

and structure of blue pine (*Pinus wallichiana* JACKSON) in Bhutan. For. Ecol. and Management **105**: 45–53 (1998). — MILLAR, C. I. and LIBBY, W. J.: Strategies for conserving clinal, ecotypic, and disjunct population diversity in widespread species. Genetics and conservation of rare plants. Ed. by FALK, D. A. and HOLSINGER, K. E. Oxford Univ. Press. 149–170 (1991). — MITTON, J. B.: The dynamic mating systems of conifers. New Forests **6**: 197–216 (1992). — MUONA, O.: Population genetics in forest tree improvement. Plant Population Genetics, Breeding and Genetic Resources. Ed. by BROWN, A. H. D., CLEGG, M. T., KAHLER, A. L. and WEIR, B. S. Sinauer Associates Inc., Sunderland, M.A. p. 282–298 (1990). — NEALE, D. B. and ADAMS, W. T.: Allozyme and mating-system variation in balsam fir (*Abies balsamea*) across a continuous elevational transect. Can. J. Bot. **63**: 2448–2453 (1985). — NEI, M.: F-statistics and analysis of gene diversity in subdivided populations. Ann. Hum. Genet. **41**: 225–233 (1977). — NEI, M.: Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**: 583–590 (1978). — POLITOV, D. V. and KRUTOVSKII, K. V.: Allozyme polymorphism, heterozygosity, and mating system of stone pines. International Workshop on Subalpine Stone Pines and their Environment: The Status of Our Knowledge. USDA Forest Service: Intermountain

Research Station. p. 36–42 (1994). — POLITOV, D. V., KRUTOVSKII, K. V. and ALTUKHOV, YU. P.: Genetic variability in Siberian stone pine, *Pinus sibirica* DU TOUR. III. Linkage relationships among isozymes loci. Genetika **25**: 1606–1618 (1989). — POTENKO, V. V. and VELIKOV, A. V.: Genetic diversity and differentiation of natural populations of *Pinus koraiensis* (SIEB. et ZUCC.) in Russia. Silvae Genetica **47**: 202–208 (1998). — PRICE, M. V. and WASER, N. M.: Pollen dispersal and optimal outcrossing in *Delphinium nelsoni*. Nature **277**: 294–297 (1979). — RITLAND, K.: A series of FORTRAN computer programs for estimating plant mating system. J. Heredity **81**: 235–237 (1990). — SHAW, D. V., KAHLER, A. L. and ALLARD, R. W.: A multilocus estimator of mating system parameters in plant populations. Proc. Nat. Acad. Sci. (USA) **78**: 1298–1302 (1981). — SWOFFORD, D. L. and SELANDER, R. B.: BIOSYS-1: A computer program for the analysis of allelic variation in population genetics and biochemical systematics. Release 1.7. Illinois Natural History Survey, IL (1989). — YEH, F. C. and LAYTON, C.: The organization of genetic variability in central and marginal populations of lodgepole pine *Pinus contorta* spp. *latifolia*. Can. J. Genet. Cytol. **21**: 487–503 (1979).

Genetic Variation in *Pinus banksiana* Populations From the Sudbury (Ontario, Canada) Region

By K. K. NKONGOLO¹) and W. S. GRATTON

Department of Biological Sciences, Laurentian University, Sudbury, Ontario, P3E 2C6, Canada

(Received 17th August 2000)

Summary

Genetic variation within and among jack pine (*Pinus banksiana*) populations from heavy metal – contaminated and uncontaminated areas in the Sudbury (Ontario, Canada) region was investigated using random amplified polymorphic markers (RAPDs). DNA samples from individual trees were analyzed using 15 oligonucleotides of random sequence. Eight of these primers allowed amplifications of random polymorphic (RAPD) loci. Close genetic relationships among jack pine populations from contaminated and uncontaminated sites were observed. Overall 26% of the scored loci were polymorphic. The analysis of molecular variance revealed that the within-group (among individuals) variations accounted for 63.3% of the total molecular variance. The difference among populations which explained 36.7% of total variation was statistically significant and higher than RAPD variations reported in *P. resinosa*. The jack pine trees were easily distinguished from red pine trees using RAPD markers. RAPD markers specific to *P. banksiana* were also identified.

Key words: RAPD, *Pinus*, DNA polymorphism, genetic diversity, heavy metals.

