MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 5129–5131

NEWS AND VIEWS

PERSPECTIVE

Ecological genomics in full colour

PAUL A. HOHENLOHE

Department of Biological Sciences, Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844, USA

Colour patterns in animals have long offered an opportunity to observe adaptive traits in natural populations. Colour plays myriad roles in interactions within and among species, from reproductive signalling to predator avoidance, leading to multiple targets of natural and sexual selection and opportunities for diversification. Understanding the genetic and developmental underpinnings of variation in colour promises a fuller understanding of these evolutionary processes, but the path to unravelling these connections can be arduous. The advent of genomic techniques suitable for nonmodel organisms is now beginning to light the way. Two new studies in this issue of *Molecular Ecology* use genomic sequencing of laboratory crosses to map colour traits in cichlid fishes, a remarkably diverse group in which coloration has played a major role in diversification. They illustrate how genomic approaches, specifically RAD sequencing, can rapidly identify both simple and more complex genetic variation underlying ecologically important traits. In the first, Henning *et al.* (2014) detect a single locus that appears to control in a Mendelian fashion the presence of horizontal stripes, a trait that has evolved in numerous cichlid lineages. In the second, Albertson *et al.* (2014) identify several genes and epistatic interactions affecting multiple colour traits, as well as a novel metric describing integration across colour traits. Albertson *et al.* (2014) go further, by quantifying differential expression of parental alleles at a candidate locus and by relating differentiation among natural populations at mapped loci to trait divergence. Herein lies the promise of ecological genomics – efficiently integrating genetic mapping of phenotypes with population genomic data to both identify functional genes and unravel their evolutionary history. These studies offer guidance on how genomic techniques can be tailored to a research question or study system, and they also add to the growing body of empirical examples addressing basic questions about how ecologically important traits evolve in natural populations.

Correspondence: Paul A. Hohenlohe, Fax: 208 885 7905;
E-mail: hohenlohe@uidaho.edu

Keywords: cichlid fish, coloration, genetic mapping, population genomics, RAD sequencing

Received 29 August 2014; revised 22 September 2014; accepted 25 September 2014

Genomic approaches to colour

The study of coloration has already addressed some long-standing questions in evolutionary biology, such as the degree of parallel genetic evolution underlying convergent phenotypic evolution. For instance, the taxonomic breadth of adaptive evolution in the melanocyte receptor gene *Mc1r* is striking (Manceau *et al.* 2010). Beyond a few textbook examples, however, genomics in nonmodel taxa may begin to reveal more general comparative patterns, such as the relative roles in adaptive divergence of single versus multiple genes or regulatory versus protein-coding alleles. An ideal system for this pursuit is cichlid fishes, an extraordinarily diverse group with strong evidence for the importance of coloration in recent adaptive radiations (Wagner *et al.* 2012) and a growing body of genomic resources. We can expect to find variation in the genetic basis of different cichlid coloration traits, and the sheer number of taxa at multiple stages of divergence allows direct testing of a wide range of hypotheses in the genetics of adaptation and diversification. Recent studies provide illustrative examples of how to apply genomics in this pursuit.

Henning *et al.* (2014) set out to map the genetic basis of midlateral and dorsolateral stripes, traits that are associated with piscivory and shoaling behaviour (Fig. 1). They used a laboratory F2 cross between two phenotypically divergent, but interfertile, *Haplochromis* species. Mapping resolution for quantitative trait loci in such a cross is limited by the number of recombination events, and several hundred markers are typically sufficient to fully characterize parental haplotype blocks. Accordingly, the authors used double-digest RAD sequencing (Peterson *et al.* 2012), which allows researchers to adjust downwards the number of loci and thus devote sequencing effort to multiplexing more individuals and increasing depth of coverage. This illustrates the flexibility among related RAD sequencing techniques to balance experimental trade-offs and tailor the design to a research question. Henning *et al.* (2014) produced a genetic map contained 867 SNP markers, providing more than enough resolution to detect a single genomic interval that apparently determines both traits, containing a logical candidate gene.

In part, the success of this study relied on the availability of a draft genome assembly for one of the parental species, to which the authors aligned the sequence data, and a well-annotated reference genome for a related species,

Fig. 1 Males of the parental species crossed by Henning *et al.* (2014), showing the absence (*Haplochromis sauvagei*, top) and presence (*H. nyererei*, bottom) of horizontal stripes (photographs courtesy of Ad Konings).

tilapia (*Oreochromis niloticus*). The authors also take advantage of their crossing design to estimate error rates for genotyping, adding to the growing and essential literature on sources of error and bias in RAD sequencing (e.g. Davey *et al.* 2013; Mastretta-Yanes *et al.* 2014). The authors recommend boosting sequencing coverage and describe quality control steps to improve confidence in genotyping and genetic map inference. They also test the sensitivity of their results to a range of parameters in the genotyping analysis (Catchen *et al.* 2013), an important step in analysing any genomic data set.

