Comparison of allelic diversity between native gene resource plantings and selections in open-pollinated progeny test of *Pinus radiata* D. Don.

By L. Zhang¹, N. Huanqiong², W. J. Gapare^{3),*}, S. K. Dillon³, X. Li³ and H. X. Wu^{3),4})

(Received 20th May 2014)

Abstract

Genetic diversity within radiata pine first generation of open-pollinated selections (OPS) from the native resource stands was compared with that observed in native populations to monitor potential changes in genetic diversity during domestication. Genetic diversity was estimated using 58 single nucleotide polymorphisms (SNPs) from 8 expressed genes. Nucleotide diversity maintained in first generation of selections (OPS) (mean $\pi = 0.0036$; mean $\theta_w = 0.0058$) was similar to that found within the native population material (mean $\pi = 0.0043$; mean for $\theta_w = 0.0065$). Likewise, mean values for expected heterozygosity (H_E) within and between native population material and OPS were similar (mean = 0.27 ± 0.04) and not significantly different (P = 0.068). Also, the overall distribution of allele frequency classes was not significantly between native population material and OPS. These results point to no evidence of loss of diversity in OPS due to artificial selection. One possible reason is that the domestication of the OPS is at a very early stage. Another may be that artificial selection in the OPS was based on tree growth and form, not wood properties. The genes selected in this study are mostly involved in cell wall formation, thus genetic diversity of these genes should remain stable between natural population and OPS, unless there was a significant sampling bias in the OPS. Although the SNP information suggests similarities among mainland populations, results from quantitative genetic studies found large provenance differences for growth-, morphological-, stem-form traits, and disease resistance. Determining the threshold at which genetic diversity levels will be significantly reduced during selection should help breeders to make informed decisions regarding the intensity of selection in managed breeding populations as well as gene resource pop-

Key words: radiata pine, tree breeding, single nucleotide polymorphisms (SNPs), allele frequency, genetic diversity.

Introduction

The genetic structure of natural forests has evolved over many generations via the processes of selection, migration, mutation, and random genetic drift. In commercial domesticated tree species, this genetic turnover will be altered by artificial selection and tree breeding programs. However, it is expected that the breeding process may reduce the genetic base of natural populations; domestication of agricultural crop plants has illustrated this process well (Godt et al., 2001). Few foresttree breeding programs have reached the stage of domestication where loss of genetic diversity is an issue; most have undergone less than four cycles of improvement (e.g., radiata pine) (Wu et al., 2007). Recent use of small, elite breeding populations managed for shortterm genetic gain could accelerate the loss of genetic diversity in domesticated populations of some tree species. Elite populations should be prone to allele loss through stringent selection, increased inbreeding and genetic drift given their small population sizes. Monitoring of the genetic diversity in tree breeding and production populations is needed, not only for gauging the impacts of breeding and production activities on reforestation germplasm, but also for effective genetic conservation of existing natural populations (e.g., AITKEN, 2000).

Molecular marker approaches have been used to investigate the genetic changes which occurred during tree domestication (ADAMS, 1981; SZMIDT and MUONA, 1985; Godt et al., 2001). For example, Godt et al. (2001) used isozyme markers to compare genetic diversity within white spruce seed orchard clones with genetic variation in natural populations in the source region for the clones. Similarly, Jones et al. (2006) used microsatellites markers to compare genetic diversity in the first generation of selections in Eucalyptus globulus (blue gum) breeding program in Australia with that observed in native populations to monitor potential changes in genetic diversity during domestication. Reduced genetic diversity in breeding populations in comparison to native populations has been reported in several studies of conifer species (e.g., radiata pine -MORAN and BELL, 1987; loblolly pine - WILLIAMS and HAMRICK, 1995; blue gum – Jones et al., 2006). Genetic diversity in some breeding populations of tree species could be higher than in native populations as a result of crossing between differentiated native populations or infusion from other breeding populations (Lefevre, 2004). However, these studies did not reveal any significant differences between these types of populations, whether the man-made populations were the result of phenotypic or genetic selection.

Radiata pine (*Pinus radiata* D. Don), the predominant plantation softwood species in the southern hemisphere, is native to the California coast (Año Nuevo, Monterey

Silvae Genetica 63, 5 (2014) 213

¹⁾ State Key Laboratory of Forest Genetic and Tree Breeding, Northeast Forest University, Harbin, Heilongjiang province, China.

²⁾ Yunnan Forestry Technological College, Yunnan, China.

³) CSIRO Agriculture Flagship, GPO Box 1600, Canberra, ACT 2601, Australia.

³),⁴) Umeå Plant Science Centre, Dept. Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden.

^{*)} Corresponding author: Washington J. Gapare.
Telephone: +61 2 6246 4841, Fax: +61 2 6246 4564. E-Mail: Washington.Gapare@csiro.au

and Cambria) and two offshore islands (Guadalupe and Cedros) (Fig. 1). Owing to its rapid growth and favorable characteristics for timber production, the species has been widely introduced in Argentina, Australia, China, Spain, Great Britain, Ireland, Italy, New Zealand, South Africa, Turkey and Chile. In New Zealand radiata pine is the most important exotic forest species, being the most extensively planted conifer in the country (1.8 million ha), covering 89% percent of the New Zealand forest industry base and earning foreign currency on the order of US\$ 4.65 million annually in exports (MAF, 2011).

Details on the introduction, genetic base and state of radiata pine breeding populations in Australia and New Zealand are summarised in Wu et al. (2007), Johnson et al. (2008) and GAPARE (2014). Radiata pine was introduced into Australia and New Zealand, via England, in the mid 1800's. Small scale provenance research was embarked on in Australia and New Zealand from the late 1930s through the 1950s, and there were small experimental plantings of island material in Australia. This, however, was done with no real commitment to provenance research, because of a prevailing belief that provenance would not be a real issue with radiata pine. A key change in thinking emerged after work published by Bannister (1959) that suggested that the entire geographic range of the species needed to be embraced, and should include both the Guadalupe and Cedros populations. During the 1960's two major collections were carried out – one by Forde was around 1962 and the second one by Dr. Reid Moran on Guadalupe Island, supplemented by an expedition by MORAN and LIBBY to

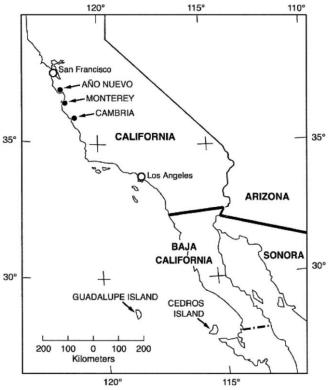


Figure 1. – Location of natural radiata pine populations: Año Nuevo, Monterey and Cambria native to California and Guadalupe and Cedros Islands in Mexico. Adapted from Ledig et al. (1998).

