Series on
Genetics, Genomics and Breeding of Crop Plants
Series Editor
Chittaranjan Kole, Clemson University, Clemson, SC, USA

Genetics, Genomics and Breeding of Conifers

Editors
Christophe Plomion • Jean Bousquet
Chittaranjan Kole

Science Publishers
CRC Press
Taylor & Francis Group

5

Genetic Mapping in Conifers

Kermit Ritland,[1,*] *Konstantin V. Krutovsky,*[2] *Yoshihiko Tsumura,*[3] *Betty Pelgas,*[4,a] *Nathalie Isabel*[4,b] and *Jean Bousquet*[5]

ABSTRACT

This chapter summarizes the history and current status of genetic mapping in conifers. We review the development of molecular markers, methods to construct genetic maps, and the resulting conifer genetic maps. Genetic maps are subdivided into (1) linkage maps of genetic markers, (2) quantitative trait loci (QTL) maps, and (3) comparative maps. Comparative maps involve alignment of marker genes and even QTLs between species. Physical mapping is also briefly discussed. Emphasis is placed up problems and approaches unique to conifers, and the involvement of new genomics technologies.

Keywords: genetic markers, genetic mapping, quantitative trait loci mapping, comparative mapping

5.1 Introduction

Genetic mapping is the ordering of specific genes or DNA fragments (genetic markers) along a chromosome, based up observed frequencies of recombination in pedigrees. It provides the approximate locations of these

[1] Department of Forest Sciences, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada; e-mail: *kermit.ritland@ubc.ca*
[2] Department of Ecosystem Science and Management, Texas A&M University, College Station, Texas 77843-2138, USA; e-mail: *k-krutovsky@tamu.edu*
[3] Forestry and Forest Products Research Institute, Tsukuba, Ibaraki 305-8687, Japan; e-mail: *ytsumu@ffpri.affrc.go.jp*
[4] Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 du P.E.P.S., P.O. Box 10380, Stn Sainte-Foy, Québec, Québec G1V 4C7, Canada;
[a] e-mail: *betty.pelgas@RNCan-NRCan.gc.ca*
[b] e-mail: *nisabel@cfl.forestry.ca*
[5] Canada Research Chair in Forest and Environmental Genomics, Centre d'étude de la forêt, Université Laval, Québec, Québec G1V 0A6, Canada; e-mail: *jean.bousquet@sbf.ulaval.ca*
*Corresponding author

entities, which can serve as DNA "landmarks" for further studies (Ott 1999). Physical mapping, in contrast, uses various molecular techniques to reassemble the actual DNA into contiguous stretches, such that numbers of bases separating genes are approximately known, as in for example the chloroplast genome (Tsumura et al. 1993) and the nuclear genome (Amarasinghe and Carlson 1998). Quantitative trait loci (QTL) mapping places the locations of putative genes underlying a quantitative trait onto a genetic map (Lander and Botstein 1989). Conifers have enormous genomes, on the order of tens of billions of nucleotides (Murray 1998). This prohibits physical mapping, and suggests that marker/QTL mapping may continue to dominate conifer genetics research (Neale et al. 1994; White et al. 2007). In addition, the conserved nature of conifer evolution places greater importance on comparing genetic and QTL maps (comparative mapping) in conifers (Krutovsky et al. 2004) and transferring information among these species.

Conifers provide unique opportunities but also problems for genetic mapping. Most notably, the gametophyte allows direct observation of the haploid product of maternal meiosis (Cairney and Pullman 2007). Secondly, conifers are outbred, and issues in data analysis arise from the fact that parents and grandparents are heterozygous for markers and QTL, requiring more complex approaches for data analysis (Liu 1998). Thirdly, the large genome size of conifers, a consequence of repeated DNA elements (Morse et al. 2009), make protocols for marker screening more complex, and the development of markers more difficult compared to most angiosperms (Kinlaw and Neale 1997). Finally, the enormous evolutionary distance between conifers and angiosperms, separated by 300 hundred million years of evolution (Savard et al. 1994), makes gene identification and annotation in conifers very difficult (Kirst et al. 2003; Ralph et al. 2008). Here, we review the current state of conifer genome mapping, with reference to current advances in genomics studies of conifers.

5.2 Types and Properties of Genetic Markers for Conifers

Over the past 20 years, the increasing availability of molecular genetic markers such as restriction fragment length polymorphisms (RFLPs), amplified fragment length polymorphisms (AFLPs), microsatellites or simple sequence repeats (SSRs), single nucleotide polymorphisms (SNPs), and conserved orthologous sets (COS), has resulted in the development of numerous genetic linkage maps in conifers. The important conifer species—many pines and spruces, Sugi and Douglas-fir—have been mapped, though marker density is still low in relation to genome size. For a typical conifer, a map with 1,000 markers would have, on average, 10–40 million nucleotide sites separating adjacent markers.

5.2.1 First Generation Markers

Before molecular markers became popular in the 1980s, isozymes or allozymes were used for molecular population genetic investigations in conifers. Isozymes are enzymes that differ in amino acid sequence but catalyze the same chemical reaction. Thus, they are representative of differences at the DNA level. Isozymes gave the first revelation about DNA variation in conifers, and for a period centering about the 1980s, many conifers were the subject of isozyme investigations. The dawn of the isozyme era was heralded by a seminal 1979 symposia on "Isozymes in Forest Genetics and Forest Insects" (Conkle 1981). The dusk of the isozyme era was after the 1990 IUFRO symposium, published in the journal New Forests, Volume 6, and in book form by Adams (1992), on the more general topic of "Isozymes in Forest Trees". These two symposia bookmark this era. Isozymes have been placed in genetic maps, but they are not numerous enough to show much linkage. Typically, 20–30 loci are the maximum number of loci that can be assayed, so that in a typical genome of 2,000–3,000 centiMorgans, few loci will be linked.

Another first generation marker used is the RFLP, a co-dominant polymorphism for the presence/absence of restriction sites. Bands were visualized via Southern blots, which require a probe or sequence complementary to the region about the polymorphism. The RFLP technique is relatively laborious to develop and implement compared to the more recent polymerase chain reaction (PCR) based methods. Neverthess, a number of conifer linkage maps were constructed using RFLP markers during 1985–1995. This marker is regarded as "first generation", as their numbers were still quite limited.

5.2.2 Second Generation Markers

RAPD markers consist of fragments generated via the PCR using a randomly selected ten base primer (Williams et al. 1990). A number of RAPD maps were constructed during the period 1992–98. "Random" refers to the fact that primers are chosen at random, without prior knowledge of any specific primer sites in the genome. Hence the step of cloning and identification of specific sequences is skipped. However RAPDs exhibit dominance, wherein heterozygotes cannot be distinguished from dominant homozygotes (which are the band phenotype; band-less phenotypes are recessive). Also, the RAPD gel band patterns often lack reproducibility, making this class of marker only reliable for studies involving controlled crosses such as genetic mapping, where segregation ratios can verify proper inheritance. A variant of RAPDs is "inter simple sequence repeats" (ISSRs), which are randomly amplified markers produced by PCR amplification with short primers that

contain both a microsatellite motif and a random sequence (Bornet and Branchard 2001). These have seen some applications for mapping.

AFLPs are a new class of dominant markers that avoid many of the pitfalls of RAPDs. Assaying for this marker involves restriction digestion of genomic DNA, then PCR amplification of a subset of these fragments (Vos et al. 1995). These markers share many of the characteristics of RAPDs, including dominance and the appearance of many loci on one gel. But the fragment patterns are more reliable, and many more fragments per gel are scoreable. AFLPs have made dense linkage maps possible. However, DNA fragments generated by this technique differ by as little as a single base, requiring use of vertical acrylamide gels or automatic fragment analyzers for clear separation.

Due to the large genome size of conifers, modifications of the AFLP technique for conifers are needed, as the standard +3/+3 primer combinations used for AFLP result in too many bands. With large genomes, one might think that one can limit the pool of selectively amplified DNA by merely increasing the number of selective nucleotides. Vos et al. (1995) found that primers with 4 or more added nucleotides actually suffered a loss of selectivity. For conifers, with their huge genome size, as a means to select subsets of fragments beyond this limit (and to also increase template concentrations), Remington et al. (1999) introduced an additional step prior to the main amplification, termed the "preamplification". It corresponds to a normal amplification, but with shorter primer combinations, usually +1/+1 or +2/+2. Numerous conifer genetic maps have been constructed with AFLPs since the end of the 1990s.

5.2.3 Third Generation Markers

The last class of markers requires *a-priori* knowledge of the DNA sequence at, or around, the marker of interest. A hybrid between second and third generation marker is characterized-sequence amplified region (SCAR), which is developed by cloning RAPD or AFLP markers, and finding the nucleotide sequence about these markers. SCARs are not suited for mapping, as the procedure is laborious; they are useful for finding candidate genes closely linked to anonymous RAPD and AFLP markers found linked to a trait of interest, but they do find themselves occasionally included in conifer genetic maps.

True third generation markers include SSRs (simple sequence repeats or microsatellites), expressed sequence tag (EST)-SSRs (microsatellites in expressed DNA regions), ESTPs (expressed sequence tag polymorphisms, a marker found in ESTs) and finally, the gold standard, the SNP (single nucleotide polymorphism, which directly reflects nucleotide polymorphism at a specific nucleotide site). For further information about markers in plants, see Ritland and Ritland (2000) and Weising et al. (2005).

The completion of the genome sequences in model species, and the accumulation of numerous EST and genomic sequences in many other species, will provide rich resources for the development of these third generation markers.

5.2.3.1 Simple Sequence Repeats

SSRs are markers that are polymorphic for the numbers of repeats of a simple motif (usually 2 to 4 bases long, for example dinucleotide repeats ATATAT….). SSRs are usually co-dominant, highly variable, and somatically stable (Morgante and Olivieri 1993). The locus is amplified by primers that flank the locus. The flanking primers are usually highly species specific. The cost of finding and designing the primers, which must be done for every species, does limit the use of this technique. Sometimes SSRs can be "transferred" to closely related species but at the risk of high null-allele frequency (when the allele does not amplify due to primer mismatches). One disadvantage of SSRs is that they cannot be multiplexed very well, making high density maps impractical.

A special class of SSRs is "EST-SSRs". These are SSRs found in EST sequences and because they are in or near coding regions, the primer regions are more conserved and better able to amplify across species, but the markers are also less polymorphic (Rungis et al. 2004; Ellis and Burke 2007).

5.2.3.2 Single Nucleotide Polymorphisms

SNPs are clearly becoming the marker of choice for species that have been subject to genomic work (either for ESTs, or for genome sequencing). This is because high-throughput genotyping is possible for SNPs, and the number of SNP loci is virtually unlimited. Common high-throughput genotyping methods include the Illumina GoldenGate and Infinium assays (*www.illumina.com*) (Pavy et al. 2008). However while the genotyping costs are much lower (typically 5–10 cent per genotype, compared to 50–100 cents for other methods), the scale of assay (96 SNPs, 480 sample minimum) makes each experiment a big budget item, on the order of many thousands of dollars.

The large number of SNPs uncovered by high-throughput technologies also presents itself with opportunities for marker transfer between species. Even if few in percentage, valuable anchor loci are provided. In the Arborea genome project, for a subset of 1,964 SNPs successfully genotyped in eastern *P. glauca*, 1,565 (80%) were found polymorphic among 11 western *P. glauca* individuals, 728 (37%) among nine *P. sitchensis*, 386 (20%) among 10 *P. abies*, and 321 (16%) among ten *P. mariana* (J. Bousquet et al. unpubl. data). Also,

about 10% of SNPs identified in loblolly pine (*Pinus taeda* L., subgenus *Pinus*, section *Pinus*, subsection *Australes*) amplified in white spruce (D Neale et al. unpubl. data); these are currently being utilized in the Treenomix II project for map synteny comparisons.

5.2.3.3 Conserved Orthologous Sets (COS) and Orthology of Maps

Comparative mapping relies on orthologous markers. The concepts of orthology and paralogy are essential to construct comparative maps. Orthologous gene pairs are directly descended from a common ancestor. Paralogous genes are separated by gene duplication events and may reside in different locations, but also be very closely linked necessitating sequencing to ascertain orthology. These concepts are essential for comparing maps between species (Gogarten and Olendzenski 1999; Koonin 2005; Theissen 2005; Pelgas et al. 2006).

COS markers are genes of low copy number within a genome, and also have low rate of evolution among species. COS markers are identified by self-BLASTing ESTs within a species, to identify genes of low copy number, then cross-BLASTing these sequences among taxa; genes of low copy number within taxa, and low divergence between taxa, are identified as COS markers (Fulton et al. 2002).

Krutovsky et al. (2006) identified COS markers for conifers using sequence comparisons between *Arabidopsis*, rice, black cottonwood, loblolly pine, white spruce, Douglas-fir, and sugi. Interestingly, almost half of the single-copy genes in the non-tree species *Arabidopsis* and rice had additional copies and homologs in poplar and conifers. However, laboratory assays indicate that the high level of evolutionary conservation of COS markers also results in lower gene diversity within populations, and less available polymorphism for mapping purposes (Liewlaksaneeyanawin et al. 2009).

In lieu of these difficulties of COS markers, with the larger number of genes available through automated genomic investigations, there is the hope that orthologous markers from the huge library of SNPs for conifers can be identified, to anchor genetic maps (Le Dantec et al. 2004; Pavy et al. 2006; Pavy et al. 2008).

5.2.4 Public Databases for Third Generation Markers

Public databases contain a wealth of *in silico* data for marker development. Expressed sequence tags (ESTs) are segments of genes expressed as messenger RNA. Hence they are most useful for identifying SNPs of putative function. For ESTs, the most intensively surveyed conifer species are pine and spruce. NCBI's Entrez Taxonomy Browser (*ncbi.nlm.nih.gov*),

as of September 2010, contained 629,815 ESTs for *Pinus* and 542,939 ESTs for *Picea*. Within *Pinus*, the numbers of ESTs are (in parenthesis) are: *P. taeda* (328,756), *P. contorta* (40,483), *P. banksiana* (36,379), *P. pinaster* (34,044), *and P. radiata* (7,538). Within *Picea*, the numbers are *P. glauca* (313,110), *P. sitchensis* (186,637), *P. engelmannii x P. glauca* (28,174) and *P. abies* (14,224). Smaller EST collections exist for other conifers including the family Cupressacea (cedars), which has 72,146 ESTs deposited, mainly for *Crytomeria japonica.* For in silico SNP development, a large number of ESTs are required, unless the deposited ESTs are used to design primers to amplify a small panel of individuals to find SNPs. For pure in silico marker development, a given gene must have at least four overlapping ESTs, in which case a SNP will be detected if two of four nucleotide sites differ in base composition (this mostly rules out sequencing error).

