NEWS AND VIEWS

PERSPECTIVE

Hybridization and the porous genome: patterns of isolation and introgression in manakins

SAM YEAMAN*†

Department of Forest Sciences and Conservation, University of British Columbia, 2424 Main Mall, Vancouver, BC, V6T 1Z4, Canada; †Biodiversity Research Center, University of British Columbia, 2212 Main Mall, Vancouver, BC, V6T 1Z4, Canada

How does one species become two? What is the relative importance of selection vs. drift? How many genes are involved, what are their effect sizes and where are they located in the genome? The answers to these questions will inevitably vary from one species to another, but by addressing these questions in a diverse array of taxa, we will gradually piece together a more complete picture of how evolution works. In this issue of *Molecular Ecology*, Parchman *et al.* (2013) have approached these questions by studying the genomic basis of differentiation along a hybrid zone between two Central American bird species, the golden-collared (*Manacus vitellinus*) and white-collared (*Manacus candei*) manakins (see Fig. 1). Their work presents an interesting window into the evolution of this hybrid zone using novel analysis tools and a nice example of how to approach the study of population genomics in nonmodel organisms.

Parchman *et al.* (2013) used genotyping-by-sequencing (GBS) to obtain large amounts of sequence from individuals sampled from 12 populations spaced along the hybrid zone, including three populations of each parental species and six admixed populations. This sequencing yielded 59 100 informative SNPs, with an average coverage of $0.63\times$ per individual. As these loci tended to be distributed relatively uniformly throughout the genome (figure 5 from Parchman *et al.* 2013), this would correspond to a density of approximately 1 SNP every 25 kb. They applied Bayesian analysis methods developed by Gompert & Buerkle (2011a) and Gompert *et al.* (2012a) to estimate individual genotypes, population allele frequencies, the divergence between parental populations at each locus (F_{ST}) and the

Correspondence: Sam Yeaman, Fax: 604-822-2727;
E-mail: yeaman@zoology.ubc.ca

pattern of introgression at each locus, as represented by two genomic cline parameters (α, β).

For those readers unfamiliar with the genomic clines approach, it works by ordering individuals from admixed populations by their hybrid index, which is the proportion of a hybrid individual's genome inherited from one parental population or species (in this case, the proportion of loci with *M. candei* ancestry). The patterns of introgression at each locus are then ascertained by comparing the locus-specific probabilities of ancestry to the hybrid index and fitting a cline model to each locus. The α parameter represents whether the cline is shifted in one direction or the other, which occurs when one allele is highly introgressed from one parental population or the other. The β parameter represents whether the cline transitions rapidly between extremes (positive β) or plateaus at intermediate values of the hybrid index (negative β).

The main results of Parchman *et al.* (2013) were found by studying the distributions of F_{ST} and cline parameters and analysing the position in the genome of loci identified as statistical outliers. Overall, they found hundreds of loci with signatures of genetic differentiation or introgression scattered throughout the genome, and significant positive correlations between differentiation and introgression. On average, they found $F_{ST} = 0.26$, with a range of 0.09–0.61, with 1007 loci identified as high F_{ST} outliers. For the introgression parameter α, they found 623 loci with an excess of *M. candei* ancestry and 514 loci with an excess of *M. vitellinus* ancestry. For the introgression parameter β, they found 203 outlier loci with $\beta > 0$ and 220 outlier loci with $\beta < 0$. Of the F_{ST} outlier loci, 43 loci were also α outliers, while 117 loci were also β outliers, with significant positive correlations found between F_{ST} and absolute values of α ($r = 0.17$) and β ($r = 0.36$) across all loci (see figure 3 from Parchman *et al.* 2013). While these results indicate that there was a tendency for individual loci to be both highly divergent and have atypical patterns of introgression, when they aligned these loci back to large scaffolds from a draft genome of *M. vitellinus*, they found little evidence that outliers tended to cluster together in the genome. While both F_{ST} and cline parameters were highly autocorrelated at distances of <100 bp, this autocorrelation decayed to very low levels beyond approximately 5000 bp, suggesting only limited clustering of estimates along chromosomes (although autocorrelation for each of these estimates was somewhat higher within scaffolds (Moran's I: $0 < I < 0.05$) than among scaffolds ($I = 0$).

There are two major hurdles that studies must face when interpreting results from this type of outlier analysis. First, while the evolutionary response of F_{ST} has been relatively well studied, much less is known about how α and β

Fig. 1 Hybrid (*Manacus candei × Manacus vitellinus*; left) and white-collared manakin (*M. candei*; right). Photograph credits: Robb Brumfield (hybrid) and ©D. Tipling/VIREO (*M. candei*).

respond to various forms of selection and complex demography, such as reproductive isolation driven by genomic conflict or drift-based divergence following repeated episodes of divergence and secondary contact. Second, the methods used to identify outliers have not yet been thoroughly tested for complex patterns of selection and demography, and many of the outliers may be driven by complex demographic history (Klopfstein *et al.* 2006; De Mita *et al.* 2013), although these hierarchical Bayesian methods may offer an improvement over previous approaches that did not account for population structure.

