



ELSEVIER

Genomic diversity in forest trees

Outi Savolainen and Tanja Pyhäjärvi

Forest trees in general are out-crossing, long-lived, and at early stages of domestication. Molecular evolution at neutral sites is very slow because of the long generation times. Transferring information between closely related conifer species is facilitated by high sequence similarity. At the nucleotide level, trees have at most intermediate levels of variation relative to other plants. Importantly, in many species linkage disequilibrium within genes declines within less than 1000 bp. In contrast to the slow rate of neutral evolution, large tree populations respond rapidly to natural selection. Detecting traces of selection may be easier in tree populations than in many other species. Association studies between genotypes and phenotypes are proving to be useful tools for functional genomics.

Addresses

Department of Biology, FIN-90014, University of Oulu, Finland

Corresponding author: Savolainen, Outi (outi.savolainen@oulu.fi) and Pyhäjärvi, Tanja (tanja.pyhajarvi@oulu.fi)

Current Opinion in Plant Biology 2007, **10**:162–167

This review comes from a themed issue on
Genome studies and molecular genetics
Edited by Stefan Jansson and Edward S Buckler

Available online 9th February 2007

1369-5266/\$ – see front matter
Published by Elsevier Ltd.

DOI [10.1016/j.pbi.2007.01.011](https://doi.org/10.1016/j.pbi.2007.01.011)

Introduction

Forest trees are ecologically dominant in many ecosystems. Many are of economical importance and their domestication has started only recently. Tree breeding in general is slow, as most species have long generation times. Thus, it is clearly desirable to find genetic and genomic methods that could speed up the detection of functionally important regions in the genome.

Molecular population genetic studies of the patterns of variation, and of the evolutionary forces that gave rise to them are providing insights into the special evolutionary features of trees. The genetic structure of populations defines how the natural variation can be used to detect and map genomic areas of functional importance. For instance, the scope for association studies needs to be assessed against population phenotypic variation and the extent of linkage disequilibrium (LD) [1,2].

Some tree species, such as *Populus*, have very small genomes, about four times the size of the *Arabidopsis*

genome [3^{••}], but conifers have genomes more than 100 times the size of the *Arabidopsis* genome [4]. The nearly full *Populus* sequence provides access to all areas of the genome [3^{••}]. Since the first *Pinus taeda* cDNA sequences were published [5], extensive expressed sequence tag (EST) resources have also become available for many other species, including *Picea* [6]. Other genomic resources are rapidly improving.

This review examines nucleotide polymorphism of coding regions in temperate conifers and of *Populus*. We also consider the role played by different evolutionary factors in shaping this diversity, and especially the potential for detecting selection. The implications for functional genomics studies are also discussed.

Life history traits and other characteristics of trees that are relevant to population genetics

Most trees are predominantly outcrossing because of self-incompatibility mechanisms or severe early (and late) inbreeding depression [7]. Pollen flow and seed flow are typically very efficient in trees [8]. The generation times are long. Northern conifers reach reproductive maturity late (at close to 20 years) and have lifespans of up to hundreds of years. *Populus* can reach reproductive maturity in 4–6 years [3^{••}], but it can also have a very long life span. Today, many commercially important species have large natural distribution areas; for instance Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*) and larches (*Larix* spp.) cover much of Eurasia. Other tree species, such as Monterey pine (*Pinus radiata*), have small fragmented natural distributions. Studies of fossil records, chloroplast DNA (cpDNA), mitochondrial DNA (mtDNA) and allozymes indicate, however, that glacial cycles have caused range contractions and expansions for most temperate species [9,10]. Important phenotypic traits, such as timing of growth and cold tolerance, have clinal variation across environmental gradients in many widely distributed tree species [11,12]. In contrast to many crop species, most commercially important forest species are essentially found in natural populations, with few traces of domestication and artificial selection.

