

Tracing the recombination and colonization history of hybrid species in space and time

C. LEXER and K. N. STÖLTING

Unit of Ecology & Evolution, Department of Biology,
University of Fribourg, Chemin du Musée 10, CH-1700
Fribourg, Switzerland

Hybrid speciation has long fascinated evolutionary biologists and laymen alike, presumably because it challenges our classical view of evolution as a 'one-way street' leading to strictly tree-like patterns of ancestry and descent. Homoploid hybrid speciation (HHS) has been a particularly interesting puzzle, as it appears to occur extremely rapidly, perhaps within less than 50 generations (McCarthy *et al.* 1995; Buerkle *et al.* 2000). Nevertheless, HHS may sometimes involve extended or repeated periods of recombination and gene exchange between populations subject to strong divergent natural selection (Buerkle & Rieseberg 2008). Thus, HHS provides a highly interesting setting for understanding the drivers and tempo of adaptive divergence and speciation in the face of gene flow (Arnold 2006; Rieseberg & Willis 2007; Nolte & Tautz 2009). In the present issue of *Molecular Ecology*, Wang *et al.* (2011) explore a particularly challenging issue connected to HHS: they attempt to trace the colonization and recombination history of an ancient (several MYA) hybrid species, from admixture and recombination in the ancestral hybrid zone to subsequent range shifts triggered by tectonic events (uplift of the Tibetan plateau) and climatic shifts (Pleistocene ice ages). This work is important because it addresses key issues related to the origin of the standing genetic variation available for adaptive responses (e.g. to climate change) and speciation in temperate species, which are topics of great current interest (Rieseberg *et al.* 2003; Barrett & Schluter 2008; de Carvalho *et al.* 2010).

Keywords: colonization, homoploid hybrid speciation, niche shift, phylogeography, recombination, standing genetic variation

Wang *et al.* (2011) focus on three closely related conifer species. Conifers provide a very special setting for genetic studies, in which chloroplast (cp) DNA is inherited through the paternal and mitochondrial (mt) DNA through the maternal lineage. Extensive experience with the use of plant cp and mt genomes in molecular ecology (Vendramin *et al.* 1996; Duminil *et al.* 2002) implies that selected regions of these DNA molecules can easily be visualized in large numbers of individuals, which facilitates large-scale phylogeographic studies. Within the limits provided by uniparentally inherited genetic markers, this approach allowed Wang *et al.* to trace the colonization and recombination history of *P. densata* in space and time at a depth rarely (never?) seen in any other homoploid hybrid species system.

(a)



(b)



Fig. 1 (a) The homoploid hybrid species *Pinus densata* in its high-altitude environment on the Tibetan plateau (up to 4200 m elevation). (b) *P. densata* is among the most dominant and ecologically successful homoploid hybrid species known to date. Photo credit: Jian-Feng Mao.

In previous studies, the authors have shown that *P. densata* is a homoploid hybrid between two other Asian pine species, *P. tabuliformis* and *P. yunnanensis* (Wang *et al.* 1990, 2001) (Fig. 1). The existence of a third, unidentified potential parental taxon was sometimes hypothesized in early studies (Wang *et al.* 2001), although a third species may not be required to explain observed allelic diversity considering that novel, private alleles frequently arise in hybrid zones (Schilthuizen *et al.* 1999). An important aspect of these earlier studies is that both inter- and intra-locus linkage disequilibrium (LD) in *P. densata* was found to be extremely low. This provides a good indicator for the cessation of gene flow between the homoploid hybrid taxon and its parents (Wang *et al.* 2001; Ma *et al.* 2006).

