REVIEW

Towards the identification of the loci of adaptive evolution

Carolina Pardo-Diaz[1]*, Camilo Salazar[1] and Chris D. Jiggins[2]

[1]*Biology Program, Faculty of Natural Sciences and Mathematics, Universidad del Rosario, Carrera 24 No 63C-69, Bogotá 111221, Colombia; and* [2]*Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK*

Summary

1. Establishing the genetic and molecular basis underlying adaptive traits is one of the major goals of evolutionary geneticists in order to understand the connection between genotype and phenotype and elucidate the mechanisms of evolutionary change. Despite considerable effort to address this question, there remain relatively few systems in which the genes shaping adaptations have been identified.

2. Here, we review the experimental tools that have been applied to document the molecular basis underlying evolution in several natural systems, in order to highlight their benefits, limitations and suitability. In most cases, a combination of DNA, RNA and functional methodologies with field experiments will be needed to uncover the genes and mechanisms shaping adaptation in nature.

Key-words: adaptation, candidate genes, DNA mapping, expression, functional analysis, QTL, resequencing

Understanding how diversity arises and is maintained in nature has been a major question in evolutionary biology since ancient times (Lewontin 1974; Mayr 1982). Until recently, addressing this question was methodologically limited, but the recent and increasing development of molecular and genomic tools has now equipped us better to face the challenge. Consequently, much of modern research in evolution is devoted to identifying the genes shaping adaptive phenotypes (Hoekstra & Coyne 2007; Stern & Orgogozo 2009; Nadeau & Jiggins 2010; Jones *et al.* 2012; Kunte *et al.* 2014). This has and will continue to contribute to answering important questions about how evolution proceeds: Do adaptations arise gradually through many small mutations, via large leaps of major effect or both? Is the evolution of similar traits in different lineages the product of mutations in the same genes? Are particular kinds of adaptive mutations more likely than others, such as gene regulatory or protein-coding mutations? Here, we will focus on the methods used to address these questions.

The narrowing of the genomic regions controlling adaptations has employed mainly three different approaches termed *forward genetics*, *reverse genetics* and *candidate gene* (Stinchcombe & Hoekstra 2007; Nadeau & Jiggins 2010). *Forward genetics* approaches seek to identify genes underlying a known adaptive trait (Stinchcombe & Hoekstra 2007; Nadeau & Jiggins 2010; Stapley *et al.* 2010), while *reverse genetics* approaches refer to the detection of selection signatures across the genome without necessarily have a prior knowledge of the associated phenotype (Stinchcombe & Hoekstra 2007; Bonin 2008; Ellegren & Sheldon 2008; Stapley *et al.* 2010). Both *forward* and *reverse* genetics approaches are currently benefiting from the recent advances in sequencing technology, although its application still faces several challenges related to data storage, data analysis and cost that, although decreasing, can be challenging especially for organisms with large genomes (Wang, Gerstein & Snyder 2009).

Alternatively, the *candidate gene* approach relies on existing knowledge about the genes participating in the formation of the adaptive phenotype under investigation in other organisms; a correlation between trait variation and allelic polymorphism suggests the use of the candidate gene in shaping the adaptive trait (Luikart *et al.* 2003; Haag *et al.* 2005; Hoekstra *et al.* 2006; Mundy 2009). Any of the above approaches, alone or often combined, can narrow a strong set of candidate genes underlying adaptations. A further validation of these candidates can benefit from the exploration of the gene expression patterns and the application of functional assays such as knockouts, knockdowns and/or transgenics. However, a relation between a candidate gene and a presumptive adaptive phenotype does not constitute the unequivocal detection of the 'loci of evolution'. The definite connection between gene function and adaptation also requires the implementation of selection experiments that test the adaptive consequences of the individual alleles of these candidates (Fig. 1). Unfortunately, this has been rarely achieved, perhaps due to the difficulty of replicating evolutionary processes under controlled conditions (Colosimo *et al.* 2005).

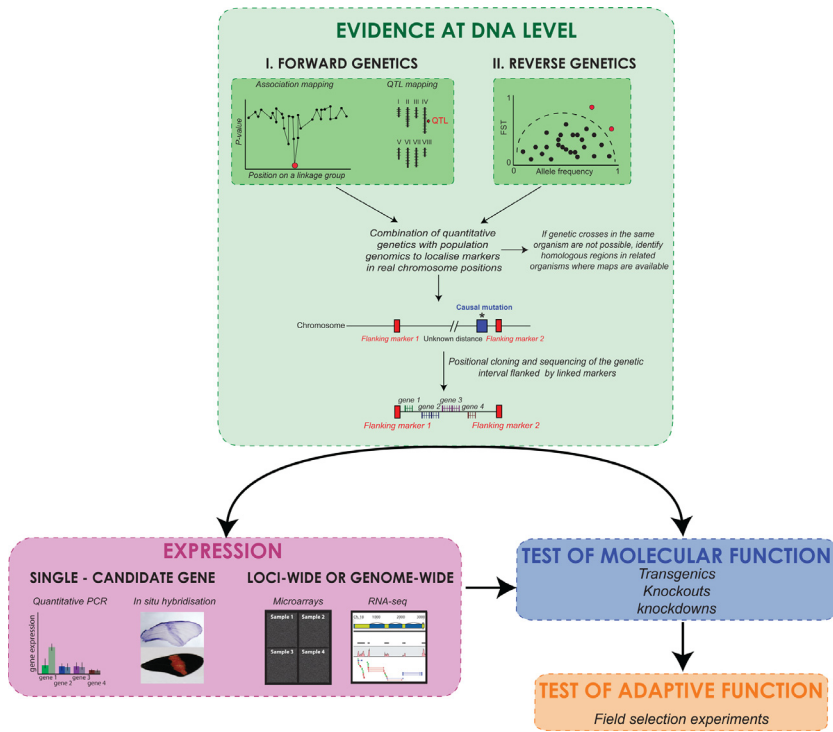**Correspondence author. E-mail: geimy.pardo@urosario.edu.co*

Fig. 1. Methodological processes useful to identify the loci underlying adaptation. Ideally, phenotype–genotype association studies, followed by the profiling of gene expression, functional tests and selection tests should be combined to identify a gene(s) as involved in shaping an adaptive trait. Evidence at the DNA level was adapted and modified from (Stinchcombe & Hoekstra 2007; Barrett & Hoekstra 2011). *In situ* hybridization shows expression of the gene *optix* in wings of *Heliconius melpomene* (Photo: Bob Reed) (Reed *et al.* 2011). Photographs of microarray and RNA-seq by Carolina Pardo-Diaz.

Narrowing down adaptations at the DNA level

FORWARD GENETICS APPROACHES

Forward genetics methods work towards finding the genetics controlling a known adaptive trait (i.e. one that is known to increase organismal fitness and/or reproduction in a particular context and where coefficients of selection have been measured and/or the response to changes in the selective agent has been quantified; see section IV: *assays of allele effects on fitness with field studies*). *Forward genetics* methods consist mainly of two different approaches: (i) association mapping using sexual populations that recombine in nature and (ii) quantitative trait loci (QTL) mapping using pedigrees, both of which survey a large number of molecular markers across the genome of individuals segregating for the adaptive trait of interest in order to identify gene regions as well as genes that are responsible for its variation (Stinchcombe & Hoekstra 2007; Nadeau & Jiggins 2010; Slate *et al.* 2010; Stapley *et al.* 2010; Martin & Jiggins 2013; Savolainen, Lascoux & Merila 2013; Wray 2013; Dittmar *et al.* 2014; Zuellig, Kenney & Sweigart 2014). Crucially in both cases, the methods involve identifying associations between the phenotypes of individuals with genetic variants. Thus, *forward genetics* is only useful where there is variation in a population for a particular trait or where such a variable population can be generated in the laboratory via crosses. This approach, however, is not useful to map differences in species that are completely isolated reproductively.

The association mapping approach, termed a 'genome-wide association study' (GWAS) when applied across the whole genome, takes advantage of the historical recombination in wild populations in order to detect non-random associations between genomic markers scattered across the genome and the adaptive trait of interest (Shimizu & Purugganan 2005; Stinchcombe & Hoekstra 2007; Hunter, Wright & Bomblies 2013). Using historical recombination increases the resolution in the detection of the locus (or loci) controlling the adaptive trait under study as the segregation would be much larger than that of a progeny of a experimental biparental population (Hwang *et al.* 2014). Thus, GWAS works by identifying non-random association of alleles between a locus with the adaptive trait (i.e. linkage disequilibrium, LD), as consequence of the action of natural selection (Long & Langley 1999; Shimizu & Purugganan 2005; Stranger, Stahl & Raj 2011). Until recently, GWAS applied genetic markers such as AFLPs, microsatellites and single copy gene markers to sample genetic variation (Table 1) (Holliday, Ritland & Aitken 2010; Holliday, Wang & Aitken 2012), but more recently, high-throughput sequencing approaches are used [such as reduced representation sequencing (Altshuler *et al.* 2000; Hohenlohe *et al.* 2010), low-coverage genotyping (Andolfatto *et al.* 2011; Elshire *et al.* 2011) and genome resequencing (Nielsen *et al.* 2011)]. Coupled with advances in bioinformatics, the use of GWAS to investigate the genetics of adaptation is increasingly popular (Hunter, Wright & Bomblies 2013). In lodgepole pines, for example, the use of restriction-site associated DNA sequencing (RAD-seq) generated markers for a GWAS of variation in serotiny (adaptation of cones to remain closed and retain seeds until a forest fire opens them) and revealed that at least 11 loci are involved in the natural variation of this trait (Parchman *et al.* 2012). Similarly, in *Arabidopsis thaliana,* a GWAS identified a *cis*-regulatory polymorphism at the *AtHKT1;1* locus as the factor controlling adaptive variation in leaf Na^+ accumulation capacity (Baxter *et al.* 2010a). Additionally, several GWASs

Table 1. Examples of natural adaptations investigated at the molecular level, the methodological approaches followed and the genes identified (if they have)

Organism	Adaptive trait	DNA evidence	Expression Profiling	Functional test	Genes identified/suggested	References
Arabidopsis thaliana	Flowering time	Forward genetics approaches–association mapping and QTL mapping	Microarray	Complementation test	*FRI* and *FLC*	Aranzana *et al.* (2005), Lempe *et al.* (2005), Bergelson & Roux (2010), Salomé *et al.* (2011)
Arabidopsis thaliana	Pathogen resistance	Forward genetics approaches–association mapping and QTL mapping	RNA blot analysis	Mutagenesis	*Rpm1*	Grant *et al.* (1995), Tornero *et al.* (2002), de Torres *et al.* (2003), Aranzana *et al.* (2005)
Gasterosteus aculeatus (threespine stickleback)	Armour plate patterning	Forward genetics approaches–association mapping and QTL mapping	qRT-PCR	Transgenics	*EDA*	Colosimo *et al.* (2005), Jones *et al.* (2012)
Gasterosteus aculeatus (threespine stickleback)	Loss of pelvic spines	Forward genetics approaches–QTL mapping	qRT-PCR and *in situ* hybridization	Transgenics	*Pitx1*	Shapiro *et al.* (2004), Shapiro, Bell & Kingsley (2006), Chan *et al.* (2010)
Peromyscus sp.	Coat colour	Forward genetics approaches–QTL mapping and reverse genetics approaches	qRT-PCR, *in situ* hybridization and antibodies	Transgenics	*Agouti*	Manceau *et al.* (2011), Linnen *et al.* (2013)
Tetrapoda	Land colonization	Reverse genetics approaches–comparative genomics	RNA-seq, *in situ* hybridization	Transgenics	Several genes	Amemiya *et al.* (2013)
Drosophila sp.	Male wing pigmentation involved in courtship	Reverse genetics approaches–genotyping of candidate genes	Immunochemistry	Transgenics	*Yellow*	Gompel *et al.* (2005), Prud'homme *et al.* (2006)
	Spatial regulation of pigmentation genes involved in wing spot formation	NA	*In situ* hybridization, Western blot, immunochemistry, microarray, qRT-PCR	RNAi screen, mutants, transgenics	*Dll*	Arnoult *et al.* (2013)
Thlaspi caerulescens	Zinc accumulation	Forward genetics approaches–association mapping and QTL mapping	Microarray and Northern blot	Yeast complementation analysis	*MT2* and *MT3*	Assunção *et al.* (2006), Hassinen *et al.* (2007)
Picea sitchensis (sitka spruce conifer)	Wood physical attributes	Forward genetics approaches–association mapping	Microarray	NA	β-expansin, Tubulin 3B, Galactosyl-transferase	Holliday, Ritland & Aitken (2010), Beaulieu *et al.* (2011)
Gasterosteus aculeatus (threespine stickleback)	Gill pigmentation	Forward genetics approaches–QTL mapping	*In situ* hybridization and allele-specific expression	NA	*Kitlg*	Miller *et al.* (2007)
Heliconius melpomene and *Heliconis erato*	Wing colour pattern	Forward genetics approaches–QTL mapping	Microarray, *in situ* hybridization, qRT-PCR	NA	*Optix, WntA*	Jiggins *et al.* (2005), Baxter *et al.* (2008, 2010b), Counterman *et al.* (2010), Reed *et al.* (2011), Hines *et al.* (2012), Martin *et al.* (2012)
Papilio polytes	Wing mimicry	Forward genetics approaches–association mapping	Immunochemistry, qRT-PCR, RNA-seq	NA	*doublesex*	Kunte *et al.* (2014)
Rattus norvegicus	Warfarin resistance	Reverse genetics approaches–microsatellite	Semi-quantitative RT-PCR	NA	Vitamin K epoxide reductase (*VKOR*)	Kohn, Pelz & Wayne (2003), Lasseur *et al.* (2006)

Table 1. (continued)

