Opinion TRENDS in Plant Science Vol.9 No.7 July 2004

Full text provided by www.sciencedirect.com
SCIENCE DIRECT

Association genetics of complex traits in conifers

David B. Neale[1] and Outi Savolainen[2]

[1]Institute of Forest Genetics, USDA Forest Service and Department of Environmental Horticulture, University of California, Davis, CA, USA
[2]Department of Biology, University of Oulu, Finland

Association studies are becoming the experimental approach of choice to dissect complex traits in many organisms from humans to model plant systems. The candidate gene based-association approach has several important advantages for complex trait dissection in many coniferous forest tree species, including random mating and unstructured populations, adequate levels of nucleotide diversity, rapid decay of linkage disequilibrium, and precise evaluation of phenotype from clonal or progeny testing. Allele discovery using association approaches should lead to more-efficient methods of marker-assisted breeding and a deeper understanding of genetic adaptation in forest trees.

The genetic dissection of complex traits to their individual gene components is a topic of great interest in biomedical [1] and agricultural science [2]. Solving the easier problem of identifying the genes coding for Mendelian traits has been highly successful [3]. Geneticists have now turned their attention to the more difficult problem of identifying the individual genes causing common and complex diseases in humans and the many agronomically important traits in crops and livestock. Model systems, such as mouse and *Arabidopsis thaliana*, are often used to overcome the experimental limitations in humans or crop plants, respectively. There is also a need to dissect complex traits in forest trees, although these long-lived species are usually regarded as difficult experimental organisms. In this review, we briefly describe some of the issues that need to be addressed to study complex traits successfully and propose that conifers are an excellent experimental system, where the nearly undomesticated random mating populations offer good opportunities for gaining general insights into the genetics of complex traits through association studies.

Approaches and pitfalls of mapping complex traits

Eric Lander and Nicholas Schork [4] have outlined the major experimental approaches used in complex trait dissection. In studies of livestock, crop plants and trees, the experimental cross approach [quantitative trait locus (QTL) mapping] is usually applied. QTL mapping in crops [2] and forest trees [5] has contributed greatly to our understanding of complex trait architecture, for example, number, position and the size of the effect of QTLs. This knowledge of QTLs can be efficiently applied to marker-assisted breeding (MAB) in inbred crops, but in out-crossing forest trees, the situations in which MAB application can be used are more limited [6]. Beyond MAB, there is a need to have a functional understanding of QTLs. This requires positional cloning of the QTL and complementation tests. This is feasible in organisms with small genomes, mutants with well-defined effects and efficient transformation systems. Furthermore, a complete genome sequence greatly facilitates positional cloning of QTLs. *Arabidopsis* has all these attributes and enables efficient QTL identification [7]. QTLs have also been identified in tomatoes and maize [8,9]. Species of the forest tree genus *Populus* also possess many of these attributes but, as yet, QTLs have not been positionally cloned in these species. By contrast, it is not possible to clone QTLs positionally in conifers. Genome sizes are large [10], there are few well-defined mutants and high-efficiency transformation is still lacking.

Complex trait dissection in humans has largely been achieved by examining the role of specific candidate genes, chosen on the basis of their physiological function, with allele sharing or association approaches. The candidate gene approach was pioneered in *Drosophila* [11]. The primary criticism of candidate-gene-based approaches in humans is the failure to reproduce results [12,13]. Less than half of the reported significant associations are repeatable in later analyses [14]. Part of the inability to reproduce results is because of the small sample sizes. Studies in admixed human populations can also result in false positive associations, where linkage disequilibrium is the result of substructure, not a true association [12]. Contrasting results among studies might also be because of locus heterogeneity (i.e. that different loci govern variation in different populations). Another limitation of the candidate-gene approach is that regulatory regions are under-represented. The highly touted solution to the candidate-gene limitation in humans is whole-genome scans using dense genetic maps based on single nucleotide polymorphisms (SNPs) [15]. This approach assumes that the genome has haplotype blocks, within which, all SNPs are highly associated. In such a case, associations with phenotype can be detected with a more-limited number of SNPs. This raises issues about the extent and the distribution of linkage disequilibrium in genomes and the

Corresponding author: David B. Neale (dbneale@ucdavis.edu).
Available online 11 June 2004

number of SNPs that will be required [16–18]. The whole-genome scan association-mapping approach has also been proposed for crops [19,20] and livestock [21] to overcome the problem of low resolution of QTL location in traditional linkage studies.