The authors acknowledge and discuss these limitations and outline some plausible interpretations of their results under the assumption that most of the outliers are driven by selection rather than demography. In the light of the large number of outlier loci distributed relatively evenly throughout the genome, they suggest that their results are consistent with the view that (i) speciation in manakins involves a porous genome where the consequences of adaptation or reproductive isolation are restricted to narrow regions of the chromosome and (ii) any adaptation or reproductive isolation is due to many alleles of small effect, a substantial proportion of which are involved in both adaptive divergence and isolation. The first of these interpretations should be relatively robust to the caveats discussed above. Because the GBS markers cover the genome at reasonably high density, it is unlikely that the study has failed to detect any clusters of genetically divergent loci larger than approximately 50 kb ('genomic islands of divergence'; Via & West 2008), so genomic islands do not seem to be an important feature in manakin speciation. Given the possibility for false positives with outlier detection, the second of these interpretations should be seen more as a provisional hypothesis about the genetic basis of reproductive isolation and adaptation, and a starting point for further exploration. It seems reasonable to expect some genetic signatures of divergent selection and reproductive isolation, as empirical evidence shows that both the maintenance of species colouration differences and the introgression of the golden collar of *M. vitellinus* into the range of *M. candei* are driven by sexual selection mediated by environmental differences (Parsons *et al.* 1993). Uy & Stein (2007) showed that the colouration patterns of each

species tended to be more conspicuous in their home range than in the range of the opposite species, and that the golden collar of *M. vitellinus* was more conspicuous in the hybrid zone. Some genetic signature of these phenotypic patterns is therefore expected. However, it is possible that the study has failed to detect some regions under selection, given the generally low levels of LD, and it is also possible that many of the observed outliers are driven by demography, as the authors acknowledge. Clearly, this is an area that is ripe for further theoretical research and testing of analysis tools (e.g. Gompert *et al.* 2012b; De Mita *et al.* 2013).

What can we conclude about the importance of genomic islands and linkage from these results? Large genomic islands may be expected if adaptive divergence occurs with high levels of gene flow between populations (Via & West 2008), but as the authors note, their failure to observe such islands does not contradict this theory, because manakins probably diverged in allopatry. High levels of gene flow during secondary contact could also be expected to result in large genomic islands, due to reduced erosion of genetic divergence around loci under strong divergent selection (Strasburg *et al.* 2012). However, manakins do not appear to have high rates of migration into the hybrid zone, as linkage disequilibrium within admixed populations is not higher than in parental populations (table 2 from Strasburg *et al.* 2012). These results therefore present a complement to theories that advocate an important role for linkage and the evolution of genomic islands in speciation: the lack of genomic islands is consistent with such theoretical predictions given the low migration rates, but linkage therefore appears relatively unimportant for adaptation and reproductive isolation in these species. It would be interesting to conduct similar comparative studies on closely related species pairs with different levels of reproductive isolation, to see whether larger clusters of differentiated loci are seen in such cases (as per Wu 2001).

As we accumulate more examples of the genetic architecture of adaptation and reproductive isolation under scenarios ranging from allopatry and secondary contact to parapatry, we will continue to learn more about the empirical importance of gene flow as a factor shaping the evolution of species. Comparative studies across 'natural experiments' with population samples spanning multiple replicate hybrid zones or environmental gradients (e.g. Jones *et al.* 2012) may provide the most powerful way to apply current analysis tools, as it is less likely that the vagaries of demography would yield the same outliers in multiple replicate populations. The study by Parchman *et al.* (2013) presents an interesting illustration of an approach that will undoubtedly become increasingly common with the increasing accessibility of genomic sequencing at the population level.

Acknowledgements

References

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

Gompert Z, Buerkle CA (2011a) Bayesian estimation of genomic clines. *Molecular Ecology*, **20**, 2111–2127.

Gompert Z, Lucas LK, Nice CC, Fordyce JA, Forister ML, Buerkle CA (2012a) Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution*, **66**, 2167–2181.

Gompert Z, Parchman T, Buerkle CA (2012b) Genomics of isolation in hybrids. *Philosophical Transactions of the Royal Society, Series B*, **367**, 439–450.

Jones FC, Grabherr MG, Chan YF et al. (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55–61.

Klopfstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, **23**, 482–490.

Parchman T, Gompert Z, Braun MJ (2013) The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. *Molecular Ecology*, **22**, 3304–3317.

Parsons TJ, Olson SL, Braun MJ (1993) Unidirectional spread of secondary sexual plumage traits across an avian hybrid zone. *Science*, **260**, 1643–1646.

Strasburg JL, Sherman NA, Wright KM, Moyle LC, Willis JH, Rieseberg LH (2012) What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Philosophical Transactions of the Royal Society, Series B*, **367**, 364–373.

Uy JAC, Stein AC (2007) Variable visual habitats may influence the spread of colourful plumage across an avian hybrid zone. *Journal of Evolutionary Biology*, **20**, 1847–1858.

Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology*, **17**, 4334–4345.

Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.