3 Spruce

Jean Bousquet[1], Nathalie Isabel[2], Betty Pelgas[1], Joan Cottrell[3], Dainis Rungis[4,5], and Kermit Ritland[4]

[1] Département des Sciences du bois et de la forêt, Université Laval, Québec G1K 7P4, Canada
[2] Ressources naturelles Canada, Service canadien des forêts, Centre de foresterie des Laurentides, 1055 rue du P.E.P.S., C.P. 10380 succ. Sainte-Foy, Québec G1V 4C7, Canada
[3] Forest Research, UK Forestry Commission, Northern Research Station, Roslin, Midlothian EH25 9SY, UK
[4] Department of Forest Sciences, University of British Columbia, Vancouver, British Columbia V6T1Z4, Canada
e-mail: kermit.ritland@ubc.ca
[5] Current address: LVMI Silava, 111 Rigas st, Salaspils, LV-2169, Latvia

3.1 Introduction

3.1.1 The spruce genus, *Picea*

The genus *Picea* consists of up to 35 species, three quarters of which are Eurasian and one quarter are North American (although reevaluation of the East Asian taxa would reduce this number, c.f. Farjon 1990; Sigurgeirsson and Szmidt 1993). The oldest recognizable spruce fossil dates back to the middle Eocene, around 45 million years ago (LePage 2001). Spruce species generally occur in the subtropical high-altitude, temperate, and boreal regions of the northern hemisphere. In the southern part of their distribution, they mainly occur in mountainous areas, while in the north, spruce occurs throughout the boreal forest, often being the dominant tree species across vast tracts of Scandinavia, Russia, Alaska, and Canada. In the mountains of southwest China and in Japan, spruce is not dominant yet shows the great species diversity, suggesting that eastern Asia is the center of origin for spruce (Wright 1955).

There is lack of agreement among taxonomists regarding the subdivision of the genus *Picea* (Schmidt-Vogt 1977). Based on morphology and crossability studies, early taxonomists divided the genus into three sections: Eupicea (or Morinda), Casicta, and Omorika. Mikkola (1969) recommended recognition of only two sections: Abies and Omorika. Part of the disagreement stems from the different traits used for morphological classification and variable results from crossability studies and because of the too little morphological and anatomical differentiation among spruce taxa (Wright 1955; Mikkola 1969; Weng and Jackson 2000). On the basis of crossability studies, Fowler (1983, 1987) further divided Omorika into two subsections: Omorikoides and Glaucoides, with white spruce, Sitka spruce, and Engelmann spruce assigned to the latter. The lack of agreement among taxonomists suggests that, in comparison to the pines (*Pinus*), the genus *Picea* is relatively monophyletic.

In work yet to be published (Presby-Germano 2003; Bouillé and Bousquet 2006), up to 35 spruce species were sequenced for chloroplast DNA (*rbcL*, *trnTLF, trnK*), mitochondrial DNA (*nad1 B/C* and *nad7 1/2* introns), and nuclear DNA (portions of 4CL). Sequence divergence was relatively low, yielding low bootstrap support for many clades in the inferred phylogenies, and phylogenies were significantly incongruent between genomes. This is indicative of recent speciation and/or more or less recent reticulate evolution. Also, mtDNA phylogenies were geographically more structured than cpDNA phylogenies, and incomplete lineage sorting is evident at nuclear loci (Bouillé and Bousquet 2005; Campbell et al. 2005). Further work involving more intensive gene coverage and population sampling is needed to resolve species groups.

In this chapter, we focus on the spruce species that have received the predominant attention from molecular mapping and breeding perspectives. These are the most economically important species: white spruce, Sitka spruce, black spruce, and Norway spruce.

White Spruce

White spruce [*Picea glauca* (Moench) Voss] is widely distributed in Canada, from the Atlantic to the Pacific coast. It is used for lumber, pulp and paper, and ranks as the second-most important conifer species for re-

Genome Mapping and Molecular Breeding in Plants, Volume 7
Forest Trees
C. Kole (Ed.)

forestation in Canada, with over 100 million seedlings planted yearly. Since the 1950s, the Canadian Forest Service and Provincial ministries in Canada (mainly British Columbia, Quebec, and New Brunswick) have invested considerable effort on white spruce breeding. Some programs are now entering the third generation of breeding (Beaulieu 1996; Tosh and Fullarton 2001; Yanchuk 2001).

White spruce is highly diverse genetically for quantitative characters (Furnier et al. 1991; Jaramillo-Correa et al. 2001), and low to moderate genetic control has been shown for a variety of growth and adaptive traits, and for wood characters (Nienstaedt 1985; Kiss and Yeh 1988; Corriveau et al. 1991; Li et al. 1993, 1997; Rweyongeza et al. 2004). These trends indicate that significant gains could be realized from selection and breeding in this species, and that genomics can provide new inputs into these programs.

Sitka Spruce

Sitka spruce [*Picea sitchensis* (Bong.)] has a narrow distribution along the mainland and offshore islands of the Pacific coast of North America, from central Alaska to northern California. In Alaska, Sitka spruce is the most important timber species (Arno and Hammerly 1977), and in 1995, southeast Alaska accounted for 70% of the total estimated North American production of Sitka spruce. Its wood has a high strength-to-weight ratio, making it useful for diverse applications including turbine blades, sailboat masts, oars, piano sounding boards, and guitars (Hosie 1969; Viereck and Little 1972; Harris 1990).

Outside its natural range, Sitka spruce has played an important role in plantation forestry, particularly in northern Europe (Hermann 1987). It is planted extensively in Britain, Ireland, and Denmark and is a minor plantation species in France and Germany (Hermann 1987). David Douglas introduced Sitka spruce to Britain in 1831; its popularity has increased since then, with plantings increasing from 27% in 1925 to 60% in 1985 (Joyce and O'Carroll 2002). Sitka spruce now accounts for almost 70% of the annual conifer planting stock (Malcolm 1997), with plantations covering over 20% of the British woodland (Cannell and Milne 1995). Breeding strategies for Sitka spruce in Britain have also been conducted (Lee 1993), with molecular approaches implemented (see below).

Sitka spruce harbors much diversity in productive traits among provenances in North America (Fletcher 1992) that are used to develop productive plantations in Europe. However, in North America, Sitka spruce has an endemic susceptibility to the white pine weevil [*Pissoides strobi* (Peck)]. Young trees are attacked and cannot maintain growth leaders, resulting in a bushy growth form. Only in the Queen Charlotte Islands, where the weevil is absent (Hall 1994), can Sitka spruce be reestablished. The reestablishment of plantation forestry in the much more fertile region of Vancouver Island (British Columbia) in the face of this pest is the primary goal of breeding for Sitka spruce in British Columbia.

Black Spruce and Ally Red Spruce

Black spruce [*Picea mariana* (Mill.) B.S.P.] is an abundant transcontinental boreal species in North America (Farrar 1995). Its natural range overlaps with white spruce; it extends across the boreal forest from Newfoundland/northern Quebec to Alaska, but it also occurs in some northern US states. Black spruce is economically very important in New Brunswick, Nova Scotia, and Quebec. In contrast, the range of red spruce [*Picea rubens* (Sarg.)] is more restricted and extends from the southern Appalachians in the United States, where it is often restricted to mountain plateaus, to the Maritime Provinces of eastern Canada (Farrar 1995). Red spruce is a minor component of commercial plantations in the Maritime Provinces.

Black spruce harbors large amounts of genetic variation in quantitative characters, which is an indication of the adaptive capacities of its populations (Khalil 1984; Beaulieu et al. 2004). In addition, various patterns of clinal variation have been reported for germination rate, survival rate, phenology, juvenile growth characters, and hardiness (Dietrichson 1969; Morgenstern 1969; Corriveau 1981; Beaulieu et al. 1989; Parker et al. 1994). Red spruce appears to be genetically less variable than the sympatric black and white spruce (Rajora et al. 2000).

Norway Spruce

Norway spruce [*Picea abies* (L.) Karst.] is economically the most important conifer tree species in Europe. Its natural distribution ranges across the Pyrenees, Alps, and Balkans, northwards to southern Germany and Scandinavia and eastwards through the Carpathian Mountains and Poland, to western Russia. It has a long history of cultivation in central Europe (Schmidt-Vogt 1977), including introductions in Belgium, Germany, and central France; more recent introductions have occurred in North America (SE Canada and NE USA), but only a few million seedlings are planted every year due to the sensitivity

of Norway spruce to white pine weevil. Breeding programs exist in several European countries, dating from the late 1940s. In the last two decades, Norway spruce has suffered severe forest decline in central Europe; as a result, focus has shifted from breeding to gene conservation and forest health. In part to guide these efforts, a Norway spruce network, EUFORGEN (http://www.bioversityinternational.org/networks/euforgen/networks/conifers/picea_abies/pabies.htm), was created.

3.1.2 Natural Hybridization

In southern British Columbia, white spruce hybridizes with Engelmann spruce (*P. engelmannii* Parry ex Engelm.), and in coastal northern British Columbia, white spruce hybridizes with Sitka spruce (Roche 1969; Sutton et al. 1991). In southern British Columbia, hybridization with Engelmann spruce is quite extensive, so much that the spruce in the region is termed "interior spruce" due to the mixed heritage of trees in this area (Kiss 1989). Black spruce and red spruce share an extensive sympatric zone in eastern Canada and northeastern USA, where they hybridize and introgress naturally (Perron and Bousquet 1997). They represent a recent progenitor-derivative species pair (Perron et al. 2000; Jaramillo-Correa and Bousquet 2003). In this region, like that of interior spruce, they form a species complex: black spruce × red spruce (*P. mariana* × *P. rubens*). As well, in northern Europe, Norway spruce exhibits introgressive hybridization with Siberian spruce (*Picea obovata* Ledeb.) (Farjon 1990; Krutovskii and Bergmann 1995), which is also considered a subspecies of *P. abies* (*P. abies* var. *obovata*) (Schmidt-Vogt 1977).

3.1.3 Molecular Genetic Variation and Mating System

Diversity at neutral nuclear loci is high for Sitka, white, black, and Norway spruce, with heterozygosity values generally exceeding 0.2 to 0.3 for isozymes (Yeh and El Kassaby 1980; Yeh and Arnott 1986; Yeh et al. 1986; Furnier et al. 1991; Krutovskii and Bergmann 1995; Isabel et al. 1995; Jaramillo-Correa et al. 2001), random amplified polymorphic DNAs (RAPDs) (Isabel et al. 1995), and expressed sequence tag polymorphisms (ESTPs) (Perry and Bousquet 1998a, b; Perry et al. 1999; Jaramillo-Correa et al. 2001). As expected, much higher diversity is observed at simple sequence repeats (SSRs) (Hodgetts et al. 2001; Rajora et al. 2001; Scotti et al. 2002).

Genetic diversity in spruce is partitioned mainly within populations: the majority (90 to 99%) occurs within populations for Sitka, white, black, and red spruce (Yeh and El-Kassaby 1980; Hawley and DeHayes 1994; Isabel et al. 1995; Rajora et al. 2000; Jaramillo-Correa et al. 2001; Perry and Bousquet 2001; Gamache et al. 2003). When maternally inherited cytoplasmic markers were surveyed in Norway and black spruce as well as in the Mexican *P. chihuahuana*, population differentiation was much higher (Sperisen et al. 2001; Bastien et al. 2003; Gamache et al. 2003; Jaramillo-Correa et al. 2004, 2006). Genetically distinct ancestral lineages could be identified from cpDNA or mtDNA polymorphisms (Vendramin et al. 2000; Sperisen et al. 2001; Jaramillo-Correa et al. 2004, 2006; Jaramillo-Correa and Bousquet 2005).