5.3 Mapping Strategies in Conifers

5.3.1 Detecting Recombination

The techniques of marker mapping date from Mendel's crosses. The pioneer of genetic mapping, Thomas Hunt Morgan, showed that recombination frequency can estimate distance separating genes; the distance over which 1% crossover frequency occurs was named by JBS Haldane as the "Morgan", and map distances are generally labeled in centiMorgans (cM) (Ott 1999).

In the context of conifer genetics, issues arise about determining linkage phase. Because conifers are heterozygous, linkage phase cannot be directly ascertained. For example, in the simplest cross, the "backcross", where for two loci A and B, with alleles A_1, A_2 and B_1, B_2, respectively, a cross of grandparent genotype $A_1A_1B_1B_1$ with $A_2A_2B_2B_2$ results in double heterozygote parent $A_1A_2B_1B_2$. Progeny from a backcross of this genotype with either grandparent genotypes may reveal recombinants $A_1A_2B_1B_1$ or $A_1A_1B_1B_2$. In the "intercross", the double heterozygote parents $A_1A_2B_1B_2$ can be crossed with another double heterozygote. As recombination can be detected in both parents in the "intercross", the data are more informative, up to twice as informative when linkage is tight (Ott 1999). However, this assumes linkage phase is known, and grandparents are homozygous.

If grandparents are not homozygous, and/or the grandparents are not genotyped, either single-heterozygote progeny or double-heterozygote progeny (but not both) can be recombinant. This is analogous to the inference of haplotypes from diploid population samples as originally investigated by Clark (1990), in that the phase is indirectly determined by reference to a population sample. Wu et al. (2002) describe how linkage phase can be inferred for outcrossing species with unknown heterozygosity of grandparents, commonly found in conifers. Margarido et al. (2007) implement this procedure in "OneMap".

5.3.2 Assembling Linkage Maps

Lander et al. (1987) popularized genetic mapping with their widely used software, MAPMAKER followed by MAPMAKER/EXP and its close descent, MAPMAKER/QTL. Since then, dozens of programs for both linkage and quantitative trait loci (QTL) mapping have been made freely available. A comprehensive list of linkage and QTL mapping software can be found at *http://linkage.rockefeller.edu/soft*. MAPMAKER starts with a two-point linkage analysis (recombination estimated between all pairs of loci). It then uses a "greedy" algorithm, which builds up linkage groups by sequentially adding markers. This does not guarantee correct orders, so various permutations of maps are done by "rippling". The most commonly used mapping program is JoinMap (Stam 1993), discussed below.

Multipoint linkage analysis takes into consideration the segregation of many linked markers simultaneously. With this approach, it becomes possible to identify individual chromosomal breakpoints and establish order with great certainty (Lathrop et al. 1985). This will become of increasing importance with the advent of high-resolution mapping of conifer genomes.

5.3.3 The Pseudo-testcross

With dominant markers, if a locus is heterozygous in one parent and null (double recessive) in the other, this mimics a testcross with 1:1 segregation ratios. This was termed a "two-way pseudo-testcross" by Grattapaglia and Sederoff (1994), and this was meant to resolve the problem with dominance of RAPD and AFLP markers. It was named "pseudo-testcross" because while it is a testcross mapping configuration, the mating configuration of the markers is not known a priori. However, in genetic mapping, one ends up with a map for the female and a second map for the male. The maps must be joined in some way.

5.3.4 Joining Maps

The integration of two or more marker genetic maps into a single unified map, named a "composite" or "consensus" map, requires common markers that segregate in two or more of the mapping populations. The "pseudo-testcross strategy" is a simple case of multiple maps (two). With dominant markers, one can infer maps for the male and female parent separately as in *Eucalyptus* (Grattapaglia and Sederoff 1994). These workers recognized that multiallelic co-dominant markers with alleles heterozygous in parents are needed as "locus bridges". Joining of maps is now a common activity in

conifers, as much of the pedigree material resides within breeding programs, which include many small pedigrees in progeny tests or diallel crosses.

Stam (1993) developed a computer program "JoinMap" that joins pairs of LGs that share the same marker(s) using either raw genetic data or recombination frequencies. The "JoinMap" algorithm estimates information about recombination in a given cross from LOD values and then combines estimates among crosses assuming a binomial sampling distribution. With more than two pedigrees, joining maps is more complicated. Hu et al. (2004) presented a likelihood approach for joining genetic maps that uses a joint likelihood function that combines information across all crosses. The main advantage of this method is substantially improved accuracy when dominant or a mixture of dominant and co-dominant markers are used.

A new approach to build verified multilocus consensus genetic maps in which shared markers are integrated into stable consensus orders was recently developed by Mester and his colleagues (Mester et al. 2003, 2004, 2006) and implemented into software (*http://www.multiqtl.com/*). The approach is based on (1) combined analysis of initial mapping data rather than manipulating with previously constructed maps, and (2) "synchronized ordering", facilitated by cycles of resampling.

However, several pitfalls exist in joining genetic maps, the most important being differences in recombination rates between pedigrees. Recombination rates can differ between crosses and individuals due to environment particularly in stressful conditions where recombination increases (Agrawal et al. 2005). It can also differ in relation to sex or age, where recombination is lower in males and in older individuals (Rose and Baillie 1979). In several pine species, significantly less recombination was observed for the female gametes than for the male gametes in radiata pine (Moran et al. 1983), loblolly pine (Groover et al. 1995) and maritime pine (Plomion and O'Malley 1996). However, Pelgas et al. (2005) observed no difference in map length between males and females in white spruce, as did Pelgas et al. (2006) and Pavy et al. (2008) for white and black spruce pedigrees. This suggests that sex-specific recombination rates may differ between conifer species. Further investigation is needed on this topic.

Another pitfall in joining maps is that markers can vary in abundance and distribution. In Norway spruce, low- and high-copy-number markers tend to occupy separate genome regions (Scotti et al. 2005). Also, microsatellites may be preferentially associated with nonrepetitive DNA (more coding DNA) in plant genomes (Morgante et al. 2002). Both of these situations indicate that joining maps with different classes of markers might be difficult, as common polymorphic markers between these marker-type classes may not be present in many parts of the genome.

5.3.5 Improving the Resolution of Maps

To get beyond the resolution of traditional marker mapping, which is of 5–10 cM resolution for mapping populations of size ca. 100, one can use larger mapping populations, or else physical mapping. Physical mapping involves the cloning and mapping (by fingerprinting) of large plasmid inserts, such as bacterial artificial chromosomes (BACs), normally 150 KB in length. In conifers, which harbor a 10–40 gigabase genome, this would require 200,000 BAC clones for a 1× coverage; ideally 2 million BACs would be needed for a 10 × coverage, as this is the typical required for a BAC tiling path (Soderlund et al. 2000). At least, the repetitive nature of the conifer genome would suggest that assembly of BAC fingerprints into a tiling path is difficult. However, suggestive data indicate that the major period of repetitive DNA activity (transposition) occured over 100 million years ago (Mya) (M Morgante et al. unpubl. data). Such a feature would actually increase the feasibility of genome assembly, since members of the same repeat class have diverged since transposition. This is a current area of research in conifer genomics—the nature of low complexity DNA in conifers and its implication for genome assembly (Nelson et al. 2008).

With high-resolution meiotic maps, a problem is that low frequency of genotyping error (1.5% or less) can influence mapping outcomes. Such an error was observed to reduce power to discriminate orders, dramatically inflate map length, and provide significant support for incorrect over correct orders (Buetow 1991). Occasional genotype errors skew estimates of recombination between closely linked loci; a similar situation occurs in paternity analysis, where just one missscored locus can invalidate the correct parent. Various workers have since dealt with this issue (Sobel et al. 2002) and new SNP genotyping methods have shown to be highly accurate, with error rate below 1% (Pavy et al. 2008).

To increase the rate that meiotic events can be detected, Gasbarra and Sillanpaa (2006), proposed pooling haploid tissue, such as conifer megametophytes, to estimate recombination rates between closely linked loci (< 1 cM). Pools of several hundred were simulated but they found that several pools were better than a single pool.

Selective mapping approach can facilitate the production of high-quality, high-density genome-wide linkage maps (Vision et al. 2000). It was demonstrated that, to construct a map with high genome-wide marker density, it is neither necessary nor desirable to genotype all markers in every individual of a large mapping population. Instead, a reduced sample of individuals bearing complementary recombinational or radiation-induced breakpoints may be selected for genotyping subsequent markers from a large, but sparsely genotyped, mapping population.

5.4 Conifer Linkage Maps

5.4.1 Overview

Genetic mapping in conifers, and in all other species for that matter, has progressed through three generations of development, corresponding to the marker categories described above. The first generation maps involved allozyme and RFLP markers, which rarely revealed genetic linkage because of their sparsity. The second generation maps involved anonymous genetic markers such as RAPDs and AFLPs; "anonymous" in the sense that we have no idea of their gene function. Nevertheless, complete genetic maps were inferred, as these markers were so much more numerous. The third generation maps involved markers of known gene function, mainly SNPs derived from genome projects. This last wave now allows incredibly detailed maps of genomes, both with numerous markers, and with markers linked to genes putatively related to adaptation and other desired traits.

As isozymes are limited in number, they did not play a significant role in linkage mapping; occasionally a few isozyme markers were added to other markers in a complete map. Significant effort into developing RFLP markers for linkage mapping has been done only for *Cryptomeria japonica, Pinus elliottii, P. taeda*, *P. radiata,* and *Pseudotsuga menziesii* (Table 5-1). Usually RFLPs were analyzed in conjunction with other markers. The most significant early generation molecular marker map was developed in loblolly pine. Devey et al. (1994a) reported an RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. Seventy three of 90 loci (including two isozymes) clustered into 20 linkage groups (LGs). Other studies are summarized in Table 5-1.

The first complete linkage maps in conifers, where the number of large LGs equalled to the haploid number of chromosomes, were made possible by the advent of RAPD and AFLP markers. In the first application of RAPD markers for conifer mapping, Tulsieram et al. (1992) mapped 47 of 61 RAPD markers into 12 LGs in white spruce. Subsequent studies are summarized in Table 5-1. Genetic maps have been constructed for ca. 12 pine species (Table 5-1: *Pinus brutia*, *caribaea, contorta, densiflora, edulis, elliottii*, *palustris*, *pinaster*, *radiata*, *strobus*, *sylvestris*, and *taeda*). Maps have been constructed for four spruce species (Table 5-1: *Picea abies*, *glauca, mariana* and *rubens*). *Cryptomeria* and *Pseudotsuga* are two other conifers that have received much attention, while isolated work has been done with *Abies, Cunninghamia, Larix* and *Taxus.* It should be noted that many of these studies used open-pollinated seed progeny of an individual tree assayed for haploid megagametophytes (this also avoids the dominance of RAPD markers). We now discuss more detailed maps made in the four most important conifer genera.

Table 5-1 Genetic linkage maps in conifers. Linkage groups include at least 3 markers. Expected coverage is the ratio between observed and predicted map sizes estimated following Hulbert et al. (1988) and Chakravarti et al. (1991); NA = not available. The haploid number of chromosomes in all species is 12 except *Cryptomeria japonica* ($n = 11$), and *Pseudotsuga menziesii* ($n = 13$). Numbers separated by "/" refer to the maternal/paternal parents of the mapping population, respectively. Marker types are defined in Section 5.2.

Species	Marker type	Markers	Linkage groups	Map length, cM	Expected coverage, %	cM per marker	Reference
Abies nordmanniana	AFLP, RAPD	556	19	1977	80	NA	Hudson (2005)
Cunninghamia lanceolata	AFLP	101/94	11	2283/2566	NA	23/27	Tong and Shi (2004)
Cryptomeria japonica	RFLP, RAPD, Isozymes	91	13	887	NA	10	Mukai et al. (1995)
	RAPD	84/119	14/21	1112/1756	40/62	13/15	Kuramoto et al. (2000)
	AFLP, CAPS	91/132	19/23	1266/1992	50/80	16/18	Nikaido et al. (2000)
	CAPS, Isozymes, SNP, RAPD, RFLP, SSR	438	11	1372	96	3	Tani et al. (2003)
Larix decidua	AFLP, ISSR, RAPD	117	17	1152	80	14	Arcade et al. (2000)
L. kaempferi	AFLP, ISSR, RAPD	125	21	1206	81	14	Arcade et al. (2000)
Picea abies	RAPD	165	17	3584	NA	22	Binelli and Bucci (1994)
	RAPD	82	13	1385	NA	24	Skov & Wellendorf (1998)
	AFLP, SSR	413	22	2198	77	9	Paglia et al. (1998)
	AFLP, ESTP, rDNA, SSR	755	12	2035	NA	3	Acheré et al. (2004)
	AFLP, IRAP, S-SAP, ESTP, SSR	203/152	27/23	2316/1669	66/79	13/ 13	Scotti et al.(2005)
P. glauca	RAPD	61	12	873	NA	14	Tulsieram et al. (1992)
	ESTP, RAPD, SCAR	144/165	19	2008/2059	73/87	9/15	Gosselin et al. (2002)
	AFLP, ESTP, SSR	802	12	1934	89	2.4	Pelgas et al. (2006)
	AFLP, ESTP, SNP, SSR	821	12	2304	98	2.8	Pavy et al. (2008)
	AFLP, ESTP, SNP, SSR	1301	12	2087	NA	1.6	Pelgas et al. (2011)
	AFLP, ESTP, SSR, COS	505	12	1835	NA	3.5	Liewlaksaneeyanawin et al. (unpubl. data)
P. mariana × *P. rubens*	AFLP, ESTP, RAPD, SSR	1124	12	1846	92	1.6	Pelgas et al. (2005)

Table 5-1 contd....

Table 5-1 contd....