From an evolutionary point of view, the nature of the genes being mapped can differ between species, with respect to the expected size of effects, the age of the alleles, the frequency spectrum of alleles at the locus, and the pattern of linkage disequilibrium around the genes [22,23]. Until now, human geneticists have mostly mapped rare disease genes. The deleterious alleles have large phenotypic effects, they might be young, and for this reason they will have extended linkage disequilibrium around the mutation [24]. In common diseases, the underlying loci might have smaller effects, and the alleles might be older [22,23]. Plant breeders are interested in genes that have been under strong selection during domestication, such as flowering time loci (*dwarf8*) [25] or loci affecting shoot architecture (*teosinte branched 1*) in maize [26]. These also tend to have large effects. Alleles that have been positively selected recently will have experienced a rapid increase in frequency. This is also likely to result in extended linkage disequilibrium around the causative SNP, as found in maize [27]. By contrast, traits that are of interest in forest trees have not been under intensive artificial positive selection or strong negative purifying selection. For instance, variation in wood quality seems to be based on many loci with small effects [28]. It is likely that the alleles are older and thus also have less disequilibrium around them. The adaptive traits responsible for local climatic adaptation, such as frost hardiness or timing of growth could have been under stronger differential selection, but QTL studies point to many loci with small individual effects in conifers [29,30], even if QTLs for larger effects have been found in other natural plant populations [31]. Thus, studies in forest trees should inform us about the genetics of standing variation under natural selection. The expected patterns of allele frequencies, patterns of disequilibrium and size of effects influence the design of association studies [32,33].

Association genetics is well suited to conifers

The fundamental difference between an association mapping study and a traditional QTL mapping study is the nature of the mapping population [12]. The traditional QTL mapping approach generates linkage disequilibrium between genetic markers and QTLs through crossing of different genotypes and creation of a segregating population (e.g. F_2 and backcross). The number of crossover events is limited and the map resolution of QTLs is determined by the size of the progeny array. In an association mapping study, individuals are selected from (preferably) non-structured populations, where recombination over many generations has broken up the linkage disequilibrium that initially existed between a marker allele and a novel QTL allele. Thus, the QTLs can be mapped at a much finer resolution. QTLs studied are also limited to the variation between the parents in the mapping cross, whereas association studies examine variation in the whole population. In conifers, the lack of suitable pedigrees also limits QTL studies.

Conifer genomes are extremely large ($>1.0 \times 10^{10}$ bp) and it will be many years before genome sequences are determined. Thus, whole genome scan association mapping is not possible. The candidate-gene-based approach is feasible and might even be more desirable given the characteristics of conifer genome variation.

The most important issues in designing and implementing an association mapping study are: (i) population structure in the natural populations, or structure generated during domestication and breeding, (ii) pattern of nucleotide diversity, (iii) linkage disequilibrium and haplotype structure and (iv) evaluation of the phenotype. We briefly discuss each of these issues with special reference to the life history characteristics and reproductive biology of conifers that make many species in this group of plants highly amenable to association mapping.