In contrast to nuclear loci, quantitative traits demonstrate much more population differentiation in boreal spruce (Li et al. 1997; Jaramillo-Correa et al. 2001; Beaulieu et al. 2004; Lagercrantz and Ryman 1990; Collignon et al. 2002; Acheré et al. 2005).

In terms of detecting the underlying loci responsible for such selective differentiation, a genomic scan of amplified fragment length polymorphism (AFLP), SSR, and ESTP markers indicated a small number of outlier loci that were more differentiated (Acheré et al. 2005). The frequency of outlier loci was higher from a genome-wide scan relying on gene SNPs (Namroud et al. 2006). Further inputs from developing genomic programs will greatly aid in identifying the specific loci underlying these adaptive differentations.

At the nucleotide level, white, black, and Norway spruce have moderate levels of intraspecific diversity (ca. 0.5%) but high levels of haplotype diversity, with H in excess of 90% (Bouillé and Bousquet 2005; Bousquet et al., personal communication). This is comparable to that found in pines (Brown et al. 2004; Pot et al. 2005). High nucleotide diversity has also been found for the ITS1, but this might be caused by interlocus divergence (Campbell et al. 2005). We also found, as in pines (Brown et al. 2004; Neale and Savolainen 2004), that linkage disequilibrium appears to decay rapidly within spruce genes. Finally, in an analysis of in silico SNP found in ESTs of white spruce, Pavy et al. (2006) found about 0.3% heterozygosity in clusters containing four or more cDNAs. These are all

examples of how data from the current spruce genome projects are helping to identify large-scale patterns of spruce genomic variation and differentiation.

With regard to outcrossing rates, boreal spruces are predominantly outcrossing (Cheliak et al. 1985; Boyle and Morgenstern 1986; Shea 1987; Chaisurisri and El-Kassaby 1994; Xie and Knowles 1994; Cottrell and White 1995; Rajora et al. 2000; Perry et al. 1999). Red spruce and subtropical taxa are exceptions, with lower outcrossing rates (e.g., Ledig et al. 2000; Rajora et al. 2000). Selection against inbreds has been noted in boreal spruces (Isabel et al. 1995) and in subtropical species (Ledig et al. 2002).

3.1.4 Cytogenetics

Chromosome number in somatic cells of spruces is $2n = 24$ (Burley 1965; Fox 1987; Murray 1998), and small supernumerary B chromosomes have been observed in some individuals of Sitka (Moir and Fox 1972; Teoh and Rees 1976; Kean et al. 1982) and white spruce (Nkongolo 1996). DNA content per diploid cell is 19 pg for Sitka (Ingle et al. 1975), 30 pg for Norway, from 17 to 40 pg for white, and from 22 to 32 pg for black spruce, depending on the sample and the method used (Murray 1998). These numbers correspond to estimated genome sizes ranging from 17 to about 40×10^9 bp (Murray 1998). The chromosomal locations of ribosomal RNA genes in Sitka spruce have been determined using fluorescent in situ hybridization. The 5s rDNA was restricted to one chromosome, whereas 18s-5.8s-26s rDNA occurred on chromosome 5 as well as four other chromosomes (Brown and Carlson 1997). The chromosomal location of large tandem repeats on the genome of white and Sitka spruce was similar, emphasizing the little divergence between the two species (Brown et al. 1998).

3.1.5 Economic and Breeding Issues

Economic Importance and Fundamental Breeding Issues

It is clear that the need for wood fiber and wood volume will increase (Brooks 1997). Modern economies must adapt to these changes. But forests are diverse, and besides providing for wood fiber and volume, they also provide for many other values, including biodiversity, recreation, employment for local communities, and the lore of native people's legends. Genetically improved planting material can, in principle, accommodate these multiple objectives. However, phenotypic assessments require that trees be grown from less than one half of their rotation age before selections for traits are made (Zobel and Talbert 1984). In the more northern climate of spruce, the rotation age of spruce is 20 to 40 years (British Columbia) to 60 to 100 years (Sweden), making selective breeding based upon traditional practices very slow. Genomics and molecular breeding can speed the progress of identifying trees adapted to future climates and provide insight into the uniqueness of conifers as the hallmark species of northern climes.

The Spruce Weevil Problem

The spruce shoot weevil (a.k.a. white pine weevil, *Pissoides strobi*), budworms, and certain bark beetle species are some of the most destructive insect pests of spruce forests worldwide (Seybold et al. 2000; Alfaro et al. 2002). Larval feeding of the spruce weevil severely damages or kills the leading shoots of susceptible host trees, resulting in reduced growth and tree deformation (Alfaro et al. 2002). In British Columbia, the weevil affects both Sitka and white spruce, but because it affects Sitka so severely, annual plantings of Sitka spruce have been reduced from 10 million seedlings to fewer than a million (King et al. 1997). The spruce weevil is endemic to North America, but as with other insect pests, it could spread to Sitka and Norway spruce forests in Europe. Genetic technologies might aid in fighting these pests.

The first evidence of substantial weevil resistance in spruce was observed in International Union of Forest Research Organization (IUFRO) provenance trials in British Columbia (Ying 1991). Since then, weevil resistance has been demonstrated to have a significant genetic component, as a large (139-family) 10-year-old test in south central British Columbia with 8 years of accumulated weevil attack gave an individual-tree heritability of >0.4 (King et al. 1997). However, to date, most mechanistic studies of conifer defense have been at the anatomical or chemical level (Trapp and Croteau 2001; Huber et al. 2004), not involving genomics technologies.

3.2 Genetic Mapping

3.2.1 First-Generation Genetic Maps in Spruce Species

We define the first generation linkage maps in forest trees as those constructed with anonymous genetic markers, mainly RAPD (Williams et al. 1990) and AFLP (Vos et al. 1995). These technologies are especially useful for species in which genomic knowledge is limited, such as in conifers, as they do not require a priori knowledge of the genome. While both RAPD and AFLP markers exhibit dominance, this handicap is overcome by the use of appropriate pedigrees and configurations such as the two-way pseudotestcross (Grattapaglia and Sederoff 1994) or by the use of haploid megagametophytes (Isabel et al. 1995), which are unique to conifers.

The first map of white spruce was produced using an array of haploid megagametophytes of a single tree mapping population (Tulsieram et al. 1992). This map consisted of 47 RAPD loci distributed on 12 linkage groups (LGs). It covered 874 cM, ca. one third of the estimated 2,700- to 2,900-cM genome length (Gosselin et al. 2002). The first map for Norway spruce (Binelli and Bucci 1994) produced a RAPD genetic linkage map from a population of 72 megagametophytes also derived from a single tree. It had 152 polymorphic loci assigned to 17 LGs covering a total distance of 3,584 cM, a puzzlingly large distance. A summary of the map characteristics is presented in Table 1.

AFLPs have become the marker of choice for anonymous marker maps owing to the large number of detected polymorphisms and the relatively low cost per locus (Schlötterer 2004). It should be mentioned that the usefulness of these markers has broken down the limits of what could be achieved in forest genetics. They have made possible the search for quantitative trait loci (QTLs) underlying specific traits in large undomesticated species such as trees (see other chapters). However, in spruce, the small megagametophyte does not yield sufficient DNA for AFLP reactions.

Paglia et al. (1998) built another single-tree genetic linkage map of Norway spruce using also a panel of 72 megagametophytes with a total of 447 segregating markers [366 AFLPs, 20 selectively amplified microsatellite polymorphic loci (SAMPLs), and 61 SSRs] covering a genetic length of 2,198 cM, which represents 77% of the estimated genome length. Out of the 447 markers, 413 were assigned to 29 LGs. Several of the SSRs developed were later found to be dominant (Pelgas et al. 2005, 2006) and, consequently, less interesting for comparative mapping.

3.2.2 Second-Generation Mapping

Second-generation maps are those that utilize both codominant markers and markers for which homology among species is clear. The first class of such markers are RFLPs, which have been around since the 1990s; but while their homology among species is high, thus allowing transfer of map information, their assay is relatively laborious. Since then, STS markers and ESTP markers (Perry and Bousquet 1998a, b; Perry et al. 1999), as well as single strand conformation polymorphisms (SSCPs) (e.g., Plomion et al. 1999), have emerged. In spruce, Fournier et al. (2002) evaluated the potential of 50 STS markers of arbitrary genes (Perry and Bousquet 1998a) for constructing maps at both intra- and interspecific levels. Several unrelated individuals of white spruce and black spruce and their corresponding progeny were screened for polymorphisms using SSCP. The use of this method of detection allowed to significantly increase the number of codominant markers for comparative genetic mapping purposes but still remained relatively labor intensive (Fournier et al. 2002).

In parallel, Gosselin et al. (2002) conducted a comparative mapping study based on an array of 100 megagametophytes for each of two individuals using dominant (RAPD) and codominant (ESTP) markers. The analysis for the first individual resulted in 165 loci (152 RAPDs, 3 SCARs, and 10 ESTPs) mapping to 23 LGs and covering 2,059 cM. For the second individual, the analysis resulted in 145 loci (137 RAPDs, 1 SCAR, and 7 ESTPs) mapping to 19 LGs and covering 2,007 cM. Both maps covered close to 90% of the entire genome. The percentage of shared loci between the two individual maps was much higher for the codominant gene-based markers than for the anonymous dominant markers (44% vs. 9%). This difference illustrates the need for increasing the number of codominant markers from expressed genes to anchor different linkage maps for a given species.

Tools for Second-Generation Genetic Maps

Second-generation genetic maps ideally utilize genomic information and involve coding or transcribed

Table 1. Linkage maps obtained for *Picea* species

Species	Number of progeny genotyped	Number and types of loci		Number of linkage groups	Software used	Genome length (cM)	Reference
P. glauca	72 gametophytes	47	RAPD	12	MAPMAKER	874	Tulsieram et al. 1992
P. abies	72 gametophytes	152	RAPD	17	MAPMAKER	3584	Binelli et Bucci 1994
P. abies	72 gametophytes	366	AFLP	29	MAPMAKER	2198	Paglia et al. 1998
		20	SAMPL				
		61	SSR				
P. glauca	1) 92 gametophytes	152	RAPD	23	JOINMAP	2059	Gosselin et al. 2002
		3	SCAR		MAPMAKER		
		10	ESTP		GMENDEL		
	2) 96 gametophytes	137	RAPD	19		2007	
		1	SCAR				
		7	ESTP				
*P. glauca**	1) 110 F1	714	AFLP	12	JOINMAP	2168	Pelgas et al. 2006
	2) 103 F1	38	SSR				
		53	ESTP				
P. glauca *	1) 88 F1	361	AFLP	13	JOINMAP	1623	Rungis et al. (pers. comm.)
	2) 88 F1	46	EST-SSR				
		81	COS				
		6	ESTP				
P. sitchensis	192 F1	134	SSR	Under construction	JOINMAP	–	J.E. Cottrell et al. unpublished
*P. mariana**	1) 118 BC1	1014	AFLP	12	JOINMAP	2319	Pelgas et al. 2005
	2) 85 F1	3	RAPD				
		53	SSR				
		54	ESTP				
*P. abies**	73 F1	661	AFLP	12	JOINMAP	2035	Acheré et al. 2004
		74	SSR				
		18	ESTP				
		2	others				
			Additional markers:				Pelgas et al. 2006
		1	SSR				
		13	ESTP				
*P. abies**	85 F1	422	AFLP, SSR, ESTP, IRAP, SSAP	13	JOINMAP	2821	Scotti et al. 2005

*Composite map

sequences, preferably with genes annotated for function. In the past 3 years, the Canadian Genomics Projects Arborea (Quebec) and Treenomix (British Columbia), funded by Genome Canada, Genome British Columbia, and Genome Quebec, as well as other projects funded by the National Science and Engineering Research Council of Canada and the Canadian Biotechnology Strategy, have accelerated the development of markers and made possible new types of genetic markers especially amenable to high-throughput genotyping and transferability across genetic maps. Here, we briefly review these innovations, which should be applicable to mapping projects with any species, plant or animal.

