Review

CellPress

The genomics of organismal diversification illuminated by adaptive radiations

Daniel Berner and Walter Salzburger

Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland

Adaptive radiation is the rapid and extensive ecological diversification of an organismal lineage to generate both phenotypic disparity (divergence) and similarity (convergence). Demonstrating particularly clear evidence of the power of natural selection, adaptive radiations serve as outstanding systems for studying the mechanisms of evolution. We review how the first wave of genomic investigation across major archetypal adaptive radiations has started to shed light on the molecular basis of adaptive diversification. Notably, these efforts have not yet identified consistent features of genomic architecture that promote diversification. However, access to a pool of ancient adaptive variation via genetic exchange emerges as an important driver of adaptive radiation. We conclude by highlighting avenues for future research on adaptive radiations, including the discovery of 'adaptation genes' based on genome scans using replicate convergent populations.

Outbursts of life

> In adaptive radiation and in every part of the whole, wonderful history of life, all the modes and all the factors of evolution are inextricably interwoven. The total process cannot be made simple, but it can be analyzed in part. (G.G. Simpson, 1953 [1])

The history of life is a cumulative story of evolutionary radiations (see Glossary); such outbursts of unprecedented forms are thought to be responsible for much of the extant and extinct organismal diversity on Earth [1–4]. For example, all cellular life has radiated out of such a simple – by way of comparison – beginning as the primordial self-replicating molecules [5]; all major animal phyla were in place by the fading of the Cambrian evolutionary radiation (commonly known as 'Cambrian explosion') [1,5]; and about every 20th extant fish species is the product of the ongoing explosive radiations of cichlids in the East African Great Lakes that, together, cover as little as ~0.031% of the total and ~0.044% of the water surface of our planet [6]. Evolutionary radiations are termed adaptive radiations (ARs) if new lifeforms emerge rapidly through the extensive ecological diversification of an organismal lineage [2,3]. Archetypal examples of ARs include Darwin's finches on the Galapagos archipelago [7]; spiders [8] and silversword plants [9] on Hawaii; anole lizards on the islands of the Caribbean [10]; threespine stickleback fish in north temperate waters [11]; and the above-mentioned cichlid fishes in the East Africa Great Lakes and in various tropical crater lakes [12]. ARs have inspired generations of naturalists and have for decades served as essential model systems in evolutionary biology. With the recent proliferation of new tools for genomic investigation and their application to ARs, an overview of the opportunities and challenges that ARs provide in understanding the genomic basis of organismal diversification seems timely. We open this review by discussing the major features of ARs, describe the biological questions that genomic investigation

Glossary

Divergence mapping: marker-based search for genome regions exhibiting exceptionally strong differentiation (typically quantified by the fixation index F_{ST}) between ecologically-different populations as a consequence of divergent natural selection on specific loci.

dN/dS ratio: index combining the frequency of non-synonymous (i.e., amino acid changing) and synonymous (i.e., not amino acid changing) nucleotide substitutions across a gene to explore the type of selection that has acted on the gene.

Ecological niche: sum of all environmental conditions (e.g., resources, antagonists) relevant to the lifestyle of an organism.

Ecological opportunity: availability of novel or under-exploited ecological niches, resulting from the colonization of a novel habitat, the extinction of antagonists, or evolutionary innovation.

Ecomorph: distinct phenotype tightly associated with a specific ecological niche.

Evolutionary radiation: increase in taxonomic diversity; in other words, in the number of phenotypically distinct populations or species, within an organismal lineage.

Genome architecture: here this term is used to subsume global characteristics of a genome, such as overall genome size, gene counts, or the number of transposable elements.

Haplotype: stretch of DNA on a chromatid, inherited as a single unit.

Historical contingency: accumulation of stochastic events in the course of the evolution of a lineage.

Introgressive hybridization: hybridization between populations that causes the exchange of genetic material (gene flow) in some regions of the genome.

Lineage: evolutionary line of descendants, independent of its taxonomic status.

Mosaic genome: a genome heterogeneous in ancestry, emerging from introgressive hybridization, lateral gene transfer, or retention of ancestral polymorphism.

Syndrome selection: selection on a combination of traits (e.g., body and jaw morphology and coloration in cichlid fish [18,27]) in a given environment.

Corresponding authors: Berner, D. (daniel.berner@unibas.ch);
Salzburger, W. (walter.salzburger@unibas.ch).

Keywords: divergence; convergence; genome sequencing; gene flow; adaptation locus.

0168-9525/

http://dx.doi.org/10.1016/j.tig.2015.07.002

Box 1. Common features of adaptive radiations

Bouts of diversification commonly considered ARs have several characteristics in common:

(i) *Extensive divergent evolution.* Diversification results in the generation of phenotypic disparity [1–3].
(ii) *Divergent natural selection.* Diversification entails performance trade-offs such that the radiating taxa perform well in their specific ecological niche but relatively poorly in the niches of other members of the radiation [1,2].
(iii) *Common ancestry.* Diversification starts from a delimited founder population [2]. Because diversification occurring within a confined area is particularly likely to meet this criterion, organisms radiating on islands or within isolated water bodies have always figured among the iconic examples of AR.
(iv) *Rapid evolution.* Diversification occurs on a relatively short time scale [83,84]. Consequently, the products of AR generally still exhibit intrinsic genetic compatibility, facilitating genetic investigation (e.g., using artificial crosses).
(v) *Speciation.* Diversification is accompanied by some degree of reproductive isolation. Restricted gene flow results directly from adaptive divergence because this implies the inferiority of hybrids and migrants between ecologically specialized populations [85–87], but many other reproductive barriers may operate (e.g., sexual isolation).
(vi) *Convergent evolution.* Diversification results in similar phenotypes in multiple independent but ecologically-similar habitats [3,88]. This phenotypic similarity often comprises a diverse suite of traits as a consequence of 'syndrome selection'.
(vii) *Evolutionary (key) innovation.* The evolution within a lineage of a phenotype of particular significance to AR because it creates access to hitherto underutilized habitats, thus triggering the subsequent diversification to different ecological niches within these habitats [1,2].

