Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Review

Adaptation in the age of ecological genomics: insights from parallelism and convergence

Kathryn R. Elmer and Axel Meyer

Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, Universitätstraβe 10, 78457 Konstanz, Germany

Parallel phenotypic diversification in closely related species is a rigorous framework for testing the role of natural selection in evolution. Do parallel phenotypes always diversify by parallel genetic bases or does selection pave many alternative genomic routes to the same phenotypic ends? In this review, we show that the advent of next-generation sequencing technologies and the growing use of genomic approaches make it increasingly feasible to answer these fundamental questions using ecological and evolutionary 'non-model' populations of vertebrates in nature. While it is generally expected, and often observed, that closely related populations or species have parallel genetic bases to parallel phenotypes, exceptions are not rare and show that alternative genetic routes can result in similar phenotypes. Ultimately, this framework may illuminate the ecological conditions, evolutionary histories and genetic architectures that result in recurrent phenotypes and rapid adaptation.

Parallelism as a framework for studying adaptive evolution

Since the Modern Evolutionary Synthesis, ecologists and evolutionary biologists have sought to understand the relative contribution of stochastic genetic processes versus natural selection in the genesis and diversification of species. Though more than 150 years have passed since the publication of *On the Origin of Species*, the importance of natural selection as a primary force, particularly in speciation, remains poorly understood (e.g. [1–3]). One reason for this continued challenge is that all species and situations are somewhat unique with respect to their environmental and evolutionary history. Thus, the noise of local effects and stochasticity can overwhelm signals. In such a situation, the best way to increase statistical power in evolutionary biology research is to study evolution in replicate and in a well-defined phylogenetic context by the comparative method [4]. That is, to focus on cases of parallel phenotypic evolution and use this as a framework to infer the genetic bases of adaptive traits (Box 1).

Famous examples of parallel phenotypic evolution in vertebrates include adaptive radiations such as Caribbean anolis lizards [5], Darwin's finches in the Galapagos [6], and cichlid fishes in African Rift [7] or Central American crater [8] lakes. Less species rich but very illustrative examples used recently for examining the genetics of adaptation include temperate freshwater fishes such as sticklebacks [9] and whitefishes [10], or replicate surface-cave populations of Mexican cave fishes [11] (Figure 1). These complexes of young sister taxa are distributed in comparable but physically isolated environments such as oceanic islands and lakes, which are the most robust geographical systems in which to study parallel evolution [4,12]. Second, the shared recent evolutionary history of these complexes means that genetic backgrounds and developmental constraints are likely to be similar between taxa, making them ideal for genetic comparisons. Third, by focusing on such recent divergence one may observe 'polymorphism in action'; that is, assess the relative importance of standing genetic variation and new mutations in the diversification process without the signal being muted by evolutionary time.

Recently, modern genomic and transcriptomic technologies have advanced to the degree that these evolutionary and ecological 'non-traditional' model systems of parallel phenotypic evolution can now be investigated at the genomic level, or can be 'genomicised' [13]. These technological advances promise an upsurge in our understanding of the genetic basis of adaptive traits.

Excitement about the genomic basis of parallel evolution and replicate adaptive radiations is palpable in many sub-disciplines of evolutionary biology [14–17]. One reason for this is that parallelism suggests that certain key elements of evolution – such as antagonistic selection pressures, common standing genetic variation, or constraints on genetic architecture – might be rate-limiting to adaptation. How 'free' are taxa (genetically-speaking) to solve their ecological problems through phenotypic innovation? Are the phenotypic potentials of evolution predetermined by genetic possibilities? It is obvious that in no case does natural selection write on a blank slate, but instead evolution proceeds piecemeal with a limited genetic repertoire that is contingent on evolutionary history and developmental constraint. Further, the parallel evolution framework promises to illuminate the genetic basis of (often rapidly

Corresponding author: Meyer, A. (axel.meyer@uni-konstanz.de)

 0169-5347/$ – see front matter © 2011 Elsevier Ltd. All rights reserved. doi:10.1016/j.tree.2011.02.008 Trends in Ecology and Evolution, June 2011, Vol. 26, No. 6

Author's personal copy

Box 1. Parallel evolution, convergent evolution: history and controversy

Generally, if similar phenotypes evolve independently in closely related lineages, then this has often been called PARALLEL EVOLUTION (e.g. stickleback ecomorphs). Convergence on similar phenotypes from divergent lineages has been called CONVERGENT EVOLUTION (e.g. echolocation in bats and dolphins), the distinction being that since their common ancestor other phenotypes or genotypes have been visited extensively en route to the contemporary one. An obvious limitation of this division is that there is no definitive cut-off between what are considered closely versus distantly related species.

A second layer of distinction has been made based on the molecular basis of the similar phenotype. Some argue that when the same developmental mechanisms, genetic pathways, genes or even nucleotide mutations are involved in the parallel phenotype then this is PARALLEL EVOLUTION. CONVERGENT EVOLUTION instead recruits novel genomic regions for the same phenotypic end (reviewed in [30,70]). A limitation of this division is that many genetic levels (e.g. mutation, gene, functional change and pathway) could be considered 'parallel' and that 'similarity' at the genetic level is vague.

It has been proposed that this distinction between parallelism and convergence, based either on evolutionary divergence or molecular genetics, is not warranted and that 'convergent evolution' is sufficient for all cases [71]. Others argue that the distinction should remain, because deep evolutionary convergence occurs by different molecular routes than 'parallel' evolution [72]. Much of biology is a grey zone and there is always a border to be disputed. This disagreement over the definition (or misunderstanding) of parallelism versus convergence is more than a hundred years old [30] and unlikely to end any time soon.

Despite its short-comings, we prefer to maintain the term 'parallel evolution' when speaking about closely related lineages because the distinction between parallel and convergent (i.e. many different phenotypes visited during its evolutionary history) is useful for formulating hypotheses and experimental approaches. 'Phenotypic' and 'genetic' are key adjectives that should be bound to 'parallelism' or 'convergence'; if they are used then much of the confusion disappears. More specifically, we advocate that, in the genomic age, it is worth explicating distinction between homologous at the same nucleotide, homologous genes, or non-homologous (mutation in different genes) evolution at the basis of parallel phenotypes (Figure 2).

evolved) adaptive traits more generally (reviewed in [18]). If a pattern emerges that parallel genetic evolution underlies parallel phenotypic evolution in adaptive radiations, then genetic parallelism could be a key component to the speed of speciation. Ideas abound but empirical studies are still sparse, especially in wild populations of vertebrates, and no consensus has emerged of how these factors operate in the context of parallel evolution [19–23].