ESTP Markers

The development of ESTP markers is still time consuming and expensive, but their use is easy and informative because they are codominant and derived from locus-specific primers (Perry and Bousquet 1998a, b; Lefort et al. 1999; Perry et al. 1999). Also, they can be indicative of coding regions (Perry and Bousquet 1998a, b). These multiallelic genetic markers are generally orthologous and have been found to be transferable from one species to another within the genus *Picea* (Perry and Bousquet 1998b; Perry et al. 1999). Consequently, they can be used for the assemblage of conspecific maps useful for interspecific comparative mapping. To accelerate the development of ESTPs, Pelgas et al. (2004) and Lamothe et al. (2006) relied on a DNA pool sequencing strategy to look for common (frequency higher than 10%) polymorphisms. This approach was found to be reliable and speeds up marker recovery for the detection of SNPs and/or indels. In spruce, it has rendered availability of more than 100 codominant markers with amenable genotyping strategies such as AGE (agarose gel), DGGE (denaturing gradient gel electrophoresis), or CAPS (cleaved amplified polymorphic DNA) (Pelgas et al. 2004, 2005, 2006; Lamothe et al. 2006).

Genomic-SSR vs. EST-SSR Markers

The first set of microsatellites (SSR) developed from genomic libraries of Sitka was the seven published by van de Ven and Kanamori (1996). Rajora et al. (2001) also developed SSR markers for spruce. However, due to the large, repetitive nature of conifer genomes, robust, single-copy SSR markers from genomic DNA are difficult to find. Based upon our own (KR) experience with several conifers, only 5 to 10% of positive clones identified in hybridization with genomic libraries eventually lead to useful SSRs. On top of this, SSRs that work within the target species rarely work in related species. For example, attempts to transfer primers developed from genomic libraries of *P. abies*, for use in Sitka, have met with little success (A'Hara and Cottrell 2004); Rungis et al. (2004) tested 101 microsatellites developed from genomic libraries of other spruces for amplification in Sitka spruce and found that only 17 amplified in Sitka.

The transcriptome (messenger RNA) is the next logical source to find single-copy conserved SSR markers. Also, since these SSRs are within the expressed portion of the genome, they are more likely to be associated with a particular gene, compared to genomically derived markers, improving their utility for QTL mapping and marker-assisted selection (MAS). To this end, Scotti et al. (2000) screened a *P. abies* cDNA library for $(AG)_n$ and $(AC)_n$ repeats and found ten and six clones, respectively, containing these repeats. They estimated that SSRs in cDNA clones occur about 20 times less frequently than in random (genomic) clones. They designed six primer pairs of which three generated clear patterns similar to genomic SSRs.

EST databases offer the opportunity to quickly identify SSR repeats of all types. From a 20,275 unigene spruce EST set, Rungis et al. (2004) identified 44 candidate EST-SSR markers. Compared to Scotti et al. (2000), this number of 44 enables a good statistical evaluation of EST-SSR diversity and their transferability to other species. Of these 44, 25 amplified and were polymorphic in white, Sitka, and black spruce; 20 amplified in all 23 spruce species tested; and the remaining 5 amplified in all except one species. The 25 EST-SSRs had about 9% less heterozygosity compared to a set of 17 genomically derived SSRs, which also showed consistent amplification across species (mean $H = 0.65$ vs. 0.72). Also, these EST-SSRs showed fewer null alleles. Additional EST-SSRs have been developed from the white spruce EST collection by Forest Research (Scotland) for their work with Sitka spruce (A'Hara and Cottrell 2004, 2007).

Ecotilling

A single-strand specific nuclease, found in extracts of celery juice, can be used to digest heteroduplex DNA and hence identify heterozygous single nucleotide polymorphism (SNP) sites in PCR products (Comai et al. 2004). This method can be used to map genes

with relative simplicity and low cost (Rungis et al. 2005). A particular nucleotide substitution does not need to be identified, and in fact, a priori knowledge of the presence of a SNP is not required, as the entire length of the PCR product is interrogated for the presence of SNPs. This feature enables application of this technique to genomes that are not well characterized, to rapidly place particular genes onto linkage maps, and will be used to map COS markers by the Treenomix project in pine and spruce (see below). While this technique is best suited for mapping markers in a backcross configuration, in the F_2 configuration, where alternative homozygotes cannot be discerned by this technique, data are still informative about linkage (Rungis et al. 2005).

Conserved Orthologous Set (COS) Markers

The advent of large-scale EST databases and genome sequences has enabled the identification of a class of genes that are especially amenable for comparative genetic mapping: "conserved orthologous set" (COS) markers. COS markers are single-copy and slowly evolving across species and are identified by self- and cross-BLASTING EST databases. They were originally developed for species of Asterids (Fulton et al. 2002) but also seem appropriate for conifers, as determined by a recent in silico analysis (Krutovsky et al. 2006)

The Treenomix project has undertaken a large-scale screening of COS markers for the white spruce-loblolly pine comparison, with the aim of mapping 1,000 COS markers in a comparative map (using ecotilling). Using the latest EST databases, annotated COS marker genes were identified by reciprocal BLAST. Primers were designed only for sites exhibiting complete sequence identity within these homologs. As of June 2006, 443 such primer pairs were tested for PCR amplification products, and 46% of these gave single bands in both species; interestingly, 31% gave single bands in these two species plus Douglas fir.

Multiple Pedigree Mapping

Another challenge with genetic and QTL mapping, especially when the number of informative markers (gene-based, multiallelic, and codominant) is limited, is that only polymorphic markers showing Mendelian segregation within a given array of progeny can be positioned. One way to alleviate this problem is to use more than one cross for a given species, with at least one parent in common to increase the probability that a given marker segregates (Beavis and Grant 1991; Tani et al. 2003). For example, the use of a common parent between mapping populations allowed Pelgas et al. (2005, 2006) to increase by around 25% the number of useful anchor markers and also to verify the consistency of macrosynteny and macrocolinearity within and between species.

If parents are not shared between pedigrees, the maps can be joined "by eye" using markers in common between maps or with the program JoinMap (Stam 1993). JoinMap starts with individual pairwise recombination estimates derived from different experiments and linearly combines them into a single estimate using weights proportional to LOD scores. However, if estimates of recombination for a pair of markers differ among pedigrees, a joint maximum likelihood procedure, given in Hu et al. (2004), has more statistical power. Combining multiple pedgrees into a single map analysis is especially appropriate in forestry, where mature progeny trials of relatively small size (10 to 40) are very common, as tree breeders tend to maximize the number of families to ensure conservation of genetic diversity. A recent example of the application of this method is a Douglas-fir AFLP map constructed from eight 40-member full-sib families (Ukrainitz et al. 2007).

3.2.3 Second-Generation Genetic Maps (as summarized in Table 1)

White Spruce Maps

Marker maps for white spruce based upon AFLPs, SSRs, ESTPs, and COS markers have recently been completed by the Canadian genome projects. As part of the Arborea project, a composite linkage map for white spruce representing an assemblage of four individual maps and delineating 12 LGs was obtained (Pelgas et al. 2006). The length of the composite map was 2,168 cM. Anonymous and locus- or gene-specific markers (714 AFLPs, 38 SSRs, and 53 ESTPs) were positioned, among which a total number of 108 anchor markers positioned (mostly SSRs and ESTPs). As part of the Treenomix project, an integrated genetic map of white spruce representing the assemblage of two pedigrees contained a total of 452 markers (361 AFLPs, 6 candidate genes, 76 COS markers, 14 orthologous markers [with loblolly pine], and 46 EST-SSR markers) mapped onto 13 major LGs (>5 markers).

The linkage map length was 1,622.7 cM (Rungis et al. pers. comm.).

Black-Red Spruce Maps

In a major effort to further compare genome structure between phylogenetically distant species in the genus and to enable comparisons with other Pinaceae, four individual linkage maps as well as a reference and a composite linkage map were constructed for the black spruce × red spruce species complex (Pelgas et al. 2005). These authors assessed the usefulness of SSR and ESTP codominant markers in combination with the use of two crosses, with one parent in common, for the assembly of a composite map. This map contained a total of 1,124 positioned markers, including 1,014 AFLPs, 3 RAPDs, 53 SSRs, and 54 ESTPs, assembled into 12 major LGs. The map length was 2,319 cM. Information deriving from a second cross contributed in this case to an increase of 24% in the number of anchor markers in comparison with the information from only a single cross.

Norway Spruce Map

Acheré et al. (2004) reported on the construction of the first saturated linkage map based on an F_1 progeny in Norway spruce. The two parental linkage maps generated were integrated into a consensus one. In total, 755 markers (661 AFLPs, 74 SSRs, 18 ESTPs, the 5S rDNA, and the early cone formation trait) were assigned and positioned onto 12 LGs. The resulting consensus map covered 2,035 cM. Additional markers (SSR and ESTP) were positioned on this map for a comparative mapping survey within the *Picea* genus (see below).

With both the same approach and species, Scotti et al. (2005) built two linkage maps covering a total length of 2,316 and 1,667 cM in the female and the male, respectively, as well as a consensus map. Different classes of markers (sequence families), sampling putatively different regions of the genome, were used for a mapping experiment. Using autocorrelation techniques, the authors found that AFLPs and SSRs were in close proximity to one another, while ESTPs did not seem to form clusters. The latter point should be taken cautiously since the number of ESTPs positioned in this study was relatively low. The final composite map consisted of 422 markers assigned to 13 LGs covering 2,821 cM (Scotti et al. 2005).

3.3 Comparative Mapping

In nonmodel species, comparing markers and genetic maps between related species is a useful strategy to transfer genomic information among species in a synergistic way. The anchoring of maps involves both the evaluation of the degree of synteny (gene content of LGs) and colinearity (gene order of LGs). To fulfill these objectives, numerous anchor markers per LG are necessary, which can recognize unambiguously orthologous chromosomal regions between species. EST-based markers are even more important for comparative mapping across species or genera and for the detection of associations between candidate genes and quantitative traits or known metabolic functions.

3.3.1 Comparative Mapping Between Spruce Species

Composite linkage maps have been produced for three *Picea* species (*P. glauca* and *P. abies* and *P. mariana* × *rubens*; Pelgas et al. 2006). In total, 88 (30 SSRs and 58 ESTPs) homologous markers were shared among these maps, and 12 homologous LGs were identified. A high level of conservation of the gene content (synteny) and macrocolinearity was observed (Fig. 1). This conservation is remarkable given that the split between these three divergent taxa occurred millions of years ago (Bouillé and Bousquet 2005).