Guadalupe and Cedros Islands in early 1964 (Hood and Libby, 1980). A Genetic Survey provenance/progeny experiment was subsequently established using seed arising from Forde's cone sampling, a seed collection made in Guadalupe by Dr. Reid Moran, two New Zealand landrace population samples as controls, and Cedros Island material from the 1964 Libby collection.

Results from the Genetic Survey showed that New Zealand land-race material was very predominantly, if not entirely from Año Nuevo and Monterey. A combination of early results from the Genetic Survey, monoterpene studies, some step-out provenance trials and evidence that edaphic adaptation was an important issue, led to the joint New Zealand/Australia 'Eldridge' expedition. As breeding programs became more sophisticated in terms of pedigree, thought was also given to expanding the genetic base by re-collecting in the original stands in California, and this was also a contributing factor to the joint New Zealand/Australia expedition. For example, a decision to make nearly 600 fresh selections in New Zealand was prompted by concerns over the scope for doing future selection among the much smaller number of '850' selections and their offspring (Shelbourne et al., 1986). A comprehensive collection was carried out in 1978 (referred in this paper as the 1978 'Eldridge collection'). The collection was done as a collaboration between New Zealand and Australia with joint funding from United States of America/Australia Cooperative Science Program, CSIRO Division of Forest Research, New Zealand Forest Research Institute, and Forest Resources Division of Food and Agricultural Organisation. Seed was collected from a total of 621 mother trees, at least 22 trees from each native population with trees selected for having at least 40 collectable cones and being well grown and of good form (ELDRIDGE, 1978) (see Fig. 1; Table 1). The 1978 'Eldridge collection' was used to establish countrywide provenance trials and large block plantings of gene-resource plantings between 1979 and 1982 at many locations in Australia, New Zealand and Chile. The main objective of the trials was to increase the genetic base of radiata pine breeding population, which was thought to be mainly of Año Nuevo origin, with less of the ancestry from Monterey and none from Cambria (BURDON et al., 2008).

In the present study we compared the selection effects on genetic diversity between native populations and a first generation breeding population. The plus trees in the breeding population were selected from open-pollinated domesticating trials of 1978 'Eldridge collection'. These domesticated trials are conserved as part of Radiata Pine Breeding Company (RPBC) genetic resources. To do so, we used samples collected in a first-generation radiata pine gene-resource plantings and in ex situ natural population material from which the trees of the breeding population originated. DILLON et al. (2010) genotyped individuals from mainland populations of the 1978 'Eldridge' collection using genetic markers called single nucleotide polymorphisms (SNPs) markers to estimate nucleotide diversity. SNPs (pronounced "snips"), are the most common type of genetic variation among organisms. Each SNP represents a difference in a single DNA building block, called a nucleotide. For

Table 1. – Seed-origin information for *Pinus radiata* populations sampled in California in 1978 (from ELDRIDGE, 1978).

Provenance	No. of mother	Latitude (°N)	Longitude (°W)	Altitude (m)	Rainfall (mm/annum)‡
	trees				
Año Nuevo	179	37°08°	122°18′	15-300	700
Monterey	244	36°37'	121°57'	5-580	500
Cambria	99	35°37'	121°09°	30-180	500
Guadalupe*	48	29°10'	118°15'	400- 1200	330-510
Cedros	51	28°22'	115°20'	380-640	150
Total	621				

⁺ Latitude and longitude are derived directly from map locations and represent approximate centre points of the island and mainland populations.

example, a SNP may replace the nucleotide cytosine (C) with the nucleotide thymine (T) in a certain stretch of DNA. The Eldridge collection provided a baseline assessment of allelic diversity within the species with which to compare open-pollinated selections (OPS) from the gene resource plantings (native population material). Our study utilizes a subset of the genes employed in DILLON et al. (2010) in which they explored genetic diversity in radiata pine using 149 SNP loci, of which ten were identified as being possible targets of diversifying selection in native population material.

Materials and Methods

Plant materials and DNA extraction

Details of the current breeding strategy for radiata pine grown in New Zealand and New South Wales, Australia are given in Dungey et al. (2009). The strategy recommended infusion of ten unrelated parents from the native-population selections from the 1978 'Eldridge collection'. Details of the collection were provided in Eldridge (1978). Growth-, form- and wood-quality data from resulting provenance/progeny trials have already been the subject of earlier publications (RAYMOND and HENSON, 2009; GAPARE et al., 2011; 2012 a & b).

Native population material

Trees sampled from each of the three mainland populations of radiata pine had been maintained ex situ as a provenance/progeny trial at Batlow, NSW, Australia (35°31'17"S; 48°08'40"E). This trial was established in 1980, from the 1978 'Eldridge collection' of radiata pine from the natural populations (Figure 1; ELDRIDGE, 1978). Needles were collected from 447 individuals, representing 447 different families, from the Año Nuevo (155), Monterey (210) and Cambria (82) provenances populations.

Open-pollinated selections (OPS)

Needle tissue samples were collected from 552 individuals of radiata pine in a provenance/progeny trial planted in Kangaroovale, NSW, Australia (Latitude 35°05'S;

Longitude 148°15'E). The trial included individuals from three provenances: Año Nuevo (185), Monterey (219) and Cambria (148). The trial was established in 1995 from material selected in gene-resource stands (1978 'Eldridge collection') of radiata pine from the naturalpopulation material planted in New Zealand in 1982. The stands had approximately 833 stems per ha. In 1994, 128 plus trees were selected in the gene-resource plantings at a rate of one tree per hectare, with at least 50 metres between selected trees. Selection criteria were based on growth, health, straightness and lack of malformation, and the selected trees also had to bear cones. Twenty cones were collected from each tree where possible, and cones were spread over 5 years of pollination. The cones were collected from pollinations postdating thinnings done in 1986 and 1994. We assumed that the material represented half-sib families, specifically that the seed parents were a random sample of their respective provenances and that each one was pollinated from a large random sample of trees from its own provenance.