Population structure and domestication history

Conifers have a mixed mating system (selfing and outcrossing) but most offspring result from outcrossing [34]. Gene flow through pollen is highly efficient. Collectively, these life history traits all lead to large effective population sizes in many commercially important conifer species, for instance *Pinus taeda*, *Pinus sylvestris*, *Pseudotsuga menziesii*, and *Picea abies*. This is reflected in that only a small proportion of variation (often less than 5%) is found between populations [35] and at the nucleotide level in, for example, *Pinus taeda* and *Pinus sylvestris* [36,37]. There are some economically important conifers that have more differentiated populations, such as *Pinus radiata* [38] and *Pinus pinaster* [39] (~10–15% of variation is found between populations). In these cases, the population structure is mostly accounted for by low divergence of isolated natural populations, not unknown patterns of admixture. Thus, the background average level of population structure at neutral loci is likely to be low in many conifers, for example, *Pinus taeda* [40] and *Pinus sylvestris* [37]. Small-scale initial studies also suggest that nuclear loci do not show strong patterns caused by the post-glacial expansion of the species, but are close to demographic equilibrium [36]. Sampling of some conifer populations might be close to the ideal of large, random-mating equilibrium populations.

The domestication and breeding of conifers is still in its infancy [41]. The first attempts at genetic selection and breeding of conifers did not begin until the mid-20th century, centuries later than the domestication of most agricultural crops. Conifers have generally not suffered diversity losses because of domestication bottlenecks. In some parts of the world, such as far northern Europe, Asia and North America, humans have had almost no impact on conifer populations [42]. Many conifer forests are the result of natural regeneration; artificially regenerated forests are the result of only one or two generations of breeding. Thus, many conifer populations retain the allelic diversity that has resulted from the forces of natural selection over a long period of evolutionary time.

This low level or lack of population structure is rare in other species. In case-control studies of humans, or in studies of crop plants [25], the data can come from different

subpopulations. For example, late- and early-flowering maize plants can differ at marker loci because of past domestication history [25]. Such substructure creates linkage disequilibrium between loci differing in allelic frequency between the subpopulations, and thus spurious associations. These problems can be avoided by inferring the population structure and/or by conducting the association analysis in unstructured populations [43]. These issues have a lesser role in conifers because, in many conifers, this structure is nearly absent, or at least known, because the natural populations have large effective sizes and there are nearly no effects of domestication history.

The lack of structure has several implications for association mapping. For traits where there is little geographic differentiation, such as wood quality, it should be possible to sample over wide areas for trait and SNP variation with little concern for substructure effects. However, forest geneticists are also interested in adaptive variation in phenotypic traits that are highly differentiated across populations, such as phenological traits related to local adaptation or frost tolerance [29,44]. Most loci in the genome are not likely to show differentiation, but it is likely that there will be higher differentiation between populations in the genomic areas governing the clines. Thus, a population genetic survey of the nature of nucleotide variation could already provide evidence for the role of the candidate gene [45]. Theoretical studies show that the exact patterns will depend on the details of the genetic architecture of the quantitative traits [46–48] and the modes of selection [49]. To confirm the associations between candidate locus genotype and phenotype, associations should be found also within the populations along a cline because there might be several genetically independent phenotypic clines.

Estimation of nucleotide diversity

Association mapping, whether it be whole-genome scan or candidate gene, requires abundant SNPs in the study population. For this reason, researchers have often conducted preliminary studies to assess nucleotide diversity in genomes. Recent surveys for a few conifer species illustrate the patterns among species and among genes (Table 1). Estimates of synonymous coding region nucleotide diversity (measured by the population mutation parameter θ) in a few crop species range from a high of $\theta = 0.0173$ in maize [50] to a low of 0.0011 in soybean [51]. Interestingly, in spite of expectations based on allozyme data and life history characteristics, estimates of nucleotide diversity from *P. taeda* and *P. sylvestris* are not among the highest (Table 1), [35]. Nevertheless, conifers appear to have ample nucleotide diversity to perform association studies, perhaps on average at least five times as much as humans do.