Neutral evolution of tree genomes is slow

At nucleotide sites not influenced by selection (neutral sites), the rate of evolution is governed by the mutation rate. The nucleotide substitution rate at synonymous sites provides a direct estimate of the mutation rate. If we know the level of divergence of two species at the DNA level and have an estimate of their time of divergence, we can obtain an estimate of the rate of nucleotide

substitution. In *Arabidopsis*, the rate of synonymous substitution has been estimated to be 1.5×10^{-8} per site per year [13]. In *Populus*, the evolutionary rate has been estimated to be only one-sixth of this rate [3^{••}]. A recent careful calibration of *Pinus* fossils resulted in an estimate of 0.7×10^{-9} /site/year, a very slow rate, in this genus [14[•]]. Thus, pine species that diverged tens of millions of years ago might differ at synonymous sites by less than 5%. However, when we convert these substitution or mutation rates to rates per generation (i.e. multiply them by a generation-time estimate, for instance 20 years), they are close to the rates estimated for many annual plant species.

Genome-wide effects on nucleotide diversity

When examining patterns of nucleotide polymorphisms, we first contrast them against simple models, and if needed, then examine additional evolutionary factors. Most plant biologists are particularly interested in the effects of natural selection on specific loci, for instance, disease resistance loci. However, it is first necessary to examine whether those factors that have a genome-wide influence could give rise to the observations. Many effects of changes in population size or migrations can resemble the patterns of polymorphism that are generated by natural selection. Caution is required in interpreting the results before inferring that selection has been operating.

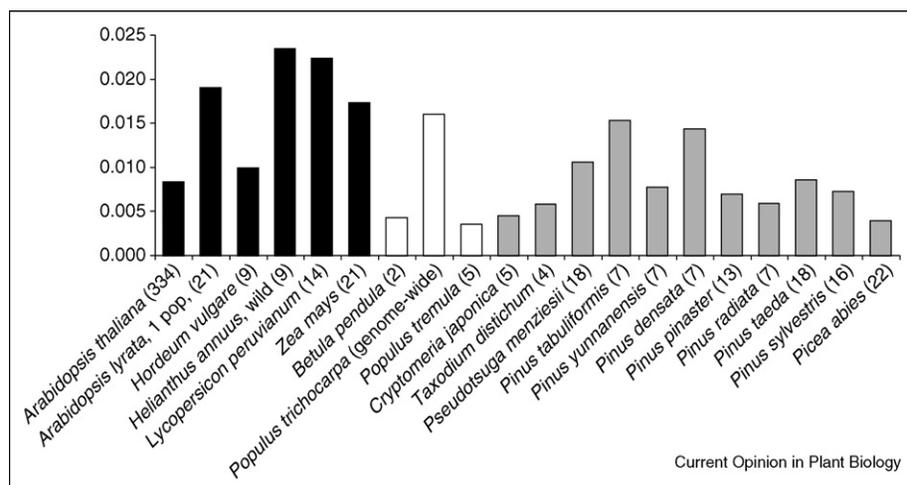
The standard neutral model considers finite populations, where mating is random and there is no population subdivision. All individuals are equally likely to survive and reproduce. Synonymous variants at nucleotide sites can be assumed to evolve according to this model. The level of nucleotide diversity in this neutral model is governed by the mutation rate and the size of the population. Other things being equal, we expect more nucleo-

tide diversity in large populations than in small ones, because random sampling (genetic drift) reduces variation in small populations more effectively. Given the mutation rates we considered above, the large effective sizes, and earlier findings on much variation at other marker loci, we would expect a high level of genetic diversity in an equilibrium population [15]. Figure 1 shows the distribution of estimates of silent nucleotide diversity in trees and other plants ([3^{••},16–27], G Géré *et al.*, pers. comm. [*P. pinaster*]). The diversity estimates of trees are highly variable, but no higher than those for other plants. This suggests that many tree populations might not be at an equilibrium situation, and that factors other than current population sizes, such as historical changes in population size, could have reduced the genome-wide diversity.

