 491–504.

Gratten J, Beraldi D, Lowder BV, McRae AF, Visscher PM, Pemberton JM *et al.* (2007). Compelling evidence that a single nucleotide polymorphism in *TYRP1* is responsible for coat colour polymorphism in a free-living population of Soay sheep. *Proc R Soc Lond B* **274**: 619–626.

Hahn MW (2006). Accurate inference and estimation in population genomics. *Mol Biol Evol* **23**: 911–918.

Hahn MW, Rockman MV, Soranzo N, Goldstein DB, Wray GA (2004). Population genetic and phylogenetic evidence for

positive selection on regulatory mutations at the factor VII locus in humans. *Genetics* **167**: 867–877.

Hansson B, Akesson M, Slate J, Pemberton JM (2005). Linkage mapping reveals sex-dimorphic map distances in a passerine bird. *Proc R Soc Lond B Biol Sci* **272**: 2289–2298.

Harr B, Kauer M, Schlotterer C (2002). Hitchhiking mapping: a population-based fine-mapping strategy for adaptive mutations in *Drosophila melanogaster. Proc Natl Acad Sci USA* **99**: 12949–12954.

Hawthorne DJ (2001). AFLP-based genetic linkage map of the Colorado potato beetle *Leptinotarsa decemlineata*: sex chromosomes and a pyrethroid-resistance candidate gene. *Genetics* **158**: 695–700.

Hawthorne DJ, Via S (2001). Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* **412**: 904–907.

Hirschhorn JN, Daly MJ (2005). Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* **6**: 95–108.

Hodges SA, Whittall JB, Fulton M, Yang JY (2002). Genetics of floral traits influencing reproductive isolation between *Aquilegia formosa* and *Aquilegia pubescens. Am Nat* **159**: S51–S60.

Hoekstra HE (2006). Genetics, development, and evolution of adaptive pigmentation in vertebrates. *Heredity* **97**: 222–234.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP (2006). A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**: 101–104.

Hoekstra HE, Drumm KE, Nachman MW (2004). Ecological genetics of adaptive color polymorphism in pocket mice: geographic variation in selected and neutral genes. *Evolution* **58**: 1329–1341.

Hoekstra HE, Nachman MW (2003). Different genes underlie adaptive melanism in different populations of rock pocket mice. *Mol Ecol* **12**: 1185–1194.

Hubert S, Hedgecock D (2004). Linkage maps of microsatellite DNA markers for the pacific oyster *Crassostrea gigas. Genetics* **168**: 351–362.

Ihle S, Ravaoarimanana I, Tautz D (2006). An analysis of signatures of selective sweeps in natural populations of the house mouse. *Mol Biol Evol* **23**: 790–794.

Jiggins CD, Mavarez J, Beltran M, McMillan WO, Johnston JS, Bermingham E (2005). A genetic linkage map of the mimetic butterfly *Heliconius melpomene. Genetics* **171**: 557–570.

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C (2000). Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* **290**: 344–347.

Kauer MO, Dieringer D, Schlotterer C (2003). A microsatellite variability screen for positive selection associated with the 'Out of Africa' habitat expansion of *Drosophila melanogaster. Genetics* **165**: 1137–1148.

Kocher TD, Lee W-J, Sobolewska H, Penman D, McAndrew B (1998). A genetic linkage map of a Cichlid Fish, the Tilapia (*Oreochromis niloticus*). *Genetics* **148**: 1225–1232.

Kohn MH, Pelz HJ, Wayne RK (2003). Locus-specific genetic differentiation at *Rw* among warfarin-resistant rat (*Rattus norvegicus*) populations. *Genetics* **164**: 1055–1070.

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M *et al.* (2004). Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *Ahaliana. Genetics* **168**: 1575–1584.

Lai CG, Lyman RF, Long AD, Langley CH, Mackay TFC (1994). Naturally-occurring variation in bristle number and DNA polymorphisms at the *scabrous* locus of *Drosophila melanogaster. Science* **266**: 1697–1702.

Lee B-Y, Lee W-J, Streelman JT, Carleton KL, Howe AE, Hulata G *et al.* (2005). A second-generation genetic linkage map of Tilapia (*Oreochromis spp*). *Genetics* **170**: 237–244.