One incongruity in synteny in these maps, involving *P. glauca* and *P. abies* and *P. mariana* × *rubens*, appears to be due to an insertional translocation, likely from white spruce (Pelgas et al. 2006). Also, a phylogenetic study of the *knox-I* gene family combined with a genetic mapping approach (Guillet-Claude et al. 2004; Pelgas et al. 2006) revealed a putative segmental duplication in spruces on LG 3, the same LG where the lack of synteny between white spruce and the two others was observed. Hence, this part of the spruce genome appears to be the site of various structural modifications that occurred at different times of the genus history.

Evolution by duplication also appears to be frequent in the genome of the Pinaceae, as a succession of three gene duplications and two chromosomal translocations can explain the chromosomal distribution of the *knox-I* gene family in the genera *Picea* and *Pinus* (Guillet-Claude et al. 2004). Mapping work in progress also indicates that several families of reg-

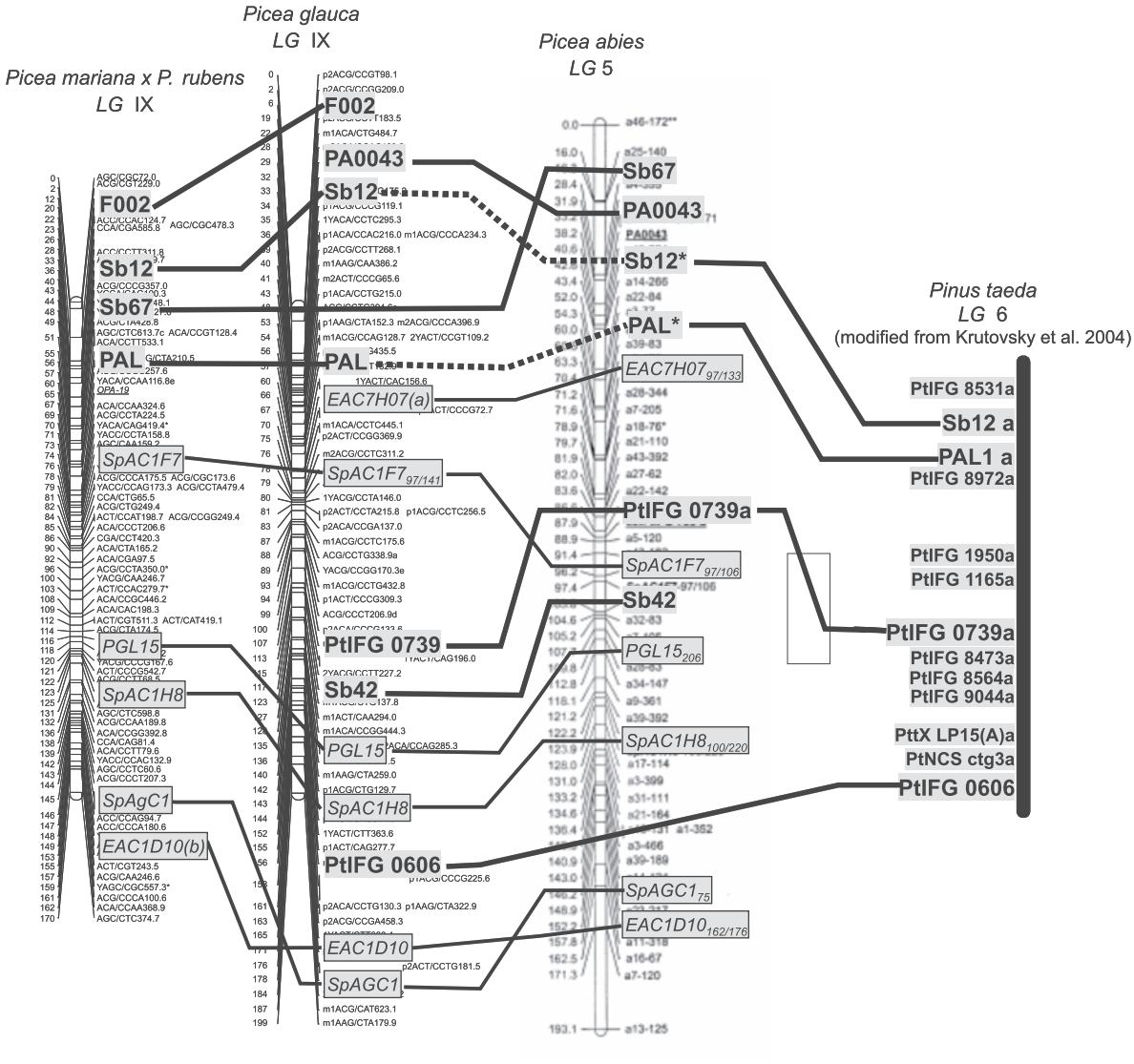

Fig. 1. Synteny between homoeologous composite LGs of *Picea mariana* × *P. rubens* complex, *P. glauca* and *P. abies* species, and *P. taeda*. Markers in *bold* are ESTPs, markers in *italic* are SSRs, and all remaining markers (for spruces) are AFLPs. Orthologous markers are connected by a *solid line*, except when they are connected with homologous anchor markers positioned onto the linkage maps of *P. abies* developed by Scotti et al. (2005; *dotted line*)

ulatory genes are clustered on the same LGs. This indicates that, perhaps, gene duplications are followed by short bursts of rapid evolution involving neofunctionalization (Guillet-Claude et al. 2004).

3.3.2 Comparative Mapping in the Pine Family

Despite their large size (6× human, 100× Arabidopsis), conifer genomes seem to be remarkably conserved. Across the Pinaceae (pine family – pine, spruce, fir, hemlock, Douglas fir, larch), chromosome number is quite conserved ($x = 12$, except $x = 13$ for Douglas fir). Consistent with this conserved chromosomal evolution, Krutovsky et al. (2004) documented synteny and colinearity between the genomes of pine and Douglas fir. This suggests that, as in the grass family, linkages and clusters of gene functions would be preserved among coniferous species.

Because of the large genome size in conifers, the absence of a conifer genome sequence, together with this putative synteny, makes comparative mapping an important tool for integrating information across the

Pinaceae. Among conifers, the largest effort in genetics and genomics has been devoted to loblolly pine, with large EST collections, rich genetic resources, and well-developed genetic and QTL maps (Neale and Wheeler 2004). Species of spruce rank second, largely because of the above-mentioned Canadian genome projects. Loblolly pine and spruce bridge the pine family, and they likely represent the most ancient lineages of the Pinaceae (Frankis 1989; Liston et al. 2003). Besides reinforcing each other, joint genomic patterns in these two genera might enable extension and integration of such genomic knowledge into other species, particularly lodgepole pine (being very proximal to loblolly pine) and Douglas fir (being intermediate between loblolly pine and spruce).

A joint map by Pelgas et al. (2006) suggests that the macrostructure of the Pinaceae genome is well conserved (Fig. 1). Moreover, this study confirmed that a chromosomal fission has presumably occurred in *Pseudotsuga* (Krutovsky et al. 2004; Pelgas et al. 2005). This may have played a central role in generating the difference in haploid chromosome number between *Pseudotsuga* ($n = 13$) and the other Pinaceae ($n= 12$). Some minor rearrangements were also observed that could only be validated by identifying genetic maps or by complete sequencing of specific genomic regions (Pelgas et al. 2006). However, several cases of putative rearrangements were dismissed after DNA sequencing to verify the orthology of the genes compared (Pelgas et al. 2005, 2006). This is an example of how caution should be taken in validating anchor marker gene orthology. Paralogy is rampant in conifers; because of their large repetitive genome, conifers harbor large gene families, giving lots of opportunities for mistaken orthology (Guillet-Claude et al. 2004).

3.4 Molecular Breeding

3.4.1 Marker-Assisted Selection

Achieving genetic gains is a slow and incremental process. The development and implementation of molecular markers predictive of phenotypic performance is expected to greatly improve the accuracy of genetic selection and accelerate the breeding process. In crops, marker-assisted selection (MAS) has been used for traits including rust or other disease resistance (e.g., Gebhardt et al. 2004; Hayden et al. 2004), phenological and adaptative traits (e.g., Saranga et al. 2001; Thornsberry et al. 2001), and crop productivity (e.g., Dirlewanger et al. 2004). Until recently, the use of MAS to shortcut the long breeding cycles in forest trees was more of a concept than a reality. The practical roadblocks include, among many others, a lack of a model system for a conifer, the lack of inbred pedigrees, and the long generation times of conifers.

Ideally, MAS should identify candidate genes for direct selection. Discovery of such genes follows two different strategies. The first involves association analysis of quantitative traits in segregating pedigrees (identifying quantitative trait loci, QTL, c.f. Lander and Botstein 1989). However, large progeny numbers are required to verify marker-trait associations (Beavis 1994). For example, in radiata pine (*Pinus radiata*), the identification and verification of QTLs involved measurements on nearly 4,435 trees from a single full-sib family planted in a commercial forestry plantation (Devey et al. 2004). The alternative strategy is association analysis, which aims to identify DNA polymorphism on genes directly underlying the phenotypic variation. The low level of linkage disequilibrium observed in softwood trees makes them good candidates for such an approach (Neale and Savolainen 2004). Association studies can be regarded as very large linkage studies of unobserved pedigrees that, in theory, permit mapping of a QTL at a higher resolution (reviewed by Cardon and Bell 2001).

Marker-Assisted Selection: Sitka Spruce

With the above concepts in mind, the Conifer Tree Breeding Team in Forest Research (FR) in Scotland has recently initiated a project aimed at using DNA technologies in their Sitka spruce breeding program. The objective is to shorten the breeding cycle and to improve the rate of genetic gain by associating microsatellite and other markers to traits of economic importance, such as wood density. In spring 2005, FR planted three large clonal tests designed to investigate the association between markers and phenotypic variation in Sitka spruce. The experiments involve the same 1,500 progeny from each of three full-sibling crosses. Each progeny (or genotype) is replicated four times at a single site, and the entire experiment is repeated at three climatically contrasting sites. The replicated clones representing each family in the field will be measured for a range of characteristics, such as wood density, stem straightness, or growth rate. It is intended that the top and bottom 150 genotypes for a given trait will be used for marker genotyp-

ing. Progress is being made with marker development (A'Hara and Cottrell 2004, 2007), and a genetic map should be available by the time the traits are measurable in the trial.

White Spruce

There are two genome projects in Canada that involve white spruce and their close relatives. The first is centered in Quebec (Arborea) and the second in British Columbia (Treenomix). The goal of the Arborea project is to identify allele-gene combinations that govern naturally occurring phenotypic variation of commercially valuable traits in natural and breeding populations of spruce. These involve investments made by the Canadian Forest Service-Laurentian Forestry Centre (CFS-LFC) for the last 30 years. In British Columbia, the goal of the Treenomix project is to use pedigrees segregating for resistance to the white pine weevil to dissect the chemical ecology of resistance to the pine weevil (and to insects in general). Likewise, the British Columbia Ministry of Forests has made large investments in breeding for weevil resistance in spruce and for general genetic gains. In both projects, large provenance and progeny tests of hundreds of open-pollinated families and full-sib families from crosses are available, involving their respective goals.