Several of these features arguably contain a subjective flavor. For instance, the extent and speed of diversification varies dramatically among ARs: only 14 species of ground and tree finches have emerged in the Galápagos archipelago within hundred thousands of generations [7]. By contrast, the younger radiations of cichlids in lakes Malawi or Victoria have each produced hundreds of endemic species [6], and white fish and sticklebacks have produced countless convergent ecomorphs within a few thousand generations [11,17,89–91]. Another ambiguity concerns the position of an AR along the speciation continuum. For example, the strength of reproductive isolation within the stickleback AR is variable, with genetic differentiation between diverging populations in close contact ranging all the way from negligible to very high ([89,92], see also [64,93] for cichlids). We thus avoid drawing a line between populations, ecomorphs, or species as the outcome of AR. Finally, many aspects of AR remain untested. For instance, performance trade-offs among divergent populations within ARs have only rarely been demonstrated, and their association to specific phenotypes often remains unclear (but see [94–97]). In particular, the concept of key innovations is controversial because tests permitting such demonstration are difficult to conceive [2].

Nevertheless, the fundamental commonality and beauty of ARs is that they highlight the power of natural selection in driving diversity in ecological function among a group of closely related organisms. Emphasis should not be placed on whether or not diversifying organismal systems qualify as AR, but instead on how such systems can be exploited to illuminate the mechanisms underlying biological diversification.

in AR can address, and highlight avenues for future research.

Key elements of ARs

AR is defined as the rapid diversification of an organismal lineage into an array of closely related species as a consequence of adaptation to different ecological niches [1–3] (a more comprehensive characterization is provided in Box 1). Closer inspection of such bouts of diversification reveals two important elements. The first is the generation of unique, distinct phenotypes specialized to different ecological niches – divergent evolution. This aspect of diversification is epitomized by Darwin's finches, which have evolved distinct beak morphologies facilitating the exploitation of distinct food resources [7]. The second element of AR concerns the repeated generation of similar phenotypes in geographically independent but ecologically similar habitats – convergent evolution. Indeed, most of the well-known ARs exhibit both divergence and convergence. For instance, *Anolis* lizards on the Greater Antilles have diverged into an array of ecomorphs adapted to specific foraging niches but, across several islands, specific niches are predictably inhabited by phenotypically similar ecomorphs [13,14] (Figure 1). Similarly striking examples of convergence nested within divergence include wing-color mimetics in passion vine butterflies (*Heliconius* [15,16]), trophic ecomorphs of white fish and threespine stickleback in northern temperate freshwaters [11,17], and cichlid fishes that have evolved extraordinary similarity in color and in body and mouth shape among several East African lakes [18], and even within single lakes [19].

The occurrence of extensive divergence and convergence in ARs offers rich opportunities for genomic explorations of two important issues in evolution. The first concerns why phenotypic divergence is vast and rapid in some lineages but not in others. Can the study of AR identify genomic or population genetic determinants of diversification rates? Second, convergent evolution represents outstanding evidence of the deterministic action of natural selection [20,21]. Molecular studies of ARs may therefore provide particularly strong insights into the genomic underpinnings and consequences of adaptation. In the following we elaborate on these two major themes.

Genomic determinants of AR

It has long been recognized that the question why some groups of organisms diversify more extensively than others is tightly connected to the pivotal concept of ecological opportunity: ARs are triggered by access to novel or hitherto under-utilized ecological niches [1,2,12], sometimes in combination with sexual selection [22]. The answer to the question about the determinants of organismal diversification is thus partly ecological and also includes historical contingency [14,23], but will remain incomplete without information from a genomic angle because exploiting ecological opportunity requires that genomes can relatively rapidly generate phenotypic modifications and innovations.

Early on, the profuse diversity displayed by some ARs already spurred speculations that the genomes of the emerging species may exhibit unusual features conducive to rapid phenotypic divergence [24–26]. The rise of next-generation sequencing technologies further fueled the expectation that such peculiarities in the genomes of radiating clades can be identified, if they exist [20,27]. To date, the quest for a genomic basis of adaptive radiation is still in its infancy. One reason is that typically only a single

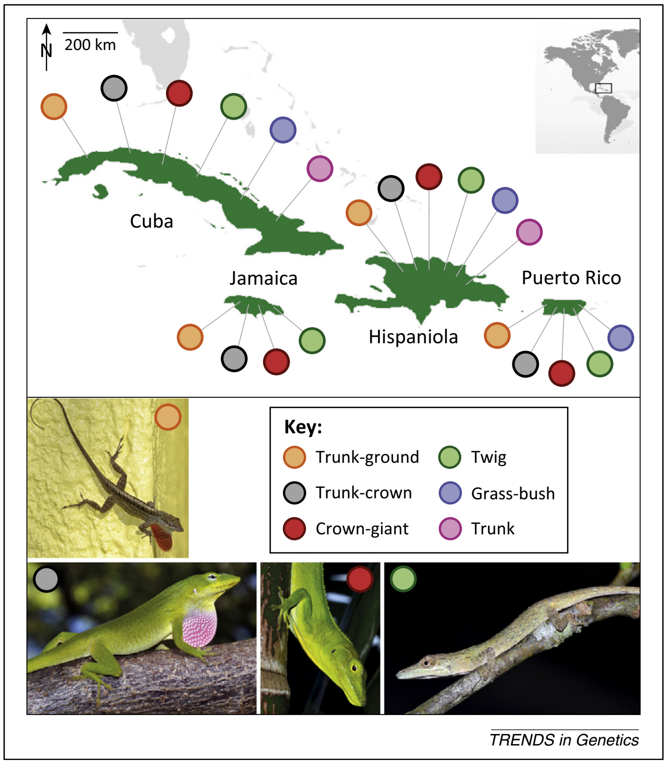


Figure 1. Divergent and convergent evolution in an adaptive radiation. On the Greater Antilles, *Anolis* lizards have diversified into four to six distinct ecomorphs (each represented by a different color) specialized for well-defined foraging niches. The outcome is phenotypic disparity within each island, and convergent evolution among islands. The photographs on the bottom illustrate the four *Anolis* ecomorphs consistently occurring on all islands. All photographs were kindly provided by Jonathan Losos.

genome or a few representative genomes within a radiation have been sequenced in sufficient quality; taxonomic outgroups providing an appropriate baseline for comparative analysis are often lacking. Nevertheless, we here summarize insights that have emerged from the first round of sequencing of 'radiating genomes' [28], emphasizing two aspects of particular significance: genome architecture and gene flow.