Organism	Adaptive trait	DNA evidence	Expression Profiling	Functional test	Genes identified/suggested	References
Cichlids	Visual pigment diversification	Reverse genetics approaches – genotyping of candidate genes	qRT-PCR on candidate genes	NA	*Opsin* genes	Hofmann *et al.* (2009)
Ovis aries (Soay sheep)	Coat colour	Forward genetics approaches – association mapping	qRT-PCR	NA	*Tyrp1*	Beraldi *et al.* (2006), Gratten *et al.* (2007)
Vertebrates	Pigmentation	Reverse genetics approaches – genotyping of candidate genes	qRT-PCR	NA	*Mc1r*	Hoekstra & Nachman (2003), Nachman, Hoekstra & D' Agostino (2003), Mundy *et al.* (2004), Rosenblum, Hoekstra & Nachman (2004), Hoekstra *et al.* (2006), Hubbard *et al.* (2010)
Chrysomela aeneicollis (montane beetle)	Adaptation to local thermal conditions	Reverse genetics approaches – genome-wide SNP typing	Western blot	NA	*Pgi– Hsp70*	Rank (1992), Dahlhoff & Rank (2000)
Chaenocephalus aceratus and *Pleuragramma antarcticum* (Antarctic notothenioid fishes)	Craniofacial skeletal morphology	Reverse genetics approaches – genotyping of candidate genes	*In situ* hybridization	NA	*col1a1, col2a1b* and *col10a1*	Albertson *et al.* (2010)
Pseudopodoces humilis (ground tit)	High altitude adaptations	Reverse genetics approaches – whole-genome sequencing	NA	NA	Adrenaline response and hormone biosynthesis genes	Cai *et al.* (2013)
Chrysemys picta bellii (western painted turtle)	Extreme anoxia and tissue freezing	Reverse genetics approaches – whole-genome sequencing	RNA-seq	NA	Tumour suppression genes, glucose transport genes and the miR-29b micro RNA	Bradley Shaffer *et al.* (2013)
Tapeworms	Parasitism adaptations	Reverse genetics approaches – whole-genome sequencing	RNA-seq	NA	Apomucin gene family, antigen B gene family, Hsp70 gene family	Tsai *et al.* (2013)
Falco peregrinus (peregrine falcon) and *Falco cherrug* (saker falcon)	Predation adaptations	Reverse genetics approaches – whole-genome sequencing	RNA-seq	NA	Olfactory receptor genes, beak development genes	Zhan *et al.* (2013)
Pteropus alecto (fruit bat) and *Myotis davidii* (insectivorous bat)	Flight and immune adaptations	Reverse genetics approaches – whole-genome sequencing	NA	NA	Repair of genetic damage genes, skin elasticity genes, muscle contraction genes, innate immunity genes	Zhang *et al.* (2013)
Ambystoma mexicanum (Mexican axolotl)	Paedomorphosis	Forward genetics approaches – QTL mapping	Microarray	NA	Thyroid hormone-response genes	Voss & Shaffer (1997), Page *et al.* (2008, 2009, 2010), Huggins *et al.* (2012)
Cichlid fish	Colour pattern	Forward genetics approaches – association mapping and QTL mapping	qRT-PCR	NA	*Pax7*	Streelman, Albertson & Kocher (2003), Roberts, Ser & Kocher (2009)
Bicyclus anynana	Eyespots	Genotyping of candidate genes	Immunohistochemistry	NA	*Distal-less*	Beldade, Brakefield & Long (2002)
Rana chensinensis and *R. kukunoris* (ranid frogs)	Adaptation to high elevation	NA	Transcriptomic analysis	NA	125 protein-coding genes	Yang *et al.* (2012)

Table 1. (continued)

Organism	Adaptive trait	DNA evidence	Expression Profiling	Functional test	Genes identified/suggested	References
Ipomoea sp.	Floral colour	Reverse genetics approaches – genotyping of candidate genes	*In situ* hybridization and Northern blot	Complementation test and enzyme assay	Genes in flavonoid biosynthesis	Durbin *et al.* (2003), Zufall & Rausher (2003, 2004)
Coregonus spp. (lake whitefish)	Growth, swimming activity, gill rakers and condition factor	Reverse genetics approaches – genome-wide SNP typing; Forward genetics approaches – QTL mapping	Microarray	NA	Several genes	Rogers & Bernatchez (2005), Derome *et al.* (2008), Renaut *et al.* (2011)
Littorina saxatilis	Local adaptation	Reverse genetics approaches – AFLPs	ESTs genomic scan	NA	Two loci	Wood *et al.* (2008), Galindo, Grahame & Butlin (2010), Westram *et al.* (2014)
Melitaea cinxia (Glanville fritillary butterfly)	Dispersal rate and flight metabolism	Reverse genetics approaches – genotyping of candidate genes	NA	NA	*Pgi*	Haag *et al.* (2005)
Anser indicus and *Chloephaga melanoptera* (geese)	High altitude adaptation	Reverse genetics approaches – candidate genes	NA	Protein assay with same mutation in heterologous human protein	Haemoglobins	Jessen *et al.* (1991)
Chicken and Japanese Quail	Plumage colour	Reverse genetics approaches – candidate genes; Forward genetics approaches – association mapping	Quantification of mRNA decay in mutant variant	NA	SLC45A2	Gunnarsson *et al.* (2007)
Lake Victoria cichlids	Light spectrum sensitivity – visual system	Reverse genetics approaches – candidate genes	NA	Protein assay	LWS	Terai *et al.* (2006)
Lakes Tanganyika and Malawi cichlids	Visual adaptation to deep-water habitats	Reverse genetics approaches – genotyping of candidate genes	NA	Protein assay	RH1	Sugawara *et al.* (2005)
Astyanax fasciatus (Mexican cave tetra)	Albinism	Forward genetics approaches – QTL mapping	NA	Cell-based functional assay	*Oca2*	Protas *et al.* (2006)
Ostrinia nubilalis (European corn borer)	Sexual isolation pheromones	Forward genetics approaches – association mapping	NA	NA	*Pher* and *Resp*	Dopman, Robbins & Seaman (2010)
Cervus elaphus (red deer)	Birth weight	Forward genetics approaches – association mapping and QTL mapping	NA	NA	One major QTL	Slate *et al.* (2002)
Peromyscus mice	Behavioural differences – burrow architecture	Forward genetics approaches – association mapping and QTL mapping	NA	NA	Four QTL	Weber, Peterson & Hoekstra (2013)
Rana temporaria (common frog)	Adaptation to altitude	Reverse genetics approaches – AFLPs	NA	NA	Eight outlier loci	Bonin *et al.* (2006)
Pinus contorta (lodgepole pine)	Cone serotiny	Reverse genetics approaches – genome-wide SNP typing	NA	NA	Eleven candidate loci	Parchman *et al.* (2012)
Triticum sp. (wheat)	Drought adaptation	Forward genetics approaches – association mapping	NA	NA	Several loci	Maccaferri *et al.* (2011)
	Body shape variation	Forward genetics approaches – association mapping	NA	NA	Several loci	Collin & Fumagalli (2011)

Table 1. (continued)

Organism	Adaptive trait	DNA evidence	Expression Profiling	Functional test	Genes identified/suggested	References
Phoxinus phoxinus (cyprinid fish European minnow)		Reverse genetics approaches–AFLPs				Bradshaw et al. (1998)
Mimulus sp.	Floral morphology	Forward genetics approaches–QTL mapping	NA	NA	Several loci	
	Floral pigmentation	Reverse genetics approaches–candidate genes	qRT-PCR	Virus-induced gene silencing (VIGS)	R2R3-MYB	Streisfeld & Rausher (2009), Streisfeld, Young & Sobel (2013)
Leptosiphon sp.	Floral characters associated with the mating system	Forward genetics approaches–association mapping and QTL mapping	NA	NA	Several loci	Goodwillie, Ritland & Ritland (2006)
Fundulus heteroclitus	Adaptation to local water temperature	NA	In vivo reporter gene expression of candidate gene	Protein activity	Lactate dehydrogenase-B gene (Ldh-B)	Schulte et al. (2000)
Drosophila yakuba and Drosophila santomea	Body colouration	NA	Expression constructs with mutant versions (in situ hybridization)	Transgenics	Tan and Yellow	Jeong et al. (2008)

have identified loci with signals of human adaptations including genes involved in immunity, cancer, infection, reproduction, healing and height (Hancock et al. 2011b; Jarvis et al. 2012; Lachance et al. 2012; Scheinfeldt & Tishkoff 2013). The power of association-based methods to detect a true association between a SNP and an adaptive trait largely depends on the phenotypic variance of the population explained by the SNP. Such phenotypic variance is determined by how strongly the alternative allelic variants differ in their phenotypic effects (effect size) and their frequency in the sample (Korte & Farlow 2013). Association-based methods are therefore biased towards detecting large-effect loci, although this is probably true of all methods described here (Rockman 2011; Martin & Jiggins 2013). This bias can be reduced by using extremely large sample sizes to maximize the genetic variance within the sample (Bodmer & Bonilla 2008; Korte & Farlow 2013). However, in some cases GWAS in humans using large sample sizes have shown that a single phenotype may be controlled by many minor effect loci that may explain only a small proportion of the trait heritability, which has limited the identification of causal variants (Rockman 2011). This limitation of GWAS is highly relevant to the study of the genetics of adaptive traits with polygenic inheritance, which may be the majority of adaptive traits (Rockman 2011; Turchin et al. 2012).

It is worth-noting that the suitability of a GWAS approach in discovering adaptation genes depends on additional factors that are highly dependent on the study system. For instance, the extent of genomic LD can introduce bias in a GWAS, so it is recommended to use high-density genotyping platforms that generate enough SNPs to examine the background level of LD in each population in order to differentiate LD outliers from genomic background (Alhaddad et al. 2013; Porto-Neto, Kijas & Reverter 2014). Also, obtaining the minimum number of markers necessary for a successful GWAS relies on factors such as genome size, number of individuals and number of groups or populations included, which all influence the genotyping effort required to get enough coverage and overlap between samples (Davey et al. 2011). Technical replicates with sufficient coverage may also be needed in order to assess the reproducibility of the genotyping technique and the quality of SNP calling, both of which can affect the proper calling of heterozygous sites and the reliability of GWAS (Hong et al. 2012; Nielsen et al. 2012).

QTL mapping is an alternative approach to understand the genetic basis of a known adaptive trait. It relies on the generation of mapping crosses to create a genetically variable and recombinant population in the laboratory and uses statistical analyses to correlate the quantitative variance of adaptive traits with molecular markers distributed across the genome and identify chromosomal regions contributing to phenotype differentiation (Ellegren & Sheldon 2008; Mackay, Stone & Ayroles 2009; Nadeau & Jiggins 2010; Stapley et al. 2010; Slate 2013) (Table 1). In sticklebacks, for example, pelvic structures and pigmentation play adaptive roles as defensive structures against predation and crypsis, respectively, and the genetic regions controlling natural adaptive variation at these two characters were both found by QTL mapping

with microsatellites (Shapiro *et al.* 2004; Miller *et al.* 2007). In a similar way, QTL mapping using AFLPs revealed that only three genomic intervals modulate most of the observed adaptive wing colour variation in the butterfly *Heliconius erato* (Papa *et al.* 2013). A growing number of studies have applied next-generation sequencing techniques to QTL mapping. In the perennial ryegrass *Lolium perenne,* the identification of three QTL that explained nearly 40% of the resistance to stem rust was achieved using RAD-seq (Pfender *et al.* 2011). Similarly, the identification of a QTL involved in spinosad resistance in the diamondback moth *Plutella xylostella* applied a RAD-seq approach (Baxter *et al.* 2011).

Although QTL mapping has been historically useful in narrowing broad regions associated with traits of interest, their power relies on obtaining large families. Unfortunately, obtaining enough recombinant offspring is not always possible due to the nature of many organisms, and thus, the power and applicability of QTL may be limited. Also, this method faces some limitations related to estimating effect size, number of loci [or Quantitative Trait Nucleotides (QTNs)] and their interactions contributing to adaptation. First, the null hypothesis for QTL mapping is the absence of a QTL and not the presence of infinitesimal QTL, so it is not so applicable for the study of adaptive traits shaped by many minimal effect loci (Rockman 2011). In line with this argument, empirical data show that QTL mapping generally finds a skewed L-distribution of effect sizes (with few large-effect loci accounting for most of the variation). It is also widely known that if QTL mapping is performed with small sample sizes, the magnitude of the identified QTL is likely to be overestimated ('Beavis effect'), the power of detecting medium to small effect QTL is limited, and the signature of several linked QTL may be blurred into a single large QTL (Mackay, Stone & Ayroles 2009; Nadeau & Jiggins 2010; Rockman 2011; Slate 2013). A recent review of QTL mapping studies showed that there is clear evidence for an upward bias in the magnitude of QTL estimates, particularly when sample sizes are small, but possibly even when they are as large as 1000 individuals (Slate 2013). Nonetheless, QTL mapping is often an important initial step towards gene discovery, which can lead to further genomic scans (such as GWAS) to establish whether QTL can be replicated (Slate 2013). This is exemplified by a study of horn morphology in Soay sheep, where a QTL was first finely mapped (Johnston *et al.* 2010) and then confirmed using an association study. This identified the gene *RXFP2* as the gene control variation in horn phenotype (Johnston *et al.* 2011).

REVERSE GENETICS APPROACHES

Reverse genetics (or genome scan) approaches refer to the genome-wide sampling of loci in order to detect regions with footprints of selection and thus detect selective (adaptive) loci even without a prior knowledge of their associated phenotypic trait(s) (Luikart *et al.* 2003; Storz 2005; Martin & Jiggins 2013). This approach benefits from the fact that no experimental crosses are needed but instead, variation in natural populations can be used (Schlötterer 2003; Stinchcombe & Hoekstra 2007) and also, it should be less biased towards large-effect loci (as it identifies actual targets of selection) (Martin & Jiggins 2013).

The *reverse genetics* strategy involves a comprehensive sampling of independent loci across the entire genome and the application of statistical analyses to identify loci with genetic variation indicative of selection that are possibly involved in adaptation (Balding & Nichols 1995; Nicholson *et al.* 2002; Luikart *et al.* 2003; Vitalis *et al.* 2003; Joost *et al.* 2007; Foll & Gaggiotti 2008; Coop *et al.* 2010; Nunes *et al.* 2012; De Mita *et al.* 2013; Frichot *et al.* 2013; Martin & Jiggins 2013; de Villemereuil *et al.* 2014). These statistics can be applied both at intra- and inter-population level.