For example, in Quebec, mature material from a provenance-progeny test initiated by the CFS in 1976 and replicated on three sites in Quebec will be used (Beaulieu et al. 2006). In British Columbia, a large open-pollinated progeny test ("red rock") will be used in conjunction with the parents of the breeding population, which are also ca. 40 years old. These breeding populations provide genetic material for which "mature" traits, such as wood quality and lifetime fitness, can be measured (King et al. 1997). In both projects, seedling progenies and their replicated lines are being established, which in their young age are adequate to test candidate genes with phenology traits (e.g., timing of bud set).

3.4.2 Somatic Embryogenesis and Genetic Transformation

Since its demonstration in Norway spruce in 1985 (Hakman et al. 1985), somatic embryogenesis (SE) is the most versatile method for vegetative propagation of conifers. SE is divided into five stages: (1) initiation of embryogenic tissue from mature or immature explants, (2) proliferation of embryogenic tissue, (3) embryo maturation, (4) germination, and (5) conversion to plants. The cryopreservation of embryogenic tissue during stage 2 offers an opportunity to select a posteriori elite individuals for clonal plantations (reviewed by Cyr and Klimaszewska 2002). Over the last decade, most of the studies on SE have focused on the proliferation and the maturation processes to get embryos of high quality for both spruce and pine species (reviewed by Stasolla et al. 2002). Recently, new SE protocols have also opened up the opportunity to bring about this technology to a large-scale delivery (Sutton et al. 2004). In addition, SE has become a primary enabling technology for genetic engineering (see below) because embryogenic tissues are amenable to genetic transformation via cocultivation with *Agrobacterium tumefaciens* (Levée et al. 1997).

An impediment to tree improvement is the time required for each progeny to reach sexual maturity. Genetic engineering could circumvent this problem by allowing the transfer of single gene traits into superior genotypes, leading to the integration of desired traits. Accordingly, genetic transformation offers an attractive alternative to breeding, especially for insect and disease resistance for which natural variants with tolerance/resistance phenotype might be nonexistent (reviewed by Pena and Séguin 2001). The first stable transformation in a commercially important conifer species (*P. glauca*) was achieved via particle bombardment of embryogenic tissue (Ellis et al. 1993). This protocol was then successfully adapted for the transformation of other spruce species (Charest et al. 1996). In Sitka spruce, transient expression of the *uidA* gene using expression of the *GUS A* gene in embryogenic cell lines was observed following *Agrobacterium*-mediated transformation, but no stable transformants could be recovered (Drake et al. 1997).

Agrobacterium-mediated transformation is the method of choice for gene transfer in conifers (reviewed by Nehra et al. 2005), primarily owing to the ease of transformation, high efficiency, and clean integration of T-DNA into the host genome. Combined with rapid selection of transgenic embryogenic tissues and highly efficient seedling regeneration via somatic embryo maturation, the production of transgenic spruce has now become routine (Klimaszewska et al. 2001). Nevertheless, this technology is perceived by many as too premature to be commercially implemented.

Many questions and concerns remain regarding genetically modified (GM) trees. In Quebec, the CFS has been conducting some confined trials with GM trees but on a limited basis – only involving the testing of gene function. In British Columbia, testing in any open environment is prohibited. Commercial plantations of GM trees are in fact not permitted in Canada (http://www.inspection.gc.ca/english/ plaveg/bio/pbobbve.shtml.). The bigger questions about the impacts of GM trees revolve around their deployment in operational forestry (Strauss et al. 2001; van Frankenhuyzen and Beardmore 2004). What is the effect of gene flow from GM to native trees of the same species? What are the benefits of GM trees to the local economy? Should we introduce foreign genes into a natural ecosystem?

3.5 Genomics

3.5.1 cDNA Libraries, EST, and SNP Collections

Genome projects in both British Columbia and Quebec have been involved with the construction of cDNA libraries implicating a large variety of tissues. In Quebec, 16 cDNA libraries were constructed, followed by the sequencing and processing of 71,424 spruce ESTs, of which 49,101 were of high quality (at least 100 contiguous nucleotides with a *Phred* score over 20) and retained for the final assembly of 16,578 consensus sequences (average of 797 bp) (Pavy et al. 2005). Half of the clones were sequenced in both directions to accelerate gene discovery, and 45% of the clones were completely sequenced, at least in one direction. A database named SpruceDB is publicly accessible and contains all high-quality ESTs (Pavy et al. 2005). A second-generation publicly accessible database, ForestDB, accommodates ESTs from multiple species (white spruce, loblolly pine, and poplar transcriptomes), and multiple EST assemblies have recently been made available, including SNP tables for each of white spruce contig, encompassing over 12,000 SNPs (Pavy et al. 2006).

In British Columbia, 20 white and Sitka spruce cDNA libraries have been constructed, including normalized and full-length versions, with 184,364 spruce ESTs now submitted to GenBank (average 661 bases per EST). In addition, 6,464 full-length sequences have also been obtained (average length 1,123 bases; not yet submitted), with two complete reads per clone to ensure quality. Using an additional 29,493 ESTs from Genbank, assembly resulted in 26,336 clusters of size at least two, and 52,316 singletons. However, these values are dependent on minimum overlap (set at 40) and minimum percent match (set at 94). There were 12,364 clusters with four or more members, of which 34,658 SNPs were identified. A database of these SNPs and pine SNPs is at http://treenomix2.forestry.ubc.ca/public/spruce_pine/index.htm.

3.5.2 Microarrays

Likewise, these two genome projects have been involved with the design, manufacture, and use of cDNA microarrays. In British Columbia, three successively larger cDNA chips, from 9.7 K to 16.7 K to 21.8 K cDNA clones, have been designed and spotted. The first 9.7 K microarrays were used to investigate the response of Sitka spruce shoot tips to the spruce budworm, and this work represented the first large-scale study of insect-induced transcripts in a gymnosperm (Ralph et al. 2006). In Quebec, a spruce unigene of set 11 K cDNA clones was resequenced and used to fabricate cDNA chips, with the objective of investigating regulatory genes and genes related to wood formation. The microarrays are also being used by various collaborators; customized cDNA microarrays can be obtained, and further collaborations are encouraged.

3.5.3 Proteomics

The complete set of proteins in a particular tissue or in an organism is termed the proteome. In the first proteomic study of conifer insect defense, Lippert et al. (2007) as part of the Treenomix project, studied changes in the proteome of Sitka spruce bark tissue in response to feeding by white pine weevils or mechanical wounding. Two-dimensional polyacrylamide gels coupled with high-throughput tandem mass spectrometry were used to detect and identify induced changes in protein abundance and protein modification. The spruce EST and full-length cDNA sequences mentioned above were essential to identifying proteins from mass spectrometry data; 72% of proteins were identified. Significant changes of protein levels were observed as early as 2 h following

the onset of insect feeding. Among the insect-induced proteins are a series of related small heat shock proteins, other stress response proteins, proteins involved in secondary metabolism, oxidoreductases, and a protein to which no homology can be found in other plant species. These proteins and their promoters are good candidate genes for SNP markers to use in molecular breeding.

3.5.4 Transgenic Spruce Lines

To enable functional studies of the biological role of candidate regulatory genes, a high-throughput DNA transformation platform has been established in Quebec in collaboration with CFS using the white spruce line PG-653 (A. Séguin et al. personal communication). A number of gene constructs were tested for each of 37 different conifer genes by focusing on gain-of-function experiments, resulting in the production of 350 uniform and stable transgenic spruce cell lines, further analyzed by quantitative PCR (qPCR). Various unusual phenotypes could be recovered, and gene expression profiling of wild type and transgenic plants is underway (J. Mackay, personal communication).

3.5.5 Genome Composition and Bacterial Artificial Chromosome (BAC) Libraries

Genome Composition

It has long been suspected that the large genome size of spruce and other conifers is due to repetitive DNA. To characterize the types of repeats, Jurman et al. (1999) sequenced 120 clones of Norway spruce and found a large proportion of the clones to be represented by long interspersed repeats (copia- and gypsy-like LTR-retrotransposons), while SINE-like elements were rare; SSRs were not very abundant. Miniature inverted-repeat transposable elements and long interspersed nuclear elements were not represented. Further information about genome structure of noncoding regions will be obtained by sequencing of BAC clones (discussed below).

Pavy et al. (2005) analyzed ca. 15,000 white spruce transcripts and found that ca. 84% of the transcripts had homology to published gymnosperm sequences but only 68% had homologs with rice or *Arabidopsis* (e-value < 1e-10). This is not surprising given the distinct evolution of gene families in gymnosperms (Guillet-Claude et al. 2004). About 70% of transcripts had a significant match to proteins in the Uniref100 database. Gene ontology (GO) terms representing 16 molecular functions could be ascribed to about 40% of these transcripts, with catalytic activity ranking first (Pavy et al. 2005). For transcription factors, a hidden Markov model (HMM) search against *Arabidopsis* and PFAM identified ca. 400 white spruce sequences that represented putative spruce transcription factors (Pavy et al. 2005). The distribution of transcripts with SNPs among GO classes largely reflected the relative abundance of GO classes (Pavy et al. 2006). Algorithms were developed to identify ORFs from contig assemblies and, hence, estimate rates of synonymous and nonsynonymous SNPs. From 3,374 distinct ORFs with reliable SNPs ($P_{Polybayes}$ over 0.95), about 40% of SNPs were nonsynonymous and 60% synonymous, and the ratio of nonsynonymous to synonymous SNPs per site was found to be 0.17 (Pavy et al. 2006), surprisingly the same as that reported for 240 ORFs in *Arabidopsis* (Zhang et al. 2002).

A 3 × Unpooled Spruce BAC Library

Due to the large genome size of conifers, BAC libraries have been extremely limited for conifers. However, a library for *Pinus pinaster* is reportedly being expanded to 3× coverage (Plomion et al. this volume). More dramatic are the plan for a 10× arrayed library for *P. taeda* (loblolly pine), consisting of 1.7 to 2 million clones (Peterson 2005), which would make it the largest library ever constructed. When coverage reaches 3×, isolation of individual genes is planned (Plomion et al. this volume).

In spruce, a 3× unarrayed library is currently being constructed for the Treenomix project by Bio S&T (Montreal), and halfway through (June 2006), insert sizes are averaging ca. 150 kb. In the Treenomix project, about ten randomly selected BAC clones will be shotgun-sequenced to obtain a picture of genome structure, another ten clones enriched for coding genes will be sequenced to get a picture of upstream regulatory sequence, and a final ten sequenced for targeted genes also sequenced in loblolly pine sequenced clones to get a picture of microsynteny and local rearrangements. To target genes, BACs are stored as pools, about 1,000 BACS per pool, arrayed in 96-well plates; each plate is pooled for rows and columns and

screened by PCR to identify wells containing the target gene; colonies from the target well are grown and individually rearrayed for a second PCR screen (see Isidore et al. 2005).

3.5.6 High-Throughput SNP Genotyping

Tests of the Illumina High-Throughput Genotyping Technology

In the framework of the Arborea and Treenomix projects, proof-of-principle experiments have been conducted to test the Illumina SNP bead-array genotyping technology, which is based on a highly multiplexed (either 96, 384, 768, or 1536 SNPs) primer-extension approach (Fan et al. 2003). The technology involves an allele-specific extension step using allele-specific oligos (ASOs), followed by ligation to a locus-specific oligo (LSO) carrying, for each SNP, a unique address sequence. Amplification is done with two labelled universal primers (Cy3 and Cy5) complementing each ASO, and a third universal primer matching the LSO. The dye-labelled DNA products are hybridized onto a plate array and a bead array reader is used for scanning fluorescence signals. With this assay, homozygous DNA should be detected by a signal in either green (Cy3) or red (Cy5) channel and heterozygous DNA in both channels.