Introduction

The genus *Pinus* is among the most widely distributed and prominent genera of trees in the world and includes many of the most economically valuable species of forest trees (CRITCHFIELD and LITTLE, 1966; STRAUSS and DOERKSEN, 1990). A number of investigators have been studying range-wide, and regional variation, differences between mature, regenerating

and seed source plantations as well as relationships to environmental variables in jack and lodgepole pine (DANCIK and YEH, 1983; YEATMAN, 1984; KNOWLES, 1985; GORDON, 1994). Allelic heterogeneity tests have indicated no differences between mature, natural young stands and seed zone plantations. Several allozyme studies have reported existence of genetic variability in ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta* var *latifolia*), and jack pine (*P. banksiana*) populations (O'MALLEY et al., 1979; YEH and LAYTON, 1979). Similar analyses revealed that the levels of genetic variability were lower in jack and red pines compared to other pine species (DANCIK and YEH, 1983; MOSSELER et al., 1991). RAPD analysis of red pine populations from Newfoundland confirmed the low levels of genetic variability (MOSSELER et al., 1992).

On the other hand, there is great concern that air pollution and its conversion products alter the genetic structure of forest tree populations by processes which are assumed to have selective effects (SCHOLZ, 1986). Preliminary investigations indicate that environment markedly influences the relationship between biochemical variation and quantitative characters as well as the degree of expression of these characters. Several authors have reported various responses to air pollution in trees growing in contaminated areas. Enzymatic studies of Norway spruce revealed genetic differences between groups of sensitive trees in polluted areas (BERGMANN and SCHOLZ, 1987).

¹) Correspondence: K.K. NKONGOLO
Email: knkongolo@nickel.laurentian.ca

Higher heterozygosity in tolerant plants of European beech, scots pine, trembling aspen and red maple have been reported (FARRAR et al., 1977; MÜLLER-STARCK, 1985; BERRANG et al., 1986; and GEBUREK et al., 1987).

Environmental studies of mining pollution in Sudbury, Ontario, Canada have focussed in the past on the effects of fumigation and metal deposition. Reports provide information of landscape degradation, soil toxicity, acidification and plant metal accumulation, but knowledge of genetic changes in plants growing in these contaminated areas is limited. The main objective of this study was to estimate genetic variation within and among different *P. banksiana* populations from heavy metal contaminated and uncontaminated areas in the Sudbury region.

Materials and Methods

Site characterization

Jack pine cones were collected from two tailings and six separate sampling plots contaminated for more than 40 years with heavy metals and located within 15 km of smelters (Fig. 1). Two uncontaminated sites, each located approximately 100 km to 150 km North-West and North-East from Sudbury were used as reference sites. The levels of heavy metal accumulation in soil and vegetation were measured (GRATTON et al., 2000).

Tree core samples were collected using an increment borer from each individual site. Each individual core was air dried, mounted on a wooden backing, sanded to clarify ring structure and examined under binocular microscope to identify cross-datable ring sequence. Annual rings were cross dated using ring patterns methods (YAMAGUCHI, 1991).

Genetic materials

For each sampling site, ten trees were randomly selected and 50 cones per tree were collected. Five seedlings from each cone were used for the genetic analysis. Three additional accessions of *P. banksiana* from different provenances were provided by the Canadian Forest Service, Fredericton. This includes accessions 7333410, 702047, and 7010380 from Audrey Lake

(Ontario), Riviere Manicouagan (Québec), and Grand Lake (Ontario), respectively.

Seeds were placed in clear polycarbonate "Petawawa germination boxes" containing wet "Kimpak" cellulose paper and kept in a germinator at 25°C. About 250 ten-day-old seedlings per samples were collected and roots and seed debris were discarded. Seedlings were weighed, frozen in liquid nitrogen and stored at -80°C until use for DNA extraction.

DNA extraction and amplification of RAPD markers

Total DNA was isolated from individual seedlings as previously described (MOSSELER et al., 1992; NKONGOLO et al., 1999). For the three accessions from Audrey Lake (Ontario), Grand Lake (Ontario) and Manicouagan (Québec), DNA was isolated from bulks of several seedlings. DNA prepared from all the pine samples was used for PCR reactions. RAPD variations within and among populations were assessed. The PCR reactions were individually primed with different primers described in table 1 according to NKONGOLO et al. (1998). These oligonucleotide primers were synthesized by GIBCO BRL. In a 25- μ L volume, 100 ng to 200 ng plant DNA, 0.5 μ M primer and 50 μ M each of dATP, dCTP, dGTP, dTTP were mixed with 10 X reaction buffer (Perkin Elmer) and 0.625 units of Taq DNA polymerase (Perkin Elmer). Samples were amplified on a DNA thermal cycler (Perkin-Elmer). After an initial denaturation at 94°C, 42 cycles consisting of 45 s denaturation at 94°C, 60 s annealing at 55°C, and 90 s extension at 72°C were performed prior to a final extension of 5 min at 72°C, and subsequent cooling to 4°C. For analysis on 0.9% Gibco agarose gels in TAE buffer, 7 μ L of the PCR reaction was mixed with 3 μ L 1 X gel loading buffer (MANIATIS et al., 1989).