Candidate genes and SNP genotyping

Genomic DNA was extracted from approximately 80 mg fresh needle tissue from the two sets of material, using a modified CTAB protocol (Doyle and Doyle, 1990), and further purified on QIAquick® PCR purification columns. Eight genes (Table 1) that possibly influence wood properties were available for this study (DIL-LON et al., 2010). These genes are involved in the development of secondary xylem, including primary cell wall modification, cell skeleton development cellulose synthesis, lignin pathway, cell wall structural proteins. Prior to genotyping, haplotype tagging SNPs were selected from 10 amplicon alignments from 8 candidate gene covering 7465 bp of DNA sequence (Table 2). The resulting SNP data set was also applied in a separate study of genephenotype association in radiata pine (DILLON et al., 2010).

PCR reaction was carried out in a final volume of 20 μ l, including 2.0 μ l bulked DNA template, 0.4 μ l 10 mM dNTPs, 4.0 μ l phire 5 × buffer (plus 1.5 mM Mg²⁺), 0.4 μ l Phire Tag, 1.0 μ l forward primer and reversed primer

[‡] Annual rainfall figures are approximate and average.

^{*} Substantial occurrence of fog during spring, summer and autumn with fog are especially important in complementing the normal precipitation in drier years. Rainfall figures from Oberbauer (2006).

respectively. A touch down cycling program was used: 98°C 3 min; 10 cycles of 98°C 5s, 65°C 5s (decrease 1°C per cycle), 72°C 3 min; 35 cycles of 98°C 5s, 56°C 5s, 72°C 3 min; then 72°C 3 min and hold at 4°C. All PCR products were added with 3' A overhangs, and ligated to vector (pGEM-T vector). These recombinants were electro-transformed into E. coli cells (DH5 α) and cultured in LB medium (37°C, 200rp, 40min). White colonies were picked up from the LB-ampicillin plates. The colonies were then tested using PCR, and only colonies which were verified in PCR were finally used for sequencing. Colony PCR products were treated with ExoSap (GE Healthcare, UK) to degrade primers and denature proteins. Two µl of the treated PCR product was used in sequencing reaction with 12.5 µl DSW, 0.25 µl Big Dye, $3.5 \mu l \ 5 \times Buffer \ and \ 1 \mu l \ 3.2 \ um \ primer. After a quick$ start at 96°C 30s, the reaction was cycled 25 times at 96°C 30s, 50°C 15s, 60°C 4 min, followed by holding at 4°C. DNAs were sequenced using an ABI 3730 capillary sequencer (Applied Biosystems, CA).

$Statistical\ analysis$

DNA sequences were verified manually and were then assembled into contigs using the Bioedit software (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Multiple sequences were aligned using the ClustalW software followed by adjustment using the MEGA5 (http://www.megasoftware.net). SNPs were identified in each of the 10 amplicons. Nucleotide diversity was estimated using the DNA sequence polymorphism (DNASP) software (version 5.10) (http://www.ub.edu/dnasp).

Insertions and deletions were excluded from analysis. The level of polymorphisms for each gene was estimated as nucleotide diversity (π) , based on the average number of nucleotide differences between sequences (NEI, 1978), and the weighted average for total nucleotide diversity (θ_{w}) , based on the number of segregating sites (WATTER-SON, 1975) were estimated for each gene, each provenance and also for all three provenances combined. Analysis of molecular variance (AMOVA) was used to test for statistical significance among and within provenances using the software GenAlEx (PEAKALL and Smouse, 2012). The $\boldsymbol{H}_{\scriptscriptstyle E}$ and $\boldsymbol{H}_{\scriptscriptstyle O}$ corrected for small samples according to Nei (1978) were estimated for each population and the homogeneity of sample variance within population was tested using the Bartlett's test and the t-test in R environment (R Development Core TEAM, 2007). The deviation of genotype frequencies from Hardy-Weinberg equilibrium estimated using the within-population fixation index statistic $(F_{\rm IS})$, and observed allele frequencies for each SNP were calculated in R environment (R Development Core Team, 2007). These analyses were performed for native-population material (each population separately and all three native populations combined) and OPS separately.

Results

Patterns of nucleotide diversity in open-pollinated selections (OPS)

A total of 58 SNPs were identified in all gene fragments (7465bp) from 8 selected genes (*Table 2*). The

Table 2. – Details of 10 genes and number of SNPs and relative nucleotide diversity estimates in radiata pine open-pollinated selections.

Gene	Name/Amplicon	L (bp)	Coding (bp)	Noncodin g (bp)	#SNPs Coding	#SNPs Noncoding	Total SNPs	π	$(\hat{ heta}_{ m W})^{}$	D
α-tubulin	αtub-1	763	285	478	1	1	2	0.0016	0.0056	-2.2233**
Cinnamyl-alcohol dehydrogenase	cad	788	210	578	1	11	12	0.0063	0.0117	-2.0132*
Cellulose synthase 1 (P. radiata)	cesa3-2F	875	377	498	5	6	11	0.0051	0.0044	-0.4181
Cellulose synthase 1 (P. radiata)	cesa3-2R	879	262	617	3	5	8	0.0039	0.0051	-0.0552
Cellulose synthase 3 (<i>P. radiata</i>)	cesa1-4F	520	278	242	0	2	2	0.0026	0.0051	-1.1048
Cellulose synthase 3 (<i>P. radiata</i>)	cesal-4R	674	409	265	1	3	4	0.0037	0.0065	-1.4882
Katanin	fra2	752	125	627	0	3	3	0.0031	0.0068	-1.0855
Phenylalanine ammonia lyase 1 (<i>P. taeda</i>)	pal l	503	503	0	6	0	6	0.0022	0.0058	-1.8943*
S-adenosyl-methionine synthase 2 (<i>P. taeda</i>)	sam2	833	411	422	0	5	5	0.0028	0.0046	-0.9463
xyloglucan endotransglycosylase 1 (P. taeda)	xet1	878	443	435	1	4	5	0.0047	0.0093	-2.1056*
•	Total	7465	3303	4162	18	40	58	-	-	-
	Average	747	330	416	2	4	6	0.0036	0.0065	-1.3335

L – length of base pairs sequenced; π – Nucleotide diversity based on the average number of nucleotide differences between sequences; θ_w – weighted average for total nucleotide diversity based on the number of segregating sites; * – significant at P < 0.05; ** – significant at P < 0.01; D – Tajima's D-statistic.