Efficient determination of haplotype and rapid decay of linkage disequilibrium

The extent of linkage disequilibrium is crucially important to association studies because it determines the physical distance over which marker by trait associations will exist and thus the marker density that is required. Estimates of haplotype frequencies and linkage disequilibrium can be obtained easily in selfing organisms, where double heterozygotes with unknown phase are rare [20]. In outcrossing species, haplotype determination usually depends on either isolating both alleles of an individual using various methods, such as progeny testing, generating lines homozygous for certain chromosomes, cloning, or allelic-specific PCR. Alternatively, bioinformatics methods can be used [52]. In conifers, the haploid storage tissue of the seeds (megagametophyte) has the same genotype as the egg cell. This makes haplotype determination by direct sequencing and estimating linkage disequilibrium straightforward. It also allows the paternal and maternal haplotype components of the zygote to be determined. Further, the haploid nature aids in distinguishing between members of gene families because two different sequences must originate from different loci.

Estimates of disequilibrium in the human genome vary across genomic areas and populations [53–56], but might extend over 50 kb. In plants, there are fewer studies of linkage disequilibrium in populations and again estimates vary considerably among species. In the annual selfing species *Arabidopsis*, a genome-wide estimate of the extent of linkage disequilibrium is ~250 kb, when sampled over the global distribution area of the species [57]. In local population samples, linkage disequilibrium might even extend over whole chromosomes [57]. This finding agrees with earlier data on linkage disequilibrium in selfing plants, where even loosely linked or unlinked markers showed evidence of linkage disequilibrium within populations [58]. In the selfing crop species soybean, there is little decline in linkage disequilibrium over a distance of ~50 kb or more [51]. In other plant species, estimates of linkage disequilibrium are restricted to within individual candidate genes. In outcrossers, linkage disequilibrium decays much more rapidly than in selfers, as expected. David Remington *et al*. [59] found that linkage disequilibrium varies among genes in maize but on average decays

Table 1. Estimates of nucleotide diversity (θ or π when indicated) in a few conifer species

Species	Number of genes (type)	θ (average)	θ (silent)	θ (non-synonymous)	Ref
Pinus taeda	19 (wood)	0.00407	0.00658	0.00108	http://dendrome.ucdavis.edu/adept
Pinus taeda	28 (disease)	0.00489	0.00588	0.00183	http://dendrome.ucdavis.edu/adept
Pinus taeda	16 (drought)	0.00460	0.00700	0.00187	http://dendrome.ucdavis.edu/adept
Pinus sylvestris	1 (*pal1*)	–[a]	0.00490	0.00030	[36]
Pinus sylvestris	2 (phytochrome)	–	0.00300	0.00050	[37]
Pseudotsuga menziesii	12 (cold related)	0.00853	0.01256	0.00463	http://dendrome.ucdavis.edu/dfgp
Cryptomeria japonica	7 (various)	–	0.00380^{π}	–	[70]

[a] –, not reported in original papers.

within ~1500 bp. Maud Tenaillon *et al.* [50] found that linkage disequilibrium decayed even more rapidly (~100–200 bp) in a more-diverse set of maize germplasm. These findings emphasize how the decay of linkage disequilibrium depends on sampling, an important aspect in designing experiments.

Estimates of linkage disequilibrium vary among several genes in *P. taeda*, but on average, linkage disequilibrium, measured by the correlation coefficient between nucleotide sites (r^2), decays to less than 0.20 within ~1500 bp (Figure 1). In *P. sylvestris*, in the few genes studied to date, there is little linkage disequilibrium. Sites a few hundred bp apart in the *pal1* gene were completely uncorrelated [36]. Population genetics theory predicts that the larger the population recombination rate ($4N_ec$, the product of effective population size and recombination rate per bp), the less linkage disequilibrium should be expected. The low linkage disequilibrium findings discussed above can be accounted for by the large effective population sizes of conifers. Outcrossing plants with smaller disjunct populations, such as an outcrossing relative of *A. thaliana*, *Arabidopsis lyrata*, might have a more-limited number of haplotypes and probably more-extensive linkage disequilibrium within populations [60].