RAPD analysis

RAPD assays of each population were performed at least three times each. Only reproducible amplified fragments were scored. For each sample, the presence and absence of fragments were recorded as 1 or 0, respectively, and treated as discrete characters. Pairwise comparison of banding patterns and JACCARD's similarity coefficients were evaluated using

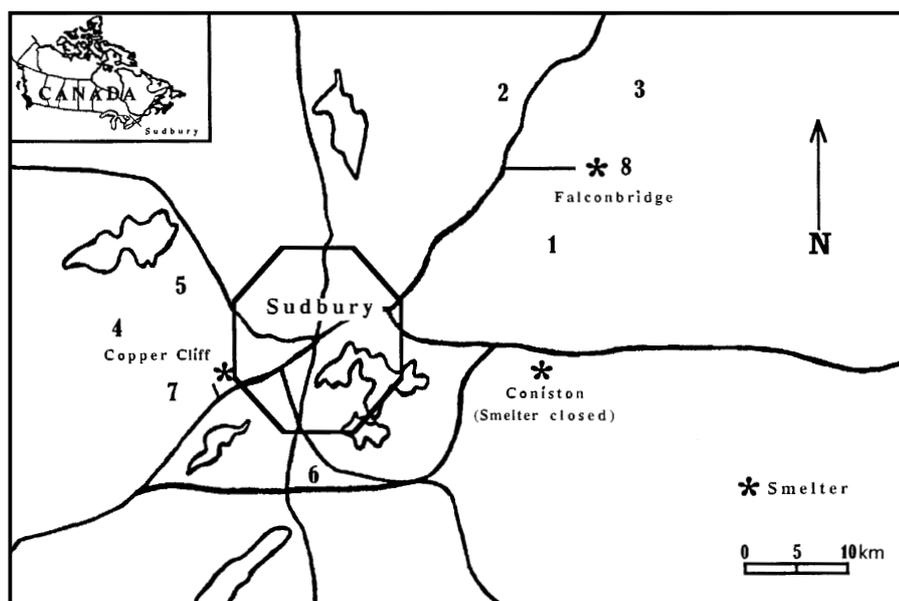


Fig. 1. – Location of Jack pine (*Pinus banksiana*) sampling sites (1 to 8) within the Sudbury region. Uncontaminated sites (9 and 10) were located outside the Sudbury area.

RAPDistance Programs, version 1.04 (ARMSTRONG et al., 1994). The analysis of molecular variance (AMOVA) was applied to estimate variance components for RAPD phenotypes. Variation was partitioned among individuals (within regions) and between regions. Levels of significance were determined using nonparametric, permutational methods with WINAMOVA program (EXCOFFIER et al., 1992). Amplified products from DNA from the three accessions from Audrey Lake (Ontario), Grand Lake (Ontario), and Riviere Manicouagan (Québec) were not included in the AMOVA since they were from bulks of several seedlings. These products were used only to confirm the presence of RAPD markers specific to jack pine.

Results

The levels of heavy metal accumulation in soil and plants from the ten contaminated and uncontaminated sites surveyed in the Sudbury region have been reported (GRATTON et al., 2000). Higher concentration of copper, nickel, lead and cadmium were measured in vegetation within 15 km of smelters (GRATTON et al., 2000). Core analysis revealed that populations located around smelting operations in Sudbury were 28 years to 55 years old. Trees from INCO Ltd. tailing and Falconbridge Ltd. property were much younger. The trees at these sites were transplanted. Trees from Temagami and Low water lake control sites were 38 years and 82 years old, respectively. Analysis of ring pattern of cores collected around smelter operation revealed three distinct sensitive ring patterns specially during the 1960 and early 1970's. Narrow rings were observed on most of the sampling cores from Sudbury for 1961 to 1962, 1966 to 1967, and 1971 periods.

Seven of the ten populations surveyed were selected based on the ecological differences for molecular analysis. DNAs extracted from trees growing in five contaminated sites (1, 3, 5, 6, and 7) and two control sites (9 and 10) were analyzed to estimate RAPD variations within and among different *P. banksiana* populations (Tables 1 and 2). Conditions have been optimized to allow reproducible amplification of RAPD amplification as described by NKONGOLO et al. (1998). Reactions parameters including MgCl₂ concentrations and temperatures significantly influenced yield and the types of amplification product synthesized. In the present study all the primers used have a GC content > 50% and the optimal annealing temperature was 55°C. The best RAPD products were obtained with 42 cycles.