The theory above also predicts that there will be variation in the level of polymorphism across different loci, depending on the mutation rate. The divergence between *Cryptomeria japonica* and *Taxodium distichum* predicts levels of variation within *Cryptomeria* for most loci (Figure 2a, based on results from [28]). Loci that are much more or much less variable than predicted on the basis of divergence might be influenced by other evolutionary factors (see circled dot in Figure 2a).

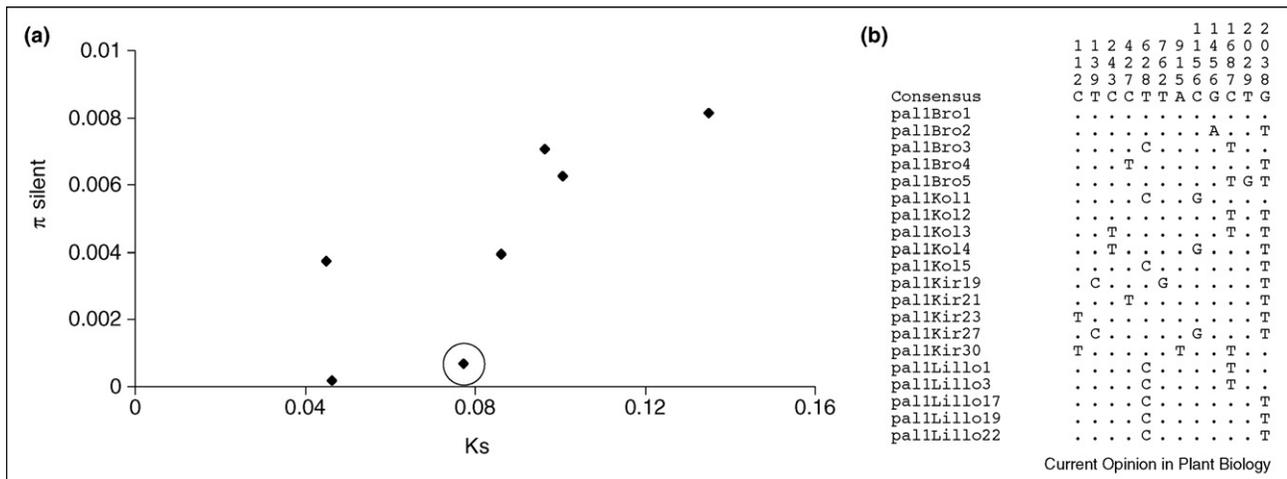
Closely linked nucleotide sites, such as sites within a single gene, are not expected to evolve independently. The statistical correlation between variant frequencies at two nucleotide sites (i.e. linkage disequilibrium, LD) is greater between closely linked sites than between those that are further apart. Figure 2b shows a small example of a set of sequences across 2000 bp of the phenylalanine ammonia-lyase gene *pall* of *Pinus sylvestris* trees from different parts of Europe [15]. The different nucleotide

Figure 1



Silent site nucleotide diversity (synonymous, noncoding, or both) estimates per base pair in other plants (black), in angiosperm trees (white) and in conifers (black). The number of loci is in parentheses. See text for references.

Figure 2



(a) Synonymous nucleotide divergence between *Cryptomeria japonica* and *Taxodium distichum* (x-axis) and nucleotide diversity in *Cryptomeria japonica* (y-axis), based on data from [28]. Regression analysis: $F = 8.49$, $P = 0.03$, $r^2 = 0.63$. Circled point has low diversity relative to divergence. **(b)** *Pinus sylvestris* haplotypes of the *pal1* gene from northern Finland (Kol), southern Finland (Bro) and Russia (Kir) and Spain (Lil) showing variable sites, based on data from [15].

sites are at least nearly independent of each other. To date, this kind of pattern of low LD and rapid decay (within 1000 bp or faster) has been found in most loci in *Pinus taeda* [29], *Populus tremula* [30], and *P. abies* [31[•]]. This is in strong contrast to the situation in highly inbred species such as *Arabidopsis thaliana* [32], where LD extends over distances of more than 20 kb in worldwide samples [33], but resembles the situation found in maize [34].