Species	Marker type	Markers	Linkage groups	Map length, cM	Expected coverage, %	cM per marker	Reference
P. mariana	AFLP, ESTP, SNP, RAPD, SSR	835	12	1850	98	2.2	Pavy et al. (2008)
Pinus brutia	RAPD	13	6	164	NA	NA	Kaya and Neale (1995)
	AFLP, SAMPL, ESTP, SSR	1111	12	1770	97	1.6	Kang et al. (2010)
P. caribaea	AFLP, SSR	109	27	1658	88	16	Shepherd et al. (2003a)
P. contorta	RAPD	225	16	2287	95	15	Li and Yeh (2001)
P. densiflora	AFLP	152	19	2341	82	18	Kim et al. (2005)
P. edulis	AFLP	338	22	2012	85	9	Travis et al. (1998)
P. elliottii	RAPD	73	13	782	64	11	Nelson et al. (1993)
	RAPD	91	13	953	62	16	Kubisiak et al. (1995)
	ESTP, Isozymes, RAPD, RFLP	154	15	1115	NA	7	Brown et al. (2001)
	AFLP, SSR	78	23	1170	82	15	Shepherd et al. (2003a)
P. palustris	RAPD	133	16	1635	85	15	Nelson et al. (1994)
	RAPD	122	18	1368	81	13	Kubisiak et al. (1995)
P. lambertiana	SNP	399	19	1231	NA	3.1	Jermstad et al. (2010)
P. pinaster	RAPD	263	13	1380	90	9-10	Plomion et al. (1995b)
	Isozymes, RAPD	463	12	1860	93	8.3	Plomion et al. (1995a)
	AFLP, Isozymes, RAPD			1873	93		Costa et al. (2000)
	AFLP, ESTP, SSR	1182	12	1994	NA	10	Ritter et al. (2002)
	AFLP	620	12	1441	NA	NA	Chagné et al. (2002)
	AFLP, ESTP	326	12	1639	NA	NA	Chagné et al. (2003)
P. radiata	RAPD, RFLP, SSR	195	14	1382	NA	7	Devey et al. (1996)
	RFLP, SSR	173	14	1223	75	7	Devey et al. (1999)
	RAPD, SSR	172	19	1117	56	NA	Kuang et al. (1999)
	AFLP, RAPD, SSR	194	21	1144	77	12	Wilcox et al. (2001)
P. strobus	RAPD, SSR, STS	101	12	745	58	14	Echt and Nelson (1997)

P. sylvestris	RAPD	261	14	2639	NA	10	Yazdani et al. (1995)
	AFLP	94/155	15	796/1335	77/86	18/17	Lerceteau et al. (2000)
	AFLP	188/245	12/15	1696/1719	86/99	9/7	Yin et al. (2003)
	AFLP, ESTP, SSR	120/112	21/16	929/1452	66/85	9/12	Komulainen et al. (2003)
P. taeda	Isozymes, RFLP	75	10	632	NA	NA	Devey et al. (1994a)
	AFLP	508	12	1528	99	9	Remington et al. (1999)
	Isozymes, RAPD, RFLP	357	20	1359	82	4	Sewell et al. (1999)
	RFLP, SSR	223	20	1281	75	4	Devey et al. (1999)
	ESTP, Isozymes, RFLP	235	12	1227	NA	5	Brown et al. (2001)
	ESTP, Isozymes, RFLP	302	12	1274	NA	NA	Krutovsky et al. (2004)
	ESTP, Isozymes, RAPD, RFLP, SNP, SSR	373	12	1228	NA	4	Eckert et al. (2009)
	ESTP, Isozymes, RAPD, RFLP, SSR	462					Echt et al. (2011)
P. thunbergii	AFLP, RAPD	207	20	2085	77-78	10	Hayashi et al. (2001)
Pseudotsuga menziesii	RFLP, RAPD	141	17	1062	NA	7.5	Jermstad et al. (1998)
	RAPD	210	16	2279	91	10	Krutovsky et al. (1998)
	RAPD	132	13	2143	NA	NA	Carlson et al. (2007)
	ESTP, Isozymes, RAPD, RFLP, SSR, STS	376	22	1859	NA	NA	Krutovsky et al. (2004)
	AFLP	120	19	939	NA	9	Ukrainetz et al. (2008a)
Taxus brevifolia	RAPD	41	17	306			Göçmen et al. (1996)

Figure 5-1, based upon a meta-analysis of Table 5-1, shows how the types of markers used in genetic maps have changed in the past 20 years. In general, the numbers of maps have declined in the past five years. From this graph, it is evident that RAPD and ISSR markers predominated during 1995–2005, but their use has declined, as they cannot be transferred among pedigrees to build additional maps. AFLPs had a big impact during 2000–2005; again, as they are anonymous markers, but their transferability is limited. Isozymes (the hobbit in the corner) and RFLPs have had a constant impact, but their numbers are still limited. SSRs and SNP/ESTP/STS markers are obviously the markers of choice for future mapping. They have seen increasing useage. These are all sequence based markers that can be transferred among pedigrees and species.

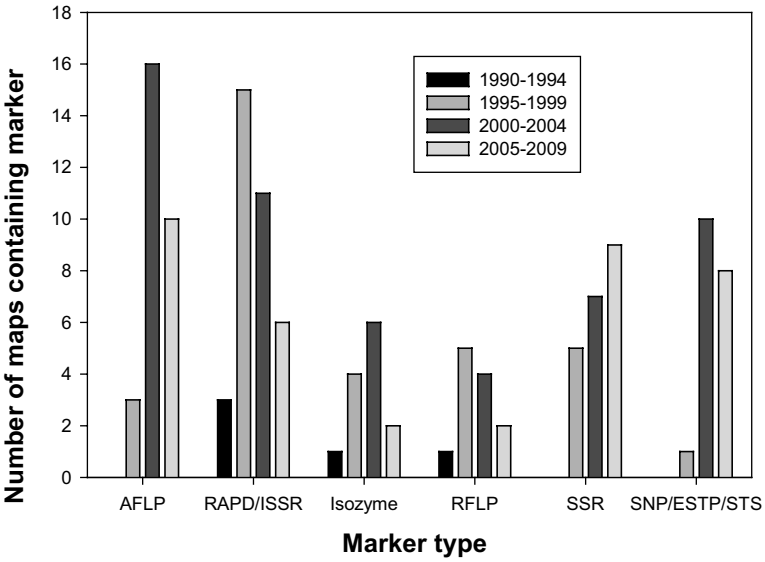

Figure 5-1 Trends in conifer maps over the past 20 years. I. Usage of the various classes of genetic markers for conifer genetic mapping.

Figure 5-2, also based upon Table 5-1, shows how the number of markers used in conifer genetic mapping has increased in the past 20 years. Figure 5-2a shows that the number of markers has obviously increased as expected, but the variance of the number of markers has also increased. This does not include plans from spruce and pine genome projects to radically increase marker number to 5,000 and above. But total map length has remained almost constant (Fig. 5.2b) as markers separated by 30 cM or less (ca. 100 markers total) are sufficient to cover genome length. High density marker maps will be of main use for assembling contigs from genome sequencing projects.

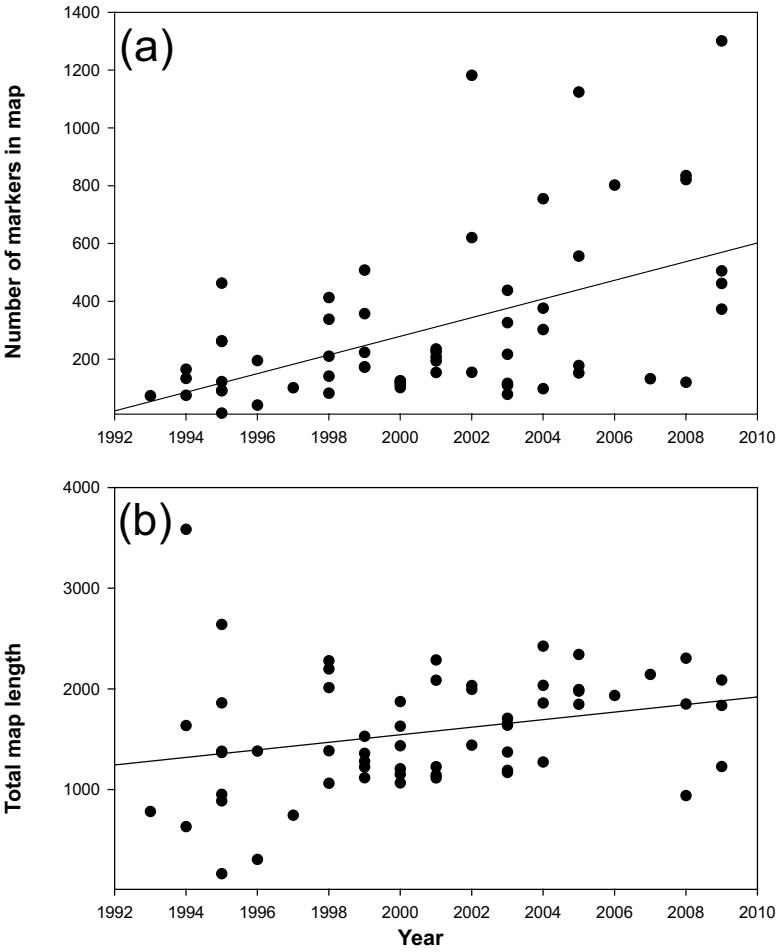

Figure 5-2 Trends in conifer maps over the past 20 years. II. (a) Numbers of markers linked to maps, (b) Total map length explained by markers.

5.4.1.1 Loblolly Pine

Pinus taeda genome maps generally contain ca. 300 loci based mostly on the two standard pedigrees: *base* (Devey et al. 1994b) and *qtl* (Groover et al. 1994). Mapping data for previously reported SSR, RFLP and ESTP markers were combined with new SSR markers to generate a loblolly pine consensus map of 462 markers covering 1,380 cM across 12 LGs, using both the *qtl* pedigree (n = 171) and *base* pedigree (n = 98) (Echt et al. 2011). Of the 234 mapped SSR loci, 171 were newly developed, 81 of which were derived from

EST sequence data. Marker data were obtained for an additional 50 new EST-SSR loci that did not segregate in either mapping population but which were polymorphic in population surveys. One hundred and ninety four mapped loci were given a functional GO assignment; 242 mapped loci were assigned to a NCBI UniGene cluster. Unigene and GO assignments, along with linkage data, aided in identifying duplicated and paralogous marker loci on the map. This species may serve as a reference map in comparative mapping with other pines and even other members of the Pinaceae family such as spruce and Douglas-fir.

5.4.1.2 Spruce

Linkage mapping in spruce (*Picea* spp.) has been directed toward three species of major economic importance: *Picea abies*, a European species, and *P. glauca* and *P. mariana*, both primarily North American species. The first saturated composite map for white spruce was reported by Gosselin et al. (2002), who used 165 RAPD, SCAR and ESTP markers to join maps from two individuals. They noted that co-dominant markers were needed to join the maps. In Norway spruce, Acheré et al. (2004) developed the second map, involving 755 markers. Interestingly, 150 of these markers were tested for their pattern of population differentiation differing from neutral expectations, and nine of these markers were found to be "outliers", or genes that showed excessive population divergence, compared to the majority of markers, suggesting they were linked to QTLs for adaptation (Acheré et al. 2005).

More recently, the Arborea project in Canada has constructed several linkage maps involving both individual and composite maps for white spruce and black spruce. A map for the black spruce × red spruce species complex was constructed (Pelgas et al. 2005), and for white spruce alone (Pelgas et al. 2006). Most notably, Pavy et al. (2008) assembled a white spruce linkage map with markers assayed via the Illumina GoldenGate SNP genotyping platform. The resulting composite map had 821 loci including 461 AFLPs, 12 SSRs, 31 ESTPs and 317 gene SNPs, and map coverage was > 98%. This map also positioned genes with SNPs involved in among-population differentiation of eastern white spruce; 50 outlier SNPs were identified (Namroud et al. 2008); these genes are putatively involved in adaptive differentiation. An expanded white spruce composite map containing 836 gene loci has recently been published (Pelgas et al. 2011).

The most recent white spruce gene composite map emerging from the Arborea project integrates two pedigrees of 500 progeny and has an increased resolution of 0.9 cM with 2,255 positioned loci including 455 AFLPs, 12 SSRs and 1,788 gene SNPs. The map covers 2,065.4 cM over 12

LGs. The average gene density is 1.16 cM. The current published spruce map has 826 genes; the largest number of mapped genes in a conifer species.

5.4.1.3 Douglas-fir

In *Pseudotsuga menziesii*, the most recent marker development has focused on ESTP and SNP markers (Krutovsky et al. 2004), which together with SSR markers, have added to the RFLP and RAPD linkage maps (Jermstad et al. 1998). The most recently published genetic map of Douglas-fir consists of 376 markers, including 172 RFLP, 77 RAPD, 2 isozyme, 20 SSR, 4 sequence tagged site (STS), and 101 expressed sequence tag (EST) markers (Krutovsky et al. 2004). This map is organized into 22 LGs that have three or more linked markers and spans 1,859 cM. Several hundred SNP markers were developed recently (Eckert et al. 2009), and their mapping is under way. When enough markers are mapped, the number of LGs should coalesce into 13, corresponding to the 13 chromosome pairs in Douglas-fir. It would be valuable to map additional ESTP, EST-SSR and SNP markers to create a high-density map that can be used for QTL, candidate gene and physical mapping to facilitate eventual complete Douglas-fir genome sequencing.

5.4.1.4 Sugi

Sugi (*Cryptomeria japonica*) has been planted widely throughout Japan over an area of 4.5 million ha, accounting for 44% of all the Japanese artificial forest. A second generation linkage map for Sugi was constructed by integrating linkage data from two unrelated third-generation pedigrees. The progeny segregation data of the first pedigree, which involved a cross between half-sibs, were derived from cleaved amplified polymorphic sequences (CAPS), SSRs, RFLPs, and SNPs (Tsumura et al. 1997; Iwata et al. 2001). The data of the second pedigree, which involved a self-pollinated individual, were derived from CAPS, isozyme markers, morphological traits, RAPDs, and RFLPs. The co-dominant DNA markers such as CAPS, RFLP and SNP were developed from ESTs and cDNA clones from several kinds of cDNA libraries (Ujino-Ihara et al. 2000; Ujino-Ihara et al. 2005). More than 95 % of the markers were gene-based markers.

Using JoinMap, linkage analyses were done for the first pedigree assuming cross-pollination, and for the second pedigree assuming selfing. Four hundred and thirty eight markers were assigned to 11 large LGs (corresponding to the 11 chromosomes of *C. japonica*), 1 small LG, and 1 non-integrated LG from the second pedigree; their total length was

1,372.2 cM (Tani et al. 2003). On average, the consensus map showed one marker every 3.0 cM. PCR-based co-dominant DNA marker such as CAPS, microsatellite and SNP were distributed over all LGs and represented about a half of mapped loci.

5.4.2 Genome Sizes

Besides providing a linear map of markers along a genome, mapping experiments can also provide estimates of genome size, in terms of map units. Hulbert et al. (1988) gave the first estimate of genome size based upon observed recombination between randomly selected pairs of markers. Chakravarti et al. (1991) improved this with a maximum likelihood method for estimating genome size. Many conifer mapping studies have provided estimate of genome size from either method; estimates range from ca. 2,000 to 3,000 map units. Relatively few numbers of markers can estimate genome size, as long as some are linked.

Genome size can also be estimated by flow cytometry, in terms of picograms (pg) of DNA per nucleus, which can be translated into millions of base pairs using the relationship 1 pg = 978 million base pairs. This gives an idea of how many nucleotides separate linked markers. Genome size in the Pinaceae ranges from 5.8 to 32.2 pg with 20 pg (20 billion base pairs) a rough average (Murray 1998); this is 100 times larger than *Arabidopsis thaliana* (0.18 pg).

Genome evolution in the gymnosperm lineage of seed plants has given rise to many of the most complex and largest plant genomes; however the elements involved are poorly understood. Most of the enormous genome complexity of pines can be explained by divergence of retrotransposons (Morse et al. 2009); however the elements responsible for genome size variation are yet to be identified. This is currently a very active area of research in conifer genomics.