Eight of the 15 primers evaluated revealed multibanded fingerprints clearly scorable (Figure 2). A total of 58 RAPD markers varying in size between 220 bp and 2800 bp were generated. The total number of polymorphic loci detected varied between primers (Table 1). Overall 26% of the scored loci were polymorphic. The between-population variance contributed 36.70% of the total molecular variance; the within-population variance accounts for 63.30%. Using a nonparametric test with 1000 permuted matrices, we found that the between-population difference was significant (Table 2). No single locus appears to be specific to contaminated sites.

The genetic distance scale runs from 0 (identical) to 1 (different for all criteria). The genetic distance estimates are summarized in table 3. In the present study, the genetic distance values varied from 0.02 to 0.71 (Table 3). Populations from sites 5 and 6 around INCO smelters were very similar. Trees from INCO tailing (site 7) and Control site Cartier (site 10), were also genetically close. Populations from site 1 around Falconbridge smelter and population from site 5 around INCO smelter were genetically distant.



Fig. 2. – RAPD amplification patterns with primer P-9. a) Lane 1 contains 1-kbp DNA ladder, lanes 2-8 contain amplified products from individual jack pine (*Pinus banksiana*) trees from different sites in Sudbury; lane 9 contains amplified products from red pines (*P. resinosa*) used as reference. Arrows indicate RAPD markers specific to jack pine and arrowhead shows a RAPD marker specific to red pine. Note the low level of RAPD variation among jack pine samples.

Table 1. – Attributes of random oligonucleotide primers used for generating RAPD markers from trees sampled from seven populations of jack pine (*Pinus banksiana*).

Primer identification	Nucleotide sequence (5' to 3')	No. of fragments amplified	Fragment size range (bp)	No. of polymorphic fragment scored
1	ACGACGTAGG	0	–	--
2	CCGGCCTTCC	4	745-1265	0
3	CCGGCTGGAA	8	300-2036	0
4	GAGGGCGTGA	0	–	--
5	GCTCCCCAC	0	–	--
6	CGATGGCTTT	0	–	--
7	TAGCCCGCTT	0	–	--
8	GTAGACGAGC	8	220-1220	2
9	GTGCGTCCTC	6	395-1370	2
10	GTTCTCGTGT	0	–	--
11	AACACACGAG	0	–	--
23	CCCGCCTTCC	7	400-1650	2
63	TTCCCGCCCC	9	300-2800	4
184	CAAACGGCAC	7	350-1542	3
146	ATGTGTTGCG	9	500-1650	2

Table 2. – Analysis of molecular variance (AMOVA) for RAPD variation for 70 individuals of *Pinus banksiana* from metal contaminated and uncontaminated areas in the Sudbury region.

Source of variation	df	MS	Variance component	% Total	P-value
Populations within region	6	0.756	0.0645	36.70	<0,001
Individuals within populations	63	0,111	0,1113	63,30	<0,001

Discussion

Loss of rare alleles, lower heterozygosity and directional selection have been concerns of plant populations subjected to air pollution (BERGMANN and SCHOLZ, 1989). Most of the forest ecosystems within the Sudbury region have been destroyed by logging, erosion, and air pollution. Although there were significant differences among sites analyzed for the accumulation of heavy metals in soil and foliage, environmental conditions within the Sudbury area have improved considerably during the last 30 years (DUDKA et al., 1995; GRATTON et al., 2000). Vascular and nonvascular plants such as conifers, birches and lichens have re-invaded semi-barren landscapes.

In the present study, jack pine populations from contaminated and uncontaminated sites were compared using RAPD markers. RAPD variation among jack pine trees from seven different populations were observed. Fifteen polymorphic loci were reliably amplified with 8 primers. RAPD analysis showed low population differentiation. No population-specific RAPD markers were identified. Most of the molecular variance resulted from within-population differences. This low genetic differentiation among jack pine populations contrasts with ecological differences among sites. The seven populations analyzed were from reforested sites and natural stands and were located up to 200 km apart. The accumulation of heavy metals and distribution and composition of vegetation were significantly different among the sites (GRATTON et al., 2000). The ring patterns from trees from contaminated sites were distinct from trees from uncontaminated sites. These reduced tree ring widths during 1960's and early 1970's are consistent with elevated sulfur dioxide concentrations.