Here we review recent advances in understanding the genetics of adaptation by natural selection within the framework of parallel evolution (see Box 1 for our definition of 'parallel evolution'). We ask, do parallel phenotypes commonly have parallel genetic bases? That is, to what extent are parallel phenotypes associated with structural or regulatory variation in homologous genes, or even homologous mutations, in different populations and species? While this is certainly a simplification of the link between genotype and phenotype, the overall point is to recognise the diversity of ways genomes evolve parallel phenotypes. Ultimately, the genetic bases of parallel phenotypes tell us something about the independence of evolution under natural selection. Modern genomic techniques allow us to address these old evolutionary questions in new ways.

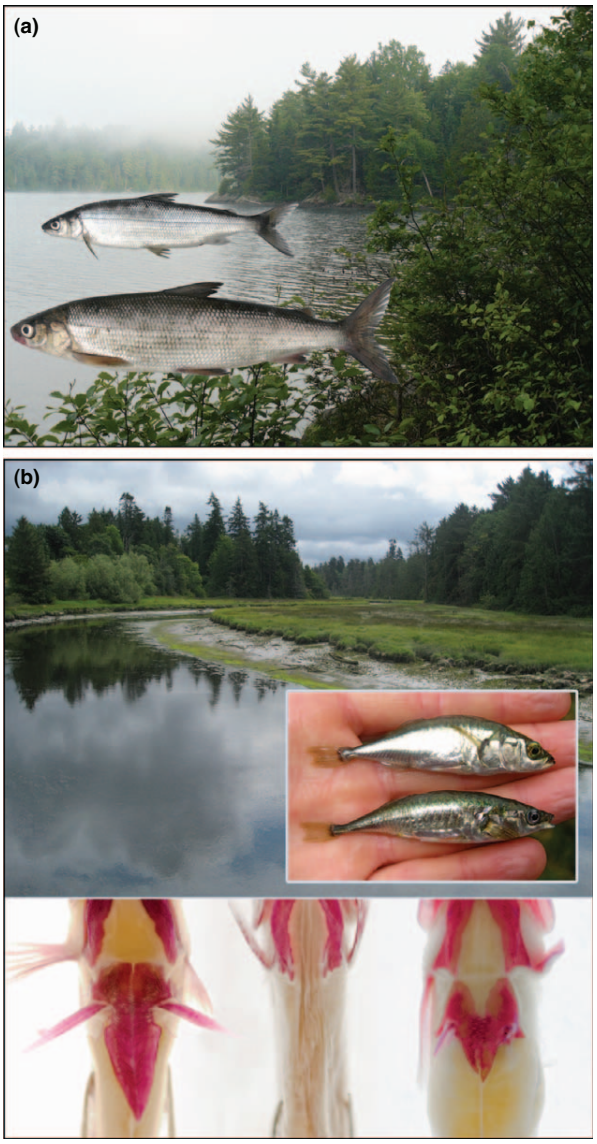

Figure 1. (a) Whitefishes (*Coregonus clupeaformis*) have repeatedly evolved 'dwarf' (above) and 'normal' (below) ecologically relevant parallel phenotypes in freshwater lakes colonised since the last glaciation (see typical habitat image behind). See Box 4 for more details about the genetic bases of these phenotypes. **(b)** As threespine stickleback fishes (*Gasterosteus aculeatus*) repeatedly leave the marine environment to invade freshwater habitats such as streams and lakes (see example habitat image), phenotypes with reduced armoured plating (top fish-in-hand) relative to ancestral plating (lower fish-in-hand) evolve repeatedly and in parallel. Freshwater sticklebacks also lose their pelvic armour with freshwater colonisation. The lower photograph shows a ventral view of the pelvis (stained for visualisation) of a marine stickleback with robust pelvic armour (left), an extreme pelvic reduction (middle), and a transgenic stickleback individual (right) with rescued pelvic armour, which confirmed that mutations at *Pitx1* were the genetic basis of this recurrent phenotype [49]. See Box 3 for more details.
Whitefish images courtesy of L. Bernatchez. Stickleback habitat and in-hand images courtesy of R. Barrett. Pelvis images courtesy of Y.F. Chan.

We focus on vertebrate 'non-model' systems of closely related lineages that have been primarily studied via populations in natural settings.

The genetic origin of parallel phenotypes

Parallel phenotypes could derive fundamentally from variability in at least three different basic patterns of molecular evolution across taxa: (i) the same mutation in the same