Tree populations have large distribution ranges and, to date, many species have been found to have rather uniform frequencies of different alleles across the range (compare the populations for *pal1* shown in Figure 2b). Within Europe, most nucleotide diversity resides within populations, and less than 5% is found between populations for *P. sylvestris* [11] and *Populus tremula* [30], whereas *P. abies* shows slightly more divergence between central and northern Europe [31[•]]. For comparison, in a group of Central European populations of *Arabidopsis lyrata*, 17% of the variation was between the populations, indicating a much higher divergence [35].

The very slow neutral processes in the large long-lived tree populations are an important feature of tree populations [29,36]. After population size changes, populations reach the new equilibria between mutation and genetic drift at a rate governed by $4N$ (where N is the population size). Before the new equilibrium is reached, the population's genetic makeup differs from the equilibrium expectation. A population-size bottleneck first eliminates rare alleles, whereas a population expansion phase results in an increased number of nucleotide sites that have low frequency variants. Such consequences of population size

changes have genome-wide effects. For example, an excess of rare variants (along with other evidence) in the nuclear genes of *P. abies* was interpreted as a consequence of expansion after an ancient bottleneck [31[•]].

Detecting selection in tree genomes

Natural selection leaves many traces in the genome. It is not easy to detect these signals, however, because many demographic events can result in similar patterns of polymorphism. For instance, the spread of a new favorable allele (selective sweep) is accompanied by a reduction of variability in areas surrounding the selected site [33], as has been found, for instance, in the maize *teosinte branched 1 (tb1)* gene [37]. The reduced variability at one of the loci in Figure 2a may be due to such a selective sweep. However, such reduced diversity could also be due to a demographic departure from the standard neutral model. An excess of rare alleles can reflect the spread of a new favorable allele, but it can also follow from an expansion phase of the population. A high frequency of nucleotide sites that have intermediate frequencies of variants could be due to selection that maintains many alleles, such as at self-incompatibility loci, but it could also be due to a recent bottleneck.

Natural selection can also cause an increase in the local level of disequilibrium, as has been found in *A. thaliana* [32,33], but LD also varies because of variation in recombination rates or due to population structure. Differential selection between populations in their local conditions can give rise to high differentiation, but this can also result from the isolation of populations. Currently, the best way to detect the effects of selection on a locus is to demonstrate that its pattern of polymorphism differs

Table 1

Detecting selection in trees: number of loci studied and kinds of selection detected.

Species	Number of loci	Selection		References
		Number of locally adapted loci	Positive/directional	
<i>Betula pendula</i>	2			[24]
<i>Populus tremula</i>	10	1	6	[3**,25]
<i>Cryptomeria japonica</i>	7		1	[26,28]
<i>Taxodium distichum</i>	4			[26]
<i>Pseudotsuga menziesii</i>	18		1	[20]
<i>Pinus tabuliformis</i>	7			[19]
<i>Pinus yunnanensis</i>	7			[19]
<i>Pinus densata</i>	7		1	[19]
<i>Pinus pinaster</i>	8		2	[22]
<i>Pinus radiata</i>	8		1	[22]
<i>Pinus taeda</i>	35		1	[21,29]
<i>Pinus sylvestris</i>	16			([11,15]; T Pyhäjärvi, unpublished)
<i>Picea abies</i>	22			[31*]
Total	151	1	13	8

significantly from the genome-wide polymorphism [38*]. Attempts to do this are being made using the very extensive datasets on populations of humans [39], maize [40] and *A. thaliana* [41]. These analyses require multi-locus datasets that are rarely available in trees.