Several aspects of linkage disequilibrium are relevant for association mapping. First, the range of linkage disequilibrium within the genome limits the possibilities of genome-wide testing and the accuracy of fine mapping. Extended linkage disequilibrium, as found in human population isolates such as the Finns, or in selfing plants, enables marker associations to be detected even with relatively sparse maps [24,57]. When more-accurate mapping is required, a population with less linkage disequilibrium, such as the African population, can be used, where less linkage disequilibrium might exist than in European-derived human populations [53] (but see [23]). In humans, the existence of a few haplotype blocks over large stretches of the genome might make whole-genome scans more practical [61]. Similar considerations might be applicable to some crop plants, such as soybeans with limited haplotypic diversity and high linkage disequilibrium [51].

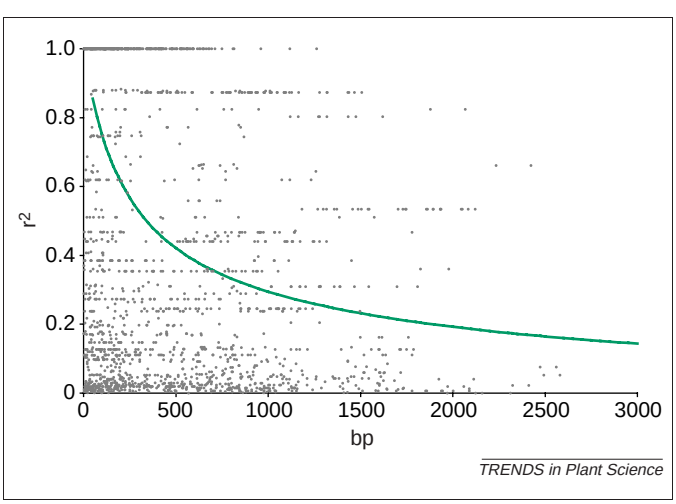

Figure 1. Average rate of decay of linkage disequilibrium, measured by the correlation coefficient between nucleotide sites (r^2), in *Pinus taeda* based on 19 candidate genes. Decays to less than 0.20 within ~1500 bp.

For conifers, the situation is different. The available data suggest that linkage disequilibrium declines within the length of an average-sized gene (Figure 1). Genome-wide association studies will not be possible because of the enormous SNP marker density required. Thus, only a candidate-gene approach is possible.

The possibilities for rational choice of candidate genes are improving as more public EST databases become available for conifers [62]. However, the EST databases might not include genes expressed at very low levels. This problem should be partly alleviated as the functional roles of genes in *Arabidopsis* and other plants are elucidated. Examining the role of a candidate gene might require considering the effects of different members of a gene family [63] to be certain that the functionally important variable gene family member is included in the study. In addition, the EST-based candidate-gene approach concentrates on the coding regions. Recent studies suggest that phenotypic effects in human diseases are likely to be caused by variants in transcribed sequences ~60% of the time [3]. However, there is also increasing evidence that promoter sequence variation underlies QTLs [64–66]. In *Drosophila*, a study of well-defined candidate loci detected many SNPs responsible for variation in antibacterial immunity in both coding and 5′ regions [67]. However, an extensive survey of all non-synonymous SNPs at a strong candidate locus *Delta* did not find the predicted associations between the locus and bristle number variation, presumably because all relevant regulatory sequences were not included [68]. Once candidate genes are adequately chosen, conifers might represent a rare opportunity to map the causative nucleotide because there is so little linkage disequilibrium even between closely spaced SNPs.

Precise evaluation of phenotypes

One of the greatest experimental challenges in association mapping in humans is evaluating the phenotype. Precise evaluation of the phenotype is difficult in most animal systems because of problems in minimizing environmental variation. In inbred crop plants, inbred lines and recombinant inbred lines enable replicated genetic tests to be established. In maize, for example, Jeffry Thornsberry *et al.* [25] used a population of 92 inbred lines to test for associations between the *Dwarf8* gene and flowering time. They were able to establish replicated genetic tests in multiple years, seasons and field sites to test for associations. In forest trees, genetic values of individual trees can be estimated by progeny testing [41] or clonal testing [69]. In *Pinus taeda*, *Pinus radiata* and *Pseudotsuga menziesii*, members of an association population can easily be cloned by methods such as somatic embryogenesis and rooted cuttings. When clones are produced, multiple phenotypes can be evaluated, measured over time and space. Destructive sampling of some individuals of the clone is possible without losing the clone for future phenotypic evaluations. Clonal tests can be established at multiple sites to test for G × E (genotype × environment) interactions and can also be measured over many years to address developmental questions. Clonal testing or progeny testing enables precise and thorough evaluation of

phenotypes, which is not always possible in animal systems or even in all crop plants.