This low level of genetic variability appears to be a common characteristics of jack pine populations as the polymorphism observed in populations from other regions in Ontario, Quebec, and New Brunswick was also limited (NKONGOLO, unpublished data). Overall the level of RAPD variations in jack pine populations were much lower than those observed by WU et al. (1998) in California pines, *P. muricata*, *P. attenuata*, and *P. radiata*. The level of polymorphism across population within each of these species varied from 70% to 85% whereas in the present study, only 26% of loci observed in *P. banksiana* populations were polymorphic. Moreover, in the California pines, populations within each species showed high differentiation from one another (WU et al., 1998). The level of polymorphism observed in *P. banksiana* is however relatively higher than in *P. resinosa* which is the least variable of all the pines species investigated (MOSSELER et al., 1992). This low level of genetic variation detected with RAPD markers is consistent with previous allozymes analyses. MOSSELER et al. (1991), DANCİK and YEH (1983), and SIMON et al. (1986) reported that the level of genetic variability in jack and red pines detected by allozyme analysis was much lower compared to other pine species.

Although, both isozymes and RAPD allow the analysis of genetic variability in plant species, fundamental differences exist between these two methods. Isozyme analysis reflects alterations in the DNA sequence through changes in amino acid composition (HAMRICK, 1989). These changes will often alter the protein charge thereby producing a change in electrophoretic mobility which is useful in evaluating levels of variation between individuals and populations on the basis of gene loci coding for specific enzymes (WEEDEN and WENDEL, 1989).

Table 3. – Distance matrix generated from pine RAPD data used in neighbour-joining analysis (RAP-Distance version 1.04).

Populations ⁺	1	3	5	6	7	9	10
10	0,30	0,20	0,30	0,18	0,09	0,24	0,00
9	0,57	0,43	0,53	0,42	0,43	0,00	
7	0,13	0,24	0,53	0,34	0,00		
6	0,48	0,29	0,02	0,00			
5	0,71	0,38	0,00				
3	0,50	0,00					
1	0,00						

⁺) Populations 1 and 3 located around Falconbridge Ltd. smelting operations; populations 5 and 6 located around INCO Ltd. smelting operations; population 7 is from a INCO Ltd tailing; populations 9 and 10 are from control sites in Tamegami and Low water lake near Cartier, respectively.

Although isozyme markers have been extremely informative in population genetic studies, the limited number of isozyme (approximately 30) reflects only a small number of the genome.

On the other hand, RAPD is a DNA based molecular marker. Genomic DNA is amplified using randomly constructed oligonucleotides as primers. Unlike isozymes, RAPD is relatively easy to apply and the number of loci that can be examined is essentially unlimited. Since the primers consisted of random sequences, and do not discriminate between coding and non-coding regions, the technique samples the genome more randomly than conventional methods (LYNCH and MILLIGAN, 1994). RAPD results in amplification of specific portions of template DNA that binds DNA primers while isozymes involve the differentiation of amino acid sequences caused by alteration in DNA.

Microsatellite markers which are co-dominant and more polymorphic than RAPD markers could be more informative in revealing differences among populations (ECHT et al., 1998 and HICKS et al., 1998). However, the level of genetic differentiation found in red pine using chloroplast microsatellites was lower than generally reported for other pines supporting RAPD data (ECHT et al., 1998; MORAN et al., 1988; and STRAUSS et al., 1993). Moreover, genetic diversity among red pine populations detected with microsatellites was also limited (ECHT et al., 1998).

One of the plausible explanations of this low level of genetic variability among jack pine populations is a possible bottleneck due to the last glaciation. The entire area of the present-day distribution of jack pine is thought to have been covered by ice during the last glacial stages. Geological and paleobotanical evidences from fossil pollen depositions indicate that jack pine survived glaciation in an extensive refugium centred on the Appalachian Highlands of Eastern North America (YEATMAN, 1967). Upon recession of the Wisconsin icecap, jack pine northward is thought to have happened rapidly (DAUBENMIRE, 1978).

Evolution of jack pine in North America could have possibly followed the same pattern of red pine which despite increases in populations numbers and mutations, have not produced much detectable genetic variation specially at nuclear level (MOSSELER et al., 1992; DE Verno et al., 1997; and ECHT et al., 1998). Self-pollination and the lack of continuous forest may also have contributed to a loss of genetic variation through inbreeding in small populations such as those found in Sudbury.

Despite this low level of variability within and among jack pine populations, genetic distances were estimated. Close genetic relationships among the jack pine populations from contaminated and uncontaminated sites were observed. DNA from trees from sites 5 and 6 around INCO showed high similarity in amplification pattern. This suggests that they may have a common origin. The jack pine population from the uncontaminated site located 100 km North-East in Cartier (site 10) was closely related to all the populations surveyed. Trees from this uncontaminated site 10 showed much closer relationship with those from the INCO tailing (site 7) despite the big age difference between them. This suggests that jack pine seeds used for revegetation of this INCO tailing are from a population genetically similar to population from site 10.