With respect to ecology, the clear elevational differences among the three taxa (*P. densata*: Tibetan plateau at 2700–4200 m; both parental species: lower elevations in north-central and south-western China) speak for strong ecological divergence with regard to temperature and other eco-variables correlated to altitude. In fact, the three species are significantly differentiated for numerous fitness-related ecological traits (Mao *et al.* 2009) and a recent niche modelling exercise revealed a distinct niche shift in the hybrid taxon (Mao & Wang 2011). This clear ecological divergence linked to elevation (made possible by the uplift of the Tibetan plateau initiated ca. 20 MYA ago) strongly suggests that ecological barriers were the primary drivers of HHS in *P. densata*, and this timeframe was confirmed by the coalescence times for several genetic loci. Geographic barriers may have contributed but are unlikely to be substantial in high-gene flow species such as wind-pollinated temperate trees (Lexer *et al.* 2005). A survey of reproductive fitness components also indicated the absence of strong fertility barriers between the hybrid and its parents (Mao *et al.* 2009), as one might expect for *Pinus* species. So even if minor chromosomal rearrangements are suggested by molecular cytogenetic work based on ribosomal DNA (Liu *et al.* 2003), their contribution to reproductive isolation among *P. densata* and its parents appears negligible.

In this study, Wang *et al.* greatly extend their earlier phylogeographic work (Song *et al.* 2003) to track the evolutionary history of this enigmatic hybrid taxon, based on an unprecedented sampling effort in the geographically challenging south-eastern Tibetan Plateau. At first sight, their results may seem inconspicuous: they find pronounced geographic patterns for maternally inherited, seed dispersed mtDNA in the hybrid taxon (*P. densata*) and weaker structure for paternally inherited, pollen-dispersed cpDNA, as one might expect (Petit *et al.* 2005). Only a second, deeper look reveals the full picture of the *mosaic-like colonization and hybridization history* of this homoploid hybrid pine, sketched backward in time here.

Patterns of cpDNA sharing indicate that populations of *P. densata* have experienced pollen flow from both parents in different parts of their range. Notably, central populations of the hybrid taxon experienced pollen flow from its parent *P. yunnanensis* only recently, whereas the western portion of the hybrid species' range is highly differentiated

for cpDNA and has thus been isolated from both parents for much longer. Maternal mtDNA lineages provide clues on processes dating even further back in time: mtDNA variation in the hybrid taxon is structured into three major haplogroups, thus pointing to the presence of multiple differentiated lineages likely resulting from independent hybrid speciation events, as suspected previously (Wang *et al.* 2001). Greatly elevated mtDNA diversity and differentiation in the north-eastern range of *P. densata* points to the location of the *ancestral hybrid zone* from which this hybrid taxon emerged. The demographic stability of populations in this region is supported by the genetic signature inherent in mtDNA and cpDNA mismatch distributions, which are incompatible with recent population expansion.

Wang *et al.* also tested for mtDNA recombination using a four-gamete test (Hudson & Kaplan 1985) under simplifying assumptions. Recombination is known to occur in plant mt genomes (Houliston & Olson 2006) and has been shown in other conifers (Jaramillo-Correa & Bousquet 2005). Here, the authors used this prior knowledge in a clever way to gain further insights into the ancient hybridization history of this Asian hybrid pine. Assuming the absence of technical artefacts, they detected a likely recombination event between mtDNA haplotypes derived from each parental species, involving two ancient mitotypes found at high frequencies in the inferred ancestral hybrid zone. Thus, by combining information from two organellar genomes with contrasting modes of inheritance, Wang *et al.* were able to trace the evolutionary history of this hybrid pine, including identification of the putative ancestral hybrid zone and inference of recurrent westward colonization events across the Tibetan plateau. Inference of the colonization process represents in itself another highly interesting aspect of this study.

Examination of allelic diversities from east to west revealed several *unexpected shifts* in mt DNA haplotype frequencies, with haplotypes that were rare or even absent in the ancestral hybrid zone reaching high frequencies in more recently colonized westerly localities. These unexpected shifts were plausibly explained by the ‘surfing’ of rare alleles into newly colonized territories during range expansion (Klopfstein *et al.* 2006; Excoffier *et al.* 2009). The characteristic spatial patterns of haplotype groups, and of ‘surfing’ events of haplotypes within groups, were used to put forward the hypothesis that westward expansion of *P. densata* across the Tibetan plateau likely occurred in *more than a single wave*. Keeping in mind limitations in sample sizes, these findings may inspire other students interested in the origin and maintenance of biological diversity in mountain systems. Although the impact of mountain building processes on macroevolutionary patterns of diversity is increasingly well understood (Hoorn *et al.* 2010), we know little about its impact on exactly those population genetic processes that represent the actual ‘engine’ of divergence and speciation.