We compiled a list of reports of selection on individual loci in trees (Table 1). Of the studied loci, 15% have been reported to be under some kind of selection, whereas in *A. thaliana*, selection was reported at 38% of the genes [38*]. Among the many possible causes for such a difference, apparent patterns of selection due to demography may be less likely in the large random mating populations of forest trees than in other species. The mechanisms of selection in general are not known. A pattern of clinal variation at the *PhyB2* locus in *Populus tremula*, against a background of low differentiation at other loci, could be related to timing of growth [42]. Different genes might govern similar traits in different species: the homolog of *PhyB* in *Pinus sylvestris* showed no evidence of clinal variation [11].

While the molecular traces of selection are still poorly understood in trees, studies on quantitative trait variation have demonstrated that selection can be highly efficient in large tree populations. For instance, many northern European populations of trees (e.g. willows, birches, spruces and pines) have evolved genotypes whose growth is arrested at very long day lengths [43]. Such locally adapted ecotypes must have become frequent during the post glacial recolonization of the northern areas, just a few thousand years ago. The evolution in response to natural selection is very rapid compared to the slow rate of neutral evolution.

Even more can be learnt about selection when the nucleotide polymorphism can be correlated with phenotypic

variation. The large, random-mating, low LD populations of trees are very well suited for association genetic studies [1,2]. The first cases of genetic polymorphisms associated with wood quality in eucalypts and loblolly pine have recently been described [44*,45]. Sequence studies aid in finding interesting areas in the genome, correlations with phenotype might help us to understand (at least current) patterns of selection.

Conclusions: the promise of natural variation

The neutral genome of species with long life spans evolves very slowly, and the large populations reach equilibrium at a slow rate compared to shorter lived plants. Thus, the genomes of long-lived trees retain the consequences of demographic events for very long times. In contrast, natural selection can have very rapid effects on the large tree populations. In tree species that have rather low average differentiation between populations and low LD, it may be easier to detect the effects of natural selection, than in species that have more complex population history and structure.

Detailed studies of tree population genomics are at an early phase. Increasing genomic resources and greater sampling across multiple populations and species will allow a much better understanding of the genome-wide patterns, against which to search for patterns caused by selection. Research on sequence variation, combined with work on natural phenotypic variation, will significantly contribute to dissecting the genetics of adaptations of tree populations.