Within a single population, measures of genetic diversity and distribution of genetic polymorphism (or allele frequency spectrum, AFS) with statistics such as π and Tajima's D are expected to be influenced by selection fixing an advantageous allele. This fixation also leads to a reduction of genetic diversity in the surrounding sequences creating a selective sweep, due to genetic hitchhiking. The location of such selective sweeps can be detected by evaluating patterns of linkage disequilibrium (LD) (Kim & Nielsen 2004). The detection of hard sweeps (i.e. new mutations that arise and quickly go to fixation) is facilitated by the pattern of strong reduced nucleotide polymorphism at the selected locus and its neighbour regions (Pritchard, Pickrell & Coop 2010; Olson-Manning, Wagner & Mitchell-Olds 2012). In contrast, detecting the signal of soft sweeps (when the selected allele has been already segregating in the population before being swept) is more difficult because the selected haplotype may be impossible to differentiate from the genetic background (Pritchard, Pickrell & Coop 2010; Olson-Manning, Wagner & Mitchell-Olds 2012). Furthermore, the signal of a selective sweep may be lost fairly rapidly over time especially in large populations. According to simulation studies, this limitation can be overcome with the combination of LD scans with AFS scans (Pavlidis, Jensen & Stephan 2010). Nonetheless, false positives remain a problem. For example, Pavlidis *et al.* (2012) used simulated data under neutrality in *Drosophila* and still detected false positive sweeps with misleading biological functions (Pavlidis *et al.* 2012). Due to these constraints, genomic scans of selective sweeps are less commonly used as a first step in identifying adaptive loci, yet they are useful to confirm sweep signals at loci involved in adaptations (Hohenlohe *et al.* 2010; Nadeau *et al.* 2013).

Demographic factors can also affect nucleotide polymorphism and produce patterns of summary statistics (like π and Tajima's D) easily confused with those expected by the action of selection (Stinchcombe & Hoekstra 2007). However, it is expected that demography affects the genome as a whole. The problem can therefore be ameliorated to some degree by using genome-wide data to evaluate the demographic history of the species and provide a neutral expectation against which particular regions can be tested for the influence of selection (McVicker *et al.* 2009; Pool *et al.* 2010; Li & Durbin 2011). Selective sweeps also influence haplotype structure, and inferences based

on more complex measures, not based on SNPs alone, are likely to be more powerful (Gusev *et al.* 2009, 2011; Pavlidis *et al.* 2013).

Inter-population measures can also be used to detect footprints of selection across the genome. The most common approach is to apply tests based on F_{ST} to reveal loci where the fixation of alternative alleles in each population results in greater differences in allele frequency than expected under neutral evolution (Beaumont & Nichols 1996; Bonhomme *et al.* 2010; Narum & Hess 2011; Martin & Jiggins 2013; de Villemereuil *et al.* 2014; Lotterhos & Whitlock 2014). Alternatively, there are methods to detect selection based on correlations of genetic data with environmental changes (Joost *et al.* 2007; Coop *et al.* 2010; De Mita *et al.* 2013). Simulation studies under a monogenic scenario have shown that F_{ST}-based methods are able to consider complex population history and structure (when known) and are less prone to detect false positives; however, they fail to detect true selection outliers when selection is not strong and their results are strongly dependent on the demographic model implemented (De Mita *et al.* 2013; Lotterhos & Whitlock 2014). Additionally, it has been recently pointed out that the use of relative measures in detection of outliers (e.g. F_{ST}) may be misleading when searching regions in the genome involved in adaptation and speciation (Noor & Bennett 2010; Martin *et al.* 2013; Cruickshank & Hahn 2014). In some genomic regions, for example, genetic divergence between species of recent origin may lead to a decrease in genetic variability and recombination rate as consequence of the speciation process, which can be misinterpreted as signatures of positive or negative selection acting within species. Similarly, regions of the genome with restricted gene flow compared to the genomic background (measured with F_{ST}) are usually interpreted as islands of divergence (Turner, Hahn & Nuzhdin 2005; Ellegren *et al.* 2012). However, this signal also may be the result of positive or background selection that acted, in the past, on the ancestral population (Noor & Bennett 2010; Cruickshank & Hahn 2014). Thus, the misinterpretation of F_{ST} patterns can confuse processes of adaptation with those of speciation. It has been suggested that absolute measures of genetic divergence (that do not rely on allelic frequencies), such as D_{XY} distance, may be better, but these also fail to detect some regions known to be under selection (Cruickshank & Hahn 2014). A related factor to consider is the relationship between recombination and selection. Divergence and low diversity is commonly higher in regions with low recombination (e.g. centromeres), which may result from natural selection or just be a consequence of limited recombination (Turner & Hahn 2010). There may also be a positive relationship between recombination and mutation, which would increase the variation available for selection to act upon (Cutter & Payseur 2013). However, to date, studies are still needed that investigate how recombination rate varies across the genome and in candidate regions for adaptation and speciation.

Environment–genetic correlation methods, on the other hand, are more powerful and work well under different strengths of selection. On the other hand, they may have a high false discovery rate (FDR) if genetic correlations between populations are not accounted for (De Mita *et al.* 2013). In some cases, correct estimation of the environmental variables can be challenging, potentially limiting application of this latter method. In polygenic scenarios, where multiple loci of small effect underlie a single trait, both environment–genetic correlation methods and F_{ST}-based methods show lower power of detection compared to major locus genetic architecture (de Villemereuil *et al.* 2014). Demographic effects (such as a high level of hierarchical population structure) and its correlation with an environmental variable underlying the selective pressure also affect the power of detection in these polygenic cases (de Villemereuil *et al.* 2014).

Despite the limitation of environment–genetic correlation methods and F_{ST}-based methods, both have helped in identifying outlier loci associated with adaptations in natural populations. Examples include the identification of whole-genome outlier SNPs associated with climatic adaptations in *Arabidopsis thaliana* (Hancock *et al.* 2011a), transcriptome-derived SNPs correlated with stress tolerance in coral reefs (Lundgren *et al.* 2013), outlier AFLP loci correlated with insecticide resistance in mosquitoes (Paris & Despres 2012), SNPs associated with adaptation to coastal environments in *Senecio lautus* (Roda *et al.* 2013) and SNPs associated with climatic adaptations in humans (Hancock *et al.* 2011b). A combination of population genetics and environmental correlations can help to reduce the number of false positives (de Villemereuil *et al.* 2014). This has been rarely done, although one exception is the identification of loci involved in host plant use in the large pine weevil *Hylobius abietis* (Manel, Conord & Despres 2009). Yet, 'new more general and robust likelihood test are needed that are flexible enough to accommodate departures from classical demographic models' (de Villemereuil *et al.* 2014).

Regardless of the method used, it is fundamental to validate that outliers are genuinely implicated in adaptation (Luikart *et al.* 2003; Barrett & Hoekstra 2011). This can be achieved by combining population genetics both within and between populations, and/or complementing *forward genetics* with *reverse genetics* (Stinchcombe & Hoekstra 2007; Butlin 2010; Hohenlohe, Phillips & Cresko 2010). The latter combination has been possible in wild populations of threespine sticklebacks where the confident identification of the *Eda* gene, responsible for the adaptive reduction of armour plates, was achieved by combining QTL mapping with SNP typing in wild populations (Colosimo *et al.* 2005) and also in *Heliconius* butterflies where the identification of the transcription factor *optix* controlling red adaptive wing pattern variation was possible thanks to the application of AFLP mapping (Baxter *et al.* 2008) followed by population genetic analyses on the focal region (Baxter *et al.* 2010b; Counterman *et al.* 2010; Reed *et al.* 2011).

CANDIDATE GENES

The knowledge of candidate genes derived from other organisms can be combined with these approaches to identify genes that are associated with a particular adaptive phenotype in a different species (Table 1). Thus, either in crosses or in naturally varying phenotypes, candidate genes can be examined for

evidence of their association with the trait of interest (Shimizu & Purugganan 2005; Stinchcombe & Hoekstra 2007). In this way, the variation in plumage colour in natural populations of the flycatcher *Monarcha castaneiventris* was shown to be associated with a single mutation of the previously identified pigmentation gene melanocortin-1 receptor (*MC1R*) (Mundy 2005; Uy *et al.* 2009). It is important to be selective at the time of applying the candidate gene approach. If there are many candidate genes associated with a phenotype, this might not be as fruitful as when there is a handful of strong candidates (Luikart *et al.* 2003). Furthermore, a major drawback of a candidate gene approach is that the literature becomes biased towards a few well-known candidate genes, which may not therefore be representative of their actual importance in evolutionary change (Mundy 2005). In addition, this approach makes the assumption that the genes that matter for evolution are necessarily few and of large effect.

As genomic technologies become more widely available and more information on gene interactions and pathways exists, researchers are moving away from a simplistic candidate gene approach and are applying larger studies that evaluate, at once, the role of multiple genes in a pathway suspected to affect the formation of the adaptive trait (candidate pathway approach) (Suh & Vijg 2005). For example, the flavonoid pathway has been a model system in plants that has helped understanding the genetics underlying flower coloration and other evolutionary processes, including the role of gene duplication in the evolution of novel phenotypes (Des Marais & Rausher 2008), causes of evolutionary rate variation among genes (Lu & Rausher 2003) and the relative importance of coding vs. regulatory mutations in the evolution of ecologically relevant traits (Wessinger & Rausher 2012). Flower coloration is an adaptive trait (Kopp 2009) caused by anthocyanin pigments, whose production requires at least six sequential reactions catalysed by six different enzymes in the anthocyanin pathway (Rausher 2006). The pathway candidate approach in flowers has not only led to identification of the particular enzymes involved in synthesizing red/orange, blue/magenta and blue/purple pigments (Zufall & Rausher 2003; Rausher 2006) but also helped in the identification of the transcriptional complex, composed by bHLH and MYB domain transcription factors, responsible for natural variation in flower coloration among many plant species (Rausher 2006; Kopp 2009). Despite the potential of the candidate pathway analysis in the study of natural adaptation, to date, it is poorly applied to evolutionary studies and largely remains restricted to studies on the genetics of human diseases (Suh & Vijg 2005).

CHARACTERIZING REGIONS NARROWED BY FORWARD OR REVERSE GENETICS

Outlier loci have been identified in a wide range of species (Bonin *et al.* 2006; Minder & Widmer 2008; Apple *et al.* 2010), but fewer studies have moved from their detection to the characterization of underlying QTNS, genes or networks controlling adaptation (Minder & Widmer 2008; Wood *et al.* 2008; Paris *et al.* 2010; Midamegbe *et al.* 2011; Rockman 2011; Ku-

nte *et al.* 2014). Following up on a particular outlier locus can be time-consuming and technically demanding, but will be a necessary step in order to find the genes or regulatory elements involved in adaptation. One of the most popular approaches is a library-based search of genomic regions flanked by outliers, followed by positional cloning using BACs and sequencing of the genetic interval (Butlin 2010; Nunes *et al.* 2012). This has been successfully applied in the identification of several wing colour pattern loci in *Heliconius* butterflies (Baxter *et al.* 2010b; Counterman *et al.* 2010), in *Papilio polytes* to find mimicry 'supergenes' (Kunte *et al.* 2014) and, in the marine gastropod *Littorina saxatilis*, it has been used to pinpoint candidate loci for local adaptation (Wood *et al.* 2008).

Pathway analysis (where multiple 'outlier' SNPs are analysed jointly) also offers an interesting, yet underutilized, tool to discover gene sets likely involved in the formation of a trait of interest (Pan *et al.* 2014). This is because although the detection of multiple outlier SNPs associated with a trait (with GWAS, for example) offers an insight into its underlying genetics, this alone may not be very informative in the case of quantitative polygenic traits, because individual SNPs only account for a small part of the trait variance (Mokry *et al.* 2013). However, pathway analysis is challenging, because a large number of SNPs per individual need to be considered in predictive models. Machine learning methods such as multidimensional reduction (MDR), support vector machines (SVM), neural networks (NN) and random forest (RF) are capable of dealing with this dimensionality problem in a flexible manner and can effectively select important variables from irrelevant ones (Goldstein *et al.* 2010; Gonzalez-Recio & Forni 2011; González-Recio, Rosa & Gianola 2014; Yang & Charles Gu 2014). In particular, RF analysis has been particularly useful in pathway analysis because interactions are implicitly modelled (De Lobel *et al.* 2010; Chung & Chen 2012) and it is straightforward to understand and interpret (Goldstein *et al.* 2010; González-Recio, Rosa & Gianola 2014). Nonetheless, this methodology has been primarily applied to the study of complex diseases in humans, considering a small number of SNPs (Chang *et al.* 2008; Ballard *et al.* 2010). Its application to large SNP data sets (such as those produced by next-generation techniques) is more complicated and requires the modification of certain standard assumptions (Goldstein *et al.* 2010; Chen & Ishwaran 2012). Still, this methodology offers an interesting option to the study of genetic pathways shaping natural adaptations. Pioneering work on this includes the search for loci involved in environmental adaptation in *Senecio lautus*. By comparing genomes of phenotypically contrasting parapatric populations, researchers assessed genetic association at different levels, from SNPs to physiological pathways (Roda *et al.* 2013).

Alternatively, when 'outlier' markers associated with a trait have been identified, they can be mapped back to a reference genome or to available linkage maps (in the same species or a closely related one), to detect whether they fall in or near protein-coding genes possibly affecting the trait. For example, in the rainbow and steelhead trout, species where no reference genome exists but linkage maps do, GWAS coupled with SNP

mapping discovered that migration in these species has a complex quantitative genetic basis, resulting from many loci of small effect (Hecht *et al*. 2013). A similar approach was used to study the genetics of the adaptive natural variation in female abdominal pigmentation in *Drosophila melanogaster* and determined that variation in this trait is under the control of *cis*-regulatory regions of the genes *tan* and *bric-à-brac* (Bastide *et al*. 2013). With the increasing availability of whole-genome sequences in a wide variety of taxa and the possibility to develop genomic resources at a reasonable cost in species that lack them, it is becoming more feasible to apply this strategy in non-model species.

Gene expression profiling

When DNA variation associated with phenotypic change does not occur in protein-coding regions, examining the expression pattern of genes provides an important complementary method to test for regulatory change (Rockman & Kruglyak 2006; Hoekstra & Coyne 2007; Hofmann *et al*. 2009). Although the detection of one or more differentially expressed genes provides information about potential candidates involved in the production of the trait of interest, this does not necessarily imply that all (or any) of them bear the causal variants. Differential expression may result from gene regulation due to upstream mutation(s), which may lie elsewhere in the genome (Stern & Orgogozo 2008; Kopp 2009; Stern & Orgogozo 2009). Thus, a combination of DNA data with expression data is often needed to determine whether the trait has a *cis*-regulatory basis (i.e. when differential expression and DNA polymorphism associated with phenotype both reside in the same locus) or it is *trans*-regulated. To this end, allele-specific expression (ASE) assays testing single or multiple genes offer a useful alternative to uncover the respective contributions of *cis*- and *trans*-regulatory variation (Knight 2004; Gilad, Rifkin & Pritchard 2008; Main *et al*. 2009; Wittkopp 2011).