Do elements of genome architecture promote AR?

One possibility is that the rapid radiation of some lineages is facilitated by specific aspects of genome architecture. Indeed, variation in structural features has been identified in the genomes of the green anole [29], threespine sticklebacks [30], Darwin's finches [31,32], cichlids [33], and *Heliconius* butterflies [34] – either by comparing several genomes within the AR or by comparison to some other genome(s) outside the AR (Table 1). For instance, the genomes of the green anole and the cichlids, which belong to taxonomically highly-diverse ARs, reveal a relative expansion of mobile genetic elements presumably related to diversification in gene expression [29,33] (note that also the *Heliconius* genome shows a higher content of DNA transposons compared to the silkworm [35]). Analyses of the genomes of stickleback [30] and cichlids [33,36] further suggest a predominant role of regulatory evolution. Finally, in sticklebacks, structural variation is abundant [37], and adaptive divergence between ecomorphs has been shown to involve chromosomal inversions [30] – a feature not yet studied in detail in other ARs. Although these results are interesting in their own right, they have been collected piecemeal and lack a formal comparative framework, and it therefore remains unclear if and how they are related to the rate of diversification. For instance, regulatory changes predominate generally in adaptive diversification between closely related taxa [38–40], not only within some ARs, and examples for a role of inversions in adaptation are accumulating in many non-radiating organismal systems as well [41–43].

So far, the most explicit attempt to compare genomic features of members of ARs against a less diversified outgroup has been made in cichlids [33]. This analysis was based on a total of five genomes, comprising a representative of a more ancestral and 'non-radiating' lineage (the Nile tilapia, *Oreochromis niloticus*) on the one hand, and four species representing the ARs in East African lakes Victoria (*Pundamilia nyererei*), Malawi (*Metriaclima zebra*) and Tanganyika (*Neolamprologus brichardi* and *Astatotilapia burtoni*, which also occurs in rivers draining into that lake) on the other hand. When compared to tilapia and other teleosts, the four East African cichlid genomes showed a higher number of gene duplications (see also [44]), an elevated genome-wide non-synonymous to synonymous (dN/dS ratio), and diversification in gene regulation by novel microRNAs in addition to the above-mentioned

Table 1. Conspicuous features in the genomes of adaptively-radiating clades

Adaptive radiation	Number of genomes	Reference genome	Genomic features						
			Gene duplications	Mobile elements	Regulatory changes	Accelerated coding evolution	Novel miRNAs	Structural changes (inversions)	Hybridization/ introgression/ILS[a]
Anolis lizards	1	[29]	n/r[b]	**Yes**[c]	Suggested	n/r	n/r	n/r	n/a[d]
East African cichlid fish	5 (>100)[e]	[33]	**Yes**[f]	**Yes**[c,f]	**Yes**[g]	**Yes**[f]	**Yes**[c]	n/r	**Yes**[g]
Darwin's finches	1 (>100)[e]	[31,32]	n/r	n/r	n/r	n/r	n/r	n/r	**Yes**[g]
Heliconius butterflies	1 (>100)[e]	[34]	n/r	**Yes**[c]	Suggested	n/r	**Yes**[c]	n/r	**Yes**[g]
Stickleback fish	1 (>100)[e]	[30]	n/r	n/r	**Yes**[g]	n/r	n/r	**Yes**[g]	**Yes**[g]

[a] Incomplete lineage sorting leading to the retention of ancient polymorphisms.

[b] Not reported (for genome-wide comparisons).

[c] In a comparison with (an) unrelated and hence genetically distinct outgroup(s).

[d] Because only a single genome is available, no such analysis is possible to date.

[e] The number in brackets indicates the number of re-sequenced genomes in addition to the reference genome.

[f] In a comparison with a closely related but 'non-radiating' outgroup.

[g] From comparisons among the members of ARs.

abundance of divergence in non-coding elements and in transposable element insertions [33]. Taken together, these findings would suggest that the genomes of rapidly-diversifying cichlid lineages are more dynamic overall. A caveat to this conclusion is that the alleged 'non-radiating' Nile tilapia belongs to a clade still containing at least 60 species – an admittedly small number in the collective of cichlids, but a massive number compared to other vertebrate ARs. Furthermore, this clade includes about a dozen species that radiated rapidly within the tiny crater lake Barombi Mbo in Cameroon [45,46], highlighting that this clade also has high diversification potential.

Overall, the somewhat disappointing – although perhaps not unexpected – insight from the first generation of genome sequencing in ARs is that specific genome-wide patterns in genome architecture causally related to rapid and extensive diversification are lacking. However, further genomic analyses explicitly tailored to comparing the genomic architecture of radiating versus closely related but non-radiating taxa are clearly needed before more definitive conclusions can be drawn. An obvious opportunity for such comparative work is provided by Darwin's finches, where the genomes of all species within the AR plus outgroup taxa have recently been sequenced [31]. It is well possible, however, that increased efforts along this line of investigation will still fail to uncover general genomic features, and that the determinants of organismal diversification need to be explored otherwise.

The pivotal role of ancient variation and gene flow

Traditionally, AR is conceived as a process in which a small founder population invades a new habitat such as an island with abundant resources, followed by incipient diversification into so far underutilized ecological niches (note that, from the point of view of a fish, a lake – or a river system for that sake – are what an island is for terrestrial taxa). Consequently, one would expect that both colonization and radiation must occur in the face of relatively modest genetic variation, with founder events and strong drift as important processes [1,24]. This view is being refuted by the emerging genome-wide sequence data: populations arising during ARs are genetically fairly diverse, and the genomes of these species contain adaptive allelic variants that originated long before the actual species or populations have formed [31,34,47–50]. In addition, it turns out that polymorphisms shared between the genomes of radiating clades are highly abundant. For example, nearly 40% of all single-nucleotide polymorphisms (SNPs) are shared among ground and tree finches on the Galapagos Islands [31]. Similarly, among 180 SNPs ascertained in Lake Malawi cichlids, more than 50% also exhibit polymorphisms in cichlid lineages that radiated elsewhere in East Africa [47]; and a genome-wide phylogenetic analysis focusing on three members of the most species-rich African cichlid lineage, the haplochromines, revealed that more than 40% of all SNPs support genealogies that are in conflict with the species tree [33]. In *Heliconius*, finally, large fractions of the genome are shared between co-occurring species but not between geographically separated species [51].