	Native material	Open-pollinated selections (OPS)	Native material	Open-pollinated selections (OPS)	Native material	Open-pollinated selections (OPS)
Population	Nucleotide	Nucleotide	Nucleotide	Nucleotide	Expected	Expected
	diversity	diversity	diversity	diversity	heterozygosity	heterozygosity
	$(\hat{\pi})$	$(\widehat{\pi})$	$(\widehat{\theta w})$	$(\widehat{\theta w})$	(\widehat{IIE})	(\widehat{IIE})
Año Nuevo	0.0032	0.0030	0.0042	0.0047	0.26 ± 0.04	0.25 ± 0.05
Monterey	0.0058	0.0032	0.0064	0.0070	0.28 ± 0.05	0.27 ± 0.06
Cambria	0.0043	0.0038	0.0069	0.0078	0.29 ± 0.03	0.29 ± 0.05
Weighted	0.0043	0.0036	0.0058	0.0065	0.28 ± 0.04	0.27 ± 0.05
Mean						

Table 3. - Genetic parameter estimates in native material and open-pollinated selections of radiata pine.

number of polymorphisms varied greatly across genes with an average of 6 SNPs per locus (range: 2–12). The majority of SNPs were in the noncoding regions (40 of 58 SNPs). The remaining 18 SNPs were in the coding regions, with an average of 2 SNPs per gene (range: 0-6). The average SNP density ranged from 1 SNP per 184 bp in coding regions to 1 SNP per 104 bp in noncoding regions. Estimates of nucleotide diversity at each amplicon in open-pollinated selections (OPS) are summarised in Table 2. These estimates also indicate the level of polymorphisms for each gene estimated as nucleotide diversity (π) , based on the average number of nucleotide differences between sequences (Nei, 1978), and the weighted average for total nucleotide diversity (θ_{w}) , based on the number of segregating sites (WATTER-SON, 1975) were estimated for each gene, each provenance and also for all three provenances combined.

Nucleotide diversity was also heterogeneous across candidate genes. A likelihood-ratio test indicated that π (based on the average number of nucleotide differences between sequences (Nei, 1978)) varied significantly across genes (chi-square = 15.517 which is greater than the critical value of 3.8415, p = 0.0138 < 0.05), with the same result occurring for $\boldsymbol{\theta}_w$ (weighted average for total nucleotide diversity based on the number of segregating sites) (Watterson, 1975)) (chi-square = 16.123 which is greater than the critical value of 3.8415, p = 0.021 <0.05). As expected, $\boldsymbol{\theta}_w$ was larger than $\boldsymbol{\pi},$ at each population level (Table 3) due to the fact that the SNPs were sampled from genic regions, but also illustrating an excess of rare SNP alleles relative to expectations under neutrality. Nucleotide diversity (π) estimates for the 10 amplicons, ranged from 0.0016 to 0.0063, with a mean of 0.0036, while Watterson's $\theta~(\theta_w)$ estimates for the 10 amplicons ranged from 0.0046 to 0.0117 with a mean of .0.0065. Negative Tajima's D values appeared in the all amplicons, reflecting their low-frequency of SNP alleles, with four amplicons registering significant Tajima's D $(Table\ 2).$

Comparison of nucleotide diversity between native population material and OPS

Nucleotide diversity estimates for native population material and OPS by provenance are presented in *Table 3*. Nucleotide diversity for OPS was highest overall for the Monterey provenance and lowest for all statistics for Año Nuevo; however differences were not significant for π (P = 0.217). Diversity estimates in OPS were generally

lower than the corresponding estimates for the native material, except for π_{w} values that were higher for OPS than native material (Table 3). The mean values for expected heterozygosity (H_E) within and between native population material and OPS were similar (general mean = 0.27 ± 0.04) and not significantly different (P = 0.068). Estimates of H_O and F_{ST} (details not presented) within and between native population material and OPS did not show any significant statistical difference (t-test; P > 0.05). Also, the overall distribution of allele frequency classes was not significantly different between native population material and OPS (chisquare = 9.04, and 5.76, d.f = 7; P = 0.24 and 0.41, d.f = 7), respectively. When considering each SNP individually, none showed significant differences in allele frequencies between the native population material and OPS after correction for multiple testing using the false discovery rate (FDR) (STOREY and TIBSHIRANI, 2003) at a relaxed confidence level of Q < 0.10.

Discussion

Our native-population material represents effectively unselected germplasm collected from the wild ('Eldridge collection'), and it best reflects the marker variation in radiata pine. Nucleotide diversity estimated from eight selected genes involved in xylem cell wall development and plant water relations, suggest low levels of nucleotide diversity in both native-population material and open-pollinated selections (OPS). A plausible hypothesis would be the concomitant effects of the smaller population effective size combined to the existence of natural and/or artificial selection acting on the set of genes sequenced (e.g., Pot et al., 2005). This result is consistent with our previous knowledge regarding the native populations analysed, where marker diversity is low compared to functional diversity (e.g., BURDON et al., 1992b). The natural range of radiata pine is extremely small and fragmented (e.g., MORAN and Bell, 1987; Karhu et al., 2006) reflecting possible effects of genetic drift. Low nucleotide diversity might therefore be expected in radiata pine, at least within populations, under neutral model of evolution. For example, Dillon et al. (2013) identified some neutral SNP loci from a set of candidate genes used in this study.