5.4.3 Physical Mapping Opportunities

Physical mapping complements genetic mapping. Unfortunately the large physical genome size of conifers as just described prohibits most of these approaches. Approaches that are free from constraints from large genome size involve hybridization of certain genes to chromosomes. Earlier works used fluorescence *in situ* hybridization (FISH) experiments to identify location and distribution of ribosomal RNA. In Sitka spruce, 5s rDNA was found to be restricted to one chromosome, whereas 18S-5.8S-26S rDNA repeats occurred on five chromosomes (Brown and Carlson 1997). Both distribution and location of large tandem repeats on the genome of white spruce and Sitka spruce were comparable (Brown et al. 1998). A reference

karyotype was presented recently for loblolly pine based on FISH and using 18S–28S rDNA, 5S rDNA, and an *Arabidopsis*-type telomere repeat sequence, *A*-type TRS signals (Islam-Faridi et al. 2007). Statistically, only seven of the 12 loblolly pine chromosomes could be distinguished by their relative lengths. However, the position and relative strength of the rDNA and telomeric sites made it possible to uniquely identify all of the chromosomes, providing a reference karyotype for use in comparative genome analyses. A dichotomous key was developed to aid in the identification of loblolly pine chromosomes and their comparison to chromosomes of other *Pinus* spp. A cytomolecular map was developed using the interstitial 18S–28S rDNA and *A*-type TRS signals. A total of 54 bins were assigned, ranging from three to five bins per chromosome. This is the first report of a chromosome-anchored physical map for a conifer that includes a dichotomous key for accurate and consistent identification of the loblolly pine chromosomes.

Recently, bacterial artificial chromosome (BAC) hybridization has been developed as an alternative to rDNA hybridization, which allows very specific identification of chromosomes, and such methods would be fruitful to apply to conifers, particularly the Pinaceae, as chromosomal morphology is hardly distinguishable among the dozen or so chromosomes. This method has been used in many plant species (Zhang et al. 2004) but not in conifer.

The normal activity of physical mapping is to construct a library of inserts, then to construct "tiling paths" to obtain an ordered set of clonal inserts that span the entire genome. For coverage of a conifer genome (5–10×), about two million BAC clones are needed, too large for practical work. Nevertheless, BACs are useful for conifers, and there are currently BAC libraries available for white spruce and loblolly pine. The spruce library is unarrayed and about 5× coverage, while the loblolly pine library is arrayed and about 8× coverage (Liu et al. 2009). Currently, both random BACs and targeted BACs (BAC identified as having a gene of interest) are being sequenced from both libraries (J MacKay et al. unpubl. data; DG Peterson et al. unpubl. data; K Ritland et al. unpubl. data).

5.5 Conifer Comparative Mapping

Alignments of genetic or QTL maps among species demonstrate the evolutionary conservation of gene linkages among species. An early paradigm was set by work with the grass family (Gale and Devos 1998). Conserved chromosomal number in the pines family (Pinaceae) suggested that similar comparisons could be made in pine family members. The "Conifer Comparative Genomics Project" organized by David Neale and his colleagues at UC Davis has verified that such approaches can be used in conifers (e.g., Krutovsky et al. 2004). The end goal is to transfer information between species about co-localization of QTL and candidate genes among species. In genome sequencing projects, it also predicts the reliability that

related conifer genomes can be "resequenced", once a reference genome is sequenced.

To facilitate the identification of orthologous markers for comparative mapping, sequence-based gene markers such as ESTPs and SNPs are best because they are usually orthologous across congeneric species, and more reliable than anonymous markers. Hidden paralogy is the ghost of map construction (Huynen and Bork 1998; Remm et al. 2001; Pelgas et al. 2006). To reduce the risk of paralogous amplification, primer pairs should be designed with a primer matching in the 3′ UTR gene region (e.g., Perry and Bousquet 1998; Brown et al. 2001; Chagné et al. 2003; Pavy et al. 2008). In conifers, resequencing from megagametophyte DNA indicates paralogous polymorphisms by the presence of double peaks on sequence chromatograms (Pelgas et al. 2004).

Until recently, limited numbers of orthologous markers were available for useful map comparisons. SNPs are virtually in unlimited number. Because they can be annotated and are dense along linkage maps, SNPs can better determine gene orthology, and serve as anchor markers for intra- and interspecific map comparisons (Pelgas et al. 2006; Pavy et al. 2008).

5.5.1 Pine Species Comparisons

Historically, the most extensive genetic maps have involved loblolly pine. Detailed comparative maps are needed to study conifer genome evolution and to leverage genomic information of adaptive and economic traits from the relatively well-studied species, such as loblolly pine, to other conifers. Most comparative maps among *Pinus* species are within the subgenus *Pinus* and based on comparisons of ESTP markers. They contain 41 common loci between *P. taeda* and *P. sylvestris* (Komulainen et al. 2003) and 32 common loci between *P. taeda* and *P. pinaster* (Chagné et al. 2004). Both of these studies used prior published *P. taeda* maps (Krutovsky et al. 2004). Recently, maps from the two subgenera of *Strobus* and *Pinus* could be compared, based on neaerly 400 gene SNPs (Jermstad et al. 2010). All 19 linkage groups of *P. lambertiana* co-aligned with the 12 linkage groups of *P. taeda*, providing a basis for integrated structural genomics approaches across pine subgenera.

5.5.2 Spruce Species Comparisons

The first comparative map of white spruce (Pelgas et al. 2006) revealed remarkable synteny with black spruce (*P. mariana*) and Norway spruce (*P. abies*); identical LGs and conservation of gene content and gene order was found. One breakdown of synteny between *P. glauca* and the other taxa involved an inter-chromosomal rearrangement of an insertional translocation. Analysis of marker colinearity also revealed a putative segmental duplication. This three-species comparison showed that genome

comparisons among *Picea* species can provide a platform for transfer of genomic information across species of spruce.

More recently, a detailed analysis of synteny and macro-colinearity between *P. glauca* and *P. mariana*, using 215 anchor markers, consisting mainly SNPs, found that 98% of the anchor genes were in synteny (Pavy et al. 2008). Translocations were validated in the case of previously reported *PgMyb4*, and three new translocations involving three genes were indicated. However, the sequencing of haploid megagametophytes for these genes indicated that these new cases were likely false positives, involving paralogous variation. Macro-colinearity was also well conserved among homologous LGs between species, with 82% of syntenic anchor markers positioned in the same order. Exceptions to colinearity involved small inversions also observed between individual maps within species, indicating that that most of these inversions were artefacts.

Figure 5-3 shows a relatively high density genetic map for both white and black spruce (LGs III-VI only), with the maps also aligned between the two species. Map distances in centiMorgan are indicated with a scale on the left side. The composite map of each species was obtained by first assembling two parental datasets for each species, using JoinMap (Stam 1993); then maps were aligned between species using common markers. There are five types of markers in these maps: SNPs (bold), ESTPs (bold and underlined), SSRs (bold and italics), RAPDs (italics and underlined) and AFLPs (others). Typically AFLPs are the most in such maps with several types of markers, but they are not useful for joining maps between species (the loci are named after the primer combination used and the band migration distance). Syntenic marker loci between spruce species are indicated in black, and these are typically gene-based markers. These syntenic markers are identified with a red solid line (colinear markers) or a red dashed line (non-colinear markers). Orthologous markers positioned onto non-homologous LGs are indicated in white with red background and paralogous markers are identified in white with blue background. Overall, there is a remarkable preservation of gene order between white and black spruce, and the exceptions may be mistaken cases of orthology and merit further investigation.

5.5.3 Pine Family Species Comparisons

The first intergeneric comparative map in conifers was constructed between loblolly pine and Douglas-fir with ESTP and RFLP markers (Krutovsky et al. 2004). Comparison of Douglas-fir and loblolly pine maps revealed 10 LGs (LG1–LG10) in loblolly pine that shared 2–10 orthologous markers with 12 apparently syntenic LGs in Douglas-fir based on 46 orthologous markers. The comparisons revealed extensive synteny and colinearity of

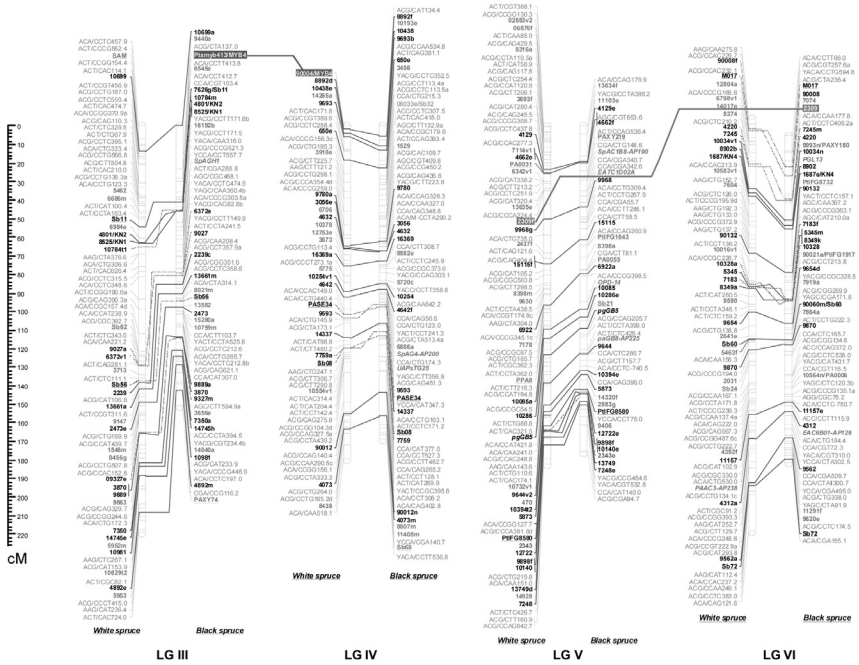

Figure 5-3 Comparison of homologous linkage groups between white spruce (*Picea glauca*) and black spruce (species complex *Picea mariana* × *P. rubens*).

Color image of this figure appears in the color plate section at the end of the book.

gene order between the two genomes, consistent with the hypothesis of conservative chromosomal evolution among even distantly related species in the Pinaceae family. This study established a working framework that the Pinaceae can be viewed as a single genetic system.

Homology of Pinaceae LGs was more recently extended to three spruce species (Pelgas et al. 2006). Between spruce and loblolly pine, 26 of 29 anchor markers were in synteny, identifying 11 homologous LGs. In this study, orthology of anchor gene markers was checked by extensive resequencing of single haploid megagametophytes in the various species. For the three exceptions to synteny, sequencing of megagametophytes indicated at least two cases of paralogy, while the third case remained dubious, implicating a conserved gene family. Between spruce and Douglas-fir, synteny could be assessed with 20 anchor markers, of which just one proved to be paralogous after megagametophyte resequencing. Of the remaining markers, three were not in synteny, including two markers on LG13 of Douglas-fir, confirming that the supernumerary chromosome of Douglas-fir is the result of fission (Krutovsky et al. 2004; Pelgas et al. 2005). The remaining marker, in synteny between spruce and lodgepole pine, was translocated to a different LG in

Douglas-fir, thus indicating that chromosome rearrangements occurred in the lineage leading to Douglas-fir. This study established rigorous criteria for determining orthology of genetic markers among species, and only after this criteria is met, can we make reliable inferences about chromosomal rearrangements among species.

Figure 5-4 shows a recent syntenic map for Douglas-fir, loblolly pine and Norway spruce. This was identified as LG6 of loblolly pine, as the high level of synteny and conservation of gene order allows homologous LGs among pine species to be identified (Neale and Krutovsky 2004). Orthologous comparative mapping markers are underlined and shown in bold (this is based upon unpublished data kindly provided by Craig Echt, USDA Forest Service, Southern Institute of Forest Genetics, Saucier, Mississippi, USA [for pine] and by Michela Troggio, IASMA Research and Innovation Centre, San Michele, Italy [for spruce]). Overall, the alignment of maps between species separated by over 100 million years of evolution is remarkable

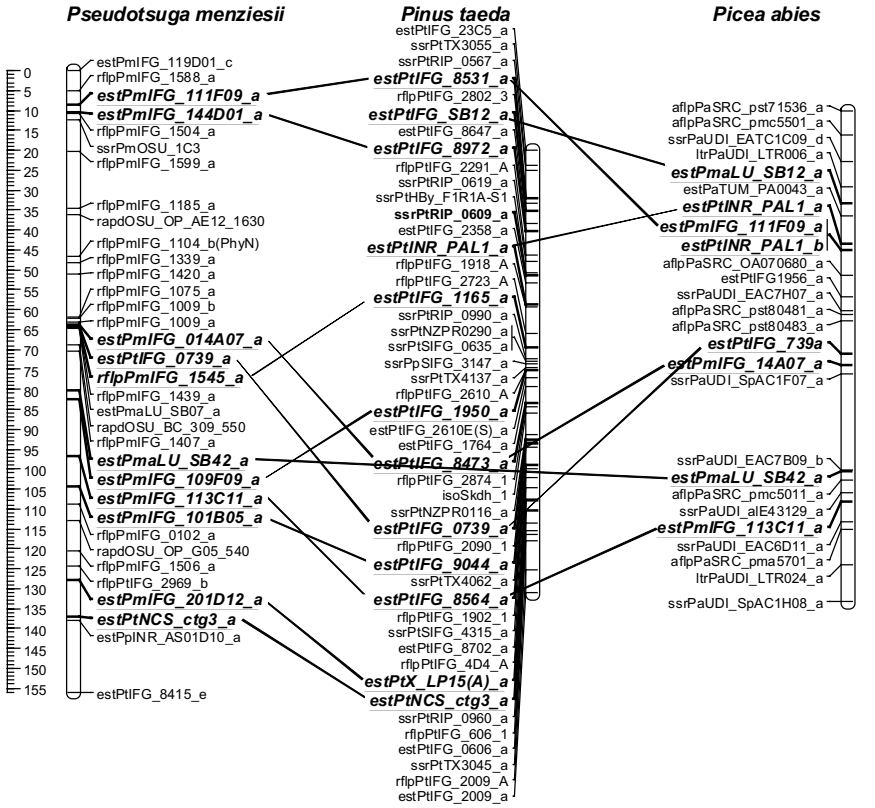

Figure 5-4 Comparison of homologous linkage groups between Douglas-fir (*Pseudotsuga menziesii*), loblolly pine (*Pinus taeda*) and Norway spruce (*Picea abies*).

for the plant kingdom, and suggests that the pine family (Pinaceae) can be viewed as one genetic system, allowing genomic information to be readily transferred, in contrast to angiosperm species with even one-tenth evolutionary separation.

However, on these maps, there are three instances of apparent segmental inversions, two between Douglas-fir and pine, and one between pine and spruce. A case where a pair of markers is reversed is likely due to mistaken orthology. However, between Douglas-fir and pine, four markers are involved with an apparent rearrangement (involving orthologous markers 3–6 in pine, which are linked to Douglas-fir). To have four, instead of two, markers involved in an apparent inversion provide much stronger evidence of true orthology. This suggests that the genetic system is less homologous in Douglas-fir, as indeed its time since evolutionary divergence is greater than between pine and spruce, and that there are limits to the transfer of genomic information between conifer taxa.