The emergence of such mosaic genomes during ARs can be attributed to two processes: incomplete lineage sorting of ancestral polymorphisms, and gene flow among radiating populations. In Darwin's finches, introgressive hybridization appears to be common and is the major factor determining genomic diversity. Moreover, ancient alleles influencing beak morphology have been exchanged between species by gene flow [31], a pattern also seen in *Heliconius* [34,52–54]. Likewise, in sticklebacks, a pool of ancient genetic variation is maintained in the ancestral marine population by gene flow across the marine–freshwater habitat boundary, allowing repeated 'recycling' of adaptive variation during the rapid diversification in freshwater habitats [30,50,55,56]. Similarly, in cichlids, new taxa have been suggested to emerge via hybridization [57–59], and even the exchange of genetic material across otherwise independent (and geographically isolated) species flocks via riverine populations appears to be common [47,60].

Collectively, a pivotal message emerging from recent genomic analyses of ARs is that the presence of extensive ancient genetic variation, and the diffusion of this variation through gene flow, represent an important determinant of diversification; rapidly- and extensively-diversifying lineages seem to be those having access to a pool of alleles useful for the adaptation to novel ecological niches. Gene flow within ARs also provides a powerful explanation for the widespread occurrence of convergent evolution: if adaptive genetic variants disperse easily, it becomes likely that the repeated adaptation to a specific ecological niche will be based on shared genetic variants and the associated phenotypes [28,47]. The availability of such genetic variants will pre-empt adaptation via alternative genetic changes, and thus represents a form of genetic bias.

The genetic basis of phenotypic divergence in ARs

Ever since G.G. Simpson [1] emphasized the importance of 'distinctive new adaptive types' in triggering adaptive radiation, the identification of such evolutionary (key) innovations has been a major focus of adaptive radiation research. With progress in molecular genetic tools, emphasis has partly shifted to investigations of the molecular basis of adaptive divergence – an endeavor of course not restricted to ARs. The promise is to gain a general understanding of several longstanding issues in evolution, such as the number of genes involved in adaptive divergence and how they are arranged in the genome; the examination of fitness consequences and contribution to reproductive isolation of adaptive variants; and the reconstruction of the history of these variants and what type of specific mutations they represent. Given that in ARs trait–environment correlations often allow a phenotype to be connected to an ecological function [2], ARs offer particularly rich opportunities to study the genetics of the phenotypes underlying adaptation and reproductive isolation. In this section we highlight some progress – and limitations – in uncovering the genetic basis of phenotypes important to AR, and emphasize how convergent evolution in particular can provide a powerful resource for the detection of adaptation genes.

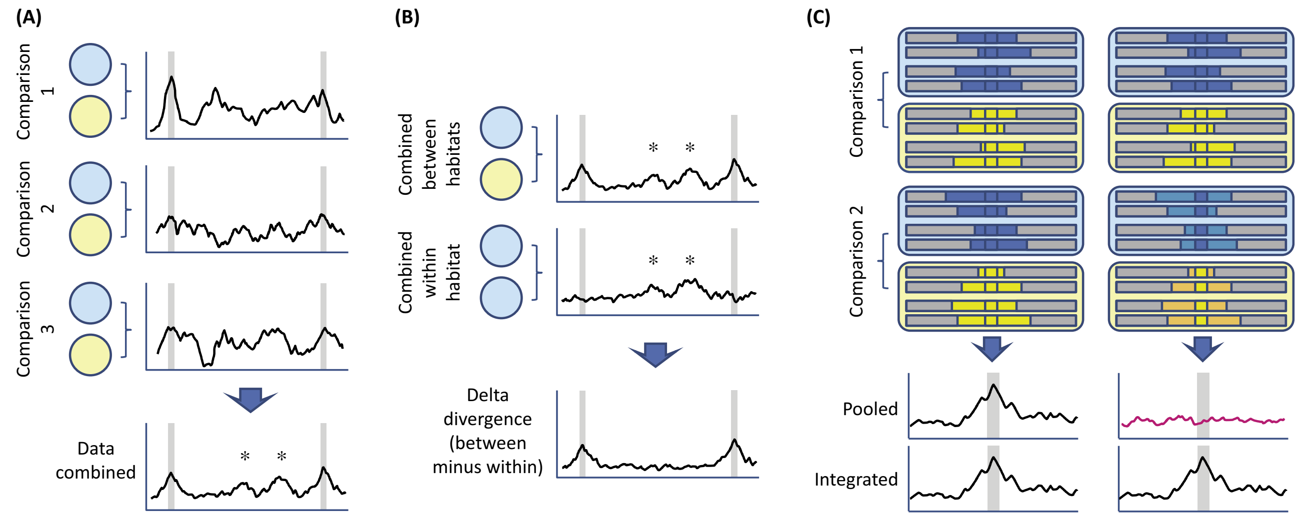
Finding the loci of adaptive radiation – examples and limitations

Driven by combinations of genetic techniques such as genetic mapping, comparative gene expression surveys, and genome scans, information about the genetic basis of phenotypic diversification in AR is accumulating. Well-known examples are the *Eda* and *Pitx1* loci in stickleback. Allelic variation at these genes is responsible for much of the adaptive divergence in defensive armor among stickleback ecomorphs [50,61–63]. Another example is an opsin gene important to vision in Lake Victoria cichlids; divergence at this gene concurrently promotes adaptation to different light environments and drives sexually-based reproductive isolation [64,65]. The evolution of cichlid fish egg-spots, an evolutionary innovation in the form of ovoid color patterns on the male anal fins of haplochromines related to mating, was connected to a transposable element insertion in the *cis*-regulatory region of the pigmentation gene *fhl2b* [66]. Moreover, in both stickleback and cichlids, genes involved in the diversification of foraging structures (teeth, gill rakers, oral and pharyngeal jaws) are being uncovered [62,67–69]. In Darwin's finches, the *ALX1* locus [31] is involved in beak shape differences, and in *Heliconius* butterflies, diversification in mimetic wing patterns is based in part on the *optix* gene [70].