Acknowledgements

We acknowledge financial support from TREESNIPS (QLRT-2001-01973), the Finnish Graduate School in Population Genetics and the Biosciences and Environment Research Council of Finland. Katri Kärkkäinen, Helmi Kuitinen, and Martin Lascoux kindly provided constructive comments on the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Neale DB, Savolainen O: **Association genetics of complex traits in conifers**. *Trends Plant Sci* 2004, **9**:325-330.
 2. Gonzalez-Martinez SC, Krutovsky KV, Neale DB: **Forest tree genomics and adaptive evolution**. *New Phytol* 2006, **170**:227-238.
 3. Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al.*: **The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray)**. *Science* 2006, **313**:1596-1604.
 - The first sequence of a tree genome, including areas of the genome outside coding regions, is an enormous resource for studies of molecular evolution and population genomics.
 4. Wakamiya I, Newton RJ, Johnston JS, Price HJ: **Genome size and environmental factors in the genus *Pinus***. *Am J Bot* 1993, **80**:1235-1241.
 5. Allona I, Quinn M, Shoop E, Swope K, St.Cyr S, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R *et al.*: **Analysis of xylem formation in pine by cDNA sequencing**. *Proc Natl Acad Sci USA* 1998, **95**:9693-9698.
 6. Pavy N, Parsons LS, Paule C, MacKay J, Bousquet J: **Automated SNP detection from a large collection of white spruce expressed sequences: contributing factors and approaches for the categorization of SNPs**. *BMC Genomics* 2006:7.174.
 7. Williams CG, Savolainen O: **Inbreeding depression in conifers: implications for breeding strategy**. *Forest Sci* 1996, **42**:102-117.
 8. Koski V: **A study of pollen dispersal as a mechanism of gene flow in conifers**. *Comm Inst For Fenn* 1970, **70**:1-78.
 9. Huntley B, Birks HJB: *An Atlas of Past and Present Pollen Maps of Europe: 0-13 000 years ago*: Cambridge University Press; 1983.
 10. Petit RJ, Aguinalgalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M *et al.*: **Glacial refugia: hotspots but not melting pots of genetic diversity**. *Science* 2003, **300**:1563-1565.
 11. García-Gil MR, Mikkonen M, Savolainen O: **Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris***. *Mol Ecol* 2003, **12**:1195-1206.
 12. Howe GT, Aitken SN, Neale DB, Jermstad KD, Wheeler N, Chen THH: **From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees**. *Can J Bot* 2003, **81**:1247-1266.
 13. Koch MA, Haubold B, Mitchell-Olds T: **Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae)**. *Mol Biol Evol* 2000, **17**:1483-1498.
 14. Willyard A, Syring J, Gernandt DS, Liston A, Cronn R: **Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for *Pinus***. *Mol Biol Evol* 2007, **24**:90-101.
 - A careful calibration of the fossil data results in improved estimates of the mutation rates of conifers.
 15. Dvornyk V, Sirviö A, Mikkonen M, Savolainen O: **Low nucleotide diversity at the *pal1* locus in the widely distributed *Pinus sylvestris***. *Mol Biol Evol* 2002, **19**:179-199.
 16. Morrell PL, Lundy KL, Clegg MT: **Distinct geographic patterns of genetic diversity are maintained in wild barley (*Hordeum vulgare* ssp. *spontaneum*) despite migration**. *Proc Natl Acad Sci USA* 2003, **100**:10812-10817.
 17. Roselius K, Stephan W, Stadler T: **The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species**. *Genetics* 2005, **171**:753-763.
 18. Schmid KJ, Ramos-Onsins S, Ringys-Beckstein H, Weisshaar B, Mitchell-Olds T: **A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism**. *Genetics* 2005, **169**:1601-1615.
 19. Ma X-F, Szmidi AE, Wang X-R: **Genetic structure and evolutionary history of a diploid hybrid pine *Pinus densata* inferred from the nucleotide variation at seven gene loci**. *Mol Biol Evol* 2006, **23**:807-816.
 20. Krutovsky KV, Neale DB: **Nucleotide diversity and linkage disequilibrium in cold-hardiness- and wood-quality-related candidate genes in Douglas Fir**. *Genetics* 2005, **171**:2029-2041.
 21. Gonzalez-Martinez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB: **DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L.** *Genetics* 2006, **172**:1915-1926.
 22. Pot D, McMillan L, Echt C, Le Provost G, Garnier-Géré P, Cato S, Plomion C: **Nucleotide variation in genes involved in wood formation in two pine species**. *New Phytol* 2005, **167**:101-112.
 23. Tenaillon M, Sawkins M, Long AD, Gaut RL, Doebley J, Gaut B: **Patterns of DNA sequence polymorphism along chromosome 1 of maize**. *Proc Natl Acad Sci USA* 2001, **98**:9161-9166.
 24. Järvinen P, Lemmetyinen J, Savolainen O, Söpanen T: **DNA sequence variation in *BpMADS2* gene in two populations of *Betula pendula***. *Mol Ecol* 2003, **12**:369-384.
 25. Ingvarsson PK: **Molecular population genetics of herbivore-induced protease inhibitor genes in European Aspen (*Populus tremula* L., Salicaceae)**. *Mol Biol Evol* 2005, **22**:1802-1812.
 26. Kado T, Ushi Y, Yoshimaru H, Tsumura Y, Tachida H: **Contrasting patterns of DNA variation in natural populations of two related conifers, *Cryptomeria japonica* and *Taxodium distichum* (Cupressaceae sensu lato)**. *Genes Genet Syst* 2006, **81**:103-113.
 27. Liu A, Burke JM: **Patterns of nucleotide diversity in wild and cultivated sunflower**. *Genetics* 2006, **173**:321-330.
 28. Kado T, Yoshimaru H, Tsumura Y, Tachida H: **DNA variation in a conifer, *Cryptomeria japonica* (Cupressaceae sensu lato)**. *Genetics* 2003, **164**:1547-1559.
 29. Brown GR, Gill GP, Kuntz RK, Langley CH, Neale DB: **Nucleotide diversity and linkage disequilibrium in loblolly pine**. *Proc Natl Acad Sci USA* 2004, **101**:15255-15260.
 30. Ingvarsson PK: **Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., Salicaceae)**. *Genetics* 2005, **169**:945-953.
 31. Heuertz M, De Paoli E, Källman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N: **Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce (*Picea abies* [L.] Karst.)**. *Genetics* 2006, **174**:2095-2105.
 - First effort in trees to examine in depth the demographic history based in sequence variation.
 32. Nordborg M, Borewitz JO, Bergelson J, Berry C, Chory J, Hagenblad J, Kreitman M, Maloof J, Noyes T, Oefner PJ *et al.*: **The extent of linkage disequilibrium in *Arabidopsis thaliana***. *Nat Genet* 2002, **30**:190-193.
 33. Nordborg M, Hu TT, Ishino Y, Juahveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Galdstone J, Goyal R *et al.*: **The pattern of polymorphism in *Arabidopsis thaliana***. *PLoS Biol* 2005, **3**:1289-1299.
 34. Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES: **Structure of linkage disequilibrium and phenotypic associations in the maize genome**. *Proc Natl Acad Sci USA* 2001, **98**:11479-11484.
 35. Clauss MJ, Mitchell-Olds T: **Population genetic structure of *Arabidopsis lyrata* in Europe**. *Mol Ecol* 2006, **15**:2753-2766.

