

NEWS AND VIEWS

PERSPECTIVE

Butterfly genomics sheds light on the process of hybrid speciation

NICOLA NADEAU

Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, UK

How common is hybridization between species and what effect does it have on the evolutionary process? Can hybridization generate new species and what indeed is a species? In this issue, Gompert *et al.* (2014) show how massive, genome-scale data sets can be used to shed light on these questions. They focus on the *Lycaeides* butterflies, and in particular, several populations from the western USA, which have characteristics suggesting that they may contain hybrids of two or more different species (Gompert *et al.* 2006). They demonstrate that these populations do contain mosaic genomes made up of components from different parental species. However, this appears to have been largely driven by historical admixture, with more recent processes appearing to be isolating the populations from each other. Therefore, these populations are on their way to becoming distinct species (if they are not already) but have arisen following extensive hybridization between other distinct populations or species (Fig. 1).

Keywords: adaptation, butterflies, evolution, ecological genetics, genomics, hybridization, speciation

Received 4 July 2014; revised 23 July 2014; accepted 28 July 2014

Their data set must be one of the largest outside of humans, with over one and half thousand butterflies genotyped at over 45 thousand variable nucleotide positions. It is this sheer amount of data that has allowed the researchers to distinguish between historical and more recent evolutionary and demographic processes. This is because it has allowed them to partition the data into common and rare genetic variants and perform separate analyses on these. Common genetic variants are likely to be older while rare variants are more likely to be due to recent mutations. Therefore, by splitting the genetic variation into these components, the researchers were able to show more admixture among common variants, while rare variants showed less admixture and clear separation of the populations. The extensive geographic sampling of individuals, including

overlapping distributions of several of the putative species, also allowed the authors to rule out the possibility that the separation of the populations was simply due to geographical distance.

The authors have developed a new programme for detecting population structure and admixture, which does the same job as STRUCTURE (Pritchard *et al.* 2000), identifying genetically distinct populations and admixture between these populations, but is designed to be used with next generation sequence data. They use the output of this model for another promising new method to distinguish between contemporary and historical admixtures. They fixed the number of source populations in the model at two and estimated the proportion of each individual's genome coming from these two populations. Therefore, an individual can either be purely population 1, or population 2 or some mixture of the two (they call this value q , the same parameter exists in STRUCTURE). They then compared this to the level of heterozygosity coming from the two source populations in the individual's genome. If an individual is an F1 hybrid of two source populations, then it would have a q of 0.5 and also be heterozygous at all loci that distinguish the parental populations. On the other hand, if it is a member of a stable hybrid lineage, it might also have a q of 0.5 but would not be expected to be heterozygous at these loci, because over time the population would become fixed for one or other of the source population states either by drift or selection (Fig. 2). This is indeed what they find in the hybrid populations. They tend to have intermediate q values, but the level of heterozygosity coming from the source populations (which they call Q_{12}) was consistently lower than expected.

Overall, the results support several of the populations as being stable hybrid lineages. Nevertheless, the strictest definitions of hybrid species specify that the process of hybridization between the parental species must be instrumental in driving the reproductive isolation of the new species from both parental populations (Abbott *et al.* 2013). This is extremely hard to demonstrate conclusively because it requires us to first of all identify the isolating mechanisms that operated in the early evolution of the species and then to show that these were caused by the hybridization event itself. One advantage of the *Lycaeides* system is that the species appear to be in the early stages of divergence, so barriers to gene flow that are operating currently are likely to be those that are driving the species divergence. While there is some evidence that hybridization gave rise to traits that allowed the new populations to colonize new environments (Gompert *et al.* 2006; Lucas *et al.* 2008), there is clearly further work to be carried out in this direction.

One of the rare examples of homoploid hybrid speciation (hybrid speciation without a change in chromosome

Correspondence: Nicola Nadeau, Fax: +44 (0)114 222 0002; E-mail: n.nadeau@sheffield.ac.uk



Fig. 1 There has been extensive historical admixture between *Lycaeides* species with some new species arising from hybrid populations. (Photo credits: Lauren Lucas, Chris Nice, and James Fordyce).