Our estimates of nucleotide diversity for native population material ($\pi = 0.0043$) and for OPS ($\pi = 0.0036$)

were higher than that reported for the New Zealand breeding population (0.0019) (Pot et al., 2005). While our study and that of PoT et al. (2005) was each limited to a restricted set of genes, it appears nucleotide diversity in trees is generally low and variable. For example, various other nucleotide diversity estimates in trees suggested that they are highly variable, ranging from 0.0002 to 0.0107 (e.g., Brown et al., 2004; Krutovsky and Neale, 2005; Heuertz et al., 2006). This supports the contention that conifers are characterised by a low to intermediate levels of nucleotide diversity (e.g., SAVOLAINEN and PYHÄJARVI, 2007; PYHÄJARVI et al., 2007). We are not aware of any published studies that have estimated loss of nucleotide diversity from generation to generation in a domesticated tree species and therefore we are not able to contrast our results with comparable scenarios.

No significant nucleotide differences were detected between populations or between native material and OPS. That is expected due to the fact that the markers were sampled from genic regions. Another plausible reason for not detecting significant nucleotide differences could be due to selective constraints on the markers and small size of the marker used. For example, Brown et al. (2004) estimated the recombination rate in loblolly pine to be 2-3 cM/Mb. Thus there is a high chance that although we used 58 SNPs they are actually linked and should be considered as many as the number of genes from which the SNPs were selected from, i.e.,10 (see Table 2). Of interest is that Año Nuevo was marginally the lowest for all statistics (Table 3), which fits with the smallest size of the population (BURDON et al., 1997). Moreover, this population is known to have had nearextinctions and then recolonisations towards its margins (Burdon et al., 1997). Both native-population material and OPS had low levels of polymorphisms, as measured by expected heterozygosity ($H_{\rm E}$ = 0.27) (Table 3). Similarly, there were no statistically significant differences in expected heterozygosity between these two groups of populations. Not surprisingly, these low estimates are a reflection of the low total genetic diversity as measured by molecular markers in radiata pine compared to most other gymnosperms (LOVELESS and HAMRICK, 1984). However, other surveys using nuclear genetic markers have revealed moderate to high levels of genetic variability in natural and introduced populations of radiata pine in Australia. For example, levels of genetic diversity, measured with microsatellites, were similar to those of native populations (i.e. $H_{\scriptscriptstyle\rm E}$ = 0.78 for Australian plantations and 0.77 for native populations) (Devey et al., 2002; Bell et al., 2004; Karhu et al., 2006). Indeed, evidence for discrepancies in population differentiation estimates due to the limited number of loci and their sparse distribution on the genome has been reported (Nybom, 2004).

The significant negative Tajima's *D*-values observed at some genes may result from a past selection event on these genes, or may be a recent one (e.g., Pot et al., 2005; Dillon et al., 2010). Negative D indicates deleterious selection. A negative selection cause a low diversity in both population native and OPS. Thus the excess of rare frequency polymorphisms would be consistent with

a hitchhiking event in the radiata pine populations. Selection can be effective in altering gene frequencies if there is a strong correlation between the phenotype and the genotype, and more so if the character is affected by a small number of genes (FALCONER and MACKAY, 1996). However, most of commercial traits of interest are thought to be affected by a large number of genes (LYNCH and WALSH, 1998), and there is evidence that DBH, various branching properties, wood density in radiata pine are controlled by a large number of genes located on several linkage groups, each one having small genetic effects (WILCOX et al., 1997; DEVEY et al., 2004). For example, DEVEY et al. (2004) reported eight QTL positions for corewood density and two for DBH, with percent variance accounted for by the markers ranging from 0.8 to 3.6%. A similar pattern is also emerging for gene polymorphisms related to various wood-quality traits in radiata pine association genetic studies, with percent of phenotypic variance explained by individual marker loci being low (2 to 6.5%) (DILLON et al., 2010).

These results point to no evidence of loss of diversity in OPS due to artificial selection, but also deleterious selection. However, we realise that they are just from a single generation of non-intensive selection and need to be interpreted with caution. We do not envisage getting data for more than this generation anytime soon. No loss in diversity indicates the domestication of the OPS is at a very early stage. The OPS were selected as plus trees in the gene resource plantings. Perhaps nearly as important is the fact that a very high proportion of genetic diversity in conifers is found within populations and not distributed between them (e.g., Brown and MORAN, 1981). According to ADAMS (1981) and SZMIDT and MUONA (1985) the loss of genetic diversity in gymnosperms during the process of domestication has been relatively minor and may actually have increased if intra-population hybrids were produced within the gene resource stands. Ellstrand and Marshall (1985) emphasized that in a naturally outcrossing species maintained in outcrossing open-pollinated breeding populations, there will be little loss of genetic diversity under domestication. In marked contrast to agricultural crops such as barley, there has been a significant loss in genetic diversity during domestication from wild progenitors to modem cultivars (Brown and Clegg, 1983; MORRELL and CLEGG, 2007). This loss is a reflection of the fact that almost 50% of diversity is distributed among populations for inbred species (LOVELESS and Hamrick, 1984). For example, Naeem et al. (2011) reported genetic diversity among populations to be 40% compared to 60% within populations of Asian barley. However, maize provides an example of a domesticated form which maintained a high level of genetic diversity, although it is highly differentiated from its wild ancestor (Lefevre, 2004). There is also the consideration that the pollination system of maize represents, in itself, quite a strong outbreeding mechanism.

The native population material was sampled in one of the trials from a series that had many families, and at least the seed parents can be assumed unrelated (see Materials and methods). The trial series probably provides some of the best estimates of genetic parameters due to the large number of founder parents (GAPARE et al., 2012). In contrast, the OPS were based on a population which is a generation removed from the base and has been subject to both natural and artificial (phenotypic on growth and form) selection in New Zealand. The native-population material would be subject to neighbourhood-inbreeding effects, whereas land-race stocks are often free of such effects. Also, neither natural nor artificial selection would be able to achieve much truncation of genetic variation, given the likely heritability of the composite growth-and-form trait that would have been involved and the segregational variance should be essentially unaffected given the apparently polygenic inheritance. The OPS were part of the conservation plantings (GAPARE et al., 2011) with radiata pine hybrids readily produced in this wind-pollinated species. The provenance of origin at least on the female side is maintained, but on the male side is subject to contamination from surrounding plantings, but is mainly assumed to be of the same provenance. While better buffered from contamination (unquantified), the OPS are in one sense effectively an advanced insight into the genetic architecture of the next generation of conservation plantings.