In conclusion, despite the large ecological difference among the heavy metal-contaminated and uncontaminated sites, and clear effect of pollution on trees from contaminated sites, the seven *P. banksiana* populations analyzed showed low differentiation from one another. Nevertheless, the among-population difference was significant and higher than RAPD variations reported in *P. resinosa*.

This study confirms data from other reports (NKONGOLO, 1999; O'KEEFE et al., 1998; WU et al., 1998) which indicated that RAPD's markers are suitable for first estimates on genetic variation of conifer populations. Reproducibility and the dominance nature of the markers could limit their application in population genetics. If a locus is polymorphic, with the recessive allele present at low frequency, almost all copies of that allele will be carried in heterozygotes and will go undetected by the RAPD assay, resulting in that locus being scored as monomorphic. RFLP analysis would be useful to compensate this loss of information since RFLP markers are codominant. But our current study using several pine species (unpublished data) confirms previous reports which indicated that RFLP yields similar frequencies of polymorphism to that obtained with RAPDs in different plant species (HALLDEN et al., 1994; DOS SANTOS et al., 1994). Most recently, simple sequence repeats (SSRs) also known as microsatellites have been used to identify polymorphism in pine (HICKS et al., 1998). But SSR marker development is labourious and present some serious constraints. All the SSR markers used to characterize plant genotypes should be unlinked. Thus a larger number of SSRs must be identified, sequenced, and evaluated.

Acknowledgements

Technical assistance of Mr. LUC TAILLEFER, Mr. DANNY and Mr. NORMAND GRATTON with sample collections is appreciated. Support for this study was received from Laurentian University Research Fund (Project 202274).

Literature

- ARMSTRONG, J. S., GIBBS, A. J., PEAKALL, R. and WEILLER, G.: The RAP-Distance Package. [Http://life.anu.edu.au/molecular/software/rapd.html](http://life.anu.edu.au/molecular/software/rapd.html) (1994). — BERGMANN, F. and SCHOLZ, F.: The impact of air pollution on the genetic structure of Norway spruce. *Silvae Genet.* **36**, 80–83 (1987). — BERGMANN, F. and SCHOLZ, F.: Selection effects of air pollution in Norway spruce (*Picea abies*) populations. In: SHOLZ, H. R. GREGORIUS, and D. RUDIN (Editors). Genetic effects of air pollutants in forest tree populations. Springer-Verlag, Berlin. pp. 143–160 (1989). — BERRANG, P., KARNOSKY, D. F., MICKLER, R. A. and BENNETT, J. P.: Natural selection for ozone tolerance in *Populus tremuloides*. *Can. J. For. Res.* **16**: 1214–1216 (1986). — CRITCHFIELD, W. B. and LITTLE, E. L.: Geographic distribution of the pines of the world. U.S. Dept. Agric., Misc. Publ. No. 991, Washington, D.C. (1966). — DANCİK, B. P. and YEH, F. C.: Allozyme variability and evolution of lodgepole pine (*Pinus contorta* var. *latifolia*) and jack pine (*P. banksiana*) in Alberta. *Can. J. For. Res.* **25**: 57–64 (1983). — DAUBENMIRE, R.: Plant geography. Academic Press, New York. pp. 338 (1978). — DE Verno, L. and MOSSELER, A.: Genetic variation in red pine (*Pinus resinosa* AIT.) revealed by RAPD and RAPD/RFLP analysis. *Can. J. For. Res.* **27**: 1316–1320 (1997). — DOS SANTOS, J. B., NIENHUIS, J., SKROCH, P., TVANG, J. and SLOCUM, M. K.: Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. *Theor. Appl. Genet.* **87**: 909–915 (1994). — DUDKA, S., PONCE-HERNANDEZ, R. and HUTCHINSON, T. C.: Current levels of total element concentrations in the surface layers of Sudbury's soils. *Sci. Total Environ.* **162**: 161–171 (1995). — ECHT, C. S., DE Verno, L. L., ANZIDE, M. and VEDRAMIN, G. G.: Chloroplast microsatellites reveal population genetic diversity in red pine, *Pinus resinosa* AIT. *Mol. Ecol.* **7**, 307–316 (1998). — EXCOFFIER, L., SMOUSE, P. E. and QUATTRO, J. M.: Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction sites. *Genetics* **131**: 479–491 (1992). — FARRAR, J. F., RELTON, J. and RUTTER, A. J.: Sulphur dioxide and the scarcity of *Pinus sylvestris* in the industrial pennines. *Environ. Pollut.* **14**: 63–68. GEBUREK, T., SCHOLZ, F., KNABE, W. and VORNWEG, A.: Genetic studies by isozymes gene loci on tolerance and sensitivity in air polluted *Pinus sylvestris* field trial. *Silvae Genetica* **36**: 49–53 (1987). — GORDON, A. G.: The interface of population genetics, ecology and silviculture: an overview for the boreal forest. Proc. of the 1994 Society of American Foresters/Canadian Institute of Forestry Convention, September 18 to 22, Alaska (1994). — GRATTON, W. S., NKONGOLO, K. K. and Spiers, G. A.: Heavy metal accumulation in soil and jack pine (*Pinus banksiana*) needles in Sudbury, Ontario. *Bull. of Environ. Contam. and Toxicol.* **64**: 550–557 (2000). — HALLDEN, C., NILSSON, N. O., RADING, I. M. and SALL, T.: Evaluation of RFLP and RAPD markers in comparison of *Brassica napus* breeding lines. *Theor. Appl. Genet.* **88**: 123–128 (1994). —