5.6 Quantitative Trait Loci Mapping in Conifers

The last aspect of mapping in conifers involves identifying genes underlying quantitative traits along the marker maps. The co-segregation of genetic markers with phenotypes within pedigrees can reveal individual genes underlying quantitative traits. The ultimate objective of QTL mapping is to infer the "genetic architecture" of the quantitative trait, e.g., the numbers of gene loci controlling the trait, the magnitudes of their effects, and their location in LGs, epistatic interactions, and gene-by-environment interactions. While the idea of using markers to study quantitative traits dates from Sax (1923), who used single-locus morphological markers as categories for continuous traits, the landmark paper that provided the modern paradigm is Lander and Botstein (1989), who considered the multiple marker mapping of QTL mapping.

QTL mapping involves associating alternative marker alleles with phenotypes in segregating progenies. The major issue in conifers is that parents should be heterozygous for both genetic markers and QTLs. Separate QTL maps (but not marker maps) need to be constructed for each parent. However, if a given marker is heterozygous in both parents, the QTL cannot be assigned to a parent, unless there is a priori knowledge about linkage. Issues about QTL mapping in outbred pedigrees are discussed in Williams (1998).

Candidate genes can also be used as marker loci in QTL mapping. For example, Wheeler et al. (2005) used 29 putative cold-hardiness candidate genes for mapping cold-hardiness related traits in Douglas-fir, and Pot et al. (2006) used 10 candidate genes involved in the biosynthesis and deposition of the secondary cell wall in maritime pine. Recently, Pelgas et al. (2011) used 836 candidate genes as marker loci for QTL mapping of

various adaptive traits. However, location of the candidate gene within a QTL interval is not proof of causality; further testing using functional or association genetic approaches are required for proof that such a candidate gene underlies the quantitative trait.

5.6.1 Crossing Designs for QTL Analysis

There are several possible experimental designs for QTL detection. In conifers, breeding programs offer genetic material for QTL analyses. The most common test is a progeny test, where a small number (8–20) of open-pollinated progeny are grown, which can estimate the genetic value of the female parent (White et al. 2007). This number is too small to estimate QTL effect in any single family, and variation among parents for QTL content adds complexity. In outbred conifers, each parent will have different QTLs. Ideally, large (> 100) full sib families are needed for reliable inference of QTL, in order to avoid the bias of inference of QTL effect due to small family size (Beavis 1998). However this ignores variation of QTL among individuals in the larger population.

For QTL mapping, the two major designs are the "inter-specific F_1" design, and the "three-generation full-sib pedigree" design. Interspecific F_1 designs are rare if non-existent in conifers as they are based upon hybridization between subspecies that are usually fixed for alternative QTL and alternative markers. The three-generation design has been employed for Douglas-fir and loblolly pine. An intermediate situation is often encountered: factorial crossing designs with 10–50 progeny per family (a complete factorial design is where N males are individually crossed with M females, resulting in NM families). This design is used to estimate general and specific combining abilities on both the male and female side (Verhoeven et al. 2005).

QTLs found in one pedigree may not exist in other pedigrees. "Validation" of QTLs is the replication of the finding on a second population. In association genetic studies, validation in other populations is a requirement. In QTL studies, this is a difficult task as emphasized by Williams et al. (2007). They point out that a given QTL may not be polymorphic in the second pedigree, and that other segregating QTLs can cause gene interactions that obscure the QTL in another pedigree. In conifers, replicate pedigrees are few due to the long generation times.

The density of markers needed for QTL mapping need not be that high. Darvasi et al. (1993) found that QTL detection probability for a map with 10 cM spacing of markers was virtually the same as that for a map with an infinite number of markers. Since SNPs usually have just two alleles per locus, a larger number of SNPs are needed to obtain the ideal 10 cM marker spacing. SSR markers are usually highly heterozygous and if on the order of 100 markers are used; their distribution is sufficiently dense such that a

given individual is usually heterozygous for at least one locus over a small (10 cM) genome interval.

5.6.2 QTL Traits of Interest

As in the choice of markers for genetic mapping, the phenotypic traits of interest need to be identified. In conifers, the two main phenotypic traits targeted in breeding programs are growth characteristics and wood quality. Total volume, height and ring width are usually used as growth measures. Wood quality is defined in terms of end-uses, and often involves several traits related to wood density, chemical composition and fiber properties. In the area of tree adaptation, phenological traits (timing of bud set and bud burst), as well as cold-hardiness, are traits of interest.

New technologies are increasing the types and numbers of quantitative traits that can be examined, and thus studied for their QTL architecture. At the wood quality level, traits such as stem straightness, stiffness, wood specific gravity, fiber coarseness, and microfibril angle can be measured with x-ray diffraction, the SilvaScan technology, or near-infrared technology (Byram et al. 2005). At the gene level, microarray technologies allow monitoring of a vast number of gene transcripts, whose expression levels are regarded as quantitative traits. Genes involved with the lignin biosynthetic pathway are often of interest, as these genes are putatively involved with wood quality and perhaps phenology. Wood cellulose carbon isotope composition, $\delta^{13}C$, is another important trait of interest, as it is regarded as a time integrated estimate of water use efficiency. A vast number of metabolites can also be assayed via gas chromatography, especially when interfaced with mass spectrometry or high performance liquid chromatography. Like gene expression, metabolite levels can also be considered a quantitative trait; however, they are not directly tied to a gene locus like gene expression levels are. Considering global climate change it becomes very important to study genetic control of adaptive traits such as phenology, cold-hardiness and drought resistance related traits.

5.6.3 QTL Maps

5.6.3.1 Loblolly Pine

In the first QTL map for a conifer, Groover et al. (1994) inferred male and female QTL maps in loblolly pine from a full-sib family of 177 progeny assayed for RFLPs. Five genome regions contained one or more RFLP loci for wood specific gravity. In an analysis of male-female QTL homology, they inferred that the male can have a different QTL segregating at the same locus than the female, and that these alleles can have epistatic interactions.

Following this original work, Knott et al. (1997) analyzed the same data for evidence of multiple QTL in the same linkage interval, finding discordant results with Groover et al. (1994). Kaya et al. (1999) used the pedigree of Groover et al. (1994), termed "*qtl*", as well as second pedigree, "*base*", used previously by Devey et al. (1994b). Thirteen height and eight diameter QTLs were detected, suggesting control by few genes of large effect. However, a given QTL was rarely expressed in multiple years or multiple genetic backgrounds.

A series of works then ensued with the "*qtl*" pedigree. Sewell et al. (2000) used the "*qtl*" pedigree to infer physical traits of wood: wood specific gravity (*wsg*), volume percentage of latewood (*vol%*) and microfibril angle (*mfa*), in both earlywood and latewood. Nine unique QTLs were detected for wood specific gravity, five for volume percentage of latewood, and five for microfibril angle (*mfa*). Most QTL for specific gravity were specific to either earlywood or latewood, whereas each *mfa* QTL occurred in both earlywood and latewood. Sewell et al. (2002) found eight unique chemical wood property QTLs, with differences among populations for QTL. Brown et al. (2003) stressed that verification of QTL is necessary, comparing inferred QTL among populations and within populations for different years. They found that QTL expressed within pedigrees were more stable than QTL expressed among pedigrees.

An unusual approach to QTL mapping, which takes advantage of the conifer megagametophyte, was undertaken by Gwaze et al. (2003). As megagametophytes are haploid, QTL haplotypes can be traced from the offspring back to individual founders in outbred pedigrees by combining founder-origin probabilities with fully informative flanking markers. A large QTL accounting for 11.3 % of the phenotypic variance in the growth rate was detected in a loblolly pine pedigree; the QTL haplotype was traced from offspring to its founder, GP_3.

5.6.3.2 Maritime Pine

Some of the earliest conifer QTL studies also occurred in *Pinus pinaster*. Plomion et al. (1996a) assayed 126 F_2 progeny for RAPD markers, including assay of megagametophytes to determine the linkage phase of the parents. Height growth components related to the initiation (controlled by the apical meristem) and elongation of shoot cycles (controlled by the subapical meristem) were mapped to different chromosomes, suggesting that the activity of these meristems is controlled by separate genetic mechanisms. Plomion et al. (1996b) further studied this cross to find a major QTL for delta 3-carene, a monoterpene, which is a constituent of turpentine. In addition, a qualitative approach found that the ''C'' locus that controls the relative

quantity of delta 3-carene was associated with RAPD markers near the major QTL. This was the first study of co-localization of QTL.

Markussen et al. (2003) found 10 QTLs for height or diameter and 40 QTLs for seven wood parameters in *P. pinaster*. They found that two SSR markers linked to QTL also were linked in a QTL mapped for *P. taeda* (Devey et al, 1999); such markers could be used for comparative QTL studies. Using a second *P. pinaster* three-generation pedigree, Brendel et al. (2002) found four QTLs for $\delta^{13}C$ (the first time found in a tree) and two QTLs for ring width, but they did not co-locate with the $\delta^{13}C$ QTL. On the same pedigree, Pot et al. (2006) detected 54 QTLs. QTL for different traits in the same map position also showed genetic correlations as estimated by traditional quantitative genetic analyses. Chagné et al. (2003) compared QTL maps of Maritime pine and loblolly pine, using 32 common mapped ESTP markers. The positions of two QTLs controlling wood density and cell wall components were conserved between the two species. This was the first ever comparison of QTL maps between conifer species.

5.6.3.3 Radiata Pine

In *Pinus radiata,* efforts for QTL mapping were directed towards eventual use for marker-assisted selection (MAS; the use of specific allelic variants detected in mapping population for tree improvement in unrelated populations). In the first investigation (Emebiri et al. 1998a), haploid megagametophytes were assayed, then progeny of the same individuals grown up to evaluate traits for QTL analysis. This is not a pseudo-testcross design, but rather it evaluates QTLs from the female parent only. From 222 RAPD markers, stem diameter, volume and height were compared at 5 months, and at 1, 2 and 3 years of age. In a second study, four QTLs for stem growth efficiency were found, which accounted for 8.5–36.4% of the population variance (Emebiri et al. 1998b). In a third study, wood density was evaluated at three stages (Kumar et al. 2000). The results suggested that early selection can be used in order to increase juvenile wood density, although the putative QTLs detected in this study need to be verified in an independent population.

Devey et al. (2004) mapped QTL for juvenile wood density (JWD) and diameter at breast height (DBH) using a large full-sib family. The percent variance accounted for by several QTL ranged from 0.78% to 3.58%, suggesting a genomic architecture of many genes with small effect. Two unrelated "bridging" families were chosen to identify markers for MAS. Four markers showed consistent association with JWD, providing the first basis for MAS in a conifer.

5.6.3.4 Scots Pine

In *Pinus sylvestris*, Lerceteau et al. (2000) generated both male and females using the two-way pseudo-testcross strategy. On the female size, 12 QTLs were detected, the largest for frost hardiness. A cluster of QTLs for tree height, trunk diameter and volume was located on one LG. On the male map, four QTLs for trunk diameter and volume were detected. Yazdani et al. (2003) also adopted the pseudo-testcross method, and found QTLs for shoot elongation; growth cessation and cold acclimation were found on both maps. Their study concluded that major QTLs control growth rhythm and autumn cold acclimation.

5.6.3.5 Pine hybrids

In the only QTL study of a conifer hybrid (*Slash pine x Caribbean Pine*), a pseudo-testcross QTL detection strategy was used to identify QTLs for wood density, secondary growth, and dry wood mass in a pedigree of size 133 (Shepherd et al. 2003b). Twelve QTLs were identified that clustered into four LGs in the slash pine parent and in only one group in the Caribbean pine parent. QTLs that influenced density and ring width did not co-locate, suggesting independent inheritance of these characters. Two other pedigrees were more recently mapped for QTLs for adventitious rooting (Shepherd et al. 2006). Most small to moderate effect QTL were congruent between the two pedigrees, while a large effect QTL was found only in one pedigree, and was postulated to be a between-species effect. Targeting between-species effects for improvement in synthetic hybrid populations may increase the efficacy and predictability of hybrid breeding.

5.6.3.6 Douglas-fir

A series of studies used a three-generation pedigree to examine various classes of traits for QTL in Douglas-fir (*Pseudotsuga menziesii*). Jermstad et al. (2001a) genotyped 192 progeny for 74 evenly distributed RFLP markers found by Jermstad et al. (1998). Thirty three QTL for timing of spring bud flush were found, and measurements for each of 3 years and 2 test sites showed that several QTLs influence the timing of bud flush over multiple years within sites but not between sites, indicating major QTL of consistent effect within sites but interactions with environment between sites. Using the same material, Jermstad et al. (2001b) found 11 and 15 QTLs affecting fall and spring cold-hardiness, respectively. Three different shoot tissues phenotyped for spring hardiness showed similar QTL, while different tissues phenotyped for fall hardiness showed little QTL similarity, supporting previous reports that spring tissues are more synchronized than fall tissues.

Jermstad et al. (2003) again used the same pedigree and markers, but for additional individuals totaling 460, to investigate QTL interactions of many of the above traits with photoperiod, moisture stress, winter chilling, and spring temperature. In the first investigation of QTL interaction with environment, they found two QTL-by-treatment interactions for growth initiation traits, and several QTL-by-treatment interactions for growth cessation traits. Finally, Wheeler et al. (2005) evaluated QTL for cold-hardiness via artificial freezing and various cold injury assessment methods in two pedigrees of size 170 and 383. Six QTL were found in the first pedigree, eight in the second, of which four were shared between the pedigrees; 17 of 29 putative cold-hardiness candidate genes identified from ESTs were located within the QTL intervals, thus identifying them as high priority for association studies. These works with Douglas-fir demonstrate a unique opportunity of working with trees: long-lived species allow "*immortal*" pedigrees that can be repeatedly phenotyped for different traits after genotyping.

Finally, QTL analyses are normally conducted in single pedigrees. In contrast, Ukrainetz et al. (2008b) examined eight full-sib families, each of size 40 progeny, for wood-related QTLs, using the software "QTL Express" (Seaton et al. 2002). They found that wood fiber and density traits both showed the lowest number of QTLs (3) with relatively small effects; wood chemistry traits showed more QTLs (7), while ring density traits large numbers of QTLs (78) and interesting patterns of temporal variation. Growth traits gave just five QTLs but of major effect. These wood quality traits are the widest suite of traits yet examined for QTL analysis in a conifer. Moreover, examination of multiple families for QTL gives a population perspective of the true extent of QTL variation.