Although the SNP information suggests similarities among mainland populations, results from quantitative genetic studies suggested provenance differences for growth, morphology, stem-form traits and disease resistance for mainland populations grown in Australia and New Zealand on a number of contrasting sites (Burdon et al., 1997; Johnson et al., 1997; Burdon et al., 1992a; Gapare et al., 2011; 2012). Genetic studies on radiata pine quantitative traits showed exceptionally high functional genetic diversity (genotypic coefficient of variation, a function of phenotypic coefficient of variation and heritability) (e.g., Burdon et al., 1997; Wu et al., 2007; GAPARE et al., 2012), despite the modest variability that has been observed for nucleotide polymorphisms. For example, provenance effects were highly significant for Dothstroma needle blight infection, with most severe infection in Cambria and least infection in Monterey (ADES and SIMPSON, 1991; BURDON et al., 1992a; GAPARE et al., 2011). The nucleotide diversity estimated in this study is not a measure of the functional diversity that we seek to maintain in the species which is needed to provide both continued gain and cope with new breeding goals that might be imposed by market shifts, climatic change and new biotic challenges. The DNA sequences analysed are associated with variation in wood properties (DILLON et al., 2013), those wood properties were not in the selection criteria for the OP material in the progeny test. There is evidence of abundant genetic variation in wood quality traits in radiata pine. For example, the genetic coefficient of variation for radiata pine wood density is 8% (e.g., Kumar et al., 2002; Gapare et al., 2012).

Conclusion

The SNPs that we identified may be of interest in future breeding programs for radiata pine, and they may exhibit a higher level of differentiation in future generations. While no significant loss in allelic diversity

was noted after one generation, subtle effects were nevertheless observed, implicating differentiation of allele frequencies at certain gene loci (DILLON et al., 2013). As for gene conservation, previous studies suggested that 30 to 70 individuals should suffice to ensure that the genes most influenced by selection (i.e., allele frequencies in the intermediate range, Namkoong et al., 2000) and of primary importance for genetic gain in the first five to ten generations would persist in the breeding population (JOHNSON et al., 2001). However, such a population size does not make it possible to maintain all very low-frequency alleles that might be important over the long term. This challenge may be overcome by conserving independent gene-resource populations. Determining the threshold at which genetic diversity levels will be significantly reduced presents an interesting approach that should allow breeders to make informed decisions regarding the intensity of selection in managed breeding populations as well as gene resource populations.

Acknowledgements

This research was jointly funded by the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Forest and Wood Products Australia (FWPA), Radiata Pine Breeding Coop (RPBC) and the Southern Tree Breeding Association (STBA). Thanks are extended to a large number of people who assisted at various stages of the work, backdating to 1978, especially the late Dr. K. G. ELDRIDGE. Special thanks to Drs. ROWLAND BURDON, JEREMY BRAWNER and SARO THAVAMANIKUMAR, for their insightful comments on earlier draft of this paper.

References

Adams, W. T. (1981): Population genetics and gene conservation in Pacific Northwest conifers. *In:* Scudder, G. G. E., Reveal, J. L. (eds) Evolution today, Proc 2nd Int Cong Syst Evol Biol. Hunt lnst Bot Document, Pittsburgh, pp 401–415.

ADES, P. K. and J. A. SIMPSON (1991): Variation in susceptibility to Dothistroma needle blight among provenances of *Pinus radiata* var. *radiata*. Silvae Genet **40**: 6–13.

AITKEN, S. N. (2000): Conserving adaptive variation in forest ecosystems. J Sustain For 10: 1–12.

Bell, J. C., M. Powell, M. Devey and G. F. Moran (2004): DNA profiling, pedigree lineage analysis and monitoring in the Australian breeding program of radiata pine. Silvae Genet. **53** (3): 130–134.

Brown, A. H. D. and M. T. Clegg (1983): Isozyme assessment of plant genetic resources. *In:* Isozymes: current topics in biological and medical research, vol 11. Alan R Liss, New York, pp 285–295.

Brown, A. H. D. and G. F. Moran (1981): Isozymes and the genetic resources of forest trees. *In*: Conkle, M. T. (ed) Isozymes of North American forest trees and forest insects. USDA, Berkeley, pp 1–10.

BROWN, G. R., G. P. GILL, R. J. KUNTZ, C. H. LANGLEY and D. B. NEALE (2004): Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc Nat Acad Sci USA 101: 15255–15260.