HAMRICK, J. L.: Isozymes and the analysis of genetic structure in plant populations. *In: Isozymes in plant biology. Edited by SOLTIS D. E. and P. S. SOLTIS.* Dioscorides Press, Portland, Oregon. pp. 87–105 (1989). — HICKS, M., ADAMS, D., O'KEEFE, S., MACDONALD, E. and HODGETTS, R.: The development of RAPD and microsatellite markers in lodgepole pine (*Pinus contorta* var. *latifolia*). *Genome* **41**: 797–805 (1998). — KNOWLES, P.: Comparison of isozyme variation among natural stands and plantations: jack pine and black spruce. *Can. J. For. Res.* **15**, 902–908 (1985). — LYNCH, M. and MILLIGAN, B. G.: Analysis of population genetic structure with RAPD markers. *Mol. Ecol.* **3**: 91–99 (1994). — MANIATIS, T., FITSCH, E. F. and SAMBROOK, J.: *Molecular Cloning: A laboratory Manual.* 2nd edition, Vol. I, II, III. Cold Spring Harbor Laboratory, Long Island, N.Y. (1989). — MORAN, G. F., BELL, J. C. and ELDRIDGE, K. G.: The genetic structure and the conservation of the five natural populations of *Pinus radiata*. *Can. J. For. Res.* **18**: 506–514 (1988). — MOSSELER, A., EGGER, K. N. and HUGHES, G. A.: Low levels of genetic diversity in red pine confirmed by random amplified polymorphic DNA markers. *Can. J. For. Res.* **22**: 1332–1337 (1992). — MOSSELER, A., INNES, D. J. and ROBERTS, B. A.: Lack of allozymic variation in disjunct Newfoundland populations of red pine (*Pinus resinosa*). *Can. J. For. Res.* **21**: 525–528 (1991). — MÜLLER-STARCK, G.: Genetic differences between tolerant and sensitive beeches (*Fagus sylvatica* L.) in an environmentally stressed adult forest stand. *Silvae Genet.* **34**: 241–246 (1985). — NKONGOLO, K. K.: RAPD and cytological analyses of *Picea* spp. from different provenances: genomic relationships among taxa. *Hereditas* **130**: 137–144 (1999). — NKONGOLO, K. K., KLIMASZEWASKA, K. and GRATTON, W. S.: DNA yields and optimization of RAPD patterns using spruce embryogen-

ic lines, seedlings and needles. *Plant Mol. Biol. Rep.* **16**: 285–384 (1998). — O'MALLEY, D. M., ALLENDORF, F. W. and BLAKE, G. M.: Inheritance of isozyme variation and heterozygosity in *Pinus ponderosa*. *Biochem. Genet.* **17**: 233–250 (1979). — SHOLZ, F.: Genetic effects of air pollutants. *Proc. 18th. IUFRO World Cong., Ljubljana, Vol. I*, 286–294 (1986). — SIMON, J.-P., BERGERON, Y. and GAGNON, D.: Isozyme uniformity in populations of red pine (*Pinus resinosa*) in the Abitibi region, Quebec. *Can. J. For. Res.* **16**: 1133–1135 (1986). — STRAUSS, S. H. and DOERKSEN, A. H.: Restriction fragment analysis of pine phylogeny. *Evolution* **44**: 1081–1096 (1990). — STRAUSS, S. H., HONG, Y. P. and HIPKINS, V. D.: High levels of population differentiation for mitochondrial DNA haplotypes in *Pinus radiata*, *P. miricata*, and *P. attenuata*. *Theor. Appl. Genet.* **86**, 605–611 (1993). — WEEDEN, N. F. and WENDEL, J. F.: Genetics of plant isozymes. *In: Isozymes in plant biology. Edited by D. E. SOLTIS and SOLTIS, P. S.* Dioscorides Press, Portland, Oregon. pp. 46–72 (1989). — WU, J., KRUTOVSKII, K. and STRAUSS, S. H.: Nuclear DNA diversity, population differentiation, and phylogenetic relationships in the California closed-cone pines based on RAPD and allozyme markers. *Genome* **42**: 893–908 (1998). — YAMAGUCHI, D. K.: A simple method for cross-dating increment cores from living trees. *Can. J. For. Res.* **21**: 414–416 (1991). — YEATMAN, C. W.: Biogeography of Jack pine. *Can. J. Bot.* **67**: 2201–2211 (1967). — YEATMAN, C. W.: P. N. F. I. Genetics and breeding: Genetics of jack pine: 1982–1983. *Proc. 19th C. T. I. A. Mtg., Toronto, Ont., Aug. 22 to 26, 1983.* Pp. 138–140 (1984). — YEH, F. C. and LAYTON, C.: The organization of genetic variability in central and marginal populations of lodgepole pine *Pinus contorta* var. *latifolia*. *Can. J. Genet. Cytol.* **21**: 2201–2211 (1979).