The biggest limitation of this study—the absence of information from nuclear genomic markers—also points to a potentially highly rewarding avenue for future research.

Future work on *P. densata* and other homoploid hybrid species should now turn to large-scale nuclear genomic surveys to trace the history of *nuclear admixture and recombination* in space and time. Knowledge of genomic distributions of recombination break points or *junctions* would enable molecular ecologists to infer the timing of the initial admixture and recombination events that triggered HHS and to unravel the migration episodes that have shaped today's "genomic mosaics" until reproductive isolation was (close to) complete. The basic analytical tools for this purpose have long been available (Fisher 1954; Chapman & Thompson 2002) and are currently being refined, driven primarily by progress in human medical genetics (reviewed by Buerkle & Lexer 2008; Winkler *et al.* 2010). Ultra High Throughput Sequencing (UHTS) methods to trace genome-wide recombination events in non-model species are now emerging as well (Baird *et al.* 2008; Hohenlohe *et al.* 2010). With the availability of new technologies affording ever greater genomic coverage and sequencing depth, it is likely that sufficient numbers of single nucleotide polymorphisms (SNPs) will soon be available for speciation genomic studies despite the complexity of *Pinus* genomes.

Recent verbal models and reviews of HHS (Jiggins *et al.* 2008; Nolte & Tautz 2009) suggest that this process is much less restricted than previously assumed (Buerkle *et al.* 2000). Also, there are multiple routes in addition to HHS through which hybridization and genetic exchange may impact on the ecology & evolution of wild species of animals and plants (Arnold 2006). Results on tree hybrid zones further indicate that the novel variation provided by genomic admixture not only affects the hybridizing species themselves but also the communities and ecosystems that depend on them (Whitham *et al.* 2006). Tracing the recombination and colonization history of homoploid hybrid lineages would thus seem like a worthwhile goal for current biology. This study on a Tibetan hybrid pine shows what can be achieved with traditional phylogeographic tools. We look forward to seeing new, exciting studies that address these issues with the complete 'next generation' tool box now available to students of the ecology and evolution of non-model species.

Acknowledgements

We thank Loren Rieseberg, Mike Arnold, and Remy Petit for numerous stimulating discussions over the years, Remy Petit and Xiao-Ru Wang for critical reading of the manuscript, and Jian-Feng Mao for photographs of *Pinus densata* and its habitat.

References

Arnold ML (2006) *Evolution Through Genetic Exchange*. Oxford University Press, Oxford.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD Markers. *PLoS ONE*, **3**, 1–7.

Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, **23**, 38–44.

Buerkle CA, Lexer C (2008) Admixture as the basis for genetic mapping. *Trends in Ecology and Evolution*, **23**, 686–694.

Buerkle CA, Rieseberg LH (2008) The rate of genome stabilization in homoploid hybrid species. *Evolution*, **62**, 266–275.

Buerkle CA, Morris RJ, Asmussen MA, Rieseberg LH (2000) The likelihood of homoploid hybrid speciation. *Heredity*, **84**, 441–451.

Chapman NH, Thompson EA (2002) The effect of population history on the lengths of ancestral chromosome segments. *Genetics*, **162**, 449–458.

de Carvalho D, Ingvarsson PK, Joseph J *et al.* (2010) Admixture facilitates adaptation from standing variation in the European aspen (*P. tremula* L.), a widespread forest tree. *Molecular Ecology*, **19**, 1638–1650.

Duminil J, Pemonge MH, Petit RJ (2002) A set of 35 consensus primer pairs amplifying genes and introns of plant mitochondrial DNA. *Molecular Ecology Notes*, **2**, 428–430.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285–298.