Perspective

Conifers are difficult genetic organisms in many respects because of long generation times, large genomes, lack of well-defined mutants and inefficiencies in transformation. However, their relatively undomesticated condition and large unstructured natural populations provide an ideal situation for association genetics. Linkage disequilibrium is expected to be limited in such populations, making a candidate-gene-based approach optimal and certainly more feasible and cost-effective than a genome-scan approach. In large conifer populations, alleles determining common phenotypes should be in detectable frequencies. Efficient phenotypic evaluation will permit the detection of the expected alleles of small effects. Conifers, unlike humans and domesticated crops and animals, provide a near-ideal population type to understand the functional significance of allelic diversity on the broad distribution of phenotypes found in nature. Conifers might even be a useful system for identifying quantitative trait nucleotides. Association mapping studies in conifers should greatly enhance our understanding of the architecture of complex traits in plants and possibly even humans. Along the way, useful tools should be developed for tree breeders that can facilitate the development of better wood and paper products but also decrease the need for harvesting and exploitation of natural forests.

Acknowledgements

D.B.N. acknowledges the support of NSF Plant Genome Grant # 9975806 and USDA/IFAFS Grant # 2001–52100–11315 (http://dendrome.ucdavis.edu/adept). O.S. acknowledges funding from the Biosciences and Environment Research Council, from the Academy of Finland and from TREESNIPS (QLK-CT-2002–01973).