- Burdon, R. D. (1997): Genetic diversity for the future: Conservation or creation and capture? Pp. 237–246. *In:* Burdon, R. D., J. M. Moore (eds) IUFRO '97 Genetics of Radiata Pine: proceedings of NZFRI IUFRO Conference, December 1–4, and Workshop December 5 1997, Rotorua, NZ. New Zealand Forest Research Institute, FRI Bull. No. 203.
- Burdon, R. D., M. H. Bannister, H. A. I. Madgwick and C. A. Low (1992a): Genetic survey of *Pinus radiata*. 1. Introduction, description of experiment and basic methodology. NZ J For Sci 22: 119–137.
- Burdon, R. D., M. H. Bannister and C. A. Low (1992b): Genetic survey of *Pinus radiata*. 2: Population comparisons for growth rate, disease resistance, and morphology. N Z J For Sci 22: 138–159.
- Burdon, R. D., P. Broekhuizen and J. A. Zabkiewicz (1997): Comparison of native-population and New Zealand land-race samples of *Pinus radiata* using cortical oleoresin monoterpenes. Pp 50–56. *In:* Burdon, R. D., Moore, J. M. (eds) IUFRO '97 Genetics of Radiata Pine: proceedings of NZFRI IUFRO Conference, December 1-4, and Workshop December 5 1997, Rotorua, NZ. New Zealand Forest Research Institute, FRI Bull. No. 203.
- Burdon, R. D., M. J. Carson and C. J. A. Shelbourne (2008): Achievement in forest tree improvement in Australia and New Zealand *Pinus radiata* in New Zealand. Aust For **71**: 263–279.
- Burley, J. (2001): Genetics in sustainable forestry: the challenges for forest genetics and tree breeding in the new millennium. Can J For Res 31: 561–565.
- CHARLESWORTH, D. and J. H. WILLIS (2009): The genetics of inbreeding depression. Nature Rev Genet 10: 783–796.
- Devey, M. E., J. C. Bell, T. L. Uren and G. F. Moran (2002): A set of microsatellite markers for fingerprinting and breeding applications in *Pinus radiata*. Genom **45**: 984–989.
- Devey, M. E., S. D. Carson, M. F. Nolan, A. C. Matheson, C. T. Riini and J. Hohepa (2004): QTL associations for density and diameter in *Pinus radiata* and the potential for marker-aided selection. Theor Appl Genet 108: 516–524.
- DILLON, S. K., M. NOLAN, W. LI, C. BELL, H. X. Wu and S. G. SOUTHERTON (2010): Allelic variation in cell wall candidate genes affecting solid wood properties in natural populations and land races of *Pinus radiata*. Genet **185**: 1477–1487.
- DILLON, S. K., M. NOLAN, P. MATTER, J. BRAGG, W. J. GAPARE and S. G. SOUTHERTON (2013): Signatures of adaptation and genetic structure among the mainland populations of *Pinus radiata* D. Don inferred from SNP loci. Tree Genet. Genomes **9**: 1447–1463.
- DOYLE, J. J. and J. L. DOYLE (1990): Isolation of plant DNA from fresh tissue. Focus 12: 13–15.
- DUNGEY, H. S., J. T. BRAWNER, F. BURGER, M. CARSON, M. HENSON, P. JEFFERSON and A. C. MATHESON (2009): A new breeding strategy for Pinus radiata in New Zealand and New South Wales. Silvae Genet 58: 28–38.
- ELDRIDGE, K. G. (1978): Refreshing the genetic resources of radiata pine plantations In. Division of Forest Research: Genetics Section Report Number 7 CSIRO. 1-120.
- ELLSTRAND, N. C. and D. L. MARSHALL (1985): The impact of domestication on the distribution of allozyme variation within and among cultivars of radish, *Raphanus sativus* L. Theor Appl Genet **69**: 393–398.

- FALCONER, D. S. and T. F. C. MACKAY (1996): Introduction to Quantitative Genetics, 4th edn. Longman, London and New York.
- FAY, J. C. and C. I. Wu (2000): Hitchhiking under positive Darwinian selection. Genetics 155: 1405–1413.
- Gapare, W. J., B. S. Baltunis, M. Ivkovich, C. B. Low, P. Jefferson and H. X. Wu (2011): Performance differences among *ex situ* native-provenance collections of Pinus radiata D. Don. 1: potential for infusion into breeding populations in Australia and New Zealand. Tree Genet. Genomes 7: 409–419.
- Gapare, W. J., M. Ivkovich, G. W. Dutkowski, D. J. Spencer, P. Buxton and H. X. Wu (2012a): Genetic parameters and provenance variation of *Pinus radiata* D. Don. 'Eldridge collection' in Australia 1: growth and form traits. Tree Genet. Genomes 8: 391–407.
- Gapare, W. J., M. Ivkovich, S. K. Dollin, F. Chen, E. Evans and H. X. Wu (2012b): Genetic parameters and provenance variation of *Pinus radiata* D. Don. 'Eldridge collection' in Australia 2: wood properties Tree Genet. Genomes 8: 895–910.
- Godt, M. J. W., J. L. Hamrick, M. A. Edwards-Burke and J. H. Williams (2001): Comparisons of genetic diversity in white spruce (*Picea glauca*) and jack pine (*Pinus banksiana*) seed orchards with natural populations. Can J For Res 31: 943–949.
- Grattapaglia, D. and M. Kirst (2008): *Eucalyptus* applied genomics: from gene sequences to breeding tools. New Phyto **179**: 911–929.
- HARTL, D. L. and A. G. CLARK (1997): Principles of Population Genetics, 3rd edn. Sinauer Associates, Sunderland, Massachusetts, USA.
- HEUERTZ, M., E. DE PAOLI, T. KÄLLMAN, H. LARSSON, I. JURMAN and M. MORGANTE et al. (2006): Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway Spruce [*Picea abies* (L.) Karst]. Genet 174: 2095–2105.
- Jaramillo-Correa, J. P., J. Beaulieu and J. Bousquet (2001): Contrasting evolutionary forces driving population structure at ESTPs, allozymes and quantitative traits in white spruce. Mol Ecol 10: 2729–2740.
- JOHNSON, I. G., P. K. ADES and K. G. ELDRIDGE (1997): Growth of natural Californian provenances of *Pinus radiata* in New South Wales, Australia. N Z J For Sci **27**: 23–38.
- Johnson, R., B. St. Clair and S. Lipow (2001): Genetic conservation in applied tree breeding programs. *In:* Proceedings of the ITTO Conference on *In Situ* and *Ex Situ* Conservation of Commercial Tropical Trees, pp. 215–230. ITTO, Yokohama, Japan.
- Jones, T. H., D. A. Steane, R. C. Jones, D. Pilbeam, R. E. Vaillancourt and B. N. Potts (2006): Effects of domestication on genetic diversity in *Eucalyptus globules*. For Ecol Manage **234**: 78–84.
- KARHU, A., C. VOGL, G. F. MORAN, J. C. BELL and O. SAVOLAINEN (2006): Analysis of microsatellite variation in *Pinus radiata* reveals effects of genetic drift but no recent bottlenecks. J Evol Bio 19: 167–175.
- Kremer, A., V. Le Corre and S. Mariette (2000): Population differentiation for adaptive traits and their underlying loci in forest trees: theoretical predictions and experimental results. *In:* Matyas, C. (Ed.), Forest Genetics and Sustainability. For Sci **63**: 59–74.
- KRUTOVSKY, K. V. and D. B. NEALE (2005): Nucleotide diversity and linkage disequilibrium in cold-hardinessand wood quality-related candidate genes in Douglas fir. Genet 171: 2029–2041.

