GENETIC LINKAGE ANALYSIS IN *SALIX ERIOCEPHALA* SUGGESTS LINKAGE CONSERVATION ABOVE THE SPECIES LEVEL

F. A. Aravanopoulos

Faculty of Forestry, University of Toronto, Toronto, Canada, and
Laboratory of Forest Genetics and Tree Breeding, Department of Forestry and Natural Environment, Aristotle University of Thessaloniki, GR–54006 Thessaloniki, Greece, phone: +30-31-992778; e-mail: aravanop@for.auth.gr

Received February 20, 2000; accepted April 2, 2001

ABSTRACT

Genetic linkage was