5.6.3.7 Norway spruce

Markussen et al. (2004) employed bulked segregant analysis and AFLP markers to compare Norway spruce (*Picea abies*) individuals with high and low wood density. Of 107 polymorphic AFLP markers, 15 markers showed significant linkage to wood density, and two of these were found to predict wood density in unrelated full-sib families. Markussen et al. (2005) extended this strategy to compare individuals with high and low extractives content. Of 14 polymorphic AFLP markers were detected between the pools, one marker was linked to low extractives content and subsequently verified as above. Recently, a full-sib family of size 250 has been assayed for *Heterobasidion* (root rot), with the objective of mapping QTLs and identifying candidate genes conferring reduced susceptibility to *Heterobasidion* spp. (Jenny Arnerup and Jan Stenlid, Univ of Uppsala, pers. comm.).

5.6.3.8 North American Spruce Species

No QTL mapping studies have been conducted in spruce until recently. In the Quebec Arborea genome project, two pedigrees of white spruce, of size 395 and 740, have been established and genotyped for 768 and 1,536 gene SNPs, respectively, using the Illumina GoldenGate assay. Experiments on different sites involving clonal propagation of root cuttings have been used to evaluate genotype-by-environment interactions for growth and adaptive traits (Pelgas et al. 2011). About 34 QTL clusters each explaining generally below 15% of phenotypic variance were found for bud flush, bud set and height growth, with about 20% of these replicated between mapping populations and 50% of them with spatial or temporal stability. At least three occurences of overlapping QTLs were noted, indicative of potential pleiotropic effects. On a smaller scale, a black spruce pedigree of size 283 is being studied for wood quality and phenology traits (J Prunier et al. unpubl. data). As the genes have already been mapped in both this pedigree and in the white spruce pedigrees, this will offer an excellent opportunity to assess QTL homology across species.

The British Columbia Treenomix genome project has worked with two factorial crosses from the spruce weevil resistance breeding program (see Alfaro et al. 2004). In the first, involving Interior spruce, 369 progeny in 3 × 2 factorial were genotyped for 253 informative SNP markers using the Illumina GoldenGate assay (I Porth et al. unpubl. data). Over 300 metabolites were also assayed (R Dauwe et al. unpubl. data). The second cross, involving Sitka spruce, is currently being assayed. An approach called "genetical genomics" may also identify previously unidentified networks of genes unique to conifers.

5.6.3.9 Sugi

Yoshimaru et al. (1998) mapped QTLs for growth, flowering and rooting ability in Sugi (*Crypomeria. japonica*). Growth is one of the most important traits for timber-producing woody species and also for carbon dioxide fixation to mitigate global warming. QTLs for juvenile growth, including height and diameter of basal area, were mapped. Flowering is essential for reproduction, but is not necessary for timber production. If the expression of flowering could be controlled, it would be useful not only for breeding but also for forestry and the environment. QTLs for male and female flowers have been mapped at two locations each, respectively. The rooting ability of this species is very important for clonal forestry in the southwestern part of Japan, especially in Kyushu Island. QTLs for rooting ability were found but there were not highly significant in the family used in the study (Yoshimaru et al. 1998). Wood quality QTLs, specifically modulus of elasticity (an important indicator of wood strength), have also been mapped in Sugi (Kuramoto et al. 2000).

Recently, pollinosis (human allergies to pollen) has become a serious social problem; 10 to 20% of Japanese have pollinosis to pollen from Sugi because of a large plantation, which now has matured to flowering. As a countermeasure, the male-sterile lines of *C. japonica* are planned to be used for reforestation. Some male-steriles seem to be controlled by a single recessive locus (Taira et al. 1993). To determine the location of the locus on the linkage map, co-dominant DNA markers have been used for mapping of the gene, using SSRs (Moriguchi et al. 2003; Tani et al. 2004), EST-SSRs (Y Moriguchi et al. unpubl. data), and SNPs (T Ujino-Ihara et al. unpubl. data). After the genome location of this male-sterile gene is found, a selective marker will be developed and used for selection of the male-sterile individuals from the plantation forests and plus trees as breeding materials.

5.7 Prospects

In a seminal review, Remington and Purugganan (2003) stated that future research in plants should expand the number of traits that are intensively studied and make greater use of QTL mapping in wild plant taxa, especially those undergoing adaptive radiations, while continuing to draw on insights from model plants. Conifers are inherently non-domesticated (e.g., wild plant taxa) and the resources provided by breeding programs and genome projects will provide rich resources for testing of candidate gene-trait associations in wild populations, genetic mapping in hybrid zones, and microarray analyses of gene expression.

In conifers, comparative analyses of genetic maps will continue to be a fertile ground for future studies. In sunflower species, a comparative study showed that in the face of extensive hybridization and gene flow, species integrity is maintained (Strasburg et al. 2009). There are many examples of hybrid zones in conifer species, such as the hybridization between Englemann spruce and white spruce in British Columbia. There have been no such studies in conifers that compare patterns of genetic divergence and diversity along chromosomal segments, which can reveal divergent selection for speciation. In conifers, few studies involving "genome scans" have been done (but see Namroud et al. 2008).

Another approach possible for conifers is to use "hitchhiking mapping" to identify regions of recent selective sweeps, due to adaptive divergence. This method starts from a genome scan using a randomly spaced set of molecular markers followed by a fine-scale analysis in the flanking regions of the candidate regions under selection. In fish, the hitchhiking approach identified a selective sweep around candidate locus *Stn90* (Makinen et al. 2008). Fine scale genome maps will help identify candidate loci for adaptation in conifers, particularly those involved with strong ecological

gradients, such as that found Sitka spruce from coastal California to coastal Alaska (see Mimura and Aitken 2007).

Yet another new avenue for using QTL maps is "genetical genomics", which combines genetic mapping with gene expression analysis. It uses variation of gene expression induced by segregation within mapping populations to infer interactions among expressed genes or metabolites. Gene networks, and even directed gene networks, can be inferred by the joint analysis of marker genotypes and gene expression and metabolite levels (Rockman 2008). In the Treenomix II project, two genetical genomic studies are nearing completion. These involve a 22K member cDNA microarray, hundreds of assayed metabolites, and scores for weevil resistance in both white spruce pedigree (I Porth et al. unpubl. data; R Dauwe et al. unpubl. data) and a Sitka spruce pedigree (S Verne et al. unpubl. data).

Recently a number of "next-generation" sequencing technologies have been invented, which can sequence fragments of DNA at astoundingly higher rates compared to Sanger sequencing. These include the Illumina/Solexa, ABI/SOLiD, 454/Roche, Pacific Biosciences/SMRT and Helicos (Morozova and Marra 2008). To date, these technologies have been applied mostly in non-marker contexts, such as whole-genome sequencing (Bentley et al. 2008), targeted resequencing (Gnirke et al. 2009), discovery of transcription factor binding sites, transcript and non-coding RNA expression profiling, and other functional genomic studies (Eveland et al. 2008). These technologies should greatly facilitate genotyping of mapping populations for mapping through direct and parallel sequencing of multiple individuals.

Finally, and last but not least, for the past several years, there has been an initiative to sequence a conifer genome, starting with the seminal paper of Neale et al. (1994). There are several initiatives such as the Pine Genome Initiative (*http://pinegenomeinitiative.org/*) and the International Conifer Genome Initiative (*http://www.pinegenome.org*). It is not clear what strategy is the best, and current initiatives are exploring alternatives. Fine scale genetic mapping will clearly enable the assembly of contigs based upon shotgun sequencing (for example, in the monkeyflower genome project, John Willis pers. comm.). A current goal of the Arborea project is to map 10,000 genes in white spruce (J Bousquet pers. comm.). Other workers in the USA, Canada and Spain have embarked upon exploratory BAC sequencing and gene enrichment of the repetitive genome to discover the structure of conifer genomes, using "gene space" explorations developed such as for maize (Liu et al. 2007). These approaches will interface with genetic mapping to help assemble the first conifer genome.

References

Acheré V, Faivre-Rampant P, Jeandroz S, Besnard G, Markussen T, Aragones A, Fladung M, Ritter E, Favre J-M (2004) A full saturated linkage map of *Picea abies* including AFLP, SSR, ESTP, 5S rDNA, and morphological markers. Theor Appl Genet 108: 1602–1613.

Acheré V, Favre J-M, Besnard G, Jeandroz S (2005) Genomic organization of molecular differentiation in Norway spruce (*Picea abies*). Mol Ecol 14: 3191–3201.

Adams WS, Strauss SH, Copes DL, Griffin AR (1992) Population Genetics of Forest Trees. Klewer, Dordrecht, The Netherlands.

Agrawal AF, Hadany L, Otto SP (2005) The evolution of plastic recombination. Genetics 171: 803–812.

Alfaro RI, vanAkker L, Jaquish B, King J (2004) Weevil resistance of progeny derived from putatively resistant and susceptible interior spruce parents. For Ecol Manag 202: 369–377.

Amarasinghe V, Carlson JE (1998) Physical mapping and characterization of 5S rRNA genes in Douglas-fir. J Hered 89: 495–500.

Arcade A, Anselin F, Faivre Rampant P, Lesage MC, Pâques LE, Prat D (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. Theor Appl Genet 100: 299–307.

Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: AH Paterson (ed) Molecular Dissection of Complex Traits. CRC Press, Boca Raton, Florida, USA, pp 145–162.

Bentley DR, Balasubramanian S, Swerdlow HP, et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53–59.

Binelli G, Bucci G (1994) A genetic linkage map of *Picea abies* Karst., based on RAPD markers, as a tool in population genetics. Theor Appl Genet 88: 283–288.

Bornet B, Branchard M (2001) Nonanchored Inter Simple Sequence Repeat (ISSR) markers: reproducible and specific tools for genome fingerprinting. Plant Mol Biol Rep 19: 209–215.

Brendel O, Pot D, Plomion C, Rozenberg P, Guehl J-M (2002) Genetic parameters and QTL analysis of delta C-13 and ring width in maritime pine. Plant Cell Environ 25: 945–953.

Brown GR, Carlson JE (1997) Molecular cytogenetics of the genes encoding 18*s*-5.8*s*-26*s* rRNA and 5s rRNA in two species of spruce (*Picea*). Theor Appl Genet 95: 1–9.

Brown GR, Newton CH, Carlson JE (1998) Organization and distribution of a *Sau*3A tandem repeated DNA sequence in *Picea* (Pinaceae) species. Genome 41: 560–565.

Brown GR, Kadel III EE, Bassoni DL, Kiehne KL, Temesgen B, van Buijtenen JP, Sewell MM, Marshall KA, Neale DB (2001) Anchored reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomics. Genetics 159: 799–809.

Brown GR, Bassoni DL, Gill GP, Fontana JR, Wheeler NC, Megraw RA, Davis MF, Sewell MM, Tuskan GA, Neale DB (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.

Buetow KH (1991) Influence of aberrant observations on high-resolution linkage analysis outcomes. Am J Hum Genet 49: 985–994.

Byram TD, Myszewski JH, Gwaze DP, Lowe WJ (2005) Improving wood quality in the western gulf forest tree improvement program: the problem of multiple breeding objectives. Tree Genet Genomes 1: 85–92.

Cairney J, Pullman GS (2007) The cellular and molecular biology of conifer embryogenesis. New Phytol 176: 511–536.

Carlson JE, Traore A, Agrama HA, Krutovsky KV (2007) Douglas-fir. In: C Kole (ed) Genome Mapping and Molecular Breeding in Plants, vol 7: Forest Trees. Springer, Berlin, Tokyo, pp 199–210.

Chagné D, Lalanne C, Madur D, Kumar S, Frigério J-M, Krier C, Decroocq S, Savouré A, Bou-Dagher-Kharrat M, Bertocchi E, Brach J, Plomion C (2002) A high density genetic map of maritime pine based on AFLPs. Ann For Sci 59: 627–636.

Chagné D, Brown G, Lalanne C, Madur D, Pot D, Neale D, Plomion C (2003) Comparative genome and QTL mapping between maritime and loblolly pines. Mol Breed 12: 185–195.

Chagné D, Chaumeil P, Ramboer A, Collada C, Guevara A, Cervera MT, Vendramin GG, Garcia V, Frigerio J-M, Echt C, Richardson T, Plomion C (2004) Cross-species transferability and mapping of genomic and cDNA SSRs in pines. Theor Appl Genet 109: 1204–1214.

Chakravarti A, Lasher LK, Reefer JE (1991) A maximum likelihood method for estimating genome length using genetic linkage data. Genetics 128: 175–182.

Clark AG (1990) Inference of haplotypes from PCR-amplified samples of diploid populations. Mol Biol Evol 7: 111–122.

Conkle MT (1981) Proceedings of the Symposium on Isozymes of North American Forest Trees and Forest Insects, 27 July, 1979, Berkeley, California. Gen Tech Rep PSW-GTR-48: Pacific Southwest Forest and Range Exp Stn, Forest Service, US Department of Agriculture, Berkeley, CA, USA.

Costa P, Pot D, Dubos C, Frigerio JM, Pionneau C, Bodenes C, Bertocchi E, Cervera M-T, Remington DL, Plomion C (2000) A genetic map of Maritime pine based on AFLP, RAPD and protein markers. Theor Appl Genet 100: 39–48.

Darvasi A, Weinreb A, Minke V, Weller JI, Soller M (1993) Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics 134: 943–951.

Devey ME, Fiddler TA, Liu BH, Knapp SJ, Neale DB (1994a) An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. Theor Appl Genet 88: 273–278.

Devey ME, Fiddler TA, Liu BH, Knapp SJ, Neale DB (1994b) An RFLP linkage map for loblolly pine based on a 3-generation outbred pedigree. Theor Appl Genet 88: 273–278.

Devey ME, Bell JC, Smith DN, Neale DB, Moran GF (1996) A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers. Theor Appl Genet 92: 673–679.

Devey ME, Sewell MM, Uren TL, Neale DB (1999) Comparative mapping in loblolly and radiata pine using RFLP and microsatellite markers. Theor Appl Genet 99: 656–662.

Devey ME, Carson SD, Nolan MF, Matheson AC, Te Riini C, Hohepa J (2004) QTL associations for density and diameter in *Pinus radiata* and the potential for marker-aided selection. Theor Appl Genet 108: 516–524.

Echt CS, Nelson CD (1997) Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). Theor Appl Genet 94: 1031–1037.

Echt CS, Saha S, Krutovsky KV, Wimalanathan K, Erpelding JE, Liang C, Nelson CD (2011) An Annotated genetic map of loblolly pine based on microsatellite and cDNA markers. BMC Genetics (in press).

Eckert AJ, Pande B, Ersoz ES, Wright MH, Rashbrook VK, Nicolet CM, Neale DB (2009) High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (*Pinus taeda* L.). Tree Genet Genomes 5: 225–234.

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99: 125–132.

Emebiri LC, Devey ME, Matheson AC, Slee MU (1998a) Age-related changes in the expression of QTLs for growth in radiata pine seedlings. Theor Appl Genet 97: 1053–1061.

Emebiri LC, Devey ME, Matheson AC, Slee MU (1998b) Interval mapping of quantitative trait loci affecting NESTUR, a stem growth efficiency index of radiata pine seedlings. Theor Appl Genet 97: 1062–1068.