- Kumar, S., K. J. S. Jayawickrama, J. Lee and M. Lausberg (2002): Direct and indirect measures of stiffness and strength show high heritability in a wind-pollinated radiata pine progeny test in New Zealand. Silvae Genet **51**: 256–261.
- LANDE, R. (1995): Mutation and conservation. Conserv Bio 9: 782–791.
- LATTA, R. G. (1998): Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. Am Nat 151: 283–292.
- LEFEVRE, F. (2004): Human impacts on forest genetic resources in the temperate zone: an updated review. Forest Ecol Manage 197: 257–271.
- LOVELESS, M. D. and J. L. HAMRICK (1984): Ecological determinants of genetic structure in plant populations. Annu Rev Ecol Syst 15: 65–95.
- Lynch, M. (1995): A quantitative-genetic perspective on conservation issues. *In:* J. C. Avise and J. L. Hamrick (eds). Conservation genetics: case histories from nature. New York: Chapman & Hall: 471–501.
- Lynch, M. and B. Walsh (1998): Genetics and Analysis of Quantitative Traits. Sinauer Associates, Sunderland, MA. USA.
- MAF (2011): A National Exotic Forest Description as at 1 April 2011. Ministry of Agriculture and Forestry, Wellington, p. 64.
- MORAN, G. and C. Bell (1987): The origin and genetic diversity of *Pinus radiata* in Australia. Theo App Genet **73**: 616–622.
- MORRELL, P. L. and M. T. CLEGG (2007): Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. Proc Natl Acad Sci USA. 2007: **104**: 3289–3294.
- NAEEM, R., L. DAHLEEN and B. MIRZA (2011): Genetic differentiation and geographical relationship of Asian barley landraces using SSRs. Genet Mol Bio: **34**: 268–273.
- Namkoong, G. H., C. Kang and J. S. Brouard (1988): Tree Breeding: Principles and Strategies. Springer-Verlag, New York, NY, USA.
- Namkoong, G., M. P. Koshy and S. N. AITKEN (2000): Selection. *In:* A. Young, D. Boshier, and T. Boyle, eds. Forest Conservation Genetics: Principles and Practices, pp. 101–111. CABI Publishing, Wallingford, UK.
- NAMROUD, M. C., J. BOUSQUET, T. DOERKSEN and J. BEAULIEU (2012): Scanning SNPs from a large set of expressed genes to assess the impact of artificial selection on the undomesticated genetic diversity of white spruce. Evol App 5: 641–656.
- NEI, M. (1978): Estimation of average heterozygosity and genetic distance from a small number of individuals. Genet 89: 583–590.
- NEI, M., T. MARUYAMA and R. CHAKRABORTY (1975): The bottleneck effect and genetic variability in populations. Evol 29: 1–10.
- NYBOM, H. (2004): Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Molecular Ecology 13: 1143–1155.
- PEAKALL, R. and P. E. SMOUSE (2012): GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537–2539.

- Pelgas, B., J. Bousquet, P. G. Meirmans, K. Ritland and N. Isabel (2011): QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. BMC Genomics 12: 145.
- PFRENDER, M. E., K. SPITZE and J. HICKS (2000): Lack of concordance between genetic diversity estimates at the molecular and quantitative-trait levels. Conserv Genet 1: 263–269.
- Pot, D., L. McMillan, C. Echt, G. L. Provost, P. Garnier-Gere, S. Cato and C. Plomion (2005): Nucleotide variation in genes involved in wood formation in two pine species. Tree Physio 167: 101–112.
- Pyhäjärvi, T., M. Rosario García-Gil, T. Timo Knürr, M. Mikkonen, W. Wachowiak and O. Savolainen (2007): Demographic history as influenced nucleotide diversity in European *Pinus sylvestris* populations. Genetics 177: 1713–1724.
- R DEVELOPMENT CORE TEAM (2007): R: A Language and Environment for Statistical Computing. r Foundation for Statistical Computing, Vienna, Austria.
- RAYMOND, C. A. and M. Henson (2009): Genetic variation within the native provenances of *Pinus radiata* D. Don. 1. Growth and form to age 26 years. Silvae Genet **58**: 242–252.
- SAVOLAINEN, O. and T. PYHÄJÄRVI (2007): Genomic diversity in forest trees. Current Opinion Plant Biology 10: 162–167.
- Shelbourne, C. J. A., R. D. Burdon, S. D. Carson, A. Firth and T. G. Vincent (1986): Development Plan for Radiata Pine Breeding. New Zealand Forest Service, FRI Special Publication.
- Storey, J. D. and R. Tibshirani (2003): Statistical significance for genome-wide studies. Proc Nat Acad Sci USA **100**: 9440–9445.
- SZMIDT, A. E. and O. Muona (1985): Genetic effects of Scots pine (*Pinus sylvestris* L.) domestication. *In:* GREGORIUS, H. R. (ed) Population genetics in forestry. Springer, Berlin Heidelberg New York Tokyo, pp 241–252
- Watterson, G. A. (1975): On the number of segregating sites in genetical models without recombination. Theoret Pop Bio 7: 256–276.
- WILCOX, P. L., T. E. RICHARDSON and S. D. CARSON (1997):
 Nature of quantitative trait variation in *Pinus radiata*: insights from QTL detection experiments. *In*: BURDON,
 R. D., MOORE, J. M. (eds) Proceedings of IUFRO '97:
 Genetics of radiata pine, Rotorua, December 1997. FRI Bull 203, pp. 304–312.
- WILLIAMS, C. G. and J. L. HAMRICK (1995): Genetic diversity levels in an advanced generation *Pinus taeda* L. program measured using molecular markers. Forest Gen Resources News 23: 45–50.
- Wu, H. X., K. G. Eldridge, A. C. Matheson, M. B. Powell and T. A. McRae (2007): Achievements in forest tree improvement in Australia and New Zealand: successful introduction and breeding of radiata pine to Australia. Aust For 70: 215–225.
- Yanchuk, A. D. (2001): A quantitative framework for breeding and conservation of forest tree genetic resources in British Columbia. Can J For Res 31: 566-576.