For three reasons, however, enthusiasm about progress in identifying the loci underlying AR should be tempered. First, all the above examples include morphological traits relatively easily amenable to genetic dissection. By contrast, the genetic basis of diversification in physiological, life-history, or behavioral trait complexes important to ARs remains essentially unexplored. Second, genetic factors subjected to in-depth molecular study are generally chosen because they have a large phenotypic effect. This

Box 2. Combining divergence mapping in replicated habitat comparisons to discover adaptation loci

The reliable identification of genomic loci that have been sorted by divergent natural selection between populations occupying ecologically-different habitats is challenging. As a hypothetical example, the upper three panels in Figure IA represent separate genome scans for divergence in populations from two different habitat types (blue and yellow). The *x* axis represents the genomic position along a chromosome, the *y* axis gives the magnitude of population divergence (e.g., as expressed by F_{ST}) at genetic markers. Profiles for each of the three population comparisons reveal heterogeneity in divergence, driven by a mix of divergent selection associated with the focal habitat contrast (adaptation loci are indicated by grey vertical bars), selection unrelated to the ecological contrast of interest, and purely stochastic processes. By combining data from the replicate comparisons, adaptation loci consistently involved in divergence can be detected with higher confidence (Figure IA, bottom profile). However, genome regions influenced by habitat-unrelated selective processes (e.g., the combination of locally reduced recombination and background selection [81]) might also display elevated divergence (asterisks). Confounding signatures of this latter process can be reduced by standardizing combined comparisons between habitats by combined comparisons within a habitat (e.g., by the subtraction of divergence, yielding 'Delta divergence' *sensu* [56]) (Figure IB). Combining divergence data can be achieved in two ways: first, by pooling genomic data from multiple replicate populations within each habitat type before divergence mapping (e.g., [30]). Second, by performing genome scans for each replicate comparison separately and then integrating the divergence data across the scans [56]. Depending on the origin and age of the selected alleles, these two approaches may yield different results. For instance, if in each population pair the alleles beneficial in the contrasting habitats are identical by descent and share extensive haplotype tracts, adaptation loci will be captured both when pooling or integrating data, even at reduced marker resolution (Figure IC, left column: the horizontal bars represent a chromosome, the selected locus is in the center, haplotype tracts specific to the two alleles are shown in dark blue and bright yellow). If, by contrast, extensive recombination has dissociated each allele from its characteristic haplotype background (Figure IC, right column: different haplotype tracts shown in different colors), divergence mapping based on pooled (but not integrated) data will miss the signature of selection (pink profile) unless marker resolution is complete and tags the causative polymorphism. If adaptive alleles arose repeatedly through independent mutations (i.e., the alleles are not identical by descent among replicate habitats of the same type), even full marker resolution might not reveal a strong signal when pooling data. In any case, complementing divergence mapping with methods for detecting selective sweeps based on haplotype structure might be valuable. An example of an investigation integrating multiple divergence scans is given in Figure 2.



TRENDS in Genetics

Figure I. (A) Combining marker data from replicate convergent populations in different habitats to discover genomic regions involved in adaptive diversification. (B) Using different ecological contrast types to reduce confounding genomic signatures. (C) Two different approaches to combining marker data from replicate convergent populations.

potentially limits our ability to draw general conclusions about the genetic architecture of diversification [71]; much of the genetic variation underlying diversification during AR might not be detectable in the first place because of small phenotypic effect size [71–73]. Third, and most importantly, genome scans in ARs designed to capture molecular signatures of adaptive divergence at the whole-organism performance level – that is, beyond specific traits – reveal a daunting complexity underlying adaptive divergence. For example, *Eda* appears to be only one among dozens to hundreds of adaptation genes involved in marine–freshwater divergence in stickleback [30,56,74], beak shape in Darwin's finches is likely driven by several other genetic factors [31,75,76], and even the major color loci in *Heliconius* seem to represent composite elements combining multiple coding and regulatory variants at several individual genes [54,77,78].

Efforts to elucidate the genetic basis of adaptive divergence in AR thus suffer from the standard challenges and biases seen in other organismal systems, although the presence of extensive convergent evolution in many ARs provides a unique methodological advantage discussed in the following section. With these limitations in mind, we anticipate that ARs will continue to provide fascinating case-studies on the genetic basis of adaptation that will complement (but not replace) traditional evolutionary investigations.

Convergence as a resource for the discovery of adaptation loci

Discovering loci important to adaptive diversification is increasingly based on the application of genomic tools. One popular strategy to identify such regions is to use large genetic marker datasets to scan the genomes of populations adapted to different environments in search of genomic regions that display exceptionally high population divergence (i.e., divergence mapping [79]). Regions exhibiting such a signature are interpreted to reflect localized allele-frequency changes driven by divergent selection between habitats. However, a notorious challenge in this approach is to isolate true signatures of selection away from background heterogeneity in genomic divergence caused by purely stochastic processes. Moreover, genomic regions displaying high divergence between populations can also arise from natural selection unrelated to the ecological difference between the focal habitats [56,80–82].

Performing population genomic analyses in ARs featuring extensive convergent evolution can alleviate these difficulties. Specifically, if marker data are available for multiple replicate populations nested within specific habitat types, divergence mapping based on a combination of these data can greatly increase the power to detect adaptation loci (see Figure IA in Box 2). The reason is that repeated selection on a gene during convergent evolution will produce a consistent signature of elevated divergence between replicate populations from ecologically-different habitats. Furthermore, adjusting divergence data from population comparisons between habitats for the magnitude of divergence seen in comparisons within habitats can help to distinguish signatures of habitat-related divergent selection from those driven by selection unrelated to the focal ecological contrast (see Figure IB in Box 2). The combination of marker information from multiple convergent populations thus increases the signal to noise ratio in genomic population scans. Although still rarely applied, this approach has identified dozens of candidate loci involved in the adaptation across the marine to freshwater habitat boundary in threespine stickleback (Figure 2) [30,56]. Similarly, divergence mapping based on populations grouped by phenotype (as an alternative to habitat) has discovered genome regions underlying the AR of Darwin's finches, including genes influencing beak shape [31].