In the Arborea project, SNPs were identified in silico from contig assemblies (e.g., Pavy et al. 2006) and from resequencing genomic DNA. A 768-SNP array has been tested on pedigree and unstructured populations and the success rate was 70%, using the stringent criteria of human SNP projects such as HapMap (that is, over 98.5% call rate success for any given SNP with normal segregation). Across six white spruce natural populations emcompassing 160 individuals, observed and expected heterozgosities (H), were respectively 0.28 and 0.26 for the 534 valid SNPs. These SNPs are currently used for the mapping of candidate genes and in genomic scans involving natural populations with significant differentiation for adaptive characters (Q_{ST}) (Jaramillo-Correa et al. 2001).

In the Treenomix project, this system was tested using 384 SNPs identified in silico from a collection of 25,000 ESTs derived from a single spruce tree (PG29) used the project. Two mapping populations of size 75, both of which had PG29 as a parent, were assayed. Two hundred seventy-three of 384 SNPs (71%) appeared to work, showing 1:1 or 1:2:1 ratios. Two instances of null alleles were found (showing 2:1:1 or 1:0:1 ratios). Given the repetitive nature of conifer genomes, this represents an unexpectedly high success rate. SNP development under the Illumina system is always a two-stage process: development of an initial testing plate of a minimum of 96 SNPs, then a final "production" plate, which can be expensive for low-throughput projects.

Second-Generation SNP Plates

Since the cost per genotype is much lower with the high-throughput systems (ca. 6 cents and likely to drop), a "universal SNP plate" could be developed for spruce in order to facilitate comparative genomic studies and accelerate molecular breeding research. However, the plated SNPs must be characterized a priori at the population level in order to verify their occurrence and abundance in each targeted taxon. A large proportion of SNPs are singletons or at low frequency (Bousquet et al. personal communication). Such SNPs are likely to be useless for most applications. In addition, common SNPs are usually different between *P. glauca*, *P. mariana*, and *P. abies* (Guillet-Claude et al. 2004; Bousquet et al. personal communication). The occurrence of common SNPs could be estimated in EST databanks from contig assemblies containing sequences from a large number of distinct clones (Pavy et al. 2006). Resequencing on genomic DNA appears necessary to ascertain the occurrence and frequency in nontarget species.

As for mapping, unless SNPs are prescreened for high heterozygosity, there will not be sufficient numbers of segregating SNPs in any one cross to warrant the cost of the plate. The above-mentioned Treenomix test provides a preliminary indication of how SNPs detected in EST databases might lead to a good SNP plate for mapping. In a survey of 27 individuals across the white spruce range, mean heterozygosity (h) for the 273 SNPs was 0.24, and h for SNPs where all three parents were heterozygous in the two pedigrees was 0.36. (It should be remembered that these h values are conditioned on the observation of heterozygosity in the tree PG29; true values of h would be considerably lower.) With an average h of 0.36, a SNP panel of 384 highly heterozygous loci would segregate for about 200 loci for either of the two parents of a pedigree. However, it should be remembered that these SNP polymorphisms are largely species specific.

3.6 Summary

In spruce, as in most coniferous tree species, probably the greatest hindrance for effective molecular genetic dissection and understanding of traits important for tree breeding is the lack of appropriate pedigree materials. Most pedigrees are small (10 to 30 progeny), meant to test breeding values of individual trees in traditional breeding programs; also the long generation time of spruce (ca. 20 years) inhibits investment in multigenerational pedigrees, let alone the breeding program itself. At least pedigree material has been developed that now makes it possible to invest major efforts in genomics in order to speed up breeding cycles using molecular tools. BC_1 and F_1 of large progeny size, and clonally replicated, have been developed by the CFS and Forest Research (Scotland) for further QTL mapping experiments.

As well, as in all species, anonymous markers have been of limited value for molecular breeding. With the recent availability of catalogs of gene sequences, new cohorts of genetic markers based on candidate genes thought to be involved in specific pathways or characters, such as wood formation and pest resistance, can and are being developed for spruce. The emergence of new methodologies for discovering and assaying SNPs, and for obtaining anchor loci for comparative maps, promises rapid advances in the near future. Using the newly developed markers, the first composite genetic map for each of both white spruce and black spruce has already been assembled (Pavy et al. 2007). In addition, the first comparisons with other Pinaceae have been made. These results have opened new areas of research for spruces not possible before and helped increase the speed of MAS development and implementation for spruce.

Finally, it would seem that genomics can add a new dimension to tree breeding, providing new targets for breeding via better understanding of the genetics of traits such as disease and pest resistance, cold and drought tolerance, and even CO_2 sequestration. One can imagine that genomics might provide novel methods to monitor the health of existing forest stock and make revised predictions about the longer-term sustainability of conifer forests in our landscape. However, the recent investments in forest genomics invite a naively optimistic vision of these new applications. Our actual level of genomics knowledge is still primitive, and, like any new area, the ideas and hypotheses of interest in genomics will change rapidly, as will the potential applications.

References

Acheré V, Faivre-Rampant P, Jeandroz S, Besnard G, Markussen T, Aragones A, Fladung M, Ritter E, Favre JM (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 JM, Besnard G, Jeandroz S (2005) Genomic organization of molecular differentiation in Norway spruce (*Picea abies*). Mol Ecol 14:3191–3201

A'Hara SW, Cottrell JE (2004) A set of microsatellite markers for use in Sitka spruce (*Picea sitchensis*) developed from *Picea glauca* ESTs. Mol Ecol Notes 4:659–663

A'Hara SW, Cottrell JE (2007) Characterization of a suite of 40 EST-derived microsatellite markers for use in a Sitka spruce (*Picea sitchensis* (Bong.) Carr) marker aided selection program. Silvae Genet (in press)

Alfaro RI, Borden JH, King JN, Tomlin ES, McIntosh RL, Bohlmann J (2002) Mechanisms of resistance in conifers against shoot infesting insects. In: Wagner MR, Clancy KM, Lieutier F, Paine TD (eds) Mechanisms and Deployment of Resistance in Trees to Insects. Kluwer, Dordrecht, pp 101–126

Arno SF, Hammerly RP (1977) Northwest Trees. Mountaineers, Seattle, WA

Bastien D, Favre JM, Colligon AM, Sperisen C, Jeandroz S (2003) Characterization of a mosaic minisatellite locus in the mitochondrial DNA of Norway spruce (*Picea abies* (L.) Karst.). Theor Appl Genet 107:574–580

Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. In: Proc 49th Annual Corn and Sorghum Industry Res Conf, American Seed Trade Associaction, Washington, DC, pp 250–266

Beavis WD, Grant D (1991) A linkage map based on information from four F_2 populations of maize (*Zea mays* L.). Theor Appl Genet 82:636–644

Beaulieu J (1996) Breeding program and strategy for white spruce in Quebec. Inf Rep LAU-X-11TE. Nat Res Can, Canadian Forest Service, Sainte-Foy, Quebec

Beaulieu J, Corriveau A, Daoust G (1989) Phenotypic stability and delineation of black spruce breeding zones in Quebec. Inf Rep LAU-X-85E. Canadian Forest Service, Sainte-Foy, Quebec

Beaulieu J, Perron M, Bousquet J (2004) Multivariate patterns of adaptive genetic variation and seed source transfer in black spruce (*Picea mariana*). Can J For Res 34:531–545

Beaulieu J, Watson P, Isabel N, Bousquet J, MacKay J (2006) Association mapping for wood candidate genes in white spruce. In: Plant & Animal Genome XIV Conf, San Diego

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

Bouillé M, Bousquet J (2005) Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the maintenance of genetic diversity in trees. Am J Bot 92:63–73

Bouillé M, Bousquet J (2006) Discordant mtDNA and cpDNA phylogenies of the entire genus *Picea* indicate patterns of allopatric speciation related to geographic dispersal and ancient reticulation. In: Proc IUFRO Population Genetics and Genomics of Forest Trees, Madrid, p 152

Boyle TJB, Morgenstern EK (1986) Estimates of outcrossing rates in six populations of black spruce in central New Brunswick. Silvae Genet 35:102–106

Brooks DJ (1997) The outlook for demand and supply of wood: implications for policy and sustainable management. Commonwealth For Rev 76:31–36

Brown GR, Carlson JE (1997) Molecular cytogenetics of the genes encoding 18s-5.8s-26s 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, Gill GP, Kuntz RJ, Langley CH, Neale DB (2004) Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc Natl Acad Sci USA 101:15255–15260

Burley J (1965) Karotype analysis of Sitka spruce, *Picea sitchensis* (Bong.) Carr. Silvae Genet 14:127–132

Cannell MGR, Milne R (1995) Carbon pools and sequestration in forest ecosystems in Britain. Forestry 68:361–378

Campbell CS, Wright WA, Cox M, Vining TF, Major CS, Arsenault MP (2005) Nuclear ribosomal DNA internal transcribed spacer 1 (ITS1) in *Picea* (Pinaceae): sequence divergence and structure. Mol Phylogenet Evol 35:165–185

Cardon LR, Bell JI (2001) Association study designs for complex diseases. Nat Rev Genet 2:91–99

Chaisurisri K, El-Kassaby YA (1994) Genetic diversity in a seed production population vs. natural populations of Sitka Spruce. Biodivers Conserv 3:512–523

Charest PJ, Devantier Y, Lachance D (1996) Stable genetic transformation of *Picea mariana* (black spruce) via particle bombardment. In Vitro Cell Dev Biol Plant 32:91–99

Cheliak WM, Pitel JA, Murray G (1985) Population structure and the mating system of white spruce. Can J For Res 15:301–308

Collignon AM, Van de Sype H, Favre JM (2002) Geographical variation in random amplified polymorphic DNA and quantitative traits in Norway spruce. Can J For Res 32:266–282

Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, Henikoff S (2004) Efficient Discovery of DNA polymorphisms in natural populations by ecotilling. Plant J 37:778–786

Corriveau AG (1981) Variabilité spatiale et temporelle de la croissance juvénile des provenances d'épinette noire. 18th Meeting of the CTIA, Duncan, BC, Canada

Corriveau A, Beaulieu J, Daoust G (1991) Heritability and genetic correlations of wood characters of Upper Ottawa Valley white spruce populations grown in Quebec. For Chron 67:698–705

Cottrell JE, White IMS (1995) The use of isozyme genetic markers to estimate the rate of outcrossing in a Sitka spruce (*Picea sitchensis* (Bong.) Carr.) seed orchard in Scotland. New For 10:111–122

Cyr DR, Klimaszewska K (2002) Conifer somatic embryogenesis: II. Applications. Dendrobiology 48:41–49

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

Dietrichson J (1969) The geographic variation of spring-frost resistance and growth cessation in Norway spruce (*Picea abies* (L.) Karst.). Medd Norske Skogforsøksvesen 27:91–104

Dirlewanger E, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Vosin R, Poizat C, Lafarque B, Baron O, Laigret F, Kleinhentz M, Arus P, Esmenjaud D (2004) Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid location of root-knot nematode resistance genes. Theor Appl Genet 109:827–838

Drake PMW, John A, Power JB, Davey MR (1997). Expression of the *gusA* gene in embryogenic cell lines of Sitka spruce following *Agrobacterium*-mediated transformation. J Exp Bot 48:151–155