Ercht CS, Saha S, Krutovsky KV, Wimalanathan K, Erpelding JE, Liang C, Nelson CD (2011) An annotated genetic map of loblolly pine based on microsatellite and cDNA markers. BMC Genetics (in press).

Eveland AL, McCarty DR, Koch KE (2008) Transcript profiling by 3'-untranslated region sequencing resolves expression of gene families. Plant Physiol 146: 32–44.

Fulton TM, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14: 1457–1467.

Gale MD, Devos KM (1998) Comparative genetics in the grasses. Proc Natl Acad Sci USA 95: 1971–1974.

Gasbarra D, Sillanpää MJ (2006) Constructing the parental linkage phase and the genetic map over distances < 1 cM using pooled haploid DNA. Genetics 172: 1325–1335.

Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust E, Brockman W, Fennell T, Giannoukos G, Fisher S, Russ C, Gabriel S, Jaffe DB, Lander ES, Nusbaum C (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol 27: 182–189.

Göçmen B, Jermstad KD, Neale DB, Kaya Z (1996) Development of random amplified polymorphic DNA markers for genetic mapping in Pacific yew (*Taxus brevifolia*). Can J For Res 26: 497–503.

Gogarten JP, Olendzenski L (1999) Orthologs paralogs and genome comparisons. Curr Opin Genet Dev 9: 630–636.

Gosselin I, Zhou Y, Bousquet J, Isabel N (2002) Megagametophyte-derived linkage maps of white spruce (*Picea glauca*) based on RAPD, SCAR and ESTP markers. Theor Appl Genet 104: 987–997.

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137: 1121–1137.

Groover A, Devey M, Fiddler T, Lee J, Megraw R, Mitchel-Olds T, Sherman B, Vujcic S, Williams C, Neale D (1994) Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics 138: 1293–1300.

Groover AT, Williams CG, Devey ME, Lee JM, Neale DB (1995) Sex-related differences in meiotic recombination frequency in *Pinus taeda*. J Hered 86: 157–158.

Gwaze DP, Zhou Y, Reyes-Valdes MH, Al-Rababah MA, Williams CG (2003) Haplotypic QTL mapping in an outbred pedigree. Genet Res 81: 43–50.

Hayashi E, Kondo T, Terada K, Kuramoto N, Goto Y, Okamura M, Kawasaki H (2001) Linkage map of Japanese black pine based on AFLP and RAPD markers including markers linked to resistance against the pine needle gall midge. Theor Appl Genet 102: 871–875.

Hu X-S, Goodwillie C, Ritland KM (2004) Joining genetic linkage maps using a joint likelihood function. Theor Appl Genet 109: 996–1004.

Hudson EE (2005) Development of a genetic linkage map in *Abies nordmanniana* accession 9M and assessment of disease resistance to *Phytophthora cinnamomi*. North Carolina State Univ, Raleigh, North Carolina, USA.

Hulbert SH, Ilott TW, Legg EJ, Lincoln SE, Lander ES, Michelmore RW (1988) Genetic analysis of the fungus *Bremia lactucae*, using restriction fragment length polymorphisms. Genetics 120: 947–958.

Huynen MA, Bork P (1998) Measuring genome evolution. Proc Natl Acad Sci USA 95: 5849–5856.

Iwata H, Ujino-Ihara T, Yoshimura K, Nagasaka K, Mukai Y, Tsumura Y (2001) Cleaved amplified polymorphic sequence markers in sugi, *Cryptomeria japonica* D. Don, and their locations on a linkage map. Theor Appl Genet 103: 881–895.

Jermstad KD, Bassoni DL, Wheeler NC, Neale DB (1998) A sex-averaged genetic linkage map in coastal Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco var '*menziesii*') based on RFLP and RAPD markers. Theor Appl Genet 97: 762–770.

Jermstad KD, Bassoni DL, Jech KS, Wheeler NC, Neale DB (2001a) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. I. Timing of vegetative bud flush. Theor Appl Genet 102: 1142–1151.

Jermstad KD, Bassoni DL, Wheeler NC, Anekonda TS, Aitken SN, Adams WT, Neale DB (2001b) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. II. Spring and fall cold-hardiness. Theor Appl Genet 102: 1152–1158.

Jermstad KD, Bassoni DL, Jech KS, Ritchie GA, Wheeler NC, Neale DB (2003) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir. III. Quantitative trait loci-by-environment interactions. Genetics 165: 1489–1506.

Jermstad KD, Eckert AF, Wegrzyn JL, Delfino-Mix A, Davis DA, Burton DC, Neale DB (2010) Comparative mapping in Pinus: sugar pine (*Pinus lambertiana Dougl.*) and loblolly pine (*Pinus taeda* L.). Tree Genet Genomes (in press).

Kang B-Y, Mann IK, Major JE, Rajora OP (2010) Near-saturated and complete genetic linkage map of black spruce (*Picea mariana*). BMC Genomics 11: 515.

Kaya Z, Neale DB (1995) Utility of random amplified polymorphic DNA (RAPD) markers for linkage mapping in Turkish red pine (*Pinus brutia* Ten.). Silvae Genet 44: 110–116.

Kaya Z, Sewell MM, Neale DB (1999) Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine (*Pinus taeda* L.). Theor Appl Genet 98: 586–592.

Kim Y-Y, Choi HS, Kang B-Y (2005) An AFLP-based linkage map of Japanese red pine (*Pinus densiflora*) using haploid DNA samples of megagametophytes from a single maternal tree. Mol Cells 20: 201–209.

Kinlaw CS, Neale DB (1997) Complex gene families in pine genomes. Trends Plant Sci 2: 356–359.

Kirst M, Johnson AF, Baucom C, Ulrich E, Hubbard K, Staggs R, Paule C, Retzel E, Whetten R, Sederoff R (2003) Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. Proc Nat Acad Sci USA 100: 7383–7388.

Knott SA, Neale DB, Sewell MM, Haley CS (1997) Multiple marker mapping of quantitative trait loci in an outbred pedigree of loblolly pine. Theor Appl Genet 94: 810–820.

Komulainen P, Brown GR, Mikkonen M, Karhu A, Garcia-Gil MR, O'Malley D, Lee B, Neale DB, Savolainen O (2003) Comparing EST-based genetic maps between *Pinus sylvestri*s and *Pinus taeda*. Theor Appl Genet 107: 667–678.

Koonin EV (2005) Orthologs, paralogs, and evolutionary genomics. Annu Rev Genet 39: 309–338.

Krutovsky KV, Vollmer SS, Sorensen FC, Adams WT, Knapp SJ, Strauss SH (1998) RAPD genome maps of Douglas-fir. J Hered 89: 197–205.

Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB (2004) Comparative mapping in the Pinaceae. Genetics 168: 447–461.

Krutovsky KV, Elsik CG, Matvienko M, Kozik A, Neale DB (2006) Conserved ortholog sets in forest trees. Tree Genet Genomes 3: 61–70.

Kuang H, Richardson T, Carson S, Wilcox P, Bongarten B (1999) Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. I. Genetic map with distorted markers. Theor Appl Genet 98: 697–703.

Kubisiak TL, Nelson CD, Nance WL, Stine M (1995) RAPD linkage mapping in a longleaf pine × slash pine F_1 family. Theor Appl Genet 90: 1119–1127.

Kumar S, Spelman RJ, Garrick DJ, Richardson TE, Lausberg M, Wilcox PL (2000) Multiple-marker mapping of wood density loci in an outbred pedigree of radiata pine. Theor Appl Genet 100: 926–933.

Kuramoto N, Kondo T, Fujisawa Y, Nakata R, Hayashi E, Goto Y (2000) Detection of quantitative trait loci for wood strength in *Cryptomeria japonica*. Can J For Res 30: 1525–1533.

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185–199.

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181.

Lathrop GM, Lalouel JM, Julier C, Ott J (1985) Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am J Hum Genet 37: 482–498.

Le Dantec L, Chagné D, Pot D, Cantin O, Garnier-Géré P, Bedon F, Frigerio J-M, Chaumeil P, Léger P, Garcia V, Laigret F, de Daruvar A, Plomion C (2004) Automated SNP detection in expressed sequence tags: statistical considerations and application to maritime pine sequences. Plant Mol Biol 54: 461–470.

Lerceteau E, Plomion C, Andersson B (2000) AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in *Pinus sylvestris*: a preliminary study. Mol Breed 6: 451–458.

Li C, Yeh FC (2001) Construction of a framework map in *Pinus contorta* subsp. *latifolia* using random amplified polymorphic DNA markers. Genome 44: 147–153.

Liewlaksaneeyanawin C, Zhuang J, Tang M, Farzaneh N, Lueng G, Cullis C, Findlay S, Ritland CE, Bohlmann J, Ritland K (2009) Identification of COS markers in the Pinaceae. Tree Genet Genomes 5: 247–255.

Liu BH (1998) Statistical Genomics: Linkage, Mapping, and QTL Analysis. CRC Press, Boca Raton, Florida, USA.

Liu R, Vitte C, Ma J, Mahama AA, Dhliwayo T, Lee M, Bennetzen JL (2007) A GeneTrek analysis of the maize genome. Proc Nat Acad Sci USA 104: 11844–11849.

Liu W, Magbanua ZV, Ozkan S, Chouvarine P, Bartlett BD, Peterson DG (2009) BAC libraries for two distantly related conifers, loblolly pine and bald cypress. In: Plant Anim Genome XVII Conf, San Diego, CA, USA.

Mäkinen HS, Shikano T, Cano JM, Merilä J (2008) Hitchhiking mapping reveals a candidate genomic region for natural selection in three-spined stickleback chromosome VIII. Genetics 178: 453–465.

Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: software for genetic mapping in outcrossing species. Hereditas 144: 78–79.

Markussen T, Fladung M, Achere V, Favre JM, Faivre-Rampant P, Aragones A, Silva Perez DD, Harvengt L, Espinel S, Ritter E (2003) Identification of QTLs controlling growth, chemical and physical wood property traits in *Pinus pinaster* (Ait.). Silvae Genet 52: 8–15.

Markussen T, Tusch A, Stephan BR, Fladung M (2004) Identification of molecular markers for selected wood properties of Norway spruce *Picea abies* L. (Karst.) I. Wood density. Silvae Genet 53: 45–50.

Markussen T, Tusch A, Stephan BR, Fladung M (2005) Identification of molecular markers for selected wood properties of Norway spruce *Picea abies* L. (Karst.) II. Extractives content. Silvae Genet 54: 145–152.

Mester D, Ronin Y, Minkov D, Nevo E, Korol A (2003) Constructing large-scale genetic maps using an evolutionary strategy algorithm. Genetics 165: 2269–2282.

Mester DI, Ronin YI, Nevo E, Korol AB (2004) Fast and high precision algorithms for optimization in large-scale genomic problems. Comp Biol Chem 28: 281–290.

Mester DI, Ronin YI, Korostishevsky MA, Pikus VL, Glazman AE, Korol AB (2006) Multilocus consensus genetic maps (MCGM): Formulation, algorithms, and results. Comp Biol Chem 30: 12–20.

Mimura M, Aitken SN (2007) Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. Heredity 99: 224–232.

Moran GF, Bell JC, Hilliker AJ (1983) Greater meiotic recombination in male vs. female gametes in *Pinus radiata*. J Hered 74: 62.

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3: 175–182.

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30: 194–200.

Moriguchi Y, Iwata H, Ujino-Ihara T, Yoshimura K, Taira H, Tsumura Y (2003) Development and characterization of microsatellite markers for *Cryptomeria japonica* D. Don. Theor Appl Genet 106: 751–758.

Morozova O, Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. Genomics 92: 255–264.

Morse AM, Peterson DG, Islam-Faridi MN, Smith KE, Magbanua Z, Garcia SA, Kubisiak TL, Amerson HV, Carlson JE, Nelson CD, Davis JM (2009) Evolution of genome size and complexity in *Pinus*. PLoS One 4: e4332.

Mukai Y, Suyama Y, Tsumura Y, Kawahara T, Yoshimaru H, Kondo T, Tomaru N, Kuramoto N, Murai M (1995) A linkage map for sugi (*Cryptomeria japonica*) based on RFLP, RAPD, and isozyme loci. Theor Appl Genet 90: 835–840.

Murray B (1998) Nuclear DNA amounts in gymnosperms. Ann Bot 82: 3–15.

Namroud M-C, Beaulieu J, Juge N, Laroche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. Mol Ecol 17: 3599–3613.

Neale DB, Krutovsky KV (2004) Comparative genetic mapping in trees: The group of conifers. In: H Lörz, G Wenzel (ed) Biotechnology in Agriculture and Forestry: Molecular Marker Systems in Plant Breeding and Crop Improvement. Springer, Berlin, Germany, pp 267–277 .

Neale DB, Kinlaw CS, Sewell MM (1994) Genetic mapping and DNA sequencing of the loblolly pine genome. For Genet 1: 197–206.

Nelson CD, Nance WL, Doudrick RL (1993) A partial genetic linkage map of slash pine (*Pinus elliottii* Engelm. var. *elliottii*) based on random amplified polymorphic DNAs. Theor Appl Genet 87: 145–151.

Nelson CD, Kubisiak TL, Stine M, Nance WL (1994) A genetic linkage map of longleaf pine (*Pinus palustris* Mill.) based on random amplified polymorphic DNAs. J Hered 85: 433–439.

Nelson CD, Peterson DG, Echt CS, Whetten R, Krutovsky KV, Yuceer C, Dean JF (2008) Pine physical mapping and genome sequencing. In: Plant Anim Genome XVI Conf, San Diego, CA, USA.

Nikaido AM, Ujino T, Iwata H, Yoshimura K, Yoshimura H, Suyama Y, Murai M, Nagasaka K, Tsumura Y (2000) AFLP and CAPS linkage maps of *Cryptomeria japonica*. Theor Appl Genet 100: 825–831.

Nurul Islam-Faridi M, Dana Nelson C, Kubisiak T (2007) Reference karyotype and cytomolecular map for loblolly pine (*Pinus taeda* L.) Genome 50: 241–251

Ott J (1999) Analysis of human genetic linkage, 3rd edn. Johns Hopkins Univ Press, Baltimore, Maryland, USA.

Paglia GP, Olivieri AM, Morgante M (1998) Towards second generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). Mol Gen Genet 258: 466–478.

Pavy N, Parsons LS, Paule C, MacKay J, Bousquet J (2006) Automated SNP detection from a large collection of white spruce expressed sequences: contributing factors and approaches for the categorization of SNPs. BMC Genom 7: 174.

Pavy N, Pelgas B, Beauseigle S, Blais S, Gagnon F, Gosselin I, Lamothe M, Isabel N, Bousquet J (2008) Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: application to the large and unsequenced genomes of white spruce and black spruce. BMC Genom 9: 21.