Author's personal copy

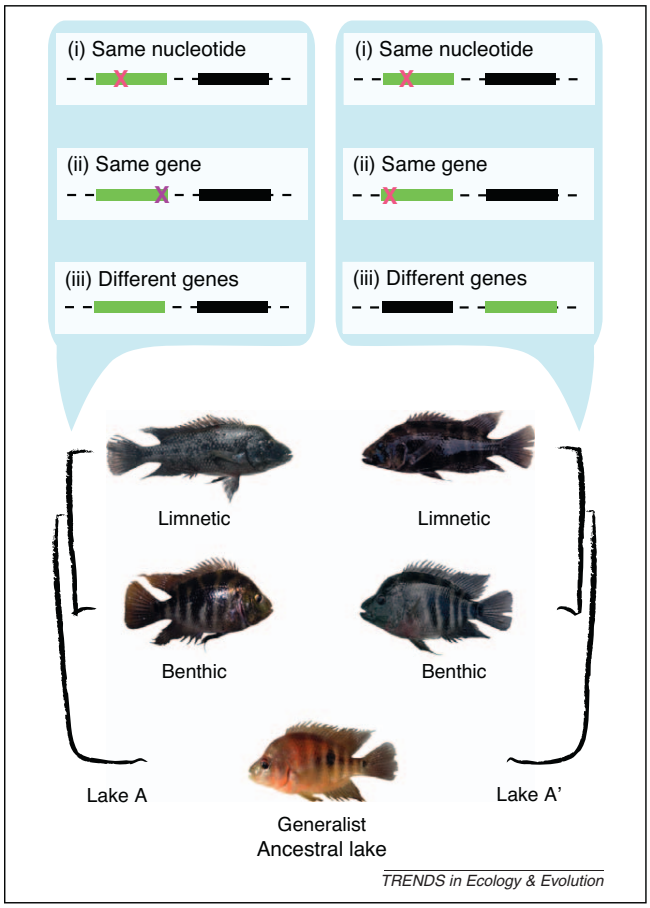

Figure 2. A schematic representation of the different genetic bases to parallel phenotypes, as discussed in the text. In this example, two lakes with similar environments (Lake A and Lake A') are independently colonised by a generalist fish (Ancestral Lake). Within the derived lake populations, the fishes speciate into parallel elongate and high-bodied phenotypes occupying limnetic and benthic niches. The genetic basis of the parallel limnetic phenotypes may be **(i)** mutations at the same nucleotide in the same gene, **(ii)** different mutations in the same gene, or **(iii)** mutations in different genes. Coloured Xs represent mutations and the gene of interest is coloured green. In all cases, the relevant mutational events could be *de novo* or drawn from standing genetic variation. This is a dramatic simplification, and mutation here includes any structural or regulatory change.

nucleotide in the same genes (homologous change); (ii) different mutations in the same gene (homologous loci); or (iii) mutations in different gene (non-homologous loci) that nonetheless result in parallel phenotypes (Figure 2). Any of these changes could be *de novo* in derived populations or the product of standing genetic variation. It is clear that various types of mutations are relevant in evolution, so we are not feeding the fire of the dichotomous structural versus regulatory variation debate (e.g. [24,25]). For the purposes of our discussion, variation can include structural variation (e.g. amino acid substitutions and gene duplications), regulatory changes outside the coding sequence (e.g. [25]) or a combination of these [20,26].

The most extreme form of genetic parallelism across lineages – the same mutation in the same nucleotide in the same gene – could be due to repeated independent mutations, but among closely related species it more likely indicates shared ancestral genetic variation. When new environments with similar selective pressures are repeatedly colonised from the same large ancestral population, it is possible that the colonising phenotypes under selection already hold the genetic variation necessary to independently evolve similar phenotypes in parallel. This evolution from standing genetic variation should be the fastest mode of adaptation, since no *de novo* mutation is required and the genetic variation should already be compatible with the background genome [27,28]. But how commonly are parallel phenotypes derived from the same standing genetic variation? At what evolutionary distance between taxa does parallel phenotypic evolution draw from standing genetic variation compared to requiring *de novo* mutation?

Even in Petri dish microbial systems [15] the relative importance of standing genetic variation versus *de novo* mutation in generating adaptive divergence remains very little studied (but see [29] for a vertebrate example). The reasons for this are broadly two-fold. First, implicating adaptive alleles initially is a major challenge requiring a combination of genetic, phenotypic and developmental approaches. Second, standing genetic variation is most reliably distinguished from *de novo* mutation by thorough sampling of the ancestral population (though inferring a selective sweep or reconstructing the allelic phylogeny are alternative approaches [27]). Until recently, this sort of extensive genotyping was almost impossible to accomplish in non-model organisms. As we discuss below, recent advances in genomic methodologies are finally helping us address this question.

Some authors argue there should be a time dependency to the genetic basis of parallel evolution. That is, variation in homologous loci and even homologous sites would be the most prevalent basis for parallel phenotypic diversification among sister taxa because of compatibility with shared genetic backgrounds (or even evidence for standing ancestral variation), while causal mutations in non-homologous loci are expected to be the slowest, most constrained mode and tend to involve evolutionarily distant species pairs (e.g. [16,25,30,31]). This underlies the blurry distinction between parallelism and convergence (Box 1) and its relevance to basic patterns of evolutionary flexibility. As we will discuss here, this role of evolutionary distance in parallelism may be intuitive but is not completely substantiated by the data.

Methods in parallel evolution of genomes

The parallel evolution framework brings increased rigour to identifying the role of natural selection in diversification and speciation. In particular, ecological genomic and population genomic techniques, validated with functional assays and genetic mapping, can elucidate the genetic basis of adaptive traits in non-model systems. Methodological possibilities have been dealt with in other recent reviews (e.g. [13,32–35]) and will not be covered here in detail. However, we do wish to outline briefly some of the most fruitful avenues for linking parallel phenotypic evolution with parallel genome evolution in non-model species, especially using 'next-generation' technologies. In order to identify the homologous and non-homologous variation that we know is important in natural populations, it will be imperative to move beyond approaches that limit exploration to laboratory populations and *a priori* genetic expectations. In the genomics era, we can embrace and describe the genetic variation found across populations in nature.

Author's personal copy

The relevant genetic variation must be characterised and localised in the genome in order to determine the genetic basis of adaptive traits. Most directly, establishing mapping crosses allows one to identify parts of the genome important in various phenotypes, both in number of loci and dominance [36,37]. Genetic mapping is the most direct way to link phenotype and genotype but is generally only possible with organisms that can be bred in the lab. However, the development of high-throughput genomics methods using numerous polymorphic markers, either anonymous or mapped, has made association mapping of phenotypes in wild populations possible [38]. Alternatively, genome scans for loci under divergent selection can prove a relatively fast way to assess the role of homologous versus non-homologous genomic changes across species or populations (reviewed in [39]). This method seeks to identify signals of selection on genomic architecture, evolution, and differentiation, rather than linking phenotype and genotype *per se*. Recently developed next-generation approaches such as sequencing restriction site associated DNA (sRAD) [40,41] or low-coverage whole genome comparisons (e.g. [17]) offer the potential to characterise thousands of loci for hundreds of individuals rapidly and relatively inexpensively. This increasing ability to seek and differentiate neutral and adaptive genes makes this an 'area to watch' [42] for major advances in coming years.

It will only be a short time until whole genome analyses of variation will be common-place for ecological and evolutionary non-model systems. However, at present other approaches to reducing the scope and complexity of genomic variation analysed are a logical and economical option. For example, one can focus on expressed and coding variation via expressed sequence tags (ESTs) using next-generation sequencing [43]. Such approaches have been recently applied to non-model organisms in the context of metapopulations [44], sympatric speciation by host shift [45], and sympatric ecological speciation [46]. In taxa lacking a published whole-genome sequence, the assembly and/or mapping of shotgun next-generation sequence data remains non-trivial, although certainly the accumulation of whole genomes from various lineages [47] will rapidly diminish this challenge. Gene expression analyses can also be conducted using next-generation sequencing (i.e. digital gene expression or RNA-Seq [48]). In the absence of mapping crosses to localise the causal gene, these methods all remain a correlation between the phenotype of interest and the putative genetic bases. Yet the fast rate of data accumulation means we are on the cusp of dissecting the genetic basis of adaptive phenotypes in many non-model systems.