A Case Study of a Provenance/Progeny Test Using Trend Analysis with Correlated Errors and SAS PROC MIXED

By C. SAENZ-ROMERO^{1,3}), E. V. NORDHEIM²), R. P. GURIES and P. M. CRUMP⁴)

Department of Forest Ecology and Management, University of Wisconsin – Madison, 120 Russell labs, 1630 Linden Drive, Madison, WI 53706, U.S.A.

(Received 30th October 2000)

Abstract

A statistical analysis was performed on data from a provenance test of *Pinus banksiana* LAMB. 1.5 year-old open-pollinated seedlings from 47 stands collected in five distinct soil-moisture classes within a pine-barren landscape, in west-central Wisconsin. A large-scale environmental gradient across the nursery bed and small-scale among-microsite variability were evident in seedling responses. We compared three analyses in terms of their capabilities for accounting for such within-experimental area variation: (A) a “standard analysis” using ANOVA for a randomized complete block design, (B) trend analysis in addition to (A), and (C) correlated errors in addition to (B). PROC GLM of the Statistical Analysis System (SAS) was used for analyses (A) and (B); PROC MIXED was necessary for analysis (C). We concluded that analysis (C) was the best option for adequately modeling the data, reducing the error variance and consequently, detecting significant differences among sets of stands grouped by soil-moisture classes. We suggest that the row and column position of each experimental unit in provenance or progeny tests with forest species be recorded in order to allow for the possibility of conducting analysis of this type. As an important caveat, we found an unexpected bimodal likelihood surface when PROC MIXED included a term for correlation among plots; this requires use of the PARMs statement when applying PROC MIXED.

Key words: Trend analysis, correlated errors, spatial autocorrelation, PROC MIXED, provenance test, progeny test, *Pinus banksiana*, genetic variation, within block variability.

Introduction

A typical experimental design for provenance and progeny testing with forest trees is the randomized complete block design (RCBD) (WRIGHT, 1976; ZOBEL and TALBERT, 1984; GUMPertz and BROWNIE, 1993), with each entry (entries can be provenances, progenies, clones or ramets) appearing in one plot per block and multiple trees (seedlings) per entry per plot (usually 4 to 10 trees per plot) (BRIDGWATER *et al.*, 1983). This model can be written:

$$[1] \quad Y_{ijk} = \mu + \beta_i + \tau_j + \beta_i \times \tau_j + \varepsilon_{ijk}$$

where Y_{ijk} = observation on the k^{th} seedling of the j^{th} entry in the i^{th} block, μ = overall mean, β_i = effect of the i^{th} block, τ_j = effect of j^{th} genetic entry, $\beta_i \times \tau_j$ = interaction of block by entry (also known as plot error), and ε_{ijk} = within plot error term, $i=1, \dots, b$, $j=1, \dots, t$, and $k=1, \dots, n_{ij}$, where b , t and n_{ij} are the

¹) Present address: Instituto de Investigaciones sobre los Recursos Naturales, Universidad Michoacana de San Nicolás de Hidalgo, Apartado Postal 2-105, 58041 Morelia Michoacán, México.

²) Also Department of Statistics, University of Wisconsin – Madison.

³) Corresponding author: Phone/fax +(52)(4)327-2350 and 51, csaenz@zeus.ccu.umich.mx

⁴) Department of Computing and Biometry, College of Agriculture and Life Sciences, University of Wisconsin – Madison, 1675 Observatory Drive, Madison, WI 53706, U.S.A.