

NEWS AND COMMENTARY

Population structure

Cohort-structured tree populations

S Jansson and PK Ingvarsson

Heredity (2010) 105, 331–332; doi:10.1038/hdy.2010.101; published online 18 August 2010

Forest tree populations are typically thought to be good approximations of ‘ideal populations’. In this issue of *Heredity*, however, Slavov *et al.* (2010) show that cryptic population structure may be more common than previously thought and that it has to be accounted for in association studies.

Forest trees as a group can clearly be regarded as winners in evolutionary terms. Unless low temperatures or arid conditions prevent colonization by plants—like in deserts or on the tundra—different forest tree species have been able to successfully compete for the resources that limit plant growth (light, nutrients and water) in the majority of terrestrial ecosystems. From the tropical rain forests to the boreal conifer forests, trees are dominating members of the ecosystem and materials produced by trees have been and are of paramount importance for many other species on Earth, including our own. For example, imagine how human settlement in most parts of the world could have occurred without access to firewood, boats, tools and houses produced from wood.

From an ecological and genetic point of view, rain forests and boreal forests are very different. A key feature of the tree flora in a rain forest is the extremely high species diversity, with a huge number of species coexisting. Even a specialist walking in the Amazon rain forest will have a hard time identifying the genus of some of the trees they encounter. The temperate forests, on the other hand, are species-poor and the vast boreal forests often contain just a handful of tree species. A day’s walk in a Northern Lapland forest may result in a species list including at most 10 species, but where the majority of these are occurring in huge numbers.

Genetic studies in forest trees have for the most part revealed abundant within-population genetic diversity, low between-population genetic differentiation and low associations between different genomic regions (linkage disequilibrium) (Ingvarsson, 2010). These factors all point to forest trees having relatively large effective population sizes, further compounded by the fact that they are also extremely long-lived

organisms. Furthermore, forest trees are typically wind pollinated and have wind-dispersed seeds, resulting in gene flow extending over great geographical distances (Sork and Smouse, 2006). For example, as the climate have warmed following the last glaciation, ‘mass invasion’ of forest trees has rapidly occurred into areas that were previously glaciated, and although relatively few generations have passed, the efficient mixing of alleles from different glacial refugia has allowed for the establishment of different genotypes adapted to many different environments (for an example in *Populus*, see De Carvahlo *et al.* (2010)).

The presence of abundant genetic variation and low linkage disequilibrium opens up great opportunities for genetic mapping in forest trees. Owing to their generation time, the utility of traditional mapping methods, like QTL mapping, is limited in forest trees as it takes a long time to develop segregating mapping populations (Neale and Ingvarsson, 2008). Therefore, association mapping (AM) has been suggested as an alternative approach for dissecting the genetic architecture in forest trees and it has been argued that, using AM, it will be possible to achieve very high resolution due to the very low extent of linkage disequilibrium.

One major issue with AM is that population structure give rise to spurious associations and so methods are needed that account for the effects of population structure (Yu *et al.*, 2006). Such population structure could be the result of, for example, admixture when previously isolated populations that have experienced different selection pressures meet. Subtle allele frequency differences between populations will then result in such loci showing associations to any phenotypic traits that also differ between populations, even if the alleles at these loci are not causal (see for example Zhao *et al.*, 2007).

Sexual propagation of trees is often very sporadic; a spruce tree in harsh environments may only flower once every decade. Furthermore, conditions for seedling establishment vary greatly from year to year, and if conditions are unfavourable when seeds are shed, no

recruitment may take place even with an abundant seed crop. The sporadic flowering and highly variable conditions for seedling establishment will therefore result in large year-to-year variation in how many new trees are recruited to a population. In the extreme case, a cohort of offspring, recruited under particularly suitable conditions, can dominate the landscape for centuries. Slavov *et al.* (2010) present a study in which population substructure in black cottonwood (*Populus trichocarpa*) is dissected. *P. trichocarpa* is the only tree species in which the completely sequenced genome has been published so far—although several others are in the pipeline—and is therefore a good model system for these studies. At two different sites, at least two coexisting subpopulations were detected in what superficially appears to be more or less continuous cottonwood stands, and the authors suggest that seedling establishment may be one critical factor explaining this cryptic substructure.