36. Bouille M, Bousquet J: **Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees.** *Am J Bot* 2005, **92**:63-73.
37. Wang RL, Stec A, Hey J, Lukens L, Doebley J: **The limits of selection during maize domestication.** *Nature* 1999, **398**:236-239.
38. Wright SI, Gaut BS: **Molecular population genetics and the search for adaptive evolution in plants.** *Mol Biol Evol* 2005, **22**:506-519.
- This review describes the status of the search for selection in plant populations.
39. Biswas S, Akey JM: **Genomic insights into positive selection.** *Trends Genet* 2006, **22**:437-445.
40. Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS: **The effects of artificial selection on the maize genome.** *Science* 2005, **308**:1310-1314.
41. Toomajian C, Hu TT, Aranzana MJ, Lister C, Tang C, Zheng H, Zhao K, Calabrese P, Dean C, Nordborg M: **A nonparametric test reveals selection for rapid flowering in the *Arabidopsis* genome.** *PLoS Biology* 2006, **4**:e137.
42. Ingvarsson PK, Garcia MV, Hall D, Luquez V, Jansson S: **Clinal variation in *phyB2*, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European Aspen (*Populus tremula*).** *Genetics* 2006, **172**:1845-1853.
43. Li C, Viherä-Aarnio A, Puhakainen T, Junttila O, Heino P, Palva ET: **Ecotype-dependent control of growth, dormancy and freezing tolerance under seasonal changes in *Betula pendula* Roth.** *Trees* 2003, **17**:127-132.
44. Gonzalez-Martinez SC, Wheeler NC, Ersoz E, Nelson DC, Neale DB: **Association genetics in *Pinus taeda* L. I. Wood property traits.** *Genetics* 2007, **175**:399-409.
- Large-scale association study of wood quality traits.
45. Thumma BR, Nolan MF, Evans R, Moran GF: **Polymorphisms in cinnamyl CoA reductase (CCR) are associated with variation in microfibril angle in *Eucalyptus* spp.** *Genetics* 2005, **171**:1257-1265.