References

1 Risch, N.J. (2000) Searching for genetic determinants in the new millennium. *Nature* 405, 847–856
2 Tanksley, S.D. (1993) Mapping polygenes. *Annu. Rev. Genet.* 27, 205–233
3 Botstein, D. and Risch, N.J. (2003) Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nat. Genet.* 33 (Suppl.), 228–237
4 Lander, E.S. and Schork, N.J. (1994) Genetic dissection of complex traits. *Science* 265, 2037–2048
5 Sewell, M.M. and Neale, D.B. (2000) Mapping quantitative traits in forest trees. In *Molecular Biology of Woody Plants* (Jain, S.M. and Minocha, S.C., eds), pp. 407–423, Kluwer Academic Publishers
6 Strauss, S.H. *et al.* (1992) Limitations of molecular-marker-aided selection in forest tree breeding. *Can. J. For. Res.* 22, 1050–1061
7 Borevitz, J.O. and Nordborg, M. (2003) The impact of genomics on the study of natural variation in *Arabidopsis*. *Plant Physiol.* 132, 718–725
8 Doebley, J. *et al.* (1997) The evolution of apical dominance in maize. *Nature* 386, 485–488
9 Fridman, E. *et al.* (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proc. Natl. Acad. Sci. U. S. A.* 97, 4718–4723
10 Murray, B.G. (1998) Nuclear DNA amounts in gymnosperms. *Ann. Bot. (Lond.)* 82, 3–15
11 Mackay, T.F.C. and Langley, C.H. (1990) Molecular and phenotypic variation in the *achaete–scute* region of *Drosophila melanogaster*. *Nature* 348, 64–66
12 Cardon, L.R. and Bell, J.I. (2001) Association study designs for complex diseases. *Nat. Rev. Genet.* 2, 91–99
13 Emahazion, T. *et al.* (2001) SNP association studies in Alzheimer's disease highlight problems for complex disease analysis. *Trends Genet.* 17, 407–413
14 Tabor, H.K. *et al.* (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat. Rev. Genet.* 3, 391–396
15 Collins, F.S. *et al.* (1997) Variations on a theme: cataloguing human DNA sequence variation. *Science* 278, 1580–1581
16 Kruglyak, L. (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.* 22, 139–144
17 Weiss, K.M. and Terwilliger, J.D. (2000) How many diseases does it take to map a gene with SNPs. *Nat. Genet.* 33, 518–521
18 Carlson, C. *et al.* (2003) Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. *Nat. Genet.* 33, 518–521
19 Buckler, E.S. *et al.* (2002) Plant molecular diversity and applications to genomics. *Curr. Opin. Plant Biol.* 5, 107–111
20 Rafalski, A. (2002) Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* 5, 94–100
21 Farnir, F. *et al.* (2002) Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees: revisiting the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14. *Genetics* 161, 275–287
22 Reich, D.E. and Lander, E.S. (2001) On the allelic spectrum of human disease. *Trends Genet.* 17, 502–509
23 Weiss, K.M. and Clark, A.G. (2002) Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* 18, 19–24
24 Hästbacka, J. *et al.* (1994) The diastrophic dysplasia gene encodes a novel sulphate transporter – positional cloning by fine structure linkage disequilibrium mapping. *Cell* 78, 1073–1087
25 Thornsberry, J.M. *et al.* (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nat. Genet.* 28, 286–289
26 White, S. and Doebley, J. (1998) Of genes and genomes and the origin of maize. *Trends Genet.* 14, 327–332
27 Palaisa, K.A. *et al.* (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell* 15, 1795–1806
28 Brown, G.R. *et al.* (2003) Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.). III. QTL verification and candidate gene mapping. *Genetics* 164, 1537–1546
29 Hurme, P. *et al.* (2000) Genetic basis of climatic adaptation in Scots pine by Bayesian QTL analysis. *Genetics* 156, 1309–1322
30 Jermstad, K.D. *et al.* (2003) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir. III. Quantitative trait loci-by-environment interactions. *Genetics* 165, 1489–1506
31 Remington, D.L. *et al.* (2003) Candidate genes, quantitative trait loci, and functional trait evolution in plants. *Int. J. Plant Sci.* 164, S7–S20
32 Clark, A.G. (2003) Finding genes underlying risk of complex disease by linkage disequilibrium mapping. *Curr. Opin. Genet. Dev.* 13, 296–302
33 Zondervan, K.T. and Cordon, R.T. (2004) The complex interplay among factors that influence allelic association. *Nat. Rev. Genet.* 5, 89–100
34 Muona, O. (1990) Population genetics in forest tree improvement. In *Plant Population Genetics, Breeding, and Genetic Resources* (Brown, A.H.D. *et al.*, eds), pp. 282–298, Sinauer
35 Hamrick, J.L. and Godt, M.J.W. (1996) Effects of life history traits on genetic diversity in plant species. *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* 351, 1291–1298
36 Dvornyk, V. *et al.* (2002) Low nucleotide diversity at the *pal1* locus in the widely distributed *Pinus sylvestsris*. *Mol. Biol. Evol.* 19, 179–199
37 García-Gil, M.R. *et al.* (2003) Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. *Mol. Ecol.* 12, 1195–1206
38 Moran, G.F. *et al.* (1988) The genetic structure and the conservation of the five natural populations of *Pinus radiata*. *Can. J. For. Res.* 18, 506–514
39 Mariette, S. *et al.* (2001) Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity* 86, 469–479
40 Al-Rabab'ah, M.A. and Williams, C.G. (2002) Population dynamics of *Pinus taeda* L. based on nuclear microsatellites. *For. Ecol. Manage.* 163, 263–271