For more low-throughput genetic methodologies, the candidate gene approach is a popular method to infer parallel genetic evolution in coding sequence or expression levels in genes of interest. This is particularly true for loci of major effect, traits with a well-elucidated genetic basis in some model vertebrates, or genes with tight correspondence to a particular physiology or phenotype (e.g. opsin genes). Particularly when hypotheses can be tested with evolutionary replicates, such as in adaptive radiations or parallel evolution (Box 2), candidate gene approaches are a powerful and accessible first-step method that can be applied to non-model organisms in nature and those which lack genomic data. By focusing on a single locus, the candidate gene approach is necessarily infused with an *a priori* assumption of parallelism. Consequently, the role of non-homologous loci can only be inferred from an absence of evidence that the homologous locus is involved. This ascertainment bias probably makes genetic parallelism, either at homologous sites or homologous genes, seem more common than it actually is. Further, there is bias in which candidate genes are focused on in the first place, and probably a reporting bias because candidate gene studies with no mutational parallelism (at the same site or the same gene) are less likely to be published. Casting the net wider to assess more of the genome, or the entire genome, using the recently developed approaches discussed above increases our ability to identify the many genetic bases to phenotypic parallelisms.

Box 2. Candidate genes and adaptive radiations of cichlid fishes

Cichlid fishes are renowned for having evolved replicate adaptive radiations (reviewed in [73]), yet the genetic bases of these parallelisms remain almost completely undetermined. The most thoroughly investigated genotype–phenotype association in cichlids is vision genes; i.e. how individuals receive the signal of their spectacularly diverse colourations and patterns. Parallel adaptation to divergent water depth and clarity environments based on mutation and expression in opsin genes has demonstrated repeated molecular evolution by divergent selection in Lake Victoria [74,75]. Entire species flocks in Lakes Tanganyika and Malawi have approximately parallel opsin gene expression profiles yet lack a parallel pattern of coding sequence variation [76]. Evolution of vision genes in a different lineage of fish, the sand goby, shows some, though not complete, parallel *de novo* coding region mutation in rhodoposin in parallel photic environments [77]. Thus, a great evolutionary lability has been identified in vision-related genes in fishes, with parallel and non-parallel structural and regulatory variation within and among species and within and among genera [78–81].

Colouration is a target of natural and sexual selection and therefore, like vision, has also been investigated with candidate gene approaches. For example, yellow egg spots on the pelvic and anal fins, thought to mimic eggs and stimulate fertilisation, have evolved independently in two lineages of African cichlids: the haplochromines and the ectodines. The xanthophore-related colour gene *colony-stimulating factor 1 receptor a* (*csf1ra*) is expressed in the egg spots of both lineages [82], suggesting a correlation between this gene and the egg spot phenotype. At least in the haplochromines, the molecular basis of the egg spots is possibly derived from a *de novo* substitution in the ligand-binding portion of *csf1ra* in the ancestral lineage [82]. Species in the neotropical Midas cichlid complex mate assortatively based on a shared intraspecific colour dimorphism (gold and dark) [83]. Increased expression of the pigmentation gene *Mc1r* was found in the gold morph but no segregating coding region polymorphism was identified across species [84], so it remains undetermined whether this replicate dimorphism is due to repeated *de novo* evolution in different lakes or a single ancestral polymorphism. Therefore, also in colouration, candidate gene approaches have identified that homologous mutations (either ancestral or *de novo*) or expression levels are associated with parallel phenotypes, but there are frequent exceptions within and among lineages.

The combined approach shows that Nature selects on many variants

Much of the data accumulated so far argues that when a broad molecular and geographical approach is used, the same mutations in the same gene, different mutations in

Author's personal copy

Box 3. Stickleback case study: parallel phenotype evolution in threespine sticklebacks has variable genomic bases

Threespine stickleback (*Gasterosteus aculeatus*) was among the first fish species to be whole-genome sequenced (http://www.ensembl.org/Gasterosteus_aculeatus/Info/Index) and – coupled with efforts from ecology, population genetics and developmental biology – these fishes have become a spearhead for research on the genetics of adaptation.

Sticklebacks are so-named because of their dorsal and pelvic spines. Freshwater populations have reduced or absent pelvic spines relative to marine populations (Figure 1). From mapping crosses, the gene *Pituitary homeobox 1* (*Pitx1*), was implicated in pelvic reduction in populations of threespine sticklebacks from across the globe ([85,86] and earlier work reviewed in [9]). A recent study [49] confirmed that, as previously suspected [87], *cis*-regulatory mutations upstream of *Pitx1* were the key to pelvic reduction. This regulatory mutation was located in a fragile genomic region subject to frequent independent deletions that disrupt an enhancer and cancel *Pitx1* expression in the pelvis [49]. This deletion was found to be of variable size (i.e. independent events) and important in 9 of 14 independent populations of fishes with a reduced pelvis phenotype. However, the *Pitx1* deletion is not *always* the origin of pelvic reduction [49] (i.e. non-homologous but to date undetermined mutational events are also important). Multiple genetic mechanisms account for pelvic reduction in ninespine sticklebacks as well [87].

Sticklebacks are unusual fishes in that they lack scales and instead have armour plating. From a highly-plated marine ancestor, in numerous freshwater environments armour plating is reduced or lost repeatedly and independently after colonisation. Allelic variation at a single locus was implicated in the parallel evolution of reduced armour plating [88]. Low plate *Ectodysplasin-A* (*Eda*) alleles affect phenotypes in various ways, including behavioural differences [89], faster growth [29,90], earlier reproduction and improved over-wintering survival in freshwater [29,91]. Experiments suggest there is strong positive selection that promotes the repeated evolution of the low-plated morph in freshwater [29] (driven in part by insect predation [92]). However, selection on *Eda* was not constant throughout the year; some early increases in low plate allele frequency later declined [29] and it might be that genes linked to *Eda* are also important [93]. Most freshwater populations have low-plate *Eda* alleles that share a common ancestry, arguing for a strong role of repeated, independent positive selection in freshwater environments from standing genetic variation [29,88,94]. However, one low-plated population in Japan did not share the mutations at *Eda* [88], suggesting that there are exceptions to the parallel genetic basis of low plated morphs.