Exploring the phylogeny of adaptation loci identified by divergence mapping can further illuminate the evolutionary history of the selected alleles within ARs. For example, around loci recruited repeatedly for marine versus freshwater stickleback divergence, populations from the two habitats show exclusive ancestry [50,56] (Figure 2, bottom), indicating that the evolution of convergent populations has used shared variants embedded in common

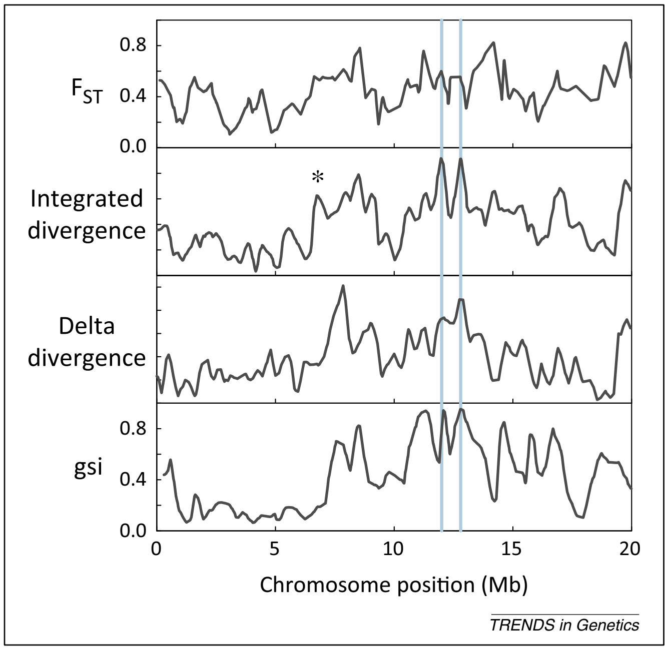


Figure 2. Example of a search for adaptation loci based on the combination of genomic data from populations showing convergent evolution. The empirical system involves multiple populations of threespine stickleback that have repeatedly adapted to freshwater habitats from a marine ancestor, and that were genotyped genome-wide by using a reduced representation approach (RADseq) [56]. The first (from top) panel displays a divergence (F_{ST}) profile for a segment of chromosome 4, resulting from a single comparison between a freshwater population and an ancestral marine sample. Divergence along the chromosome is heterogeneous, but two genes known to be involved in adaptive divergence between marine and freshwater populations (*Abcb7* and *Eda* [30]; positions indicated by blue vertical lines) show only weak divergence between the habitats and would not emerge as strong candidate adaptation loci. However, when integrating 16 replicate marine–freshwater genome scans (Figures IA,C in Box 2), these loci exhibit striking divergence between the habitats (second panel). Strong divergence at *Abcb7* and *Eda* persists even when adjusting divergence between the habitats for background divergence within a habitat (i.e., 'Delta divergence' calculated as integrated marine–freshwater divergence minus integrated divergence among all freshwater populations; see Figure IB in Box 2) (third panel), confirming their pivotal role in adaptive diversification along the focal ecological axis. By contrast, a region displaying high integrated divergence (asterisk in the second panel) disappears after the above adjustment, suggesting a signature of selection unrelated to marine–freshwater ecology. Using the same set of populations, the bottom panel displays phylogenetic structure along the chromosome, as captured by the genealogical sorting index (gsi) [98]. This reveals complete reciprocal monophyly between marine and freshwater populations at the two adaptation loci, suggesting the fixation of shared haplotype tracts around alleles identical by descent within multiple convergent populations.

haplotypes that were spread by gene flow. Similar phylogenetic patterns consistent with the diffusion of adaptive variants during convergent evolution have emerged at mimicry loci in *Heliconius* [34], and at a major beak-shape gene in Darwin's finches [31].

Although divergence mapping and complementary analyses using populations that have adaptively converged towards similar habitats offers the opportunity to identify and characterize adaptation loci with high confidence, this analytical framework has its limitations. First, adaptation genes will necessarily only be identified if they were consistently involved in convergent evolution within an AR [56]; that is, signatures of selection at loci promoting adaptation to a specific habitat within a single population, but not in others, are likely to be overlooked. This renders conclusions about the number of loci underlying adaptation based on divergence mapping across multiple population comparisons problematic. Likewise, it is possible that the loci reused with high fidelity during replicate ARs are those that have a relatively large fitness consequence, which would bias conclusions about the effect size of adaptation loci. A methodological complication is that a functional variant at an adaptation locus might be embedded in different haplotypes in populations that independently adapted to the same environment. This is possible when independent mutations at a locus generated a functionally-similar variant multiple times, or when extensive recombination dissociated a single functional variant from the haplotype background in which it arose. For this reason, it matters how genomic data are combined across replicate populations (see Figure IC in Box 2). Finally, although convergent populations can help to establish a relationship between a polymorphism in a particular genomic region and ecological divergence, a general challenge remains: the identification of the actual variant under selection, and the phenotypes and ecological consequences associated with its variation.

Concluding remarks

In this review we have described the potential of molecular investigations in ARs to improve our understanding of the mechanisms underlying biological diversification. One of the major insights that has emerged is that rapid and extensive diversification and convergent evolution is facilitated by the diffusion of ancient adaptive variation, thus blurring the distinction between incomplete lineage sorting, standing genetic variation, and gene flow via introgressive hybridization. Certainly, a wider appreciation of the highly-dynamic history, demography, and connectivity of diversifying populations is needed. The question of whether aspects of genomic architecture represent important determinants of diversification, however, remains an open issue; progress will require explicit comparative analyses based on massive sequencing efforts, ideally covering entire ARs together with appropriate outgroups. Genomic investigations of AR will also continue to uncover loci underlying adaptive divergence, thereby benefiting from the increase in the detection power of genome scans conferred by replicate, convergent populations. Such studies will contribute to our understanding of how consistently adaptive variants are re-used, where they originated, what type of mutations they represent, and how they disperse – keeping in mind that resulting insights might not apply to loci underlying adaptation in general.

Another lesson from genomic studies of AR is that the depth of biological insight is crucially dependent on the availability of robust reference genomes; insufficient genomic resources still represent a major analytical obstacle in many fascinating ARs. At the same time, identifying key phenotypes in ARs, and testing their fitness consequences using traditional field and laboratory experiments, remain as important today as they were in past decades, and should keep up with our efforts to illuminate the genomic basis of adaptive diversification.