Fisher RA (1954) A fuller theory of "junctions" in inbreeding. *Heredity*, **8**, 187–197.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA *et al.* (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, 1–23.

Hoorn C, Wesselingh FP, ter Steege H, Bermudez MA, Mora A *et al.* (2010) Amazonia Through Time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.

Houliston GJ, Olson MS (2006) Nonneutral evolution of organelle genes in *Silene vulgaris*. *Genetics*, **174**, 1983–1994.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

Jaramillo-Correa JP, Bousquet J (2005) Mitochondrial genome recombination in the zone of contact between two hybridizing conifers. *Genetics*, **171**, 1951–1962.

Jiggins CD, Salazar C, Linares M, Mavarez J (2008) Hybrid trait speciation and *Heliconius* butterflies. *Philosophical Transactions of the Royal Society B*, **363**, 3047–3054.

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.

Lexer C, Fay M, Joseph J, Nica M-S, Heinze B (2005) Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): the role of ecology and life history in gene introgression. *Molecular Ecology*, **14**, 1045–1057.

Liu ZL, Zhang D, Hong DY, Wang XR (2003) Chromosomal localization of 5S and 18S-5.8S-25S ribosomal DNA sites in five Asian pines using fluorescence in situ hybridization. *Theoretical and Applied Genetics*, **106**, 198–204.

Ma XF, Szmidt AE, Wang XR (2006) Genetic structure and evolutionary history of a diploid hybrid pine *Pinus densata* inferred from the nucleotide variation at seven gene loci. *Molecular Biology and Evolution*, **23**, 807–816.

Mao JF, Wang XR (2011) Distinct niche divergence characterizes the homoploid hybrid speciation of *Pinus densata* on the Tibetan Plateau. *American Naturalist*, **177**, 424–439.

Mao JF, Li Y, Wang XR (2009) Empirical assessment of the reproductive fitness components of the hybrid pine *Pinus densata* on the Tibetan Plateau. *Evolutionary Ecology*, **23**, 447–462.

McCarthy EM, Asmussen MA, Anderson WW (1995) A theoretical assessment of recombinational speciation. *Heredity*, **74**, 502–509.

Nolte AW, Tautz D (2010) The onset of hybrid speciation. *Trends in Genetics*, **26**, 54–58.

- Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689–701.
- Rieseberg LH, Willis JH (2007) Plant speciation. *Science*, **317**, 910–914.
- Rieseberg LH, Raymond O, Rosenthal DM *et al.* (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science*, **301**, 211–216.
- Schilthuizen M, Hoekstra R, Gittenberger E (1999) Selective increase of a rare haplotype in a land snail hybrid zone. *Proceedings of the Royal Society of London B*, **266**, 2181–2185.
- Song BH, Wang XQ, Wang XR, Ding KY, Hong DY (2003) Cytoplasmic composition in *Pinus densata* and population establishment of the diploid hybrid pine. *Molecular Ecology*, **12**, 2995–3001.
- Vendramin GG, Lelli L, Rossi P, Morgante M (1996) A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. *Molecular Ecology*, **5**, 595–598.
- Wang XR, Szmidt AE, Lewandowski A, Wang ZR (1990) Evolutionary analysis of *Pinus densata*, a putative hybrid .1. Allozyme variation. *Theoretical and Applied Genetics*, **80**, 635–640.
- Wang XR, Szmidt AE, Savolainen O (2001) Genetic composition and diploid hybrid speciation of a high mountain pine, *Pinus densata*, native to the Tibetan plateau. *Genetics*, **159**, 337–346.
- Wang B, Mao JF, Gao J, Zhao W, Wang XR (2011) Colonization of the Tibetan plateau by the homoploid hybrid pine *Pinus densata*. *Molecular Ecology*, **20**, 3796–3811.
- Whitham TG, Young WP, Martinsen GD, Gehring CA, Schweitzer JA *et al.* (2006) A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, **7**, 510–523.
- Winkler CA, Nelson GW, Smith MW (2010) Admixture mapping comes of age. *Annual Review of Genomics and Human Genetics*, **11**, 65–89.