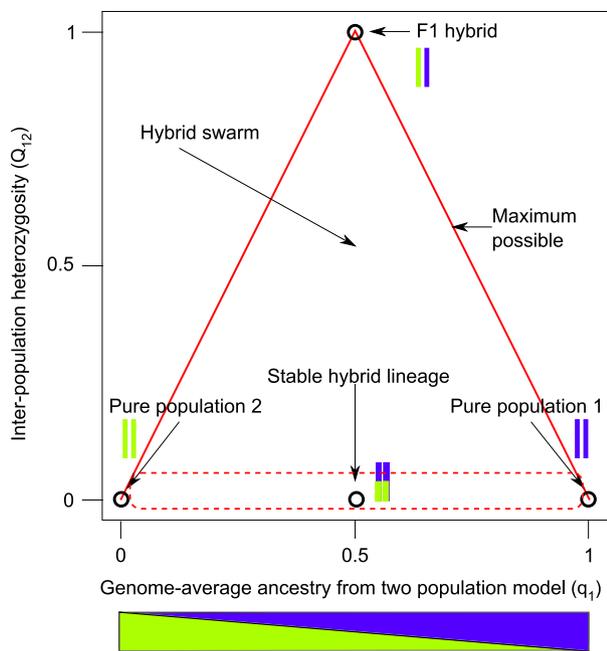


Fig. 2 The Q-matrix analysis used by Gompert *et al.* (2014) to distinguish between contemporary (hybrid swarm) and historical (stable hybrid lineage) admixture.

number) where the reproductive isolation criterion has been demonstrated, comes from the *Heliconius* butterflies. In this case, hybridization of two species has been shown to give rise to a new colour pattern that instantly becomes reproductively isolated from the parental species due to mate preference for that pattern (Mavárez *et al.* 2006). However, while

this has become a widely accepted example (Abbott *et al.* 2013), the naturally occurring 'hybrid species' in fact has derived most of its genome from one of the parental species, with largely just the colour pattern controlling locus coming from the other parent, a process that has been termed 'hybrid trait speciation' (Salazar *et al.* 2010). This distinction is an important one in terms of our understanding of the organization of biological diversity. While hybrid trait speciation will still largely fit the model of a neatly branching evolutionary tree, with perhaps only the region surrounding the single introgressed gene deviating from this model, hybrid species that end up with mosaic genomes, like *Lycaeides*, will not fit this model when considering the genome as a whole.

This distinction also more broadly applies when comparing the patterns of divergence between *Heliconius* and *Lycaeides*. These two butterfly genera have been driving forward our understanding of the prevalence and importance of hybridization at the genomic level, but they reveal different ways in which hybridization can influence the organization of biological diversity. Recent work in *Heliconius* has shown that admixture is extensive and has been ongoing over a large portion of the evolutionary history of species (Martin *et al.* 2013; Nadeau *et al.* 2013). Nevertheless, this has not obscured the clear and robust pattern of a bifurcating evolutionary tree when considering the genome as a whole (Nadeau *et al.* 2013). In contrast in *Lycaeides*, the genome-wide phylogeny clearly does not fit a bifurcating tree, resembling more of a messy shrub, with hybrid taxa falling at intermediate positions on the phylogeny (Gompert *et al.* 2014). The extent to which we need to rethink the way we describe and organize biological diversity will depend on the relative prevalence of these different outcomes of hybridization.

We are likely to see many more of these types of large sequence data sets for ecologically interesting organisms. Gompert *et al.* (2014) show that these data need not only be a quantitative advance, but can also qualitatively change our understanding of the evolutionary history of these organisms. In particular, analysing common and rare genetic variants separately may provide information that would otherwise be missed. The emerging field of 'speciation genomics' (Seehausen *et al.* 2014) should follow this lead in developing new ways of making the most of the flood of genomic data that is being generated, but also improve methods for integrating this with field observations and experiments to identify the sources and targets of selection and divergence.

References

- Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.
- Gompert Z, Fordyce JA, Forister ML, Shapiro AM, Nice CC (2006) Homoploid hybrid speciation in an extreme habitat. *Science*, **314**, 1923–1925.
- Gompert Z, Lucas LK, Buerkle CA *et al.* (2014) Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology*, **23**, 4555–4573.
- Lucas LK, Fordyce JA, Nice CC (2008) Patterns of genitalic morphology around suture zones in North American *Lycaeides* (Lepidoptera: Lycaenidae): implications for taxonomy and historical biogeography. *Annals of the Entomological Society of America*, **101**, 172–180.
- Martin SH, Dasmahapatra KK, Nadeau NJ *et al.* (2013) Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, **23**, 1817–1828.
- Mavárez J, Salazar CA, Bermingham E *et al.* (2006) Speciation by hybridization in *Heliconius* butterflies. *Nature*, **441**, 868–871.
- Nadeau NJ, Martin SH, Kozak KM *et al.* (2013) Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Molecular Ecology*, **22**, 814–826.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Salazar C, Baxter SW, Pardo-Diaz C *et al.* (2010) Genetic evidence for hybrid trait speciation in *Heliconius* butterflies. *PLoS Genetics*, **6**, e1000930.
- Seehausen O, Butlin RK, Keller I *et al.* (2014) Genomics and the origin of species. *Nature Reviews Genetics*, **15**, 176–192.

This article was written and figures prepared by N.N. except as specified in the text (photo credits).

doi: 10.1111/mec.12877