Lewontin RC, Krakauer J (1973). Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms. *Genetics* **74**: 175–195.

Lin JZ, Ritland K (1996). Construction of a genetic linkage map in the wild plant *Mimulus* using RAPD and isozyme markers. *Genome* **39**: 63–70.

Linde M, Diel S, Neuffer B (2001). Flowering ecotypes of *Capsella bursa-pastoris* (L.). Medik. (Brassicaceae) analysed by a cosegregation of phenotypic characters (QTL) and molecular markers. *Ann Bot* **87**: 91–99.

Long AD, Langley CH (1999). The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Res* **9**: 720–731.

Long AD, Lyman RF, Langley CH, Mackay TFC (1998). Two sites in the Delta gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster. Genetics* **149**: 999–1017.

Long AD, Mullaney SL, Reid LA, Fry JD, Langley CH, Mackay TFC (1995). High-resolution mapping of genetic-factors affecting abdominal bristle number in *Drosophila melanogaster. Genetics* **139**: 1273–1291.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003). The power and promise of population genomics: from genotyping to genome typing. *Nat Rev Genet* **4**: 981–994.

Mackay TFC (1995). The genetic-basis of quantitative variation – numbers of sensory bristles of *Drosophila melanogaster* as a model system. *Trends Genet* **11**: 464–470.

Mackay TFC (1996). The nature of quantitative genetic variation revisited: lessons from *Drosophila* bristles. *Bioessays* **18**: 113–121.

Mackay TFC (2001). Quantitative trait loci in *Drosophila. Nat Rev Genet* **2**: 11–20.

Marchini J, Cardon LR, Phillips MS, Donnelly P (2004). The effects of human population structure on large genetic association studies. *Nat Genet* **36**: 512–517.

Mauricio R, Stahl EA, Korves T, Tian D, Kreitman M, Bergelson J (2003). Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana. Genetics* **163**: 735–746.

Maynard Smith J, Haigh J (1974). The hitchhiking effect of a favorable gene. *Genet Res* **23**: 23–35.

McDonald SJ, Long AD (2004). A potential regulatory polymorphism upstream of *hairy* is not associated with bristle number variation in wild-caught *Drosophila. Genetics* **167**: 2127–2131.

McVean G, Spencer CCA, Chaix R (2005). Perspectives on human genetic variation from the HapMap Project. *PLoS Genet* **1**: 413–418.

Mitchell-Olds T, Schmitt J (2006). Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis. Nature* **441**: 947–952.

Moen T, Hoyheim B, Munck H, Gomez-Raya L (2004). A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. *Anim Genet* **35**: 81–92.

Nachman MW, Hoekstra HE, D'Agostino SL (2003). The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci USA* **100**: 5268–5273.

Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C (2005). Genomic scans for selective sweeps using SNP data. *Genome Res* **15**: 1566–1575.

Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H *et al.* (2005). The pattern of polymorphism in *Arabidopsis thaliana. PLoS Biol* **3**: 1289–1299.

Nordborg M, Tavare S (2002). Linkage disequilibrium: what history has to tell us. *Trends Genet* **18**: 83–90.

Nürnberger B, Hofman S, Forg-Brey B, Praetzel G, Maclean A, Szymura JM *et al.* (2003). A linkage map for the hybridising toads *Bombina bombina* and *B. variegata* (Anura: Discoglossidae). *Heredity* **91**: 136–142.

Olsen KM, Halldorsdottir SS, Stinchcombe JR, Weinig C, Schmitt J, Purugganan MD (2004). Linkage disequilibrium mapping of *Arabidopsis CRY2* flowering time alleles. *Genetics* **167**: 1361–1369.

Ometto L, Glinka S, De Lorenzo D, Stephan W (2005). Inferring the effects of demography and selection on *Drosophila melanogaster* populations from a chromosome-wide scan of DNA variation. *Mol Biol Evol* **22**: 2119–2130.

Orengo DJ, Aguade M (2004). Detecting the footprint of positive selection in a European population of *Drosophila melanogaster*: multilocus pattern of variation and distance to coding regions. *Genetics* **167**: 1759–1766.

Orr HA (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* **52**: 935–949.