The presence of population structure at very small spatial scales has implications for association studies in *P. trichocarpa*. Population substructure will have to be taken into account, even in situations where population structure has previously been thought to be virtually nonexistent. Whether or not the same will turn out to be a serious concern in other tree species remains to be elucidated. Other tree species, such as Douglas fir (Eckert *et al.*, 2009), *Eucalyptus* (Thumma *et al.*, 2009), European aspen (Ingvarsson *et al.*, 2008) or Loblolly Pine (Gonzalez-Martinez *et al.*, 2007) may have more continuous distributions and larger effective population sizes that could perhaps make analysis in these species more robust. However, it is clear that the issue of potentially cryptic population structure will need to be addressed as population structure cannot be assumed to be absent. The presence of population structure does not impose any real limit on the utility of AM in tree species in general—AM is widely used also in the highly structured human population (McCarthy *et al.*, 2008)—and tools have been developed that allow for analyses in which population structure is explicitly taken into account. What this study demonstrates is that careful studies of the extent of population structure are needed to prevent the accumulation of false positives in future AM studies of forest trees.

Conflict of interest

The authors declare no conflict of interest.

S Jansson is at the Department of Plant Physiology, Umeå Plant Science Centre, Umeå University, Umeå, Sweden and PK Ingvarsson is at the Department of Ecology and Environmental Science, Umeå Plant Science Centre, Umeå University, Umeå, Sweden.

e-mail: stefan.jansson@plantphys.umu.se

- De Carvahlo D, Ingvarsson PK, Joseph J, Suter L, Sedivy C, Macaya-Sanz D *et al.* (2010). Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Molec Ecol* **19**: 1638–1650.
- Eckert AJ, Bower AD, Wegrzyn JL, Pande B, Jermstad KD, Krutovsky KV *et al.* (2009). Association genetics of coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics* **182**: 1289–1302.
- Gonzalez-Martinez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB (2007). Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics* **175**: 399–409.
- Ingvarsson PK (2010). Nucleotide polymorphism, linkage disequilibrium and complex trait dissection in *Populus*. In: Jansson S, Bhalerao R, Groover AT (eds). *Genetics and Genomics of Populus*. Springer: NY, pp 91–112.
- Ingvarsson PK, Garcia MV, Luquez V, Hall D, Jansson S (2008). Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae). *Genetics* **178**: 2217–2226.
- McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA *et al.* (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* **9**: 356–369.
- Neale DB, Ingvarsson PK (2008). Population, quantitative and comparative genomics of adaptation in forest trees. *Curr Opin Plant Biol* **11**: 149–155.
- Slavov GT, Leonardi S, Adams WT, Strauss SH, DiFazio SP (2010). Population substructure in continuous and fragmented stands of *Populus trichocarpa*. *Heredity* **105**: 348–357.
- Sork VL, Smouse PE (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecol* **21**: 821–836. 11.
- Thumma BR, Matheson BA, Zhang D, Meeske C, Meder R, Downes GM *et al.* (2009). Identification of a *cis*-acting regulatory polymorphism in a Eucalypt COBRA-like gene affecting cellulose content. *Genetics* **183**: 1153–1164.
- Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF *et al.* (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* **38**: 203–208.
- Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C *et al.* (2007). An *Arabidopsis* example of association mapping in structured samples. *PLoS Genet* **3**: e4.

Editor's suggested reading

- García-Gil MR, Olivier F, Kamruzzahan S, Waldmann P (2009). Joint analysis of spatial genetic structure and inbreeding in a managed population of Scots pine. *Heredity* **103**: 90–96.
- Derory J, Scotti-Saintagne C, Bertocchi E, Le Dantec L, Graignic N, Jauffres A *et al.* (2009). Contrasting relationships between the diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. *Heredity* **104**: 438–448.
- Tollefsrud MM, Sonstebo JH, Brochmann C, Johnsen Ø, Skrøppa T, Vendramin GG (2009). Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European *Picea abies*. *Heredity* **102**: 549–562.
- González-Martínez SC, Huber D, Ersoz E, Davis JM, Neale DB (2008). Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. *Heredity* **101**: 19–26.