41 Zobel, B.J. and Talbert, J.T. (1984) *Applied Forest Tree Improvement*, John Wiley & Sons
42 Richardson, D.M. ed. (1998) *Ecology and Biogeography of Pinus*, Cambridge University Press
43 Pritchard, J.K. and Rosenberg, N.A. (1999) Use of unlined genetic markers to detect population stratification in association studies. *Am. J. Hum. Genet.* 65, 220–228
44 Jermstad, K.D. and Neale, D.B. (2003) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir III. *Genetics* 165, 1489–1506
45 Beaumont, M.A. and Nichols, R.A. (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. London Ser. B. Biol. Sci.* 263, 1619–1626
46 Latta, R.G. (1998) Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *Am. Nat.* 151, 292–293
47 LeCorre, V. and Kremer, A. (2003) Dynamics of genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics* 164, 1205–1219
48 Barton, N.H. (1999) Clines in polygenic traits. *Genet. Res.* 74, 223–236
49 Nordborg, M. and Innan, H. (2003) The genealogy of sequences containing multiple sites subject to strong selection in a subdivided population. *Genetics* 163, 1201–1213
50 Tenaillon, M. *et al.* (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize. *Proc. Natl. Acad. Sci. U. S. A.* 98, 9161–9166
51 Zhu, Y.L. *et al.* (2003) Single-nucleotide polymorphisms in soybean. *Genetics* 163, 1123–1134
52 Clark, A.G. (1990) Inference of haplotypes from PCR-amplified samples of diploid populations. *Mol. Biol. Evol.* 7, 111–122
53 Reich, D. (2001) Linkage disequilibrium in the human genome. *Nature* 411, 199–204
54 Pritchard, J.K. and Przeworski, M. (2001) Linkage disequilibrium in humans: models and data. *Am. J. Hum. Genet.* 69, 1–114
55 Ardlie, K.G. *et al.* (2002) Patterns of linkage disequilibrium in the human genome. *Nat. Rev. Genet.* 3, 299–309
56 Reich, D.E. *et al.* (2002) Human genome sequence variation and the influence of gene history, mutation and recombination. *Nat. Genet.* 32, 135–143
57 Nordborg, M. *et al.* (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nat. Genet.* 30, 190–193
58 Clegg, M.T. *et al.* (1972) Is the gene the unit of selection? *Proc. Natl. Acad. Sci. U. S. A.* 69, 2474–2478
59 Remington, D.L. *et al.* (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11479–11484
60 Wright, S. *et al.* (2003) Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*. *Mol. Ecol.* 12, 1247–1263
61 Gabriel, S.B. *et al.* (2002) The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229
62 Kirst, M. *et al.* (2003) Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7383–7388
63 Kinlaw, C.S. and Neale, D.B. (1997) Complex gene families in pine genomes. *Trends Plant Sci.* 2, 356–359
64 Brem, R.B. *et al.* (2002) Genetic dissection of transcriptional regulation in budding yeast. *Science* 296, 752–755
65 Cowles, C.R. *et al.* (2002) Detection of regulatory variation in mouse genes. *Nat. Genet.* 32, 432–437
66 Schadt, E.E. *et al.* (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297–302
67 Lazzaro, B.P. *et al.* (2004) Genetic basis of natural variation in *D. melanogaster* antibacterial immunity. *Science* 303, 1873–1876
68 Genissel, A. *et al.* (2004) No evidence for an association between common nonsynonymous polymorphisms in DELTA and bristle number variation in natural and laboratory populations of *Drosophila melanogaster*. *Genetics* 166, 291–306
69 Forster, G.S. and Shaw, D.V. (1988) Using clonal replicates to explore genetic variation in a perennial plant species. *Theor. Appl. Genet.* 76, 788–794
70 Kado, T. *et al.* (2003) DNA Variation in a conifer, *Cryptomeria japonica* (Cupressaceae *sensu lato*). *Genetics* 164, 1547–1559

PS2004 13th International Congress on Photosynthesis
29 August–3 September 2004
Palais des Congrès de Montréal, Canada

For more information, please see http://www.uqtr.ca/ps2004/

18th International Plant Growth Substances Conference
20–24 September 2004
Canberra, Australia

For more information, please see http://www.conlog.com.au/ipgsa2004/