Orr HA, Coyne JA (1992). The genetics of adaptation: a reassessment. *Am Nat* **140**: 725–742.

Palsson A, Gibson G (2004). Association between nucleotide variation in *Egfr* and wing shape in *Drosophila melanogaster*. *Genetics* **167**: 1187–1198.

Parsons YM, Shaw KL (2002). Mapping unexplored genomes: a genetic linkage map of the Hawaiian cricket *Laupala*. *Genetics* **162**: 1275–1282.

Payseur BA, Cutter AD, Nachman MW (2002). Searching for evidence of positive selection in the human genome using patterns of microsatellite variability. *Mol Biol Evol* **19**: 1143–1153.

Peichel CL, Nereng KS, Ohgi KA, Cole BLE, Colosimo PF, Buerkle CA *et al.* (2001). The genetic architecture of divergence between threespine stickleback species. *Nature* **414**: 901–905.

Pekkinen M, Varvio S, Kulju KKM, Kärkkäinen H, Smolander S, Viherä-Aarnio A *et al.* (2005). Linkage map of birch, *Betula pendula* Roth, based on microsatellites and amplified fragment length polymorphisms. *Genome* **48**: 619–625.

Pollinger JP, Bustamante CD, Fledel-Alon A, Schmutz S, Gray MM, Wayne RK (2005). Selective sweep mapping of genes with large phenotypic effects. *Genome Res* **15**: 1809–1819.

Pool JE, DuMont VB, Mueller CJL, Aquadro F (2006). A scan of molecular variation leads to the narrow localization of a selective sweep affecting both afrotropical and cosmopolitan populations of *Drosophila melanogaster*. *Genetics* **172**: 1093–1105.

Price AH (2006). Believe it or not, QTLs are accurate! *Trends Plant Sci* **11**: 213–216.

Pritchard JK, Stephens M, Donnelly P (2000a). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000b). Association mapping in structured populations. *Am J Hum Genet* **67**: 170–181.

Protas ME, Hersey C, Kochanek D, Zhou Y, Wilkens H, Jeffery WR *et al.* (2006). Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat Genet* **38**: 107–111.

Przeworski M, Coop G, Wall JD (2005). The signature of positive selection on standing genetic variation. *Evolution* **59**: 2312–2323.

Reed KM, Chaves LD, Hall MK, Knutson TP, Harry DE (2005). A comparative genetic map of the turkey genome. *Cytogenet Genome Res* **111**: 118–127.

Reich D, Patterson N, Jager PLD, McDonald GJ, Waliszewska A, Tandon A *et al.* (2005). A whole-genome admixture scan finds a candidate locus for multiple sclerosis susceptibility. *Nat Genet* **37**: 1113–1118.

Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T *et al.* (2003). Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**: 1211–1216.

Rockman MV, Hahn MW, Soranzo N, Loisel DA, Goldstein DB, Wray GA (2004). Positive selection on MMP3 regulation has shaped heart disease risk. *Curr Biol* **14**: 1531–1539.

Rockman MV, Hahn MW, Soranzo N, Zimprich F, Goldstein DB, Wray GA (2005). Ancient and recent positive selection transformed opioid cis-regulation in humans. *PLoS Biol* **3**: 2208–2219.

Roethele JB, Feder JL, Berlocher SH, Kreitman ME, Lashkari DA (1997). Toward a molecular genetic linkage map for the apple maggot fly (Diptera:Tephritidae): comparison of alternative strategies. *Ann Entomol Soc Am* **90**: 470–479.

Roethele JB, Romero-Severson J, Feder JL (2001). Evidence for broad-scale conservation of linkage map relationships between *Rhagoletis pomonella* (Diptera:Tephritidae) and *Drosophila melanogaster* (Diptera:Drosophilidae). *Ann Entomol Soc Am* **94**: 936–947.

Rogers J, Mahaney MC, Witte SM, Nair S, Newman D, Wedel S *et al.* (2000). A genetic linkage map of the Baboon (*Papio hamadryas*) genome based on human microsatellite polymorphisms. *Genomics* **67**: 237–247.

Rogers SM, Bernatchez L (2005). 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.