Acknowledgments

We thank the Editor R. Macrae for the invitation and for comments; C. Jiggins plus an anonymous reviewer and the members of the laboratory of W.S. for discussion and comments on the manuscript; and J. Losos for providing photographs. This work was supported, in part, from grants from the Swiss National Science Foundation (SNF) to D.B. and W.S., and from the European Research Council (ERC, CoG 'CICHLID~X') to W.S.

References

1 Simpson, G.G. (1953) *The Major Features of Evolution*, Columbia University Press
2 Schluter, D. (2000) *The Ecology of Adaptive Radiation*, Oxford University Press
3 Gavrilets, S. and Losos, J.B. (2009) Adaptive radiation: contrasting theory with data. *Science* 323, 732–737
4 Mayr, E. (2001) *What Evolution Is*, BasicBooks
5 Cowen, R. (2013) *History of Life*, Wiley-Blackwell
6 Salzburger, W. *et al.* (2014) Ecology and evolution of the African lakes and their faunas. *Annu. Rev. Ecol. Evol. Syst.* 45, 519–545
7 Grant, P.R. and Grant, B.R. (2008) *How and Why Species Multiply: The Radiations of Darwin's Finches*, Princeton University Press
8 Gillespie, R. (2004) Community assembly through adaptive radiation in Hawaiian spiders. *Science* 303, 356–359
9 Carlquist, S. *et al.* (2003) *Tarweeds and Silverswords: Evolution of the Madiinae*, Missouri Botanical Garden Press
10 Losos, J.B. (2009) *Lizards in an Evolutionary tree: Ecology and Adaptive Radiation of Anoles*, University of California Press
11 Bell, M.A. and Foster, S.A. (1994) *Evolutionary Biology of the Threespine Stickleback*, Oxford University Press
12 Fryer, G. and Iles, T.D. (1972) *The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution*, Oliver & Boyd
13 Mahler, D.L. *et al.* (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* 341, 292–295
14 Losos, J.B. *et al.* (1998) Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279, 2115–2118
15 Brown, K.S. (1981) The biology of *Heliconius* and related genera. *Annu. Rev. Entomol.* 26, 427–457
16 Turner, J.R.G. (1981) Adaptation and evolution in *Heliconius:* a defense of neo-Darwinism. *Annu. Rev. Ecol. Syst.* 12, 99–121
17 Bernatchez, L. (2004) Ecological theory of adaptive radiation: an empirical assessment from coregonine fishes (Salmoniformes). In *Evolution Illuminated: Salmon and their Relatives* (Hendry, A.P. and Stearns, S.C., eds), Oxford University Press
18 Kocher, T.D. *et al.* (1993) Similar morphologies of cichlid fish in lakes Tanganyika and Malawi are due to convergence. *Mol. Phylogenet. Evol.* 2, 158–165
19 Muschick, M. *et al.* (2012) Convergent evolution within an adaptive radiation of cichlid fishes. *Curr. Biol.* 22, 2362–2368
20 Losos, J.B. (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. American Society of Naturalists E.O. Wilson Award Address. *Am. Nat.* 175, 623–639
21 Schluter, D. and Nagel, L.M. (1995) Parallel speciation by natural selection. *Am. Nat.* 146, 292–301
22 Wagner, C.A. *et al.* (2012) Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* 487, 366–369

23 Seehausen, O. (2007) Chance, historical contingency and ecological determinism jointly determine the rate of adaptive radiation. *Heredity* 99, 361–363

24 Mayr, E. (1984) Evolution of fish species flocks: a commentary. In *Evolution of Fish Species Flocks* (Echelle, A.A. and Kornfield, I., eds), pp. 3–12, University of Maine at Orono Press

25 Stebbins, G.L. (1971) Relationships between adaptive radiation, speciation and major evolutionary trends. *Taxon* 20, 3–16

26 Salzburger, W. and Meyer, A. (2004) The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften* 91, 277–290

27 Salzburger, W. (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Mol. Ecol.* 18, 169–185

28 Jiggins, C.D. (2014) Evolutionary biology: radiating genomes. *Nature* 513, 318–319

29 Alfoldi, J. *et al.* (2011) The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* 477, 587–591

30 Jones, F.C. *et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484, 55–61

31 Lamichhaney, S. *et al.* (2015) Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518, 371–375

32 Zhang, G. *et al.* (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346, 1311–1320

33 Brawand, D. *et al.* (2014) The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513, 375–381

34 The Heliconius Genome Consortium (2012) Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487, 94–98

35 Lavoie, C.A. *et al.* (2013) Transposable element evolution in *Heliconius* suggests genome diversity within Lepidoptera. *Mobile DNA* 4, 21

36 Baldo, L. *et al.* (2011) Comparative transcriptomics of Eastern African cichlid fishes shows signs of positive selection and a large contribution of untranslated regions to genetic diversity. *Genome Biol. Evol.* 3, 443–455

37 Feulner, P.G. *et al.* (2013) Genome-wide patterns of standing genetic variation in a marine population of three-spined sticklebacks. *Mol. Ecol.* 22, 635–649

38 Wittkopp, P.J. *et al.* (2008) Regulatory changes underlying expression differences within and between *Drosophila* species. *Nat. Genet.* 40, 346–350

39 Carroll, S.B. (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134, 25–36

40 Wray, G.A. (2007) The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.* 8, 206–216

41 Lowry, D.B. and Willis, J.H. (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* 8, e1000500

42 Kunte, K. *et al.* (2014) Doublesex is a mimicry supergene. *Nature* 507, 229–232

43 Poelstra, J.W. *et al.* (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* 344, 1410–1414

44 Machado, H.E. *et al.* (2014) Gene duplication in an African cichlid adaptive radiation. *BMC Genomics* 15, 161

45 Schliewen, U.K. *et al.* (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* 368, 629–632

46 Dunz, A.R. and Schliewen, U.K. (2013) Molecular phylogeny and revised classification of the haplotilapiine cichlid fishes formerly referred to as '*Tilapia*'. *Mol. Phylogenet. Evol.* 68, 64–80

47 Loh, Y.H. *et al.* (2013) Origins of shared genetic variation in African cichlids. *Mol. Biol. Evol.* 30, 906–917

48 Elmer, K.R. *et al.* (2009) Pleistocene desiccation in East Africa bottlenecked but did not extirpate the adaptive radiation of Lake Victoria haplochromine cichlid fishes. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13404–13409