Albertson *et al.* (2014) conducted an extensive study that integrates across a number of data types to identify the genetic basis of a large number of coloration traits (Fig. 2). As in the previous study, they applied a form of RAD sequencing to a laboratory F2 cross between phenotypically divergent parental species. Their use of the standard RAD protocol (Baird *et al.* 2008) provides a higher density of markers across the genome, allowing for finer-scale mapping. They sequenced DNA from both F2 offspring and wild-caught individuals, illustrating how the same set of loci can be genotyped in both laboratory cross and population samples without ascertainment bias in the SNPs detected for either set of individuals. Combining QTL mapping and allele-specific expression assays, Albertson *et al.* (2014) implicate a regulatory change in *pax3a* that plays an important role in xanthophore-based flank coloration. Further resequencing at this locus in several species

Fig. 2 Males of the parental species studied by Albertson *et al.* (2014): *Tropheops* 'red cheek' (top) and *Labeotropheus fuelleborni* (bottom), showing divergence in a range of colour traits including integration across the body (photographs courtesy of Justin Marshall).

showed that the same SNP segregates with the phenotypic trait across the genus *Tropheops*.

The genetics of colour

These two studies add to the catalogue of examples of the genetic basis of colour pattern evolution and begin to inform us about more general trends in adaptive evolution, such as the role of single large-effect versus many smaller-effect genes. As with other cases, such as *Mc1r* or pigmentation in *Drosophila* (Kronforst *et al.* 2012), these authors found relatively simple genetic bases to the colour traits examined – Henning *et al.* (2014) identified a single Mendelian locus, and Albertson *et al.* (2014) identified from one to five loci associated with any single trait, although they also identified up to four pairwise epistatic interactions for several traits. A particular innovation by Albertson *et al.* (2014) is a novel metric to quantify the degree of integration in colour patterning across the body, which maps to two direct-effect loci and one epistatic interaction. Mapping such multivariate or composite traits may be especially useful, particularly to the extent that they are the targets of selection – in this case, the trait corresponds to colour contrast across the body of these fish.

These studies lend some support to the idea of parallel genetics underlying parallel phenotypic evolution. Albertson *et al.* (2014) identified the same nucleotide polymorphism in *pax3a* associated with coloration across spe-

cies. Henning *et al.* (2014) note that the tight correlation between midlateral and dorsolateral stripes, which traces to the single locus they identified, is common across lineages that have evolved horizontal stripes, although it remains to be seen whether the same gene or even the same polymorphism is responsible. Such parallel genetic evolution may be expected in radiations such as East African cichlids, where the rapid timescale of diversification means that much polymorphism is probably shared among related species and could respond to repeated selection, and introgressive hybridization continues to move functional alleles among species (Keller *et al.* 2013).

Future prospects for ecological genomics

Genetic mapping in a laboratory cross has obvious limitations, including the fact that only genetic variation specific to the individual parents is actually mapped. The true power of ecological genomics lies in integrating across types of data, as Albertson *et al.* (2014) have begun here, to link to phenotypes in natural populations. The multiple generations over which selection acts serve to break down linkage disequilibrium around functional loci, allowing for much finer mapping if large numbers of markers are available. Loci identified in a laboratory cross can be a launching point for further population-based studies, for instance targeted genotyping of candidate loci in population samples to test whether the genotype–phenotype association persists. This may reveal multiple independent mutations at the same locus contributing to adaptive phenotypic variation (e.g. Linnen *et al.* 2013). Broader population genomic studies – that is dense-marker surveys of genetic variation and differentiation across the genome in multiple natural populations – also complement a genetic map approach. One extension of RAD sequencing, assembly of contigs from paired-end data, can provide haplotype sequences of 800–1000 bp or more per locus (Hohenlohe *et al.* 2013), allowing coalescent-based or other analytical tools to reveal aspects of evolutionary history across the genome. By combining genomic data types in nonmodel species, and leveraging them against gene annotation and functional information from related taxa, researchers will progress rapidly towards a richer understanding of phenotypic adaptation.

References

Albertson C, Powder K, Hu Y, Coyle K, Roberts R, Parsons K (2014) Genetic basis of continuous variation in the levels and modular inheritance of pigmentation in cichlid fishes. *Molecular Ecology*, **23**, 5135–5150.

Baird NA, Etter PD, Atwood TS *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Davey JW, Cezard T, Fuentes-Utrilla P, Eland C, Gharbi K, Blaxter M (2013) Special features of RAD sequencing data: implications for genotyping. *Molecular Ecology*, **22**, 3151–3164.

Henning F, Lee HJ, Franchini P, Meyer A (2014) Genetic mapping of horizontal stripes in Lake Victoria cichlid fishes: benefits and pitfalls of using dense linkage mapping in non-model organisms. *Molecular Ecology*, **23**, 5224–5240.

Hohenlohe PA, Day MD, Amish SJ *et al.* (2013) Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology*, **22**, 3002–3013.

Keller I, Wagner CE, Greuter L *et al.* (2013) Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Molecular Ecology*, **22**, 2848–2863.

Kronforst MR, Barsh GS, Kopp A *et al.* (2012) Unraveling the thread of nature's tapestry: the genetics of diversity of convergence in animal pigmentation. *Pigment Cell & Melanoma Research*, **25**, 411–433.

Linnen CR, Poh Y-P, Peterson BK *et al.* (2013) Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, **339**, 1312–1316.

Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE (2010) Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2439–2450.

Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC (2014) Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. *Molecular Ecology Resources*, DOI: 10.1111/1755-0998.12291.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE*, **7**, e37135.

Wagner CE, Harmon LJ, Seehausen O (2012) Ecological opportunity and sexual selection together predict adaptive radiation. *Nature*, **487**, 366–369.

This paper was conceived and written by P.A.H.

doi: 10.1111/mec.12945