Samollow PB, Kammerer CM, Mahaney SM, Schneider JL, Westenberger SJ, VandeBerg JL *et al.* (2004). First-generation linkage map of the gray, short-tailed opossum, *Monodelphis domestica*, reveals genome-wide reduction in female recombination rates. *Genetics* **166**: 307–329.

Schlotterer C (2002). A microsatellite-based multilocus screen for the identification of local selective sweeps. *Genetics* **160**: 753–763.

Schlotterer C (2003). Hitchhiking mapping – functional genomics from the population genetics perspective. *Trends Genet* **19**: 32–38.

Schmid K, Törjék O, Meyer R, Schmuths H, Hoffmann M, Altmann T (2006). Evidence for a large-scale population structure of *Arabidopsis thaliana* from genome-wide single nucleotide polymorphism markers. *Theor Appl Genet* **112**: 1104–1114.

Schmid KJ, Ramos-Onsins S, Ringys-Beckstein H, Weisshaar B, Mitchell-Olds T (2005). A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism. *Genetics* **169**: 1601–1615.

Schofl G, Schlotterer C (2004). Patterns of microsatellite variability among X chromosomes and autosomes indicate a high frequency of beneficial mutations in non-African *D. simulans*. *Mol Biol Evol* **21**: 1384–1390.

Scotti-Saintagne C, Mariette S, Porth I, Goicoechea PG, Barreneche T, Bodenes C *et al.* (2004). Genome scanning for interspecific differentiation between two closely related Oak species [*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.]. *Genetics* **168**: 1615–1626.

Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng BJ, Schluter D *et al.* (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**: 717–723.

Simpson GG, Dean C (2002). *Arabidopsis*, the rosetta stone of flowering time? *Science* **296**: 285–289.

Slate J (2005). Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Mol Ecol* **14**: 363–379.

Slate J, Visscher PM, MacGregor S, Stevens D, Tate ML, Pemberton JM (2002). A genome scan for quantitative trait loci in a wild population of Red Deer (*Cervus elaphus*). *Genetics* **162**: 1863–1873.

Smith JJ, Kump DK, Walker JA, Parichy DM, Voss SR (2005). A comprehensive expressed sequence tag linkage map for Tiger Salamander and Mexican Axolotl: enabling gene mapping and comparative genomics in *Ambystoma*. *Genetics* **171**: 1161–1171.

Smith MW, O'Brien SJ (2005). Mapping by admixture linkage disequilibrium: advances, limitations, and guidelines. *Nat Rev Genet* **6**: 623–632.

Srinivasan J, Sinz W, Lanz C, Brand A, Nandakumar R, Raddatz G *et al.* (2002). A bacterial artificial chromosome-based

genetic linkage map of the nematode *Pristionchus pacificus*. *Genetics* **162**: 129–134.

Stahl EA, Dwyer G, Mauricio R, Kreitman M, Bergelson J (1999). Dynamics of disease resistance polymorphism at the *RPM*1 locus of *Arabidopsis*. *Nature* **400**: 667–671.

Stajich JE, Hahn MW (2005). Disentangling the effects of demography and selection in human history. *Mol Biol Evol* **22**: 63–73.

Staten R, Schully SD, Noor MAF (2004). A microsatellite linkage map of *Drosophila mojavensis*. *BMC Genet* **5**: 12–12.

Steiner CC, Weber JN, Hoekstra HE. Two interacting pigmentation genes underlie adaptive variation in beach mice. (in review).

Stinchcombe JR, Caicedo AL, Hopkins R, Mays C, Boyd EW, Purugganan MD *et al.* (2005). Vernalization sensitivity in *Arabidopsis thaliana* (Brassicaceae): the effects of latitude and *FLC* variation. *Am J Bot* **92**: 1701–1707.

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS *et al.* (2004). A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proc Natl Acad Sci USA* **101**: 4712–4717.

Storz JF (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol Ecol* **14**: 671–688.

Storz JF, Dubach JM (2004). Natural selection drives altitudinal divergence at the albumin locus in deer mice, *Peromyscus maniculatus*. *Evolution* **58**: 1342–1352.

Storz JF, Payseur BA, Nachman MW (2004). Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. *Mol Biol Evol* **21**: 1800–1811.