Ellis DD, McCabe DE, McInnis S, Ramachandran R, Russell DR, Wallace KM, Martinell BJ, Roberts DR, Raffa KF, McCown BH (1993) Stable transformation of *Picea glauca* by particle acceleration. Biotechnology (NY) 11:84–89

Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, et al (2003) Highly parallel SNP genotyping. Cold Spring Harb Symp Quant Biol 68:69–78

Farjon A (1990) Pinaceae: drawings and descriptions of the genera *Abies, Cedrus, Pseudolarix, Keteleeria, Nothotsuga, Tsuga, Cathaya, Pseudotsuga, Larix* and *Picea*. Koeltz, Königstein, Germany

Farrar JL (1995) Trees in Canada. Canadian Forest Service and Fitzhenry & Whiteside, Markham, Ontario, Canada

Fletcher AM (1992) Breeding improved Sitka for the 90's. Forestry Commission Bulletin No 103, Edinburgh, pp 11–24

Fournier D, Perry DJ, Beaulieu J, Bousquet J, Isabel N (2002) Optimizing expressed sequence tag polymorphisms by single-strand conformation polymorphism in spruces. For Genet 9:11–18

Fowler DP (1983) The hybrid black × Sitka spruce, implications to phylogeny of the genus *Picea*. Can J For Res 13:108–115

Fowler DP (1987) The hybrid white × Sitka spruce: species crossability. Can J For Res 17:413–417

Fox DP (1987) The chromosomes of *Picea sitchensis* (Bong.) Carr. and its relatives. Proc R Soc Edinburgh Sect B Biol Sci 93:51–59

Frankis MP (1989) Genetic inter-relationships in Pinaceae. Notes R Bot Gard Edinburgh 45:527–548

Fulton T, van der Hoeven R, Eannetta N, Tanksley S (2002) Identification, analysis and utilization of a conserved ortholog set (COS) markers for comparative genomics in higher plants. Plant Cell 14:1457–1467

Furnier GR, Stine M, Mohn CA, Clyde MA (1991) Geographic patterns of variation in allozymes and height growth in white spruce. Can J For Res 21:707–712

Gamache I, Jaramillo-Correa JP, Payette S, Bousquet J (2003) Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: Imprint of a founder effect associated with postglacial colonization. Mol Ecol 12:891–901

Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P, Schüler K (2004) Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. Mol Breed 13:93–102

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 pseudotest-cross mapping strategy and RAPD markers. Genetics 137:1121–1137

Guillet-Claude C, Isabel N, Pelgas B, Bousquet J (2004) The evolutionary implications of *knox-I* gene duplications in conifers: correlated evidence from phylogeny, gene mapping, and analysis of functional divergence. Mol Biol Evol 21:2232–2245

Hakman I, Fowke LC, von Arnold S, Eriksson T (1985) The development of somatic embryos in tissue cultures initiated from immature embryos of *Picea abies* (Norway spruce). Plant Sci 38:53–59

Hall PM (1994) Ministry of forests perspectives on spruce reforestation in British Columbia. FRDA Rep 226. Victoria, BC, Canadian Forest Service, pp 1–6

Harris AS (1990) *Picea sitchensis* (Bong.) Carr. Sitka spruce. In: Burns RM, Honkala BH (tech coords) Silvics of North America, vol 1: Conifers. Agric Handbook 654. USDA Forest Service, Washington, DC, pp 260–267

Hawley GJ, DeHayes DH (1994) Genetic diversity and population structure of red spruce (*Picea rubens*). Can J Bot 72:1778–1786

Hayden MJ, Kuchel H, Chalmers KJ (2004) Sequence tagged microsatellites for the *Xgwm533* locus provide new diagnostic markers to select for the presence of stem rust resistance gene *Sr2* in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 109:1641–1647

Hermann RK (1987) North American tree species in Europe: transplanted species offer good growth potential on suitable sites. J For 85:27–32

Hodgetts RB, Aleksiuk MA, Brown A, Clarke C, Macdonald E, Nadeem S, Khasa D (2001) Development of microsatellite markers for white spruce (*Picea glauca*) and related species. Theor Appl Genet 102:1252–1258

Hosie RC (1969) Native Trees of Canada. 7th ed. Canadian Forestry Service, Department of Fisheries and Forestry, Ottawa, Ontario, Canada

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

Huber DPW, Ralph S, Bohlmann J (2004) Genomic hardwiring and phenotypic plasticity of terpenoid-based defenses in conifers. J Chem Ecol 30:2401–2420

Ingle J, Timmis JN, Sinclair J (1975) The relationship between satellite deoxyribonucleic acid, ribosomal ribonucleic acid gene redundancy, and genome size in plants. Plant Physiol 55:496–501

Isabel N, Beaulieu J, Bousquet J (1995) Complete congruence between gene diversity estimates derived from genotypic data at enzyme and RAPD loci in black spruce. Proc Natl Acad Sci USA 92:6369–6373

Isidore E, Scherrer B, Bellec A, Budin K, Faivre-Rampant P, Waugh R, Keller B, Caboche M, Feuilllet C, Chalhoub B (2005) Direct targeting and rapid isolation of BAC clones spanning a defined chromosomal region. Funct Integr Genom 9:97–103

Jaramillo-Correa JP, Bousquet J (2003) New evidence from mitochondrial DNA of a progenitor-derivative species relationship between black spruce and red spruce (Pinaceae). Am J Bot 90:1801–1806

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

Jaramillo-Correa JP, Beaulieu J, Bousquet J (2001) Contrasting evolutionary forces driving population structure at ESTPs, allozymes, and quantitative traits in white spruce. Mol Ecol 10:2729–2740

Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004) Variation in mtDNA reveals multiple glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. Mol Ecol 13:2735–2748

Jaramillo-Correa JP, Beaulieu J, Ledig FT, Bousquet, J (2006) Decoupled mitochondrial and chloroplast DNA population structure reveals Holocene collapse and population isolation in a threatened Mexican-endemic conifer. Mol Ecol 15:2878–2800

Joyce PM, O'Carroll N (2002) Sitka spruce in Ireland. COFORD, Dublin

Jurman I, Zuccolo I, Vischi M, Morgante M (1999) Analysis of structure and organization of repetitive sequences in the Norway spruce genome (*Picea abies* K.). In: Plant & Animal Genome VII Conf, San Diego

Kean VM, Fox DP, Faulkner R (1982) The accumulation mechanism of the supernumerary (B-) chromosome in *Picea sitchensis* (Bong.) Carr. and the effect of this chromosome on male and female flowering. Silvae Genet 31:126–131

Khalil MAK (1984) All-range black spruce provenance study in Newfoundland: Performance and genotypic stability of provenances. Silvae Genet 33:63–71

King JN, Yanchuk AD, Kiss GK, Alfaro RI (1997) Genetic and phenotypic relationships between weevil (*Pissodes strobi*) resistance and height growth in spruce populations of British Columbia. Can J For Res 27:732–739

Kiss GK (1989) Engelmann × Sitka spruce hybrids in central British Columbia. Can J For Res 19:1190–1193

Kiss G, Yeh FC (1988) Heritability estimates for height for young interior spruce in British Columbia. Can J For Res 18:158–162

Klimaszewska K, Lachance D, Pelletier G, Lelu MA, Seguin A (2001) Regeneration of transgenic *Picea glauca*, *P. mariana*, and *P. abies* after cocultivation of embryogenic tissue with *Agrobacterium tumefaciens*. In Vitro Cell Dev Biol Plant 37:748–755

Krutovskii KV, Bergmann F (1995) Introgressive hybridization and phylogenetic relationships between Norway, *Picea abies* (L.) Karst., and Siberian, *P. obovata* Ledeb., spruce species studied by isozyme loci. Heredity 74:464–480

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 Genetics and Genomics 3:61–70

Lagercrantz U, Ryman N (1990) Genetic structure of Noway spruce (*Picea abies*): concordance of morphological and allozymic variation. Evolution 44:38–53

Lamothe M, Meirmans P, Isabel N (2006) A set of polymorphic EST-derived markers for *Picea* species. Mol Ecol Notes 6:237–240

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

Ledig FT, Bermejo-Velazquez B, Hodgskiss PD, Johnson DR, Flores-Lopez C, Jacob-Cervantes V (2000) The mating system and genic diversity in Martinez spruce, an extremely rare endemic of Mexico's Sierra Madre Orientale: an example of facultative selfing and survival in interglacial refugia. Can J For Res 30:1156–1164

Ledig FT, Hodgskiss PD, Jacob-Cervantes V (2002) Genetic diversity, mating system, and conservation of a Mexican subalpine relect, *Picea mexicana* Martinez. Conserv Genet 3:113–122

Lee SJ (1993) Breeding strategy for Sitka spruce in Britain. In: Lee SJ (Ed) Proc Nordic Group for Tree Breed. Forestry Research, Roslin, Scotland, pp 95–109

Lefort F, Echt C, Streiff R, Vendramin GG (1999) Microsatellite sequences: a new generation of molecular markers for forest genetics. For Genet 6:15–20

LePage BA (2001) New species of *Picea* A. Dietrich (Pinaceae) from the middle Eocene of Axel Heiberg Island, Arctic Canada. Bot J Linn Soc 135:137–167

Levée V, Lelu MA, Jouanin L, Cornu D, Pilate G (1997) *Agrobacterium tumefaciens*-mediated transformation of hybrid larch (*Larix kaempferi* T *L. decidua*) and transgenic plant regeneration. Plant Cell Rep 16:680–685

Li P, Beaulieu J, Corriveau A, Bousquet J (1993) Genetic variation in juvenile growth and phenology in a white spruce provenance-progeny test. Silvae Genet 42:52–60

Li P, Beaulieu J, Bousquet J (1997) Genetic structure and patterns of genetic variation among populations in eastern white spruce (*Picea glauca*). Can J For Res 27:189–198

Lippert D, Chowrira S, Ralph SG, Zhuang J, Aeschliman D, Ritland C, Ritland K, Bohlmann J (2007) Conifer defense against insects: Proteome analysis of Sitka spruce (*Picea sitchensis*) bark induced by mechanical wounding or feeding by white pine weevils (*Pissodes strobi*). Proteomics 7:248–270

Liston A, Gernandt DS, Vining TF, Campbell CS, Pinero D (2003) Molecular phylogeny of Pinaceae and *Pinus*. Acta Hort 615:107–114

Malcolm DC (1997) The silviculture of conifers in Great Britain. Forestry 70:293–307

Mikkola L (1969) Observations on interspecific sterility in *Picea*. Ann Bot Fenn 6:285–339

Moir RB, Fox DP (1972) Supernumerary chromosomes in *Picea sitchensis* (Bong.) Carr. Silvae Genet 21:182–186

Morgenstern EK (1969) Genetic variation in seedlings of *Picea mariana* (Mill.) B.S.P.: I. Correlation with ecological factors. Silvae Genet 18:151–161

Murray BG (1998) Nuclear DNA amounts in Gymnosperms. Annals Bot 82(Suppl A):3-15

Neale DB, Sewell MM, Brown GR (2002) Molecular dissection of the quantitative inheritance of wood property traits in loblolly pine Ann For Sci 59:595–605

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. Trends Plant Sci 9:325–330

Neale DB, Wheeler NC (2004) Mapping of quantitative trait loci in loblolly pine and Douglas-fir: a summary. For Genet 11:173–178