Studies of gene expression can be conducted at either individual candidate loci (e.g. *in situ* hybridization, reverse-transcriptase quantitative PCR, immunochemistry) or many loci at once (e.g. microarrays, RNA-seq) depending on the information and resources available (Pavey *et al*. 2010). Studying thousands of transcripts allows a detailed and unbiased description of the genes involved in shaping natural evolution, and has the potential to identify entire genetic and developmental pathways driving adaptive variation. Thus, transcriptomic approaches can catalyse the discovery of multiple components of these gene networks, from the genes regulating phenotypic 'switches' to the repertoire of molecules responding to such genes of major effect using, for example, the recently developed Weighted Gene Co-Expression Network Analysis (WGCNA). This method offers the possibility to identify networks of co-expressed genes participating in the formation of a trait while at the same time points to candidate nodal 'key' genes likely controlling phenotypic variation only using gene expression data (Oldham, Horvath & Geschwind 2006; Filteau *et al*. 2013). This approach, however, does not identify the actual mutations controlling trait variance, and therefore,

complementary DNA characterization on these candidates should also be carried out.

Microarrays have pioneered the genome-wide characterization of transcripts associated with adaptive natural variation. Almost a decade ago, microarrays were used in one of the first attempts to identify genes controlling beak morphology variation in Darwin's finches and showed that the calmodulin (CaM)-dependent pathway is a key component of the evolution of beak variation in these birds (Abzhanov *et al*. 2006). Similarly, microarrays showed that in limnetic Coregonine fishes, the parallel phenotypic evolution towards using the same ecological niche involves similar changes in expression at the same genes (Derome & Bernatchez 2006). Nonetheless, microarrays suffer from several limitations. Background levels of hybridization, differences in hybridization properties among probes and the restriction of interrogating only the transcripts included in the array are among the most common problems of this technique (Marioni *et al*. 2008).

Nowadays, high-throughput mRNA sequencing technology (RNA-seq) has the potential to overcome some of these limitations (Marioni *et al*. 2008; Wang, Gerstein & Snyder 2009). By using next-generation sequencing technologies, RNA-seq allows for a direct estimation of relative transcript abundance across the entire genome (Cheviron & Brumfield 2012) keeping the background noise low where sequences can be unambiguously mapped to unique regions of the genome (Marioni *et al*. 2008; Wang, Gerstein & Snyder 2009). This technique is not limited to detecting transcripts in organisms with a reference genome as the same RNA-seq data can be used to create a transcriptome assembly which is then used as a reference for read mapping (Grabherr *et al*. 2011), thus being particularly attractive for non-model organisms. This approach has recently been applied to naturally varying organisms such as the intertidal copepod *Tigriopus californicus*, where differences in thermal tolerance were associated with differential expression of heat-shock proteins (*Hsp*) and genes involved in ubiquitination and proteolysis (Schoville *et al*. 2012) and, in eucalyptus, RNA-seq has provided insights into the molecular mechanisms underlying the adaptation to water shortage (Villar *et al*. 2011). RNA-seq, as any other next-generation sequencing technique, presents limitations in terms of data storage, analysis and cost. Nonetheless, with RNA-seq, it is especially important to consider sequence coverage (directly related to the sequencing cost) because in organisms with large genomes and complex transcriptomes, more sequencing depth will be required for an adequate coverage (Wang, Gerstein & Snyder 2009) and/or where multiple replicates and comparisons will be needed to correctly tackle a particular trait.

Alternatively, when strong single candidate genes have been isolated at the DNA level or have been derived from transcriptomics studies, their further and complementary characterization can be done using reverse-transcriptase quantitative PCR (RT-qPCR). Several studies have applied RT-qPCR to profile transcription levels of particular candidate loci. For example, measuring the expression levels of the gene *Agouti* with RT-qPCR (implicated in producing pheomelanin in *Mus musculus*) demonstrated that *cis*-regulatory evolution at this gene

was involved in adaptive variation in cryptic colouration of deer mice (Linnen *et al*. 2009). Similarly, adaptive differential retinal sensitivity in African cichlids inhabiting clear vs. turbid water relates to differences in *opsin* gene expression measured with RT-qPCR (Hofmann *et al*. 2009). However, RT-qPCR depends on the performance and specificity of primers so only the gene of interest is quantified (Busk 2014), making optimization a time-consuming process. Also, its accuracy is strongly reliant on the use of multiple control genes (e.g. reference or housekeeping genes) for normalization and correction of the multiple variation sources. A correct choice of control genes is not a trivial task. It is desirable to use more than one, as a single control gene can lead to normalization biases. Also, it is necessary to be sure that they are equally expressed across all the tissue types and species interrogated, as it can affect the accuracy of the calculation of relative expression differences between samples (Fedrigo *et al*. 2010).

The spatial distribution of expression patterns of candidate genes can be visualized using *in situ* hybridization (ISH), and this approach has provided the foundation for much of the field of 'evo-devo'. The technique involves hybridizing an antisense RNA probe to an mRNA transcript, and it is a powerful method to characterize gene expression in tissues. The ISH procedure follows five major steps: (i) sample preparation, including fixation, mounting and ISH pre-treatment, (ii) probe preparation, (iii) hybridization, (iv) probe removal and (v) detection (Apostolopoulos 2001). Several studies have exemplified the usefulness of this technique to characterize the loci of adaptation. Shapiro *et al*. (2004) used ISH to compare profiles of expression of the gene *Pitx1* between benthic and marine sticklebacks and thus showed that a *cis*-regulatory element of *Pitx1* is responsible for pelvic size variation in fishes of the two environments (Shapiro *et al*. 2004). In the Darwin's finches, ISH patterns of expression of the genes TGFβIIr, β-catenin and Dickkopf-3 are differentially expressed in the developing pre-maxillary bone of embryos of species with different beak shapes, a trait associated with the exploitation of multiple ecological niches (Mallarino *et al*. 2011). In a similar way, ISH showed that in *Heliconius* butterflies, *cis*-regulatory evolution of the transcription factor *optix* drives the convergent evolution of red wing patterns in distantly related species (Reed *et al*. 2011). Nonetheless, ISH protocols are not easy to establish because the technique can be challenging to optimize (Abzhanov *et al*. 2008). Furthermore, it requires enough supply of organismal tissue at different developmental stages. This therefore requires raising a large number of individuals in a controlled environment or sampling enough individuals in the wild, specifically at the developmental points required (Abzhanov *et al*. 2008), which already imposes a limitation for many natural systems.

The use of immunohistochemistry provides an alternative to spatially localize the products of gene expression, and it is a much more forgiving technique than ISH (Abzhanov *et al*. 2008). It has been successfully applied to document the genetics of adaptive wing radiation in *Heliconius* (Martin *et al*. 2014) and the genes controlling male wing pigmentation in *Drosophila* (Gompel *et al*. 2005; Prud'homme *et al*. 2006). However,

immunohistochemistry depends on the development of a species-specific antibody targeting the protein of interest, which can be time-consuming and expensive, or the availability of a cross-reactive antibody in a different species (Abzhanov *et al*. 2008). Just as ISH, immunohistochemistry also requires enough supply of organismal tissue at specific developmental points, limiting its application in many natural organisms.

Nonetheless, care should be taken in design and interpretation of gene expression assays. Variation in ecologically relevant traits is sometimes due to phenotypic plasticity (Hoffman & Goodisman 2007; Bossdorf, Richards & Pigliucci 2008; Whiteman & Agrawal 2009) which can be confounded with adaptive heritable variation. In this way, if the adaptive relevance of a trait has not been experimentally tested and it turns out to be a plastic phenotype, comparing gene expression in different conditions will yield a set of genes that do not contain the causal adaptive variants. Of course this is not to deny the importance of phenotypic plasticity in adaptation, but this is a subject beyond the scope of this paper (Ghalambor *et al*. 2007; Bossdorf, Richards & Pigliucci 2008; Hughes 2012).

Assays of molecular function

The implementation of assays of molecular function such as transgenics, knockouts, knockdowns (with RNA interference (RNAi), for example) and gene replacement constitutes an important test to prove that a gene actually underlies natural variation (Shimizu & Purugganan 2005; Hoekstra & Coyne 2007; Pavey *et al*. 2012) and establish whether it is required and/or sufficient for the development of the adaptive trait (Abzhanov *et al*. 2008). Functional tests of candidate genes are well implemented in model organisms like *Drosophila*, yeast, nematodes and mice (Feder & Mitchell-Olds 2003; Heffer & Pick 2013) and provide the standard evidence for the genetic basis of trait variation. In these organisms, for example, the application of functional tests has confirmed the identification of the genes, and even mutations, controlling adaptive natural variation. In *Drosophila*, the application of a set of transgenic reporter assays found the actual mutations in the regulatory elements of the *ebony* gene controlling adaptive abdominal pigmentation in African natural populations (Rebeiz *et al*. 2009). Similarly, in the deer mouse *Peromyscus maniculatus*, the generation of *Agouti* knockouts confirmed the involvement of this gene in adaptive melanism (Kingsley *et al*. 2009).

Over the past several years, new functional protocols have been developed for a wide range of emerging organisms including *Daphnia*, wasps, crickets, ladybirds, cavefish and sticklebacks (Osanai-Futahashi *et al*. 2012; Pavey *et al*. 2012) and, although this is a field under active development, these experiments are still not feasible in all organisms and therefore impose a limitation in several non-model systems. Nonetheless, when functional assays are impossible in the target organism, it is still possible to use closely related species as 'model' organisms. For example, use of the retroviral vector RCAS in the chicken embryo implicated the genes *CaM, TGFβIIr,* β-catenin and *Dkk3* in controlling beak development which may imply a role in the evolution of Darwin's finches (Abzhanov *et al*.

2006; Mallarino *et al.* 2011). However, it is important to bear in mind that such experiments in 'model' species do not directly identify the role of natural variants. Similarly, in the sticklebacks, a functional test of the *Ectodysplasin-A* (*EDA*) gene using transgenics showed how this gene controls adaptive plate variation in natural populations (Colosimo *et al.* 2005), but the transgenic construct carried a mouse *EDA-A1* cDNA rather than the native stickleback 'complete' *EDA* allele. Although changes in plate phenotype were indeed observed, the results were variable. Three out of fourteen transgenic 'low-plated' fishes developed extra plates on their sides, but not in a consistent manner; the number and type of extra plates developed varied between and within individuals (different in each side) (Colosimo *et al.* 2005). In the future, it is hoped that experiments can be developed that more directly replicate the role of naturally occurring variants in their native species.

With the constant development of functional tools that were previously only available in more traditional 'model' organisms, now it is possible not only to pinpoint genes and mutations shaping natural adaptations but also to establish new organisms in which to study the genetics underlying evolution. Nonetheless, the task of developing more functional assays applicable to a wider range of organism is still needed. RNA interference (RNAi), for example, is a method for knocking down expression of a target gene and appeared to be easily accessible. However, in some taxa such as the Lepidoptera, it has proved to be highly problematic (Terenius *et al.* 2011). Responding to this need, recently developed approaches commonly referred as 'genome editing tools' and based on the use of engineered nucleases coupled to DNA recognition domains have been developed; these include zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and the clustered regulatory interspaced short palindromic repeats (CRISPR)/Cas9 endonuclease system (Gaj, Gersbach & Barbas Iii 2013; Wei *et al.* 2013). In all these 'genome editing tools', the DNA-binding module recognizes and binds the target gene while the nuclease module induces DNA double-strand breaks (Bassett *et al.* 2013; Gaj, Gersbach & Barbas Iii 2013; Wei *et al.* 2013). This activates either error-prone non-homologous end joining, which commonly introduces indels or frameshifts leading to the knockout of gene function, or homology directed repair, which allows the introduction of changes from single nucleotide changes to entire transgenes (Gaj, Gersbach & Barbas Iii 2013). These emerging technologies have already started to expand the ability to manipulate genes in less traditional organisms. For example, a related technique using zinc-finger nucleases (ZFNs) has allowed the mutation of genes related to the circadian clockwork in the monarch butterfly (*Danaus plexippus*) (Merlin *et al.* 2012). Interestingly, the (CRISPR)/Cas9 system is now viewed as a more attractive technical choice as their DNA recognition potential is bigger than that of ZFNs or TALENs, although given its recent development, at the moment it only has been applied to 'model' organisms (Bassett *et al.* 2013; Chang *et al.* 2013; Gaj, Gersbach & Barbas Iii 2013; Wei *et al.* 2013). Yet, the potential of (CRISPR)/Cas9 in the study of the genetics of adaptation in natural systems is very promising

since it could be combined with classical genetic approaches (such as genetic complementation) to simultaneously map natural variation and functionally test the genes harbouring causal alleles (Turner 2014). This great advance in genome editing technologies opens new opportunities to decipher and test the molecular underpinnings of adaptations in a wider range of organisms and exemplifies how active research and development of tools broadens our methodological possibilities to answer longstanding questions.

In the future, transgenic tests should ideally involve replacement of alternate natural alleles at a locus, in order to demonstrate the functional value of particular substitutions. To date, this has only rarely, if ever, been achieved. Methods that involve knockouts or experiments involving 'model' species should be seen as a complement to other DNA or RNA approaches narrowing candidate genes likely shaping adaptive variation, but not a definitive test of adaptive function.

Assays of ecological function

Although all the methods described above help to identify genes (or QTNs) contributing to adaptive phenotypic variation, it is crucial to perform field experiments either in nature or in conditions that closely mimic naturally occurring events, to evaluate that the trait of interest is indeed adaptive and, also, to test the fitness consequences of allelic substitutions at the causal genes (Barrett & Hoekstra 2011). Experimental evolution studies that link genes, phenotype and fitness have been possible under laboratory conditions and using organisms such as virus, bacteria and yeast, with very short generation times and where the replicated sequencing of whole genomes is feasible (Rokyta *et al.* 2005; Barrick *et al.* 2009; Araya *et al.* 2010; Brockhurst, Colegrave & Rozen 2011; Barrick & Lenski 2013). However, potential confounding effects or artefacts in those artificial systems make it hard to extrapolate the patterns and conclusions derived from them to a context of natural adaptation, highlighting the need to perform this experiments in natural systems (Barrett & Hoekstra 2011).