Pelgas B, Isabel N, Bousquet J (2004) Efficient screening for expressed sequence tag polymorphisms (ESTPs) by DNA pool sequencing and denaturing gradient gel electrophoresis (DGGE) in spruces. Mol Breed 13: 263–279.

Pelgas B, Bousquet J, Beauseigle S, Isabel N (2005) A composite linkage map from two crosses for the species complex *Picea mariana* × *P. rubens* and analysis of synteny with other Pinaceae. Theor Appl Genet 111: 1466–1488.

Pelgas B, Beauseigle S, Acheré V, Jeandroz S, Bousquet J, Isabel N (2006) Comparative genome mapping among *Picea glauca, P. mariana x P. rubens* and *P. abies*, and correspondence with other Pinaceae. Theor Appl Genet 113: 1371–1393.

Pelgas B, Bousquet J, Meirmans P, Ritland K, Isabel N (2011) QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. BMC Genomics (in press).

Perry DJ, Bousquet J (1998) Sequence-tagged-site (STS) markers of arbitrary genes: development, characterization and analysis of linkage in black spruce. Genetics 149: 1089–1098.

Plomion C, O'Malley DM (1996) Recombination rate differences for pollen parents and seed parents in *Pinus pinaster*. Heredity 77: 341–350.

Plomion C, Bahrman N, Durel C-E, O'Malley DM (1995a) Genomic mapping in *Pinus pinaster* (maritime pine) using RAPD and protein markers. Heredity 74: 661–668.

Plomion C, O'Malley DM, Durel CE (1995b) Genomic analysis in maritime pine (*Pinus pinaster*). Comparison of two RAPD maps using selfed seeds and open-pollinated of the same individual. Theor Appl Genet 90: 1028–1034.

Plomion C, Durel C-E, OMalley DM (1996a) Genetic dissection of height in maritime pine seedlings raised under accelerated growth conditions. Theor Appl Genet 93: 849–858.

Plomion C, Yani A, Marpeau A (1996b) Genetic determinism of delta 3-carene in maritime pine using RAPD markers. Genome 39: 1123–1127.

Pot D, Rodrigues J-C, Rozenberg P, Chantre G, Tibbits J, Cahalan C, Pichavant F, Plomion C (2006) QTLs and candidate genes for wood properties in maritime pine (*Pinus pinaster* Ait.). Tree Genet Genomes 2: 10–24.

Ralph SG, Chun HJE, Kolosova N, Cooper D, Oddy C, Ritland CE, Kirkpatrick R, Moore R, Barber S, Holt RA, Jones SJM, Marra MA, Douglas CJ, Ritland K, Bohlmann J (2008) A conifer genomics resource of 200,000 spruce (*Picea* spp.) ESTs and 6,464 high-quality, sequence-finished full-length cDNAs for Sitka spruce (*Picea sitchensis*). BMC Genom 9: 484.

Ran J-H, Wei X-X, Wang X-Q (2006) Molecular phylogeny and biogeography of *Picea* (Pinaceae): Implications for phylogeographical studies using cytoplasmic haplotypes. Mol Phylogenet Evol 41: 405–419.

Remington DL, Purugganan MD (2003) Candidate genes, quantitative trait loci, and functional trait evolution in plants. Int J Plant Sci 164: S7–S20.

Remington DL, Whetten RW, Liu B-H, O'Malley DM (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theor Appl Genet 98: 1279–1292.

Remm M, Storm CEV, Sonnhammer ELL (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. J Mol Biol 314: 1041–1052.

Ritland C, Ritland K (2000) DNA-fragment markers in plants. In: AJ Baker (ed) Molecular Methods in Ecology. Blackwell Publ, Oxford, UK, pp 208–234.

Ritter E, Aragonés A, Markussen T, Acheré V, Espinel S, Fladung M, Wrobel S, Faivre-Rampant P, Favre J-M (2002) Towards construction of an ultra high density linkage map for *Pinus pinaster*. Ann For Sci 596: 37–643.

Rockman MV (2008) Reverse engineering the genotype-phenotype map with natural genetic variation. Nature 456: 738–744.

Rose AM, Baillie DL (1979) The effect of temperature and parental age on recombination and nondisjunction in *Caenorhabditis elegans*. Genetics 92: 409–418.

Rungis D, Bérubé Y, Zhang J, Ralph S, Ritland CE, Ellis BE, Douglas C, Bohlmann J, Ritland K (2004) Robust simple sequence repeat markers for spruce (*Picea* spp.) from expressed sequence tags. Theor Appl Genet 109: 1283–1294.

Savard L, Li P, Strauss SH, Chase MW, Michaud M, Bousquet J (1994) Chloroplast and nuclear gene sequences indicate Late Pennsylvanian time for the last common ancestor of extant seed plants. Proc Nat Acad Sci USA 91: 5163–5167.

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris.* Genetics 8: 552–560.

Scotti I, Burelli A, Cattonaro F, Chagné D, Fuller J, Hedley PE, Jansson G, Lalanne C, Madur D, Neale D, Plomion C, Powell W, Troggio M, Morgante M (2005) Analysis of the distribution of marker classes in a genetic linkage map: a case study in Norway spruce (*Picea abies* karst). Tree Genet Genomes 1: 93–102.

Seaton G, Haley CS, Knott SA, Kearsey M, Visscher PM (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics 18: 339–340.

Sewell MM, Sherman BK, Neale DB (1999) A consensus map for loblolly pine (*Pinus taeda* L.) I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. Genetics 151: 321–330.

Sewell MM, Bassoni DL, Megraw RA, Wheeler NC, Neale DB (2000) Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). I. Physical wood properties. Theor Appl Genet 101: 1273–1281.

Sewell MM, Davis MF, Tuskan GA, Wheeler NC, Elam CC, Bassoni DL, Neale DB (2002) Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.) II. Chemical wood properties. Theor Appl Genet 104: 214–222.

Shepherd M, Cross M, Dieters MJ, Henry R (2003a) Genetic maps for *Pinus elliottii* var. *elliottii* and *P. caribaea* var. *hondurensis* using AFLP and microsatellite markers. Theor Appl Genet 106: 1409–1419.

Shepherd M, Cross M, Dieters MJ, Harding K, Kain D, Henry R (2003b) Genetics of physical wood properties and early growth in a tropical pine hybrid. Can J For Res 33: 1923–1932.

Shepherd M, Huang S, Eggler P, Cross M, Dale G, Dieters M, Henry R (2006) Congruence in QTL for adventitious rooting in *Pinus elliottii X Pinus caribaea* hybrids resolves between and within-species effects. Mol Breed 18: 11–28.

Skov E, Wellendorf H (1998) A partial linkage map of *Picea abies* clone V6470 based on recombination of RAPD-markers in haploid megagametophytes. Silvae Genet 47: 273–282.

Sobel E, Papp JC, Lange K (2002) Detection and integration of genotyping errors in statistical genetics. Am J Hum Genet 70: 496–508.

Soderlund C, Humphray S, Dunham A, French L (2000) Contigs built with fingerprints, markers, and FPC V4.7. Genome Res 10: 1772–1787.

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3: 739–744.

Strasburg JL, Scotti-Saintagne C, Scotti I, Lai Z, Rieseberg LH (2009) Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. Mol Biol Evol 26: 1341–1355.

Taira H, Teranishi H, Kenda S (1993) A case study of male sterility in sugi (*Cryptomeria japonica*) (in Japanese with English summary). J Jap For Soc 75: 377–379.

Tani N, Takahashi T, Iwata H, Mukai Y, Ujino-Ihara T, Matsumoto A, Yoshimura K, Yoshimaru H, Murai M, Nagasaka K, Tsumura Y (2003) A consensus linkage map for sugi (*Cryptomeria japonica*) from two pedigrees, based on microsatellites and expressed sequence tags. Genetics 165: 1551–1568.

Tani N, Takahashi T, Ujino-Ihara T, Iwata H, Yoshimura K, Tsumura Y (2004) Development and characteristics of microsatellite markers for sugi (*Cryptomeria japonica* D. Don) derived from microsatellite-enriched libraries. Ann For Sci 61: 569–575.

Theissen G (2005) Birth, life and death of developmental control genes: New challenges for the homology concept. Theory Biosci 124: 199–212.

Tong C, Shi J (2004) Constructing genetic linkage maps in Chinese-fir using F1 progeny. Acta Genet Sin 31: 1149–1156.

Travis SE, Ritland K, Whitham TG, Keim P (1998) A genetic linkage map of pinyon pine (*Pinus edulis*) based on amplified fragment length polymorphisms. Theor Appl Genet 97: 871–880.

Tsumura Y, Ogihara Y, Sasakuma T, Ohba K (1993) Physical map of chloroplast DNA in sugi, *Cryptomeria japonica*. Theor Appl Genet 86: 166–172.

Tsumura Y, Suyama Y, Yoshimura K, Shirato N, Mukai Y (1997) Sequence-tagged-sites (STSs) of cDNA clones in *Cryptomeria japonica* and their evaluation as molecular markers in conifers. Theor Appl Genet 94: 764–772.

Tulsieram LK, Glaubitz JC, Kiss G, Carlson JE (1992) Single tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. Bio/Technology 10: 686–690.
Ujino-Ihara T, Yoshimura K, Ugawa Y, Yoshimaru H, Nagasaka K, Tsumura Y (2000) Expression analysis of ESTs derived from the inner bark of *Cryptomeria japonica*. Plant Mol Biol 43: 451–457.
Ujino-Ihara T, Kanamori H, Yamane H, Taguchi Y, Namiki N, Mukai Y, Yoshimura K, Tsumura Y (2005) Comparative analysis of expressed sequence tags of conifers and angiosperms reveals sequences specifically conserved in conifers. Plant Mol Biol 59: 895–907.
Ukrainetz NK, Ritland K, Mansfield SD (2008a) An AFLP linkage map for Douglas-fir based upon multiple full-sib families. Tree Genet Genomes 4: 181–191.
Ukrainetz NK, Ritland K, Mansfield SD (2008b) Identification of quantitative trait loci for wood quality and growth across eight full-sib coastal Douglas-fir families. Tree Genet Genomes 4: 159–170.
Verhoeven KJF, Jannink J-L, McIntyre LM (2005) Using mating designs to uncover QTL and the genetic architecture of complex traits. Heredity 96: 139–149.
Vision TJ, Brown DG, Shmoys DB, Durrett RT, Tanksley SD (2000) Selective mapping: a strategy for optimizing the construction of high-density linkage maps. Genetics 155: 407–420.
Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucl Acids Res 23: 4407–4414.
Weising K, Nybom H, Wolff K, Kahl G (2005) DNA Fingerprinting in Plants: Principles, Methods, and Applications, 2nd edn. CRC Press, Boca Raton, Florida, USA.
Wheeler NC, Jermstad KD, Krutovsky K, Aitken SN, Howe GT, Krakowski J, Neale DB (2005) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. IV. Cold-hardiness QTL verification and candidate gene mapping. Mol Breed 15: 145–156.
White TL, Adams WT, Neale DB (2007) Forest Genetics. CABI Pub, Wallingford, Oxfordshire, UK; Cambridge, Massachusetts, USA.
Wilcox PL, Richardson TE, Corbett GE, Ball RD, Lee JR, Djorovic A, Carson SD (2001) Framework linkage maps of *Pinus radiata* D. Don based on pseudotestcross markers. For Genet 8: 109–117.
Williams CG (1998) QTL Mapping in outbred pedigrees. In: AH Paterson (ed) Molecular Dissection of Complex Traits. CRC Press, Boca Raton, Florida, USA, pp 81–94.
Williams CG, Reyes-Valdes MH, Huber DA (2007) Validating a QTL region characterized by multiple haplotypes. Theor Appl Genet 116: 87–94.
Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl Acids Res 18: 6531–6535.
Wu R, Ma C-X, Painter I, Zeng Z-B (2002) Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. Theor Pop Biol 61: 349–363.
Yazdani R, Yeh FC, Rimsha J (1995) Genomic mapping of *Pinus sylvestris* (L.) using random amplified polymorphic DNA markers. For Genet 2: 109–116.
Yazdani R, Nilsson JE, Plomion C, Mathur G (2003) Marker trait association for autumn cold acclimation and growth rhythm in *Pinus sylvestris*. Scand J For Res 18: 29–38.
Yin T-M, Wang X-R, Andersson B, Lerceteau-Köhler E (2003) Nearly complete genetic maps of *Pinus sylvestris* L. (Scots pine) constructed by AFLP marker analysis in a full-sib family. Theor Appl Genet 106: 1075–1083 .
Yoshimaru H, Ohba K, Tsurumi K, Tomaru N, Murai M, Mukai Y, Suyama Y, Tsumura Y, Kawahara T, Sakamaki Y (1998) Detection of quantitative trait loci for juvenile growth, flower bearing and rooting ability based on a linkage map in sugi (*Cryptomeria japonica* D. Don). Theor Appl Genet 97: 45–50.
Zhang P, Li W, Fellers J, Friebe B, Gill BS (2004) BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. Chromosoma 112: 288–299.

Chapter 5

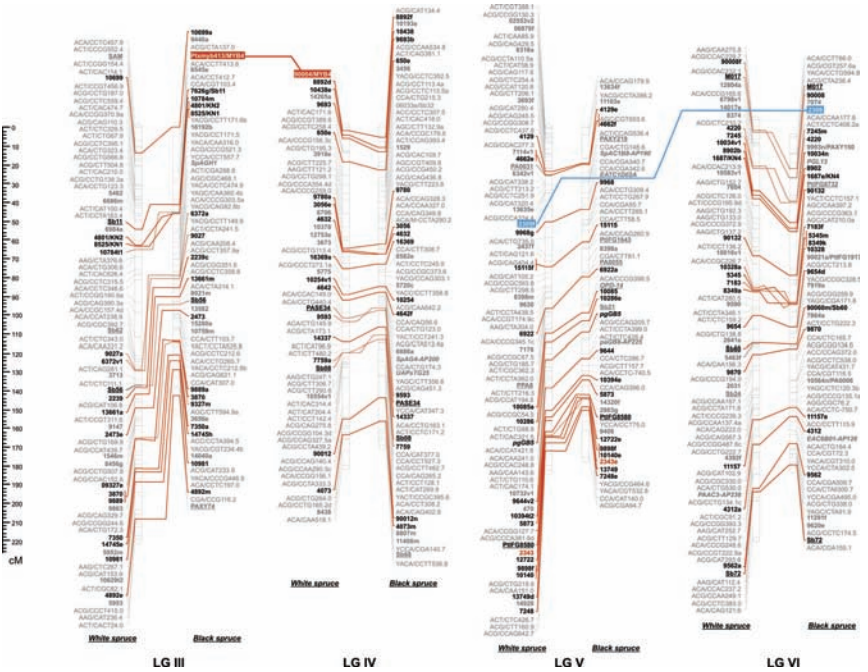

Figure 5-3 Comparison of homologous linkage groups between white spruce (*Picea glauca*) and black spruce (species complex *Picea mariana* × *P. rubens*).