Nehra NS, Becwar MR, Rottmann WH, Pearson L, Chowdhury K, Chang Shujun, Wilde HD, Kodrzycki RJ, Zhang C, Gause KC, Parks DW, Hinchee MA (2005) Forest Biotechnology: Innovative methods, emerging opportunities. Invited Review. In Vitro Cell Dev Biol Plant 41:701–717

Nienstaedt H (1985) Inheritance and correlations of frost injury, growth, flowering, and cone characteristics in white spruce, *Picea glauca* (Moench) Voss. Can J For Res 15:498–504

Nkongolo KK (1996) Chromosome analysis and DNA homology in three *Picea* species, *P. mariana, P. rubens*, and *P. glauca* (Pinaceae). Plant Syst Evol 203:27–40

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

Parker WH, Van Niejenhuis A, Charette P (1994) Adaptive variation in *Picea mariana* from northwestern Ontario determined by short-term common environment tests. Can J For Res 24:1653–1661

Pavy N, Paule C, Parsons L, Crow JA, Morency M-J, Cooke J, Johnson JE, Noumen E, Guillet-Claude C, Butterfield Y, Barber S, Yang G, Liu J, Stott J, Kirkpatrick R, Siddiqui A, Holt R, Marra M, Séguin A, Retzel E, Bousquet J, Mackay J (2005) Generation, annotation, analysis, and database integration of 16,500 white spruce EST clusters. BMC Genomics 6(144):19

Pavy N, Parsons L, 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 Genomics 7(174):14

Pavy N, Gosselin I, Gagnon F, Lamothe M, Laroche L, Isabel N, Bousquet J (2007) A SNP-based composite map of regulatory genes in the conifer white spruce and structural relationships with other plants In: Plant & Animal Genome XV Conf, San Diego, CA, USA

Pelgas B, Isabel N, Bousquet J (2004) Efficient screening for expressed sequence tag polymorphisms (ESTP) by DNA pool sequencing and denaturing gradient gel electrophoresis 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* [Mill.] B.S.P. × *P. rubens* (Sarg.) 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. abies* and *P. mariana* × *rubens*, and correspondance with other Pinaceae. Theor Appl Genet 113:1371–1393

Pena L, Séguin A (2001) Recent advances in the genetic transformation of trees. Trends Biotechnol (NY) 19:500–506

Perron M, Bousquet J (1997) Natural hybridization between *Picea mariana* and *P. rubens*. Mol Ecol 6:725–734

Perron M, Perry DJ, Andalo C, Bousquet J (2000) Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. Proc Natl Acad Sci USA 97:11331–11336

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

Perry DJ, Bousquet J (1998b) Sequence tagged-site (STS) markers of arbitrary genes: the utility of black spruce-derived STS primers in other conifers. Theor Appl Genet 97:735–743

Perry DJ, Bousquet J (2001) Genetic diversity and mating system of post-fire and post-harvest black spruce: an investigation using codominant sequence-tagged site (STS) markers. Can J For Res 31:32–40

Perry DJ, Isabel N, Bousquet J (1999) Sequence-tagged-site (STS) markers of arbitrary genes: the amount and nature of variation revealed in Norway spruce. Heredity 83:239–248

Peterson DG (2005) Reduced representation strategies and their application to plant genomes. In: Meksem K, Kahl G (eds) The Handbook of Plant Genome Mapping: Genetic and Physical Mapping. Wiley-VCH, Weinheim, Germany, pp 307–335

Plomion C, Hurme P, Frigerio JM, Ridolfi M, Pot D, Pionneau C, Avila C, Gallardo F, David H, Neutelings G, Campbell M, Canovas FM, Savolainen O, Bodénès C, Kremer A (1999) Developing SSCP markers in two *Pinus* species. Mol Breed 5:21–31

Pot D, McMillan L, Echt C, Le Provost G, Garnier-Géré P,Cato S, Plomion C (2005) Nucleotide variation in genes involved in wood formation in two pine species. New Phytol 167:101–112

Presby-Germano J (2003) Nucleotide and cytoplasmic genetic variation in *Picea:* DNA markers for evaluating past migration, introgression and evolutionary history. PhD Dissertation, University of New Hampshire, Durham, NH

Rajora OP, Mosseler A, Major JE (2000) Indicators of population viability in red spruce, *Picea rubens*. II. Genetic diversity, population structure, and mating behavior. Can J Bot 78:941–956

Rajora OP, Rahman MH, Dayanandan S, Mosseler A (2001) Isolation, characterisation, inheritance and linkage of microsatellite DNA markers in white spruce (*Picea glauca*) and their usefulness in other spruce species. Mol Gen Genet 264:871–882

Ralph SG, Yueh H, Friedmann M, Aeschliman D, Zeznik JA, Nelson CC, Butterfield YSN, Kirkpatrick R, Liu J, Jones SJM, Marra MA, Douglas CJ, Ritland K, Bohlmann J (2006) Conifer defense against insects: Microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobi*) reveals large-scale changes of the host transcriptome. Plant, Cell and Environment 29:1545–1570.

Roche L (1969) A genecological study of the genus *Picea* in British Columbia. New Phytol 68:505–554

Rungis D, Berube 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

Rungis D, Hamberger B, Berube Y, Wilkin J, Bohlmann J, Ritland K (2005) Efficient genetic mapping of single nucleotide polymorphisms based upon DNA mismatch digestion. Mol Breed 16:261–270

Rweyongeza DM, Yeh FC, Dhir NK (2004) Genetic parameters for seasonal height and height growth curves of white spruce seedlings and their implications to early selection. For Ecol Mng 187:159–172

Saranga Y, Menz M, Jiang CX, Wright RJ, Yakir D, Paterson AH (2001) Genomic dissection of genotype × environment interactions conferring adaptation of cotton to arid conditions. Genome Res 11:1988–1995

Schlötterer C (2004) The evolution of molecular markers-just a matter of fashion? Nat Rev Genet 5:63–69

Schmidt-Vogt H (1977) Die Fichte. Band I [The spruces. Vol 1]. Taxonomie-Verbreitung-Morphologie-Ökologie-Waldgesellscahften. Parey, Hamburg

Scotti I, Magni F, Fink R, Powell W, Binelli G, Hedley PE (2000) Microsatellite repeats are not randomly distributed within Norway spruce (*Picea abies* K.) expressed sequences. Genome 43:41–46

Scotti I, Magni F, Paglia G, Morgante M (2002) Trinucleotide microsatellites in Norway spruce (*Picea abies*): their features and the developement of molecular markers. Theor Appl Genet 106:40–50

Scotti I, Burelli A, Cattonaro F, Chagné F, 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 Genetics and Genomics 1:93–102

Seybold SJ, Bohlmann J, Raffa KF (2000) The biosynthesis of coniferophagous bark beetle pheromones and conifer isoprenoids: evolutionary perspective and synthesis. Can Entomol 132:697–753

Shea KL (1987) Effects of population structure and cone production on outcrossing rates in Engelmann spruce and subalpine fir. Evolution 41:124–136

Sigurgeirsson A, Szmidt AE (1993) Phylogenetic and biogeographic implications of chloroplast DNA variation in *Picea*. Nord J Bot 13:233–246

Sperisen C, Buchler U, Gugerli F, Matyas G, Geburek T, Vendramin GG (2001) Tandem repeats in plant mitochondrial genomes: application to the analysis of population differentiation in the conifer Norway spruce. Mol Ecol 10:257–263

Stasolla C, Kong L, Yeung EC, Thorpe TA (2002) Maturation of somatic embryos in conifers: morphogenesis, physiology, biochemistry, and molecular biology. In Vitro Cell Dev Biol Plant 38:93–105

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 5:739–744

Strauss SH, Difazio SP, Meilan R (2001) Genetically modified poplars in context. For Chron 77:272–279

Sutton BCS, Flanagan DJ, Gawley JR, Newton CH, Lester DT, El-Kassaby YA (1991) Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of hybrids from introgression zones. Theor Appl Genet 82:242–248

Sutton BCS, Attree SM, El-Kassaby YA, Grossnickle SC, Polonenko DR (2004) Commercialisation of somatic embryogenesis for plantation forestry. In: Walter C, Carson M (eds) Plantation Forest Biotechnology for the 21st Century, Trivandrum, India

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

Teoh SB, Rees H (1976) Nuclear DNA amounts in populations of *Picea* and *Pinus* species. Heredity 36:123–137

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES (2001) Dwarf8 polymorphisms associate with variation in flowering time. Nat Genet 28:286–289

Tosh KJ, Fullarton MS (2001) Tree improvement progress by the New Brunswick department of Natural Resources and Energy. In: Simpson JD (ed) Proc 27th Meeting of the CTIA. Natural Resources of Canada, Sault-Ste. Marie, Ontario, Canada

Trapp S, Croteau R (2001) Defensive resin biosynthesis in conifers. Annu Rev Plant Physiol Plant Mol Biol 52:689–724

Tulsieram LK, Glaubitz JC, Kiss G, Carlson JE (1992) Single tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. Biotechnology (NY) 10:686–90

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 (2007) An AFLP linkage map for Douglas-fir based upon small multiple full-sib families. Tree Genetics and Genomics (in press)

van de Ven WTG, Kanamori RJ (1996) Microsatellites as DNA markers in Sitka spruce. Theor Appl Genet 93:613–617

van Frankenhuyzen K, Beardmore T (2004) Current status and environmental impacts of transgenic forest trees. Can J For Res 34:1163–1180

Vendramin GG, Anzidei M, Madaghiele A, Sperisen C, Bucci G (2000) Chloroplast microsatellite analysis reveals the presence of population subdivision in Norway spruce (*Picea abies* K.). Genome 43:68–78

Viereck LA, Little EL Jr (1972) Alaska trees and shrubs. Agric Handbook 410. USDA Forest Service, Washington, DC

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Friters A, Pot J, Paleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Weng C, Jackson ST (2000) Species differentiation of North American spruce (*Picea*) based on morphological and anatomical characteristics of needles. Can J Bot 86:1742–1753

Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary

primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Wright W (1955) Species crossability in spruce in relation to distribution and taxonomy. For Sci 1:319–349

Xie CY, Knowles P (1994) Mating system and effective pollen immigration in a Norway spruce (*Picea abies* (L.) Karst) plantation. Silvae Genet 43:48–52

Yanchuk A (2001) The role and implications of biotechnology in forestry. For Genet Resour 30:18–22

Yeh FC, Arnott JT (1986) Electrophoretic and morphological differentiation of *Picea sitchensis*, *Picea glauca*, and their hybrids. Can J For Res 16:791–798

Yeh FC, El-Kassaby YA (1980) Enzyme variation in natural populations of Sitka spruce (*Picea sitchensis*): 1. Genetic variation patterns among trees from 10 IUFRO provenances. Can J For Res 10:415–422

Yeh FC, Khalil MAK, El-Kassaby YA, Trust DC (1986) Allozyme variation in *Picea mariana* from Newfoundland: genetic diversity, population structure, and analysis of differentiation. Can J For Res 16:713–720

Ying CC (1991) Genetic resistance to the white pine weevil in Sitka spruce. Research Note 106, British Columbia Ministry of Forests, Victoria, BC, Canada

Zhang T, Vision T, Gaut B (2002) Patterns of nucleotide substitution among simultaneously duplicated gene pairs in *Arabidopsis thaliana*. Mol Biol Evol 19:1464–1473

Zobel BJ, Talbert JT (1984) Applied Forest Tree Improvement. Wiley, New York