The most common approach to evaluate whether a trait has a direct impact on fitness in nature is by testing cause-and-effect relationships in a planned field experiment comparing different conditions (i.e. varying the suspected natural selection agent). There, phenotypic variation in the trait driven by differences in environment should be observed. Field experiments to test the adaptive value of traits in the wild come in a great variety and complexity of forms including Q_{ST}–F_{ST} comparison (Leinonen *et al.* 2013), reciprocal transplants of hybrid individuals (Lowry *et al.* 2009), controlled introduction of live organisms to new environments (Reznick *et al.* 1997; Kapan 2001; Barrett & Schluter 2008; Irschick & Reznick 2009; Gompert *et al.* 2014) and replacing of the real organisms with synthetic replicas in nature primarily to quantify the impact of predation (Irschick & Reznick 2009; Merrill *et al.* 2012; Linnen *et al.* 2013). For example, Linnen *et al.* (2013) used plasticine mice models of two different colours in the field and found that light-coloured models matching light-coloured soil were less

attacked by visually hunting predators (Linnen *et al.* 2013). Similarly, when *Anolis* lizards were introduced from the mainland into a series of islands, the introduced population evolved a different hind limb phenotype potentially as an adaption to the use of narrow surfaces (Losos, Warheitt & Schoener 1997). While these examples show that a trait affects organismal fitness in the wild, they do not tell us about how the genetic variation in the genes shaping those adaptive traits evolves in response to the experimental treatments, and thus, no connection between genotype, phenotype and fitness can be established.

In addition, QTL mapping has been useful to identify regions associated with habitat adaptation using reciprocal transplants of hybrid individuals [F2, backcross, recombinant inbred lines (RILs), near isogenic lines (NILs)] in contrasting environments (Bradshaw & Schemske 2003; Lowry *et al.* 2009). This approach allows evaluating both genotype–environment interactions and the effects of epistasis in fitness (Barrett & Hoekstra 2011).

To overcome this, when the identity of the gene underlying the formation of the adaptive trait is known (using the methods described above), one should study whether genetic variation in the causal gene evolves in the expected direction in response to differential treatments and also presents signatures of selection (Weinig *et al.* 2003). For instance, hybrid sticklebacks between a lake and a river population were transplanted into river and lake environments that differ in their parasitic diversity. Under the hypothesis that in order to survive and reproduce a host should resist to local parasites and pathogens, researchers measured allele diversity at the major histocompatibility complex (MHC), involved in the recognition of parasite-specific antigens. After one generation, diversity of MHC alleles was higher at the lake environment (which bears a broader range of parasites than the river environment), thus providing lake sticklebacks the advantage of fighting a more diverse set of parasites (Eizaguirre *et al.* 2012). In a similar experiment, researchers measured selection on natural allelic variants of the (*Eda*) locus, known to control adaptive differences in armour plates in sticklebacks (Colosimo *et al.* 2005). By transplanting marine sticklebacks harbouring both the low-plate and high-plate alleles of *Eda* into freshwater ponds and studying genotype frequency variations in one generation, researchers found that the low-plate allele was positively selected once lateral plates developed, likely because it provides a growth advantage in freshwater environments. However, the same allele was negatively selected before the plates were formed, indicating that either the *Eda* gene affects additional unmeasured traits under selection, or that tightly linked loci also have effects on fitness (Barrett & Schluter 2008). This last observation suggests that, when possible, tests of selection should be performed genome-wide using next-generation sequencing techniques in order to generate a high density of markers. This resolution can permit not only to confirm/detect selection signatures on the adaptive gene itself but also to detect the loci controlling unmeasured traits with fitness effects, and determine whether they are shaped by the same pleiotropic gene or by multiple linked loci (Barrett & Hoekstra 2011). To date, only a handful of studies have explored the genomic consequences of contemporary selection in the field using whole-genome data. In one study, researchers transplanted stick insects to native and novel host plants and measured allele frequency changes within a generation at genome-wide level (Gompert *et al.* 2014). In another study, patterns of genome-wide selection in purple sea urchins were evaluated under different ocean acidification levels (Pespeni *et al.* 2013). Both studies detect changes in allele frequencies driven by selection at multiple loci across the genome. However, as there was no previous knowledge of gene(s) underlying the formation of the adaptive traits directly affecting fitness, these results do not directly lead to conclusions about how selection affects allelic variance in a causal gene and the possible explanations for observing selection in non-causal regions of the genome (i.e. pleiotropy, indirect selection, linkage).

In addition, the combination of genomic data with selection experiments also gives the opportunity to evaluate the role of epistasis in adaptation. This has been applied to the study of laboratory adaptation using yeast, bacteria and viruses. These laboratory-based studies show a global pattern of diminishing returns epistasis (i.e. where the more mutations that accumulate, the weaker their fitness effect), which impedes the rate of ongoing adaptation relative to a null model of independent mutational effects (Chou *et al.* 2011; Khan *et al.* 2011; Kryazhimskiy *et al.* 2014). As the evolutionary patterns observed in small laboratory populations may not be the same as those contributing to natural evolution, the confirmation of these results still needs to be carried out in natural populations in order to determine the genomic effect of epistasis and its overall contribution to natural adaptation.

Conclusions

Recent research has led to a remarkable growth in our understanding of the molecular basis of adaptive evolution (Table 1). Altogether, these studies have provided important insights into the genetic basis of adaptations and also the methodological approaches needed to answer this evolutionary problem. Nonetheless, it has become clear that the characterization of the genes underlying adaptive traits is not an easy task because factors such as demography, epistasis and pleiotropy can introduce confounding effects that will complicate any clear genetic signal. Also, methodological bias can mislead findings by pointing to large-effect loci and missing the detection of genes with small effect, thus complicating the description of 'all' important variants contributing to natural adaptation. Still, the search for the loci of evolution can benefit from following an organized and complementary methodology. First, it is necessary to corroborate that a trait affects fitness in the field and is in fact adaptive. Then, the region(s) of the genome in which genotypes are correlated with adaptive phenotypes should be defined either with classical genetic tools or applying new genomic approaches. Next, when DNA polymorphism associated with phenotype in the candidate genes does not occur in protein-coding regions, the expression pattern of such genes must be analysed for in order to test whether

the trait has a *cis-* or *trans-*regulatory basis. Ultimately, functional experiments are required to prove that a gene or mutation is actually responsible for the phenotype observed. Once individual genes or SNPs have been identified, it is important to quantify their effect in the 'trait value' (i.e. how much variation in the phenotype is explained by the candidate SNPs/genes). Finally, the genetic variation in the genes shaping those adaptive traits should be evaluated in field selection experiments in order to establish a definite connection between genotype, phenotype and fitness.

A comprehensive review of the conclusions of such studies is beyond the scope of this article. However, some of the major findings include the following. First, the evolution of similar adaptive traits in different lineages commonly involves the action of the same genes (Colosimo *et al.* 2005; Nadeau & Jiggins 2010; Reed *et al.* 2011). Secondly, both *cis-*regulatory changes and coding changes contribute to adaptive variation (Mundy *et al.* 2004; Colosimo *et al.* 2005; Chan *et al.* 2010; Kunte *et al.* 2014). However, *cis-*regulatory changes may be more frequently involved in the evolution of morphological traits compared to physiological traits and in the evolution of morphological interspecific differences compared to the evolution of morphological intraspecific variation (Stern & Orgogozo 2008). Thirdly, the position of an adaptive gene in a regulatory network matters, as mutation in upstream patterning (input) genes is likely to affect the development of several body structures, while mutations in downstream (responsive) genes will influence the form of all incidences of the particular structure (Stern & Orgogozo 2008). However, *cis-*regulatory mutations in input/output genes provide great precision in evolutionary change with minimal pleiotropic effects (if any) (Stern & Orgogozo 2008; Gompel & Prud'homme 2009). Fourthly, adaptations can evolve from standing genetic variation, *de novo* mutations and adaptive introgression (Hermisson & Pennings 2005; Feldman, Brodie & Pfrender 2009; Hedrick 2013). Finally, most of the adaptations reported to date seem to arise through few initial mutations of major effect followed by many small effect mutations on minor genes (Orr 2005; Rockman 2011; Olson-Manning, Wagner & Mitchell-Olds 2012). Nonetheless, this last observation may be tainted by our experimental bias towards detecting large-effect alleles, so there is a likely ascertainment bias in the literature (Rockman 2011).

The search for the loci of evolution will be surely fuelled by the continuous increase in genomic and transcriptomics resources in natural populations, along with the development of novel methodologies applicable to such organisms. This offers exciting opportunities for testing new predictions and understanding how evolution proceeds.

Acknowledgements

Data accessibility

This manuscript does not use data.

References

Abzhanov, A., Kuo, W.P., Hartmann, C., Grant, B.R., Grant, P.R. & Tabin, C.J. (2006) The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature*, **442**, 563–567.

Abzhanov, A., Extavour, C.G., Groover, A., Hodges, S.A., Hoekstra, H.E., Kramer, E.M. & Monteiro, A. (2008) Are we there yet? Tracking the development of new model systems. *Trends in Genetics*, **24**, 353–360.

Albertson, R., Yan, Y., Titus, T., Pisano, E., Vacchi, M., Yelick, P., Detrich, H.r. & Postlethwait, J. (2010) Molecular pedomorphism underlies craniofacial skeletal evolution in Antarctic notothenioid fishes. *BMC Evolutionary Biology*, **10**, 4.

Alhaddad, H., Khan, R., Grahn, R.A., Gandolfi, B., Mullikin, J.C., Cole, S.A. *et al.* (2013) Extent of linkage disequilibrium in the domestic cat, *Felis silvestris catus*, and its breeds. *PLoS ONE*, **8**, e53537.

Altshuler, D., Pollara, V.J., Cowles, C.R., Van Etten, W.J., Baldwin, J., Linton, L. & Lander, E.S. (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature*, **407**, 513–516.

Amemiya, C.T., Alfoldi, J., Lee, A.P., Fan, S., Philippe, H., MacCallum, I. *et al.* (2013) The African coelacanth genome provides insights into tetrapod evolution. *Nature*, **496**, 311–316.

Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T.T., Mast, J., Sunayama-Morita, T. & Stern, D.L. (2011) Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Research*, **21**, 610–617.

Apostolopoulos, J. (2001) Nonradioactive *in situ* hybridization in atherosclerotic tissue. *Atherosclerosis* (ed. A. Drew), pp. 195–206. Humana Press, New York.

Apple, J.L., Grace, T., Joern, A., Amand, P.S. & Wisely, S.M. (2010) Comparative genome scan detects host-related divergent selection in the grasshopper *Hesperotettix viridis*. *Molecular Ecology*, **19**, 4012–4028.

Aranzana, M., Kim, S., 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 Genetics*, **1**, e60.

Araya, C., Payen, C., Dunham, M. & Fields, S. (2010) Whole-genome sequencing of a laboratory-evolved yeast strain. *BMC Genomics*, **11**, 88.

Arnoult, L., Su, K.F.Y., Manoel, D., Minervino, C., Magriña, J., Gompel, N. & Prud'homme, B. (2013) Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science*, **339**, 1423–1426.

Assunção, A., Pieper, B., Vromans, J., Lindhout, P., Aarts, M. & Schat, H. (2006) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. *New Phytologist*, **170**, 21–32.

Balding, D. & Nichols, R. (1995) A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. *Genetica*, **96**, 3–12.

Ballard, D., Abraham, C., Cho, J. & Zhao, H. (2010) Pathway analysis comparison using Crohn's disease genome wide association studies. *BMC Medical Genomics*, **3**, 25.

Barrett, R.D.H. & Hoekstra, H.E. (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics*, **12**, 767–780.

Barrett, R.D.H. & Schluter, D. (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**, 38–44.

Barrick, J.E. & Lenski, R.E. (2013) Genome dynamics during experimental evolution. *Nature Reviews Genetics*, **14**, 827–839.

Barrick, J., Yu, D., Yoon, S., Jeong, H., Oh, T., Schneider, D., Lenski, R. & Kim, J. (2009) Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*, **461**, 1243–1247.

Bassett, A.R., Tibbit, C., Ponting, C.P. & Liu, J.-L. (2013) Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Reports*, **4**, 220–228.

Bastide, H., Betancourt, A., Nolte, V., Tobler, R., Stöbe, P., Futschik, A. & Schlötterer, C. (2013) A genome-wide, fine-scale map of natural pigmentation variation in *Drosophila melanogaster*. *PLoS Genetics*, **9**, e1003534.

Baxter, S.W., Papa, R., Chamberlain, N., Humphray, S.J., Joron, M., Morrison, C., ffrench-Constant, R.H., McMillan, W.O. & Jiggins, C.D. (2008) Convergent evolution in the genetic basis of Müllerian mimicry in *Heliconius* butterflies. *Genetics*, **180**, 1567–1577.

Baxter, I., Brazelton, J.N., Yu, D., Huang, Y.S., Lahner, B., Yakubova, E. *et al.* (2010a) A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS Genetics*, **6**, e1001193.

Baxter, S.W., Nadeau, N.J., Maroja, L.S., Wilkinson, P., Counterman, B.A., Dawson, A. *et al.* (2010b) Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in the *Heliconius melpomene* clade. *PLoS Genetics*, **6**, e1000794.

Baxter, S.W., Davey, J.W., Johnston, J.S., Shelton, A.M., Heckel, D.G., Jiggins, C.D. & Blaxter, M.L. (2011) Linkage mapping and comparative genomics

using next-generation RAD sequencing of a non-model organism. *PLoS ONE*, **6**, e19315.

Beaulieu, J., Doerksen, T., Boyle, B., Clément, S., Deslauriers, M., Beauseigle, S. et al. (2011) Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics*, **188**, 197–214.

Beaumont, M.A. & Nichols, R.A. (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **263**, 1619–1626.

Beldade, P., Brakefield, P.M. & Long, A.D. (2002) Contribution of *distal-less* to quantitative variation in butterfly eyespots. *Nature*, **415**, 315–318.

Beraldi, D., McRae, A.F., Gratten, J., Slate, J., Visscher, P.M. & Pemberton, J.M. (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.

Bergelson, J. & Roux, F. (2010) Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nature Reviews Genetics*, **11**, 867–870.

Bodmer, W. & Bonilla, C. (2008) Common and rare variants in multifactorial susceptibility to common diseases. *Nature Genetics*, **40**, 695–701.

Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S. & SanCristobal, M. (2010) Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics*, **186**, 241–262.

Bonin, A. (2008) Population genomics: a new generation of genome scans to bridge the gap with functional genomics. *Molecular Ecology*, **17**, 3583–3584.

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*). *Molecular Biology and Evolution*, **23**, 773–783.

Bossdorf, O., Richards, C.L. & Pigliucci, M. (2008) Epigenetics for ecologists. *Ecology Letters*, **11**, 106–115.

Bradley Shaffer, H., Minx, P., Warren, D., Shedlock, A., Thomson, R., Valenzuela, N. et al. (2013) The western painted turtle genome, a model for the evolution of extreme physiological adaptations in a slowly evolving lineage. *Genome Biology*, **14**, R28.

Bradshaw, H.D. & Schemske, D.W. (2003) Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature*, **426**, 176–178.

Bradshaw, H.D., Otto, K.G., Frewen, B.E., McKay, J.K. & Schemske, D.W. (1998) Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics*, **149**, 367–382.

Brockhurst, M.A., Colegrave, N. & Rozen, D.E. (2011) Next-generation sequencing as a tool to study microbial evolution. *Molecular Ecology*, **20**, 972–980.

Busk, P. (2014) A tool for design of primers for microRNA-specific quantitative RT-qPCR. *BMC Bioinformatics*, **15**, 29.

Butlin, R. (2010) Population genomics and speciation. *Genetica*, **138**, 409–418.

Cai, Q., Qian, X., Lang, Y., Luo, Y., Pan, S., Hui, Y. et al. (2013) The genome sequence of the ground tit *Pseudopodoces humilis* provides insights into its adaptation to high altitude. *Genome Biology*, **14**, R29.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G. Jr, Shapiro, M.D., Brady, S.D. et al. (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science*, **327**, 302–305.

Chang, J.S., Yeh, R.-F., Wiencke, J.K., Wiemels, J.L., Smirnov, I., Pico, A.R. et al. (2008) Pathway analysis of single-nucleotide polymorphisms potentially associated with glioblastoma multiforme susceptibility using Random Forests. *Cancer Epidemiology Biomarkers & Prevention*, **17**, 1368–1373.

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.-W. & Xi, J.J. (2013) Genome editing with RNA-guided Cas9 nuclease in Zebrafish embryos. *Cell Research*, **23**, 465–472.

Chen, X. & Ishwaran, H. (2012) Random forests for genomic data analysis. *Genomics*, **99**, 323–329.

Cheviron, Z.A. & Brumfield, R.T. (2012) Genomic insights into adaptation to high-altitude environments. *Heredity*, **108**, 354–361.

Chou, H.-H., Chiu, H.-C., Delaney, N.F., Segré, D. & Marx, C.J. (2011) Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science*, **332**, 1190–1192.

Chung, R.-H. & Chen, Y.-E. (2012) A two-stage random forest-based pathway analysis method. *PLoS ONE*, **7**, e36662.

Collin, H. & Fumagalli, L. (2011) Evidence for morphological and adaptive genetic divergence between lake and stream habitats in European minnows (*Phoxinus phoxinus*, Cyprinidae). *Molecular Ecology*, **20**, 4490–4502.

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

Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J.K. (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 1411–1423.

Counterman, B.A., Araujo-Perez, F., Hines, H.M., Baxter, S.W., Morrison, C.M., Lindstrom, D.P. et al. (2010) Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*. *PLoS Genetics*, **6**, e1000796.

Cruickshank, T.E. & Hahn, M.W. (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, **23**, 3133–3157.

Cutter, A.D. & Payseur, B.A. (2013) Genomic signatures of selection at linked sites: unifying the disparity among species. *Nature Reviews Genetics*, **14**, 262–274.

Dahlhoff, E.P. & Rank, N.E. (2000) Functional and physiological consequences of genetic variation at phosphoglucose isomerase: heat shock protein expression is related to enzyme genotype in a montane beetle. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 10056–10061.

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M. & Blaxter, M.L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.

De Lobel, L., Geurts, P., Baele, G., Castro-Giner, F., Kogevinas, M. & Van Steen, K. (2010) A screening methodology based on Random Forests to improve the detection of gene-gene interactions. *European Journal of Human Genetics*, **18**, 1127–1132.

De Mita, S., Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J. & Vigouroux, Y. (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, **22**, 1383–1399.

Derome, N. & Bernatchez, L. (2006) The transcriptomics of ecological convergence between 2 limnetic coregonine fishes (Salmonidae). *Molecular Biology and Evolution*, **23**, 2370–2378.

Derome, N., Bougas, B., Rogers, S.M., Whiteley, A.R., Labbe, A., Laroche, J. & Bernatchez, L. (2008) Pervasive sex-linked effects on transcription regulation as revealed by expression quantitative trait loci mapping in lake whitefish species pairs (*Coregonus* sp., Salmonidae). *Genetics*, **179**, 1903–1917.

Des Marais, D.L. & Rausher, M.D. (2008) Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature*, **454**, 762–765.

Dittmar, E.L., Oakley, C.G., Ågren, J. & Schemske, D.W. (2014) Flowering time QTL in natural populations of *Arabidopsis thaliana* and implications for their adaptive value. *Molecular Ecology*, **23**, 4291–4303.

Dopman, E., Robbins, P. & Seaman, A. (2010) Components of reproductive isolation between North American pheromone strains of the European corn borer. *Evolution*, **64**, 881–902.

Durbin, M.L., Lundy, K.E., Morrell, P.L., Torres-Martinez, C.L. & Clegg, M.T. (2003) Genes that determine flower color: the role of regulatory changes in the evolution of phenotypic adaptations. *Molecular Phylogenetics and Evolution*, **29**, 507–518.

Eizaguirre, C., Lenz, T.L., Kalbe, M. & Milinski, M. (2012) Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, **15**, 723–731.

Ellegren, H. & Sheldon, B.C. (2008) Genetic basis of fitness differences in natural populations. *Nature*, **452**, 169–175.

Ellegren, H., Smeds, L., Burri, R., Olason, P.I., Backstrom, N., Kawakami, T. et al. (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature*, **491**, 756–760.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. & Mitchell, S.E. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.

Feder, M.E. & Mitchell-Olds, T. (2003) Evolutionary and ecological functional genomics. *Nature Reviews Genetics*, **4**, 649–655.

Fedrigo, O., Warner, L.R., Pfefferle, A.D., Babbitt, C.C., Cruz-Gordillo, P. & Wray, G.A. (2010) A pipeline to determine RT-QPCR control genes for evolutionary studies: application to primate gene expression across multiple tissues. *PLoS ONE*, **5**, e12545.

Feldman, C.R., Brodie, E.D. & Pfrender, M.E. (2009) The evolutionary origins of beneficial alleles during the repeated adaptation of garter snakes to deadly prey. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 13415–13420.

Filteau, M., Pavey, S.A., St-Cyr, J. & Bernatchez, L. (2013) Gene co-expression networks reveal key drivers of phenotypic divergence in lake whitefish. *Molecular Biology and Evolution*, **30**, 1384–1396.

Foll, M. & Gaggiotti, O. (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. *Genetics*, **180**, 977–993.

Frichot, E., Schoville, S.D., Bouchard, G. & François, O. (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687–1699.

Gaj, T., Gersbach, C.A. & Barbas Iii, C.F. (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, **31**, 397–405.

Galindo, J., Grahame, J.W. & Butlin, R.K. (2010) An EST-based genome scan using 454 sequencing in the marine snail *Littorina saxatilis*. *Journal of Evolutionary Biology*, **23**, 2004–2016.

Ghalambor, C.K., McKay, J.K., Carroll, S.P. & Reznick, D.N. (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, **21**, 394–407.

Gilad, Y., Rifkin, S.A. & Pritchard, J.K. (2008) Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends in Genetics*, **24**, 408–415.

Goldstein, B.A., Hubbard, A.E., Cutler, A. & Barcellos, L.F. (2010) An application of Random Forests to a genome-wide association dataset: methodological considerations & new findings. *BMC Genetics*, **11**, 49.

Gompel, N. & Prud'homme, B. (2009) The causes of repeated genetic evolution. *Developmental Biology*, **332**, 36–47.

Gompel, N., Prud'homme, B., Wittkopp, P.J., Kassner, V.A. & Carroll, S.B. (2005) Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature*, **433**, 481–487.

Gompert, Z., Comeault, A.A., Farkas, T.E., Feder, J.L., Parchman, T.L., Buerkle, C.A. & Nosil, P. (2014) Experimental evidence for ecological selection on genome variation in the wild. *Ecology Letters*, **17**, 369–379.

Gonzalez-Recio, O. & Forni, S. (2011) Genome-wide prediction of discrete traits using bayesian regressions and machine learning. *Genetics Selection Evolution*, **43**, 7.

González-Recio, O., Rosa, G.J.M. & Gianola, D. (2014) Machine learning methods and predictive ability metrics for genome-wide prediction of complex traits. *Livestock Science*, **166**, 217–231.

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**, 491–502.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I. *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech*, **29**, 644–652.

Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W. & Dangl, J.L. (1995) Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science*, **269**, 843–846.

Gratten, J., Beraldi, D., Lowder, B., McRae, A.F., Visscher, P., Pemberton, J. & Slate, J. (2007) Compelling evidence that a single nucleotide substitution in *TYRP1* is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 619–626.

Gunnarsson, U., Hellström, A.R., Tixier-Boichard, M., Minvielle, F., Bed'hom, B., Ito, S.i. *et al.* (2007) Mutations in *SLC45A2* cause plumage color variation in chicken and japanese quail. *Genetics*, **175**, 867–877.

Gusev, A., Lowe, J.K., Stoffel, M., Daly, M.J., Altshuler, D., Breslow, J.L., Friedman, J.M. & Pe'er, I. (2009) Whole population, genome-wide mapping of hidden relatedness. *Genome Research*, **19**, 318–326.

Gusev, A., Kenny, Eimear.E., Lowe, Jennifer.K., Salit, J., Saxena, R., Kathiresan, S. *et al.* (2011) DASH: a method for identical-by-descent haplotype mapping uncovers association with recent variation. *The American Journal of Human Genetics*, **88**, 706–717.

Haag, C.R., Saastamoinen, M., Marden, J.H. & Hanski, I. (2005) A candidate locus for variation in dispersal rate in a butterfly metapopulation. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2449–2456.

Hancock, A.M., Brachi, B., Faure, N., Horton, M.W., Jarymowycz, L.B., Sperone, F.G., Toomajian, C., Roux, F. & Bergelson, J. (2011a) Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*, **334**, 83–86.

Hancock, A.M., Witonsky, D.B., Alkorta-Aranburu, G., Beall, C.M., Gebremedhin, A., Sukernik, R. *et al.* (2011b) Adaptations to climate-mediated selective pressures in humans. *PLoS Genetics*, **7**, e1001375.

Hassinen, V., Tervahauta, A., Halimaa, P., Plessl, M., Peräniemi, S., Schat, H., Aarts, M., Servomaa, K. & Kärenlampi, S. (2007) Isolation of Zn-responsive genes from two accessions of the hyperaccumulator plant *Thlaspi caerulescens*. *Planta*, **225**, 977–989.

Hecht, B.C., Campbell, N.R., Holecek, D.E. & Narum, S.R. (2013) Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. *Molecular Ecology*, **22**, 3061–3076.

Hedrick, P.W. (2013) Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, **22**, 4606–4618.

Heffer, A. & Pick, L. (2013) Conservation and variation in Hox genes: how insect models pioneered the evo-devo field. *Annual Review of Entomology*, **58**, 161–179.

Hermisson, J. & Pennings, P.S. (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics*, **169**, 2335–2352.

Hines, H.M., Papa, R., Ruiz, M., Papanicolaou, A., Wang, C., Nijhout, H.F., McMillan, O.W. & Reed, R.D. (2012) Transcriptome analysis reveals novel patterning and pigmentation genes underlying *Heliconius* butterfly wing pattern variation. *BMC Genomics*, **13**, 288.

Hoekstra, H.E. & Coyne, J.A. (2007) The locus of evolution: evo-Devo and the genetics of adaptation. *Evolution*, **61**, 995–1016.

Hoekstra, H.E. & Nachman, M.W. (2003) Different genes underlie adaptive melanism in different populations of rock pocket mice. *Molecular Ecology*, **12**, 1185–1194.

Hoekstra, H.E., Hirschmann, R.J., Bundey, R.A., Insel, P.A. & Crossland, J.P. (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science*, **313**, 101–104.

Hoffman, E. & Goodisman, M. (2007) Gene expression and the evolution of phenotypic diversity in social wasps. *BMC Biology*, **5**, 23.

Hofmann, C.M., O'Quin, K.E., Marshall, N.J., Cronin, T.W., Seehausen, O. & Carleton, K.L. (2009) The eyes have it: regulatory and structural changes both underlie cichlid visual pigment diversity. *PLoS Biology*, **7**, e1000266.

Hohenlohe, P.A., Phillips, P.C. & Cresko, W.A. (2010) Using population genomics to detect selection in natural populations: key concepts and methodological considerations. *International Journal of Plant Sciences*, **171**, 1059–1071.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. & Cresko, W.A. (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e1000862.

Holliday, J., Ritland, K. & Aitken, S. (2010) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytology*, **188**, 501–514.

Holliday, J., Wang, T. & Aitken, S. (2012) Predicting adaptive phenotypes from multilocus genotypes in Sitka Spruce (*Picea sitchensis*) using random forest. *G3 (Bethesda)*, **2**, 1085–1093.

Hong, H., Xu, L., Liu, J., Jones, W.D., Su, Z., Ning, B. *et al.* (2012) Technical reproducibility of genotyping SNP arrays used in genome-wide association studies. *PLoS ONE*, **7**, e44483.

Hubbard, J.K., Uy, J.A.C., Hauber, M.E., Hoekstra, H.E. & Safran, R.J. (2010) Vertebrate pigmentation: from underlying genes to adaptive function. *Trends in Genetics*, **26**, 231–239.

Huggins, P., Johnson, C.K., Schoergendorfer, A., Putta, S., Bathke, A.C., Stromberg, A.J. & Voss, S.R. (2012) Identification of differentially expressed thyroid hormone responsive genes from the brain of the Mexican Axolotl (*Ambystoma mexicanum*). *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, **155**, 128–135.

Hughes, A.L. (2012) Evolution of adaptive phenotypic traits without positive Darwinian selection. *Heredity*, **108**, 347–353.

Hunter, B., Wright, K.M. & Bomblies, K. (2013) Short read sequencing in studies of natural variation and adaptation. *Current Opinion in Plant Biology*, **16**, 85–91.

Hwang, E.-Y., Song, Q., Jia, G., Specht, J., Hyten, D., Costa, J. & Cregan, P. (2014) A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics*, **15**, 1.

Irschick, D. & Reznick, D.N. (2009) Field experiments, introductions, and experimental evolution. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments* (eds T. Garland & M.R. Rose), pp. 173–193. Oxford University Press, Oxford.

Jarvis, J.P., Scheinfeldt, L.B., Soi, S., Lambert, C., Omberg, L., Ferwerda, B. *et al.* (2012) Patterns of ancestry, signatures of natural selection, and genetic association with stature in western African pygmies. *PLoS Genetics*, **8**, e1002641.

Jeong, S., Rebeiz, M., Andolfatto, P., Werner, T., True, J. & Carroll, S.B. (2008) The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell*, **132**, 783–793.

Jessen, T.H., Weber, R.E., Fermi, G., Tame, J. & Braunitzer, G. (1991) Adaptation of bird hemoglobins to high altitudes: demonstration of molecular mechanism by protein engineering. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 6519–6522.

Jiggins, C.D., Mavarez, J., Beltrán, M., McMillan, W.O., Johnston, J.S. & Bermingham, E. (2005) A genetic linkage map of the mimetic butterfly *Heliconius melpomene*. *Genetics*, **171**, 557–570.

Johnston, S.E., Beraldi, D., McRae, A.F., Pemberton, J.M. & Slate, J. (2010) Horn type and horn length genes map to the same chromosomal region in Soay sheep. *Heredity*, **104**, 196–205.

Johnston, S.E., McEwan, J.C., Pickering, N.K., Kijas, J.W., Beraldi, D., Pilkington, J.G., Pemberton, J.M. & Slate, J.O.N. (2011) Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*, **20**, 2555–2566.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J. *et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55–61.

Joost, S., Bonin, A., Bruford, M.W., DesprÉS, L., Conord, C., Erhardt, G. & Taberlet, P. (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, **16**, 3955–3969.

Kapan, D.D. (2001) Three-butterfly system provides a field test of mullerian mimicry. *Nature*, **409**, 338–340.

Khan, A.I., Dinh, D.M., Schneider, D., Lenski, R.E. & Cooper, T.F. (2011) Negative epistasis between beneficial mutations in an evolving bacterial population. *Science*, **332**, 1193–1196.

Kim, Y. & Nielsen, R. (2004) Linkage disequilibrium as a signature of selective sweeps. *Genetics*, **167**, 1513–1524.

Kingsley, E.P., Manceau, M., Wiley, C.D. & Hoekstra, H.E. (2009) Melanism in *Peromyscus* is caused by independent mutations in *Agouti*. *PLoS ONE*, **4**, e6435.

Knight, J.C. (2004) Allele-specific gene expression uncovered. *Trends in Genetics*, **20**, 113–116.

Kohn, M.H., Pelz, H.-J. & Wayne, R.K. (2003) Locus-specific genetic differentiation at Rw among warfarin-resistant rat (*Rattus norvegicus*) populations. *Genetics*, **164**, 1055–1070.

Kopp, A. (2009) Metamodels and phylogenetic replication: a systematic approach to the evolution of developmental pathways. *Evolution*, **63**, 2771–2789.

Korte, A. & Farlow, A. (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, **9**, 29.

Kryazhimskiy, S., Rice, D.P., Jerison, E.R. & Desai, M.M. (2014) Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science*, **344**, 1519–1522.

Kunte, K., Zhang, W., Tenger-Trolander, A., Palmer, D.H., Martin, A., Reed, R.D., Mullen, S.P. & Kronforst, M.R. (2014) *doublesex* is a mimicry supergene. *Nature*, **507**, 229–232.

Lachance, J., Vernot, B., Elbers, C.C., Ferwerda, B., Froment, A., Bodo, J.-M. *et al.* (2012) Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse african hunter-gatherers. *Cell*, **150**, 457–469.

Lasseur, R., Longin-Sauvageon, C., Videmann, B., Billeret, M., Berny, P. & Benoit, E. (2006) Warfarin resistance in a French strain of rats. *Journal of Biochemical and Molecular Toxicology*, **19**, 379–385.

Leinonen, T., McCairns, R.J.S., O'Hara, R.B. & Merila, J. (2013) Q_{ST}-F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics*, **14**, 179–190.

Lempe, J., Balasubramanian, S., Sureshkumar, S., Singh, A., Schmid, M. & Weigel, D. (2005) Diversity of flowering responses in wild *Arabidopsis thaliana* strains. *PLoS Genetics*, **1**, e6.

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

Li, H. & Durbin, R. (2011) Inference of human population history from individual whole-genome sequences. *Nature*, **475**, 493–496.

Linnen, C.R., Kingsley, E.P., Jensen, J.D. & Hoekstra, H.E. (2009) On the origin and spread of an adaptive allele in deer mice. *Science*, **325**, 1095–1098.

Linnen, C.R., Poh, Y.-P., Peterson, B.K., Barrett, R.D.H., Larson, J.G., Jensen, J.D. & Hoekstra, H.E. (2013) Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, **339**, 1312–1316.

Long, A.D. & Langley, C.H. (1999) The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Research*, **9**, 720–731.

Losos, J.B., Warheitt, K.I. & Schoener, T.W. (1997) Adaptive differentiation following experimental island colonization in Anolis lizards. *Nature*, **387**, 70–73.

Lotterhos, K.E. & Whitlock, M.C. (2014) Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, **23**, 2178–2192.

Lowry, D.B., Hall, M.C., Salt, D.E. & Willis, J.H. (2009) Genetic and physiological basis of adaptive salt tolerance divergence between coastal and inland *Mimulus guttatus*. *New Phytologist*, **183**, 776–788.

Lu, Y. & Rausher, M.D. (2003) Evolutionary rate variation in anthocyanin pathway genes. *Molecular Biology and Evolution*, **20**, 1844–1853.

Luikart, G., England, P.R., Tallmon, D., Jordan, S. & Taberlet, P. (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

Lundgren, P., Vera, J., Peplow, L., Manel, S. & van Oppen, M. (2013) Genotype – environment correlations in corals from the Great Barrier Reef. *BMC Genetics*, **14**, 9.

Maccaferri, M., Sanguineti, M.C., Demontis, A., El-Ahmed, A., Garcia del Moral, L., Maalouf, F. *et al.* (2011) Association mapping in durum wheat grown across a broad range of water regimes. *Journal of Experimental Botany*, **62**, 409–438.

Mackay, T.F.C., Stone, E.A. & Ayroles, J.F. (2009) The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics*, **10**, 565–577.

Main, B., Bickel, R., McIntyre, L., Graze, R., Calabrese, P. & Nuzhdin, S. (2009) Allele-specific expression assays using Solexa. *BMC Genomics*, **10**, 1–9.

Mallarino, R., Grant, P.R., Grant, B.R., Herrel, A., Kuo, W.P. & Abzhanov, A. (2011) Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 4057–4062.

Manceau, M., Domingues, V.S., Mallarino, R. & Hoekstra, H.E. (2011) The developmental role of *Agouti* in color pattern evolution. *Science*, **331**, 1062–1065.

Manel, S., Conord, C. & Despres, L. (2009) Genome scan to assess the respective role of host-plant and environmental constraints on the adaptation of a widespread insect. *BMC Evolutionary Biology*, **9**, 288.

Marioni, J., Mason, C., Mane, S., Stephens, M. & Gilad, Y. (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research*, **18**, 1509–1517.

Martin, S.H. & Jiggins, C.D. (2013) Genomic studies of adaptation in natural populations. *eLS*. John Wiley & Sons, Ltd, Chichester. doi: 10.1002/9780470015902.a9780470024613

Martin, A., Papa, R., Nadeau, N.J., Hill, R.I., Counterman, B.A., Halder, G. *et al.* (2012) Diversification of complex butterfly wing patterns by repeated regulatory evolution of a *Wnt* ligand. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 12632–12637.

Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F. *et al.* (2013) Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, **23**, 1817–1828.

Martin, A., McCulloch, K., Patel, N., Briscoe, A., Gilbert, L. & Reed, R. (2014) Multiple recent co-options of *optix* associated with novel traits in adaptive butterfly wing radiations. *EvoDevo*, **5**, 7.

Mayr, E. (1982) *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Belknap Press, Cambridge, Massachusetts.

McVicker, G., Gordon, D., Davis, C. & Green, P. (2009) Widespread genomic signatures of natural selection in hominid evolution. *PLoS Genetics*, **5**, e1000471.

Merlin, C., Beaver, L., Taylor, O., Wolfe, S. & Reppert, S. (2012) Efficient targeted mutagenesis in the monarch butterfly using zinc-finger nucleases. *Genome Research*, **23**, 159–168.

Merrill, R.M., Wallbank, R.W.R., Bull, V., Salazar, P.C.A., Mallet, J., Stevens, M. & Jiggins, C.D. (2012) Disruptive ecological selection on a mating cue. *Proceedings of the Royal Society B: Biological Sciences*, **22**, 4907–4913.

Midamegbe, A., Vitalis, R., Malausa, T., Delava, E., Cros-Arteil, S. & Streiff, R. (2011) Scanning the European corn borer (*Ostrinia* spp.) genome for adaptive divergence between host-affiliated sibling species. *Molecular Ecology*, **20**, 1414–1430.

Miller, C.T., Beleza, S., Pollen, A.A., Schluter, D., Kittles, R.A., Shriver, M.D. & Kingsley, D.M. (2007) *cis*-regulatory changes in *kit* ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell*, **131**, 1179–1189.

Minder, A.M. & Widmer, A. (2008) A population genomic analysis of species boundaries: neutral processes, adaptive divergence and introgression between two hybridizing plant species. *Molecular Ecology*, **17**, 1552–1563.

Mokry, F., Higa, R., de Alvarenga Mudadu, M., Oliveira de Lima, A., Meirelles, S.L., Barbosa da Silva, M.V. *et al.* (2013) Genome-wide association study for backfat thickness in Canchim beef cattle using Random Forest approach. *BMC Genetics*, **14**, 47.

Mundy, N.I. (2005) A window on the genetics of evolution: MC1R and plumage colouration in birds. *Proc Biol Sci.*, **272**, 1633–1640.

Mundy, N.I. (2009) Conservation and convergence of colour genetics: *MC1R* mutations in brown cavefish. *PLoS Genetics*, **5**, e1000388.

Mundy, N.I., Badcock, N.S., Hart, T., Scribner, K., Janssen, K. & Nadeau, N.J. (2004) Conserved genetic basis of a quantitative plumage trait involved in mate choice. *Science*, **303**, 1870–1873.

Nachman, M.W., Hoekstra, H.E. & D'Agostino, S.L. (2003) The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 5268–5273.

Nadeau, N.J. & Jiggins, C.D. (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics*, **26**, 484–492.

Nadeau, N.J., Martin, S.H., Kozak, K.M., Salazar, C., Dasmahapatra, K.K., Davey, J.W. *et al.* (2013) Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Molecular Ecology*, **22**, 814–826.

Narum, S.R. & Hess, J.E. (2011) Comparison of F_{ST} outlier tests for SNP loci under selection. *Molecular Ecology Resources*, **11**, 184–194.

Nicholson, G., Smith, A.V., Jónsson, F., Gústafsson, Ó., Stefánsson, K. & Donnelly, P. (2002) Assessing population differentiation and isolation from single-nucleotide polymorphism data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, 695–715.

Nielsen, R., Paul, J.S., Albrechtsen, A. & Song, Y.S. (2011) Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, **12**, 443–451.

Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y. & Wang, J. (2012) SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. *PLoS ONE*, **7**, e37558.

Noor, M.A.F. & Bennett, S.M. (2010) Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity*, **103**, 439–444.

Nunes, V.L., Beaumont, M.A., Butlin, R.K. & Paulo, O.S. (2012) Challenges and pitfalls in the characterization of anonymous outlier AFLP markers in non-model species: lessons from an ocellated lizard genome scan. *Heredity*, **109**, 340–348.

Oldham, M.C., Horvath, S. & Geschwind, D.H. (2006) Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 17973–17978.

Olson-Manning, C.F., Wagner, M.R. & Mitchell-Olds, T. (2012) Adaptive evolution: evaluating empirical support for theoretical predictions. *Nature Reviews Genetics*, **13**, 867–877.

Orr, H.A. (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, **6**, 119–127.

Osanai-Futahashi, M., Ohde, T., Hirata, J., Uchino, K., Futahashi, R., Tamura, T., Niimi, T. & Sezutsu, H. (2012) A visible dominant marker for insect transgenesis. *Nature Communications*, **3**, 1295.

Page, R., Voss, S., Samuels, A., Smith, J., Putta, S. & Beachy, C. (2008) Effect of thyroid hormone concentration on the transcriptional response underlying induced metamorphosis in the Mexican axolotl (*Ambystoma*). *BMC Genomics*, **9**, 1–17.

Page, R.B., Monaghan, J.R., Walker, J.A. & Voss, S.R. (2009) A model of transcriptional and morphological changes during thyroid hormone-induced metamorphosis of the axolotl. *General and Comparative Endocrinology*, **162**, 219–232.

Page, R., Boley, M., Smith, J., Putta, S. & Voss, S. (2010) Microarray analysis of a salamander hopeful monster reveals transcriptional signatures of paedomorphic brain development. *BMC Evolutionary Biology*, **10**, 199.

Pan, Q., Hu, T., Malley, J.D., Andrew, A.S., Karagas, M.R. & Moore, J.H. (2014) A system-level pathway-phenotype association analysis using synthetic feature Random Forest. *Genetic Epidemiology*, **38**, 209–219.

Papa, R., Kapan, D.D., Counterman, B.A., Maldonado, K., Lindstrom, D.P., Reed, R.D., Nijhout, H.F., Hrbek, T. & McMillan, O.W. (2013) Multi-allelic major effect genes interact with minor effect QTLs to control adaptive color pattern variation in *Heliconius erato*. *PLoS ONE*, **8**, e57033.

Parchman, T.L., Gompert, Z., Mudge, J., Schilkey, F.D., Benkman, C.W. & Buerkle, C.A. (2012) Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, **21**, 2991–3005.

Paris, M. & Despres, L. (2012) Identifying insecticide resistance genes in mosquito by combining AFLP genome scans and 454 pyrosequencing. *Molecular Ecology*, **21**, 1672–1686.

Paris, M., Boyer, S., Bonin, A., Collado, A., David, J.-P. & Despres, L. (2010) Genome scan in the mosquito *Aedes rusticus*: population structure and detection of positive selection after insecticide treatment. *Molecular Ecology*, **19**, 325–337.

Pavey, S.A., Collin, H., Nosil, P. & Rogers, S.M. (2010) The role of gene expression in ecological speciation. *Annals of the New York Academy of Sciences*, **1206**, 110–129.

Pavey, S.A., Bernatchez, L., Aubin-Horth, N. & Landry, C.R. (2012) What is needed for next-generation ecological and evolutionary genomics? *Trends in Ecology & Evolution*, **27**, 673–678.

Pavlidis, P., Jensen, J.D. & Stephan, W. (2010) Searching for footprints of positive selection in whole-genome SNP data from nonequilibrium populations. *Genetics*, **185**, 907–922.

Pavlidis, P., Jensen, J.D., Stephan, W. & Stamatakis, A. (2012) A critical assessment of story-telling: Gene Ontology categories and the importance of validating genomic scans. *Molecular Biology and Evolution*, **29**, 3237–3248.

Pavlidis, P., ZivkoviC, D., Stamatakis, A. & Alachiotis, N. (2013) SweeD: likelihood-based detection of selective sweeps in thousands of genomes. *Molecular Biology and Evolution*, **30**, 2224–2234.

Pespeni, M.H., Sanford, E., Gaylord, B., Hill, T.M., Hosfelt, J.D., Jaris, H.K. *et al.* (2013) Evolutionary change during experimental ocean acidification. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 6937–6942.

Pfender, W.F., Saha, M.C., Johnson, E.A. & Slabaugh, M.B. (2011) Mapping with RAD (restriction-site associated DNA) markers to rapidly identify QTL for stem rust resistance in Lolium perenne. *Theoretical and Applied Genetics*, **122**, 1467–1480.

Pool, J.E., Hellmann, I., Jensen, J.D. & Nielsen, R. (2010) Population genetic inference from genomic sequence variation. *Genome Research*, **20**, 291–300.

Porto-Neto, L., Kijas, J. & Reverter, A. (2014) The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genetics Selection Evolution*, **46**, 22.

Pritchard, J.K., Pickrell, J.K. & Coop, G. (2010) The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, **20**, R208–R215.

Protas, M.E., Hersey, C., Kochanek, D., Zhou, Y., Wilkens, H., Jeffery, W.R., Zon, L.I., Borowsky, R. & Tabin, C.J. (2006) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics*, **38**, 107–111.

Prud'homme, B., Gompel, N., Rokas, A., Kassner, V.A., Williams, T.M., Yeh, S.-D., True, J.R. & Carroll, S.B. (2006) Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. *Nature*, **440**, 1050–1053.

Rank, N. (1992) A hierarchical analysis of genetic differentiation in a montane leaf beetle *Chrysomela aeneicollis* (coleoptera: chrysomelidae). *Evolution*, **46**, 1097–1111.

Rausher, M.D. (2006) The evolution of flavonoids and their genes. *The Science of Flavonoids* (ed. E. Grotewold), pp. 175–211. Springer, New York.

Rebeiz, M., Pool, J.E., Kassner, V.A., Aquadro, C.F. & Carroll, S.B. (2009) Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science*, **326**, 1663–1667.

Reed, R.D., Papa, R., Martin, A., Hines, H.M., Counterman, B.A., Pardo-Diaz, C. *et al.* (2011) *Optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science*, **333**, 1137–1141.

Renaut, S., Nolte, A.W., Rogers, S.M., Derome, N. & Bernatchez, L. (2011) SNP signatures of selection on standing genetic variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (*Coregonus* spp.). *Molecular Ecology*, **20**, 545–559.

Reznick, D.N., Shaw, F.H., Rodd, F.H. & Shaw, R.G. (1997) Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science*, **275**, 1934–1937.

Roberts, R.B., Ser, J.R. & Kocher, T.D. (2009) Sexual conflict resolved by invasion of a novel sex determiner in lake malawi cichlid fishes. *Science*, **326**, 998–1001.

Rockman, M.V. (2011) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution*, **66**, 1–17.

Rockman, M.V. & Kruglyak, L. (2006) Genetics of global gene expression. *Nature Reviews Genetics*, **7**, 862–872.

Roda, F., Liu, H., Wilkinson, M.J., Walter, G.M., James, M.E., Bernal, D.M. *et al.* (2013) Convergence and divergence during the adaptation to similar environments by an Australian groundsel. *Evolution*, **67**, 2515–2529.

Rogers, S.M. & Bernatchez, L. (2005) Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeaformis*). *Molecular Ecology*, **14**, 351–361.

Rokyta, D.R., Joyce, P., Caudle, S.B. & Wichman, H.A. (2005) An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. *Nature Genetics*, **37**, 441–444.

Rosenblum, E., Hoekstra, H. & Nachman, M. (2004) Adaptive reptile color variation and the evolution of the Mc1r gene. *Evolution*, **58**, 1794–1808.

Salomé, P.A., Bomblies, K., Laitinen, R.A.E., Yant, L., Mott, R. & Weigel, D. (2011) Genetic architecture of flowering-time variation in *Arabidopsis thaliana*. *Genetics*, **188**, 421–433.

Savolainen, O., Lascoux, M. & Merila, J. (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, **14**, 807–820.

Scheinfeldt, L.B. & Tishkoff, S.A. (2013) Recent human adaptation: genomic approaches, interpretation and insights. *Nature Reviews Genetics*, **14**, 692–702.

Schlötterer, C. (2003) Hitchhiking mapping - functional genomics from the population genetics perspective. *Trends in Genetics*, **19**, 32–38.

Schoville, S., Barreto, F., Moy, G., Wolff, A. & Burton, R. (2012) Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. *BMC Evolutionary Biology*, **12**, 170.

Schulte, P.M., Glémet, H.C., Fiebig, A.A. & Powers, D.A. (2000) Adaptive variation in lactate dehydrogenase-B gene expression: role of a stress-responsive regulatory element. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 6597–6602.

Shapiro, M.D., Bell, M.A. & Kingsley, D.M. (2006) Parallel genetic origins of pelvic reduction in vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 13753–13758.

Shapiro, M.D., Marks, M.E., Peichel, C.L., Blackman, B.K., Nereng, K.S., Jónsson, B., Schluter, D. & Kingsley, D.M. (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, **428**, 717–723.

Shimizu, K. & Purugganan, M. (2005) Evolutionary and ecological genomics of *Arabidopsis*. *Plant Physiology*, **138**, 578–584.

Slate, J. (2013) From beavis to beak color: a simulation study to examine how much QTL mapping can reveal about the genetic architecture of quantitative traits. *Evolution*, **67**, 1251–1262.

Slate, J., Visscher, P.M., MacGregor, S., Stevens, D., Tate, M.L. & Pemberton, J.M. (2002) A genome scan for quantitative trait loci in a wild population of red deer (*Cervus elaphus*). *Genetics*, **162**, 1863–1873.

Slate, J., Santure, A.W., Feulner, P.G.D., Brown, E.A., Ball, A.D., Johnston, S.E. & Gratten, J. (2010) Genome mapping in intensively studied wild vertebrate populations. *Trends in Genetics*, **26**, 275–284.

Stapley, J., Reger, J., Feulner, P.G.D., Smadja, C., Galindo, J., Ekblom, R. *et al.* (2010) Adaptation genomics: the next generation. *Trends in Ecology & Evolution*, **25**, 705–712.

Stern, D.L. & Orgogozo, V. (2008) The loci of evolution: how predictable is genetic evolution? *Evolution*, **62**, 2155–2177.

Stern, D.L. & Orgogozo, V. (2009) Is genetic evolution predictable? *Science*, **323**, 746–751.

Stinchcombe, J.R. & Hoekstra, H.E. (2007) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.

Storz, J.F. (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.

Stranger, B.E., Stahl, E.A. & Raj, T. (2011) Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*, **187**, 367–383.

Streelman, J., Albertson, R. & Kocher, T. (2003) Genome mapping of the orange blotch colour pattern in cichlid fishes. *Molecular Ecology*, **12**, 2465–2471.

Streisfeld, M.A. & Rausher, M.D. (2009) Altered *trans*-regulatory control of gene expression in multiple anthocyanin genes contributes to adaptive flower color evolution in *Mimulus aurantiacus*. *Molecular Biology and Evolution*, **26**, 433–444.

Streisfeld, M.A., Young, W.N. & Sobel, J.M. (2013) Divergent selection drives genetic differentiation in an R2R3-MYB transcription factor that contributes to incipient speciation in *Mimulus aurantiacus*. *PLoS Genetics*, **9**, e1003385.

Sugawara, T., Terai, Y., Imai, H., Turner, G.F., Koblmüller, S., Sturmbauer, C., Shichida, Y. & Okada, N. (2005) Parallelism of amino acid changes at the RH1 affecting spectral sensitivity among deep-water cichlids from Lakes Tanganyika and Malawi. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 5448–5453.

Suh, Y. & Vijg, J. (2005) SNP discovery in associating genetic variation with human disease phenotypes. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **573**, 41–53.

Terai, Y., Seehausen, O., Sasaki, T., Takahashi, K., Mizoiri, S., Sugawara, T. *et al.* (2006) Divergent selection on opsins drives incipient speciation in lake Victoria cichlids. *PLoS Biology*, **4**, e433.

Terenius, O., Papanicolaou, A., Garbutt, J.S., Eleftherianos, I., Huvenne, H., Kanginakudru, S. *et al.* (2011) RNA interference in Lepidoptera: an overview of successful and unsuccessful studies and implications for experimental design. *Journal of Insect Physiology*, **57**, 231–245.

Tornero, P., Chao, R., Luthin, W., Goff, S. & Dangl, J. (2002) Large-scale structure-function analysis of the *Arabidopsis* RPM1 disease resistance protein. *Plant Cell*, **14**, 435–450.

de Torres, M., Sanchez, P., Fernandez-Delmond, I. & Grant, M. (2003) Expression profiling of the host response to bacterial infection: the transition from basal to induced defence responses in *RPM1*-mediated resistance. *Plant J*, **33**, 665–676.

Tsai, I.J., Zarowiecki, M., Holroyd, N., Garciarrubio, A., Sanchez-Flores, A., Brooks, K.L. *et al.* (2013) The genomes of four tapeworm species reveal adaptations to parasitism. *Nature*, **496**, 57–63.

Turchin, M.C., Chiang, C.W.K., Palmer, C.D., Sankararaman, S., Reich, D. & Hirschhorn, J.N. (2012) Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nature Genetics*, **44**, 1015–1019.

Turner, T.L. (2014) Fine-mapping natural alleles: quantitative complementation to the rescue. *Molecular Ecology*, **23**, 2377–2382.

Turner, T.L. & Hahn, M.W. (2010) Genomic islands of speciation or genomic islands and speciation? *Molecular Ecology*, **19**, 848–850.

Turner, T.L., Hahn, M.W. & Nuzhdin, S.V. (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, **3**, e285.

Uy, J., Moyle, R., Filardi, C. & Cheviron, Z. (2009) Difference in plumage color used in species recognition between incipient species is linked to a single amino acid substitution in the *melanocortin-1 receptor*. *The American Naturalist*, **174**, 244–254.

Villar, E., Klopp, C., Noirot, C., Novaes, E., Kirst, M., Plomion, C. & Gion, J. (2011) RNA-Seq reveals genotype-specific molecular responses to water deficit in eucalyptus. *BMC Genomics*, **12**, 538.

de Villemereuil, P., Frichot, É., Bazin, É., François, O. & Gaggiotti, O.E. (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

Vitalis, R., Dawson, K., Boursot, P. & Belkhir, K. (2003) DetSel 1.0: a computer program to detect markers responding to selection. *Journal of Heredity*, **94**, 429–431.

Voss, S.R. & Shaffer, H.B. (1997) Adaptive evolution via a major gene effect: paedomorphosis in the Mexican axolotl. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 14185–14189.

Wang, Z., Gerstein, M. & Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, **10**, 57–63.

Weber, J.N., Peterson, B.K. & Hoekstra, H.E. (2013) Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice. *Nature*, **493**, 402–405.

Wei, C., Liu, J., Yu, Z., Zhang, B., Gao, G. & Jiao, R. (2013) TALEN or Cas9 - Rapid, efficient and specific choices for genome modifications. *Journal of Genetics and Genomics*, **40**, 281–289.

Weinig, C., Dorn, L.A., Kane, N.C., German, Z.M., Halldorsdottir, S.S., Ungerer, M.C. *et al.* (2003) Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*. *Genetics*, **165**, 321–329.

Wessinger, C.A. & Rausher, M.D. (2012) Lessons from flower colour evolution on targets of selection. *Journal of Experimental Botany*, **63**, 5741–5749.

Westram, A.M., Galindo, J., Alm Rosenblad, M., Grahame, J.W., Panova, M. & Butlin, R.K. (2014) Do the same genes underlie parallel phenotypic divergence in different *Littorina saxatilis* populations? *Molecular Ecology*, **23**, 4603–4616.

Whiteman, D. & Agrawal, A. (2009) What is phenotypic plasticity and why is it important? *Phenotypic Plasticity of Insects: Mechanisms and Consequences* (ed. D.W.W.T.N. Ananthakrishnan), pp. 1–63. CRC Press, Boca Raton.

Wittkopp, P.J. (2011) Using pyrosequencing to measure allele-specific mRNA abundance and infer the effects of cis- and trans-regulatory differences. *Molecular Methods for Evolutionary Genetics* (eds V. Orgogozo & M.V. Rockman), pp. 297–317. Humana Press, New York.

Wood, H.M., Grahame, J.W., Humphray, S., Rogers, J. & Butlin, R.K. (2008) Sequence differentiation in regions identified by a genome scan for local adaptation. *Molecular Ecology*, **17**, 3123–3135.

Wray, G.A. (2013) Genomics and the evolution of phenotypic traits. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 51–72.

Yang, W. & Charles Gu, C. (2014) Random forest fishing: a novel approach to identifying organic group of risk factors in genome-wide association studies. *European Journal of Human Genetics*, **22**, 254–259.

Yang, W., Qi, Y., Bi, K. & Fu, J. (2012) Toward understanding the genetic basis of adaptation to high-elevation life in poikilothermic species: a comparative transcriptomic analysis of two ranid frogs, *Rana chensinensis* and *R. kukunoris*. *BMC Genomics*, **13**, 588.

Zhan, X., Pan, S., Wang, J., Dixon, A., He, J., Muller, M.G. *et al.* (2013) Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. *Nature Genetics*, **45**, 563–566.

Zhang, G., Cowled, C., Shi, Z., Huang, Z., Bishop-Lilly, K.A., Fang, X. *et al.* (2013) Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science*, **339**, 456–460.

Zuellig, M.P., Kenney, A.M. & Sweigart, A.L. (2014) Evolutionary genetics of plant adaptation: insights from new model systems. *Current Opinion in Plant Biology*, **18**, 44–50.

Zufall, R.A. & Rausher, M.D. (2003) The genetic basis of a flower color polymorphism in the common morning glory (*Ipomoea purpurea*). *Journal of Heredity*, **94**, 442–448.

Zufall, R.A. & Rausher, M.D. (2004) Genetic changes associated with floral adaptation restrict future evolutionary potential. *Nature*, **428**, 847–850.