Teshima KM, Coop G, Przeworski M (2006). How reliable are empirical genomic scans for selective sweeps? *Genome Res* **16**: 702–712.

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES (2001). Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet* **28**: 286–289.

Thornton K, Andolfatto P (2006). Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. *Genetics* **172**: 1607–1619.

Tian DC, Araki H, Stahl E, Bergelson J, Kreitman M (2002). Signature of balancing selection in *Arabidopsis*. *Proc Natl Acad Sci USA* **99**: 11525–11530.

Tobler A, Kapan D, Flanagan NS, Gonzalez C, Peterson E, Jiggins CD *et al.* (2005). First-generation linkage map of the warningly colored butterfly *Heliconius erato*. *Heredity* **94**: 408–417.

Vasemagi A, Nilsson J, Primmer CR (2005). Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Mol Biol Evol* **22**: 1067–1076.

Vasemagi A, Primmer CR (2005). Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Mol Ecol* **14**: 3623–3642.

Vigouroux Y, McMullen M, Hittinger CT, Houchins K, Schulz L, Kresovich S *et al.* (2002). Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proc Natl Acad Sci USA* **99**: 9650–9655.

Vitalis R, Dawson K, Boursot P (2001). Interpretation of variation across marker loci as evidence of selection. *Genetics* **158**: 1811–1823.

Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006). A map of recent positive selection in the human genome. *PLoS Biol* **4**: 446–458.

Voss SR, Shaffer HB (1997). Adaptive evolution via a major gene effect: paedomorphosis in the Mexican axolotl. *Proc Natl Acad Sci USA* **94**: 14185–14189.

Voss SR, Smith JJ, Gardiner DM, Parichy DM (2001). Conserved vertebrate chromosome segments in the large salamander genome. *Genetics* **158**: 735–746.

Wang BQ, Porter AH (2004). An AFLP-based interspecific linkage map of sympatric, hybridizing *Colias* butterflies. *Genetics* **168**: 215–225.

Watterson GA (1978). The homozygosity test of neutrality. *Genetics* **88**: 405–417.

Weinig C, Ungerer MC, Dorn LA, Halldorsdottir SS, Toyonaga Y, Mackay TFC *et al.* (2002). Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. *Genetics* **162**: 1875–1884.

Wilding CS, Butlin RK, Grahame J (2001). Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *J Evol Biol* **14**: 611–619.

Williamson SH, Hernandez R, Fledel-Alon A, Zhu L, Nielsen R, Bustamante CD (2005). Simultaneous inference of selection and population growth from patterns of variation in the human genome. *Proc Natl Acad Sci USA* **102**: 7882–7887.

Wilson LM, Whitt SR, Ibanez AM, Rocheford TR, Goodman MM, Buckler IV ES (2004). Dissection of maize kernel composition and starch production by candidate gene association. *Plant Cell* **16**: 2719–2733.

Windsor AJ, Schranz ME, Formanova N, Gebauer-Jung S, Bishop JG, Schnabelrauch D *et al.* (2006). Partial shotgun sequencing of the *Boechera stricta* genome reveals extensive microsynteny and promoter conservation with *Arabidopsis*. *Plant Physiol* **140**: 1169–1182.

Wootton JC, Feng XR, Ferdig MT, Cooper RA, Mu JB, Baruch DI *et al.* (2002). Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. *Nature* **418**: 320–323.

Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD *et al.* (2005). The effects of artificial selection on the maize genome. *Science* **308**: 1310–1314.

Yamasaki M, Tenaillon MI, Bi IV, Schroeder SG, Sanchez-Villeda H, Doebley JF *et al.* (2005). A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell* **17**: 2859–2872.

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF *et al.* (2005). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* **38**: 203–208.

Zenger KR, McKenzie LM, Cooper DW (2002). The first comprehensive genetic linkage map of a marsupial: the Tammar Wallaby (*Macropus eugenii*). *Genetics* **162**: 321–330.

Zheng LB, Benedict MO, Cornel AJ, Collins FH, Kafatos FC (1996). An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics* **143**: 941–952.

Zhong D, Pai A, Yan G (2004). AFLP-based genetic linkage map for the red flour beetle (*Tribolium castaneum*). *J Hered* **95**: 53–61.