49 Berner, D. *et al.* (2010) Constraints on speciation suggested by comparing lake–stream stickleback divergence across two continents. *Mol. Ecol.* 19, 4963–4978

50 Colosimo, P.F. *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* 307, 1928–1933

51 Martin, S.H. *et al.* (2013) Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* 23, 1817–1828

52 Pardo-Diaz, C. *et al.* (2012) Adaptive introgression across species boundaries in *Heliconius* butterflies. *PLoS Genet.* 8, e1002752

53 Kronforst, M.R. *et al.* (2013) Hybridization reveals the evolving genomic architecture of speciation. *Cell Rep.* 5, 666–677

54 Supple, M.A. *et al.* (2013) Genomic architecture of adaptive color pattern divergence and convergence in *Heliconius* butterflies. *Genome Res.* 23, 1248–1257

55 Barrett, R.D. and Schluter, D. (2008) Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23, 38–44

56 Roesti, M. *et al.* (2014) The genomic signature of parallel adaptation from shared genetic variation. *Mol. Ecol.* 23, 3944–3956

57 Salzburger, W. *et al.* (2002) Speciation via introgressive hybridization in East African cichlids? *Mol. Ecol.* 11, 619–625

58 Seehausen, O. (2004) Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19, 198–207

59 Meyer, B.S. *et al.* (2015) A tribal level phylogeny of Lake Tanganyika cichlid fishes based on a genomic multi-marker approach. *Mol. Phylogenet. Evol.* 83, 56–71

60 Schwarzer, J. *et al.* (2012) Repeated trans-watershed hybridization among haplochromine cichlids (Cichlidae) was triggered by Neogene landscape evolution. *Proc. R. Soc. B.* 279, 4389–4398

61 Cresko, W.A. *et al.* (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6050–6055

62 Berner, D. *et al.* (2014) Genetic architecture of skeletal evolution in European lake and stream stickleback. *Evolution* 68, 1792–1805

63 Chan, Y.F. *et al.* (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* 327, 302–305

64 Seehausen, O. *et al.* (2008) Speciation through sensory drive in cichlid fish. *Nature* 455, 620–626

65 Terai, Y. *et al.* (2006) Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biol.* 4, e433

66 Santos, M.E. *et al.* (2014) The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. *Nat. Commun.* 5, 5149

67 Fraser, G.J. *et al.* (2009) An ancient gene network is co-opted for teeth on old and new jaws. *PLoS Biol.* 7, e31

68 Cleves, P.A. *et al.* (2014) Evolved tooth gain in sticklebacks is associated with a cis-regulatory allele of Bmp6. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13912–13917

69 Glazer, A.M. *et al.* (2015) Genome assembly improvement and mapping convergently evolved skeletal traits in sticklebacks with genotyping-by-sequencing. *G3* Published online June 3, 2015. http://dx.doi.org/10.1534/g3.115.017905

70 Reed, R.D. *et al.* (2011) Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* 333, 1137–1141

71 Rockman, M.V. (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66, 1–17

72 Tiffin, P. and Ross-Ibarra, J. (2014) Advances and limits of using population genetics to understand local adaptation. *Trends Ecol. Evol.* 29, 673–680

73 Le Corre, V. and Kremer, A. (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Mol. Ecol.* 21, 1548–1566

74 Terekhanova, N.V. *et al.* (2014) Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. *PLoS Genet.* 10, e1004696

75 Abzhanov, A. *et al.* (2006) The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* 442, 563–567

76 Abzhanov, A. *et al.* (2004) Bmp4 and morphological variation of beaks in Darwin's finches. *Science* 305, 1462–1465

77 Counterman, B.A. *et al.* (2010) Genomic hotspots for adaptation: the population genetics of Mullerian mimicry in *Heliconius erato*. *PLoS Genet.* 6, e1000796

78 Joron, M. *et al.* (2011) Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* 477, 203–206

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

80 Roesti, M. *et al.* (2012) Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. *Mol. Ecol.* 21, 2852–2862

81 Charlesworth, B. *et al.* (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* 70, 155–174
82 Slatkin, M. and Wiehe, T. (1998) Genetic hitch-hiking in a subdivided population. *Genet. Res.* 71, 155–160
83 Gavrilets, S. and Vose, A. (2005) Dynamic patterns of adaptive radiation. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18040–18045
84 Verheyen, E. *et al.* (2003) Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. *Science* 300, 325–329
85 Nosil, P. (2012) *Ecological Speciation*, Oxford University Press
86 Rundle, H.D. and Nosil, P. (2005) Ecological speciation. *Ecol. Lett.* 8, 336–352
87 Hendry, A.P. (2004) Selection against migrants contributes to the rapid evolution of ecologically-dependent reproductive isolation. *Evol. Ecol. Res.* 6, 1219–1236
88 Rundle, H.D. *et al.* (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science* 287, 306–308
89 Berner, D. *et al.* (2009) Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution* 63, 1740–1753
90 Schluter, D. and McPhail, J.D. (1992) Ecological character displacement and speciation in sticklebacks. *Am. Nat.* 140, 85–108
91 Vonlanthen, P. *et al.* (2012) Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature* 482, 357–362
92 Hendry, A.P. *et al.* (2009) Along the speciation continuum in sticklebacks. *J. Fish. Biol.* 75, 2000–2036
93 Theis, A. *et al.* (2014) Adaptive divergence between lake and stream populations of an East African cichlid fish. *Mol. Ecol.* 23, 5304–5322
94 Schluter, D. (1995) Adaptive radiation in sticklebacks: trade-offs in feeding performance and growth. *Ecology* 76, 82–90
95 Losos, J.B. (1990) The evolution of form and function: morphology and locomotor performance in West Indian *Anolis* lizards. *Evolution* 44, 1189–1203
96 Fulton, M. and Hodges, S.A. (1999) Floral isolation between *Aquilegia formosa* and *Aquilegia pubescens*. *Proc. R. Soc. B.* 266, 2247–2252
97 Meyer, A. (1989) Cost of morphological specialization: feeding performance of the two morphs in the trophically polymorphic cichlid fish, *Cichlasoma citrinellum*. *Oecologia* 80, 431–436
98 Cummings, M.P. *et al.* (2008) A genealogical approach to quantifying lineage divergence. *Evolution* 62, 2411–2422