Therefore, there is a combination of recruiting standing genetic variation, different mutations at the homologous gene, and the involvement of different genes in the repeated parallel phenotypes of stickleback adaptations to freshwater.

Box 4. Lake whitefish case study: the genetics of adaptation without a sequenced genome

Lake whitefish (*Coregonus clupeaformis*) in northeastern North America are found with two ecomorphs, dwarf (smaller, more limnetic) and normal (benthic), that have evolved rapidly, independently and in parallel in different freshwater lakes (Figure 1). There are numerous life history, metabolic and morphological traits that differ between the two ecomorphs under divergent selection [10]. Despite lacking a sequenced genome (which exists for threespine stickleback) or a closely related model species with a sequenced genome (which exists for wild mice), research into the genetic basis of the parallel phenotype has proceeded based on QTL crosses [95], gene expression analyses [96,97], SNP-based genome scans [98] and next-generation transcriptome sequencing [99]. Parallel patterns in the expression of some key genes has been identified and strongly suggest a driving role of natural selection, particularly for genes that relate to energetics, though a number of genes (reviewed in [10]) and polymorphisms [98] also show non-parallel patterns across equivalent ecomorphs along an ecological gradient. Divergent selection is also implicated in parallel phenotype evolution across the genus, but no evidence has been found for parallel gene expression across other limnetic species [100]. In contrast to other studies discussed in this review, many of which focus on particular candidate regions and identify few genes of major effect, sympatric whitefish ecomorphs are diverged at minimally several hundred genes (1–3% of the genome) [10]. The importance of a few genes of major effect versus many genes of small effect likely derives from the complexity and subtlety of the phenotypic trait under selection.

the same gene, and even different genes (usually at lower frequency) tend to underlie parallel phenotypes among closely related species. To find the exceptions of genetic non-parallelism, one needs a multifaceted approach: minimally, by sampling multiple populations from nature, identifying the genetic variability at target genetic regions for each population, and mapping the causal gene at least once.

An excellent example of this includes major phenotypic transitions in stickleback fishes (Box 3), especially pelvic reduction whereby independent deletions affecting a homologous gene enhancer (i.e. *de novo* homologous mutation) and non-homologous mutation at an undetermined gene are the repeated, independent genetic basis of a globally distributed parallel phenotype [49] (Figure 1b). Research on whitefishes has also embraced a combined approach and identified homologous and non-homologous genetic associations to replicate parallel 'dwarf' and 'normal' phenotypes along an ecological gradient (Box 4) (Figure 1a). Thus, there is growing evidence to contradict the expectation that closely related species should evolve parallel phenotypes through the same genetic bases. Though parallelisms are common there are almost always exceptions.

Cavefishes (*Astyanax mexicanus*) provide a replicate pattern to analyse parallel phenotypic evolution because closely related populations have pairs of surface and cave morphs, which have reduced eye structures and reduced pigmentation (reviewed in [50]). Whether eye loss is driven by pressures of natural selection, an outcome of pleiotropy, or a product of stochastic drift, remains a topic of much debate [51–53]. It has been shown that eyes are a complex trait that involves the co-ordination of many unlinked loci in these fishes [54]. Genetic mapping experiments have demonstrated that for the loss-of-function trait of eye regression, across four different phenotypically parallel cave populations, three cases involve different genes and one case has a shared genetic basis due to standing genetic variation [55]. The absence of colour pigments, or albinism, in three populations of cave phenotypes was also found to be due to independent mutations at the ocular and cutaneous albinism-2 gene (*Oca2*) [56]. In two populations the albinism is due to loss-of-function alleles at *Oca2* but the deletions (i.e. events) are different in each population and it appears that in a third cave population, a different independent mutation in *Oca2* is responsible [56]. A reduced pigmentation phenotype called 'brown', which is found in individuals without albinism and co-occurs in populations with albinos, also has variable genetic bases in nature: two are parallelisms involving different mutations at the same gene (variants of *Mc1r* mutation resulting in loss-of-function that might be derived from standing genetic variation) and a third undetermined mutation at a

Author's personal copy

different gene (i.e. does not involve *Mc1r*) [57]. Therefore, many examples of parallel phenotypic evolution, in this case of eye regression and two different pigmentation phenotypes, are each a combination of parallel evolution at homologous genes by independent events: parallel evolution at homologous genes because of standing genetic variation, and also non-parallel, non-homologous evolution of different genes.

Cichlid fishes are renowned for parallel evolution of colouration and body shape [58]. Orange-blotched (OB) is a body colour pattern found within and across species of cichlids in African Rift lakes. Mapping crosses in the lab followed by association mapping of populations in Lake Malawi located the causative locus for this phenotype [59]. Because of the mutation-level parallelism and shared haplotype block across all four of the genera examined, the authors inferred a single genetic origin of the OB phenotypes in Lake Malawi [59]. However, the phenotypically parallel OB cichlid species in Lake Victoria do not share the nucleotide polymorphisms located in the Lake Malawi cichlids, which suggests a non-homologous genetic basis and independent evolutionary origin to the Lake Victoria OB species [59].

Homologous and non-homologous genetic origins are also implicated in the parallel phenotype evolution of light and dark pigmentation in wild populations of mice (reviewed in [20]). In the oldfield mouse–beach mouse species (*Peromyscus polionotus* complex) all the subspecies show a significant association between coat colour and habitat substrate colour, with lighter-coated mice being found on sand compared to field habitats [60]. Mutation in the pigmentation gene *Mc1r* plays a frequent but not ubiquitous role [61–63] and a second gene, *Agouti*, also has been implicated [64]. In a sister species (*Peromyscus maniculatus*) in Nebraska, Linnen *et al*. [65] identified that a *de novo* dominant allele at the *Agouti* locus and higher expression of *Agouti* was associated with the light ('wide-band') coat colour in populations that live on the light-coloured Sand Hills habitat compared to darker ('wildtype') conspecifics on darker substrates. No coding mutations at *Mc1r* were found to be associated with the light coat phenotype [65], despite its frequent involvement in pigmentation across disparate taxa [66]. There is a dark coated ('melanic') phenotype of *P. maniculatus* that also follows no fixed pattern of parallelism within or between genes [67]. In a New England population, the dark phenotype is due to a single recessive allele with a 125-kb deletion at the regulatory and adjacent exonic regions, which decreases *Agouti* expression. In two other North American populations, the dark phenotype was associated with either a different mutation in *Agouti* or a mutation at a gene other than *Agouti* [67]. Thus, different mutations at the same gene or different genes were associated with the dark coat for each of the phenotypically parallel populations investigated.

These recent findings indicate that when a breadth of molecular approaches are used, multiple populations are sampled from nature, and multiple loci are assessed, the pattern of mutation-level or locus-level genetic parallelism being associated with parallel phenotypes is far from clear. While often the same loci are involved, examples from stickleback fishes, whitefishes, cavefishes, cichlids and mice – among others – demonstrate that even within the same or closely related species, the many exceptions suggest we do not yet have any rule.

The future

We predict that research in the coming years will bring study of populations of non-model species in nature firmly into the era of genomics. The ability to infer the mechanisms of natural selection at the genetic level will continue to be especially powerful in the context of parallel evolution, or replicate adaptive radiations. It is hoped that further empirical research on the genetics of adaptation within the context of parallel evolution will allow us to develop firmer conclusions about the patterns and processes of parallelism. The ultimate goal of this is to determine the selective conditions and genetic architectures that result in parallel phenotypes, by any genetic means, and to identify the role genetic parallelism may play in the speed of adaptation.

To advance the field, it will be imperative to thoroughly sample multiple populations within species in order to identify whether various routes to the same phenotype even exist. Also, we require phylogenetically controlled genetic and phenotypic analyses that would allow us to test the relationship between evolutionary distance and the frequency of homologous versus non-homologous genetic bases.

With whole genome information (promoted, for example, by the 10k Genomes Project [47]) and genetic mapping data, researchers will soon be able to identify the role of particular genomic regions, such as fragile sites, inversion, or sex chromosomes, in promoting replicate patterns of phenotypic diversification among closely related species. Once they arise, such structural patterns in the genome may allow for rapid spreading of advantageous combinations of loci from standing genetic variation. If genomic architectural designs are important in parallel phenotype evolution and are conserved among closely related species, then it provides evidence for the idea that parallel, homologous genetic evolution indeed is more common in closely related species.

We have greatly simplified the genetic bases of structural and regulatory genetic evolution for the sake of illustration. Genes are not isolated from each other and developmental pathways are complicated, non-linear and multigenic processes. In particular, the roles of gene interaction (epistasis) and the importance of pleiotropic effects is largely unknown but critical territory for future research [19]. With a greater genomic understanding of non-model systems, coupled with the rise of evo-devo [68], it can be expected that we will better understand whether non-parallel genetic bases of parallel phenotypes nonetheless make the same important parallel alterations further downstream in pathways. This would not be surprising, as clearly Nature is a tinkerer.

Conclusions

The implication of this 'tinkering' is that in fact many mutational options might exist during a population's history in order for it to solve its 'ecological problems'. In

Author's personal copy

contrast to this breadth of genetic 'answers', the very existence of replicate adaptive radiations suggests there might be limited phenotypic possibilities, or adjustments to the morphological and developmental *bauplan* [69], within species complexes. Thus, one of the lessons from research on parallel phenotypic and genetic evolution is that the genomic bases might be much more labile than the eco-morphological phenotype.

Another major lesson that can be drawn from recent technical innovations is that the surprisingly large genetic variability in natural populations will be key to parallel evolution [27]. As repeatedly shown in the examples we covered here, years of dedicated crossing and QTL experiments to hone in on 'the gene' often find that, once the variability of natural populations is taken into account, the clean associations start to smudge. It seems that rarely will an answer derived from a laboratory population turn out to be a ubiquitous answer in nature.

The exciting work on the genetics and genomics of ecological and evolutionary (rather than genetic) model systems with parallel phenotypes has provided celebrated and welcome novel insights into the genetics of adaptation. Recent empirical findings indicate that there is little pattern of time dependency to genetic parallelism: even populations within a species, and species within complexes, can dramatically span standing genetic variation, *de novo* homologous mutation and *de novo* non-homologous mutation. Nonetheless, we predict that by using genomic tools and embracing the complexity of nature, the framework of parallel evolution will prove one of the most powerful contexts to reveal fundamental patterns of adaptive evolution.

Acknowledgements

We apologise to colleagues whose relevant work was not included due to oversight or space constraints. We thank Y. F. Chan, F. Henning, J. Losos, M. Pierotti, two anonymous reviewers, and especially H. Hoekstra for comments that improved the manuscript. We thank R. Barrett, L. Bernatchez, and Y.F. Chan for contributing images. KRE is funded by an NSERC postdoctoral fellowship and AM by the DFG.

References

1 Schluter, D. (2009) Evidence for ecological speciation and its alternative. *Science* 323, 737–741
2 Venditti, C. *et al.* (2010) Phylogenies reveal new interpretation of speciation and the Red Queen. *Nature* 463, 349–352
3 Lenormand, T. *et al.* (2009) Stochasticity in evolution. *Trends Ecol. Evol.* 24, 157–165
4 Schluter, D. (2000) *The Ecology of Adaptive Radiations*, Oxford University Press
5 Losos, J.B. (2009) *Lizards in an Evolutionary Tree*, University of California Press
6 Grant, P.R. and Grant, B.R. (2008) *How and Why Species Multiply: The Radiation of Darwin's Finches*, Princeton University Press
7 Salzburger, W. (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Mol. Ecol.* 18, 169–185
8 Elmer, K.R. *et al.* (2010) Local variation and parallel evolution: morphological and genetic diversity across a species complex of Neotropical crater lake cichlid fishes. *Phil. Trans. R. Soc. B.* 365, 1769–1782
9 Foster, S.A. and Baker, J.A. (2004) Evolution in parallel: new insights from a classic system. *Trends Ecol. Evol.* 19, 456–457
10 Bernatchez, L. *et al.* (2010) On the origin of species: insights from the ecological genomics of the lake whitefish. *Phil. Trans. R. Soc. B.* 365, 1783–1800
11 Culver, D.C. and Pipan, T. (2009) *The Biology of Caves and Other Subterranean Habitats*, Oxford University Press
12 Losos, J.B. and Ricklefs, R.E. (2009) Adaptation and diversification on islands. *Nature* 457, 830–836
13 Stapley, J. *et al.* (2010) Adaptation genomics: the next generation. *Trends Ecol. Evol.* 25, 705–712
14 Slot, J.C. and Rokas, A. (2010) Multiple GAL pathway gene clusters evolved independently and by different mechanisms in fungi. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10136–10141
15 Kassens, R. (2009) Toward a general theory of adaptive radiation: insights from microbial experimental evolution. *Ann. N.Y. Acad. Sci.* 1168, 3–22
16 Losos, J.B. (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* 175, 623–639
17 Turner, T.L. *et al.* (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nat. Genet.* 42, 260–264
18 Wood, T.E. *et al.* (2005) Parallel genotypic adaptation: when evolution repeats itself. *Genetica* 123, 157–170
19 Kopp, A. (2009) Metamodels and phylogenetic replication: A systematic approach to the evolution of developmental pathways. *Evolution* 63, 2771–2789
20 Manceau, M. *et al.* (2010) Convergence in pigmentation at multiple levels: mutations, genes and functions. *Phil. Trans. R. Soc. B.* 365, 2439–2450
21 Stern, D.L. and Orgogozo, V. (2009) Is genetic evolution predictable? *Science* 23, 746–751
22 Walsh, B. and Blows, M.W. (2009) Abundant genetic variation + strong selection = multivariate genetic constraints: A geometric view of adaptation. *Annu. Rev. Ecol. Evol. Syst.* 40, 41–59
23 Futuyma, D.J. (2010) Evolutionary constraint and ecological consequences. *Evolution* 64, 1865–1884
24 Hoekstra, H.E. and Coyne, J.A. (2007) The locus of evolution: Evo devo and the genetics of adaptation. *Evolution* 61, 995–1016
25 Wray, G.A. (2007) The evolutionary significance of *cis*-regulatory mutations. *Nat. Rev. Genet.* 8, 206–216
26 Stern, D.L. and Orgogozo, V. (2008) The loci of evolution: How predictable is genetic evolution? *Evolution* 62, 2155–2177
27 Barrett, R.D.H. and Schluter, D. (2008) Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23, 38–44
28 Orr, H.A. and Betancourt, A.J. (2001) Haldane's sieve and adaptation from the standing genetic variation. *Genetics* 157, 875–884
29 Barrett, R.D.H. *et al.* (2008) Natural selection on a major armor gene in threespine stickleback. *Science* 322, 255–257
30 Haas, O. and Simpson, G.G. (1946) Analysis of some phylogenetic terms, with attempts at redefinition. *Proc. Am. Phil. Soc.* 90, 319–349
31 Shapiro, M.D. *et al.* (2009) The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks. *Curr. Biol.* 19, 1140–1145
32 Ouborg, N.J. *et al.* (2010) Conservation genetics in transition to conservation genomics. *Trends Ecol. Evol.* 26, 177–187
33 Vasmägi, A. and Primmer, C.R. (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Mol. Ecol.* 14, 3623–3642
34 Storz, J.F. and Wheat, C.W. (2010) Integrating evolutionary and functional approaches to infer adaptation at specific loci. *Evolution* 64, 2489–2509
35 Metzker, M.L. (2010) Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11, 31–46
36 Stinchcombe, J.R. and Hoekstra, H.E. (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100, 158–170
37 Mackay, T.F.C. *et al.* (2009) The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10, 565–577
38 Slate, J. *et al.* (2009) Gene mapping in the wild with SNPs: guidelines and future directions. *Genetica* 136, 97–107
39 Nosil, P. *et al.* (2009) Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* 18, 375–402
40 Baird, N.A. *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3, e3376
41 Hohenlohe, P.A. *et al.* (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6, e1000862

Author's personal copy

42 Breakthrough of the year: Areas to watch. *Science* 330 (2010) 1608–1609
43 Morozova, O. *et al.* (2009) Applications of new sequencing technologies for transcriptome analysis. *Annu. Rev. Genom. Hum. G.* 10, 135–151
44 Vera, J.C. *et al.* (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Mol. Ecol.* 17, 1636–1647
45 Schwarz, D. *et al.* (2009) Sympatric ecological speciation meets pyrosequencing: sampling the transcriptome of the apple maggot *Rhagoletis pomonella*. *BMC Genomics* 10, 633
46 Elmer, K.R. *et al.* (2010) Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. *Mol. Ecol.* 19 (Suppl. 1), 197–211
47 Genome 10K Community of Scientists (2009) Genome 10K: A proposal to obtain whole-genome sequence for 10 000 vertebrate species. *J. Hered.* 100, 659–674
48 Wang, Z. *et al.* (2009) RNA-Seq: a revolutionary tool for transcriptomes. *Nat. Rev. Genet.* 10, 57–63
49 Chan, Y.F. *et al.* (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* 327, 302–305
50 Jeffery, W.R. (2009) Regressive evolution in *Astyanax* cavefish. *Annu. Rev. Genet.* 43, 25–47
51 Wilkens, H. (2010) Genes, modules and the evolution of cave fish. *Heredity* 105, 413–422
52 Wilkens, H., (2010) Variability and loss of functionless traits in cave animals. Reply to Jeffery (2010). *Heredity* advance online publication doi:10.1038/hdy.2010.105, 1–2.
53 Protas, M. *et al.* (2008) Multi-trait evolution in a cave fish Astyanax mexicanus. *Evol. Dev.* 10, 196–209
54 Protas, M. *et al.* (2007) Regressive evolution in the Mexican cave tetra Astyanax mexicanus. *Curr. Biol.* 17, 452–454
55 Borowsky, R. (2008) Restoring sight in blind cavefish. *Curr. Biol.* 18, 23–24
56 Protas, M.E. *et al.* (2006) Genetic analysis cavefish reveals molecular convergence in the evolution of albinism. *Nat. Genet.* 38, 107–111
57 Gross, J.B. *et al.* (2009) A novel role for *Mc1r* in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genet.* 5, e1000326
58 Stiassny, M.L.J. and Meyer, A. (1999) Cichlids of the Rift Lakes. *Sci. Am.* 64–69
59 Roberts, R.B. *et al.* (2009) Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science* 326, 998–1001
60 Mullen, L.M. *et al.* (2009) Adaptive basis of geographic variation: genetic, phenotypic and environmental differences among beach mouse populations. *Proc. R. Soc. Lond. B.* 276, 3809–3818
61 Hoekstra, H.E. *et al.* (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313, 101–104
62 Steiner, C.C. *et al.* (2009) The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. *Mol. Biol. Evol.* 26, 35–45
63 Mullen, L.M. and Hoekstra, H.E. (2008) Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution* 62, 1555–1570
64 Steiner, C.C. *et al.* (2007) Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol.* 5, e219
65 Linnen, C.R. *et al.* (2009) On the origin and spread of an adaptive allele in deer mice. *Science* 325, 1095–1098
66 Hubbard, J.K. *et al.* (2010) Vertebrate pigmentation: from underlying genes to adaptive function. *Trends Genet.* 26, 231–239
67 Kingsley, E.P. *et al.* (2009) Melanism in *Peromyscus* is caused by independent mutations in *Agouti*. *PLoS ONE* 4, e6435
68 Carroll, S.B. (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134, 25–36
69 Wake, D.B. (1991) Homoplasy: The result of natural selection, or evidence of design limitations? *Am. Nat.* 138, 543–567
70 Christin, P-A. *et al.* (2010) Causes and evolutionary significance of genetic convergence. *Trends Genet.* 26, 400–405
71 Arendt, J. and Reznick, D.N. (2007) Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* 23, 26–32
72 Leander, B.S. (2008) Different modes of convergent evolution reflect phylogenetic distance; a reply to Arendt and Reznick. *Trends Ecol. Evol.* 23, 481–483
73 Hulsey, C.D. (2009) Cichlid genomics and phenotypic diversity in a comparative context. *Integr. Comp. Biol.* 49, 618–629
74 Terai, Y. *et al.* (2006) Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biol.* 4, 2244–2251
75 Seehausen, O. *et al.* (2008) Speciation through sensory drive in cichlid fish. *Nature* 455, 620–626
76 O'Quin, K.E. *et al.* (2010) Parallel evolution of opsin gene expression in African cichlid fishes. *Mol. Biol. Evol.* 27, 2839–2854
77 Larmuseau, M.H.D. *et al.* (2010) High molecular diversity in the rhodopsin gene in closely related goby fishes: A role for visual pigments in adaptive speciation? *Mol. Phylogen. Evol.* 55, 689–698
78 Sugawara, T. *et al.* (2002) Natural selection of the rhodopsin gene during the adaptive radiation of East African Great Lakes cichlid fishes. *Mol. Biol. Evol.* 19, 1807–1811
79 Hofmann, C.M. *et al.* (2010) Plasticity of opsin gene expression in cichlids from Lake Malawi. *Mol. Ecol.* 19, 2064–2074
80 Hofmann, C.M. *et al.* (2009) The eyes have it: Regulatory and structural changes both underlie cichlid visual pigment diversity. *PLoS Biol.* 7, e1000266
81 Carleton, K.L. *et al.* (2010) Genetic basis of differential opsin gene expression in cichlid fishes. *J. Evol. Biol.* 23, 840–853
82 Salzburger, W. *et al.* (2007) Adaptive sequence evolution in a color gene involved in the formation of the characteristic egg-dummies of male haplochromine cichlid fishes. *BMC Biol.* 5, 51
83 Elmer, K.R. *et al.* (2009) Color assortative mating contributes to sympatric divergence of neotropical crater lake cichlid fish. *Evolution* 63, 2750–2757
84 Henning, F. *et al.* (2010) Genetic, comparative genomic and expression analyses of the Mc1r locus in the polychromatic Midas cichlid fish (Teleostei, Cichlidae *Amphilophus* sp.) species group. *J. Mol. Evol.* 70, 405–412
85 Coyle, S.M. *et al.* (2007) Parallel evolution of *PitxI* underlies pelvic reduction in Scottish threespine stickleback (*Gasterosteus aculeatus*). *J. Hered.* 98, 581–586
86 Shapiro, M.D. *et al.* (2006) Parallel genetic origins of pelvic reduction in vertebrates. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13753–13758
87 Shapiro, M.D. *et al.* (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* 428, 717–723
88 Colosimo, P.F. *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307, 1928
89 Barrett, R.D.H. *et al.* (2009) Should I stay or should I go? The *Ectodysplasin* locus is associated with behavioural differences in threespine stickleback. *Biol. Lett.* 5, 788–791
90 Marchinko, K.B. and Schluter, D. (2007) Parallel evolution by correlated response: lateral plate reduction in threespine stickleback. *Evolution* 61, 1084–1090
91 Barrett, R.D.H. *et al.* (2009) Environment specific pleiotropy facilitates divergence at the *Ectodysplasin* locus in threespine stickleback. *Evolution* 63, 2831–2837
92 Marchinko, K.B. (2008) Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. *Evolution* 63, 127–138
93 Cresko, W.A. (2008) Armor development and fitness. *Science* 322, 204–205
94 Schluter, D. and Conte, G.L. (2009) Genetics and ecological speciation. *Proc. Natl. Acad. Sci. U.S.A.* 106 (Suppl 1), 9962–9995
95 Whiteley, A.R. *et al.* (2008) The phenomics and expression quantitative trait locus mapping of brain transcriptomes regulating adaptive divergence in Lake Whitefish species pairs (*Coregonus* sp.). *Genetics* 180, 147–164
96 Derome, N. *et al.* (2006) Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchell) ecotypes. *Mol. Ecol.* 15, 1239–1249
97 Nolte, A. *et al.* (2009) Divergence in gene regulation at young life history stages of whitefish (*Coregonus* sp) and the emergence of genomic isolation. *BMC Evol. Biol.* 9, 59

Author's personal copy

98 Renaut, S. *et al.* (2011) Gradients of ecological speciation SNP signature of selection on standing genetic variation, and association with adaptive phenotypes in lake whitefish species pairs (Coregonus spp.). *Mol. Ecol.* 20, 545–559
99 Renaut, S. *et al.* (2010) Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Mol. Ecol.* 19, 115–131
100 Jeukens, J. *et al.* (2009) Candidate genes and adaptive radiation: insights from transcriptional adaptation to the limnetic niche among coregonine fishes (Coregonus spp., Salmonidae). *Mol. Biol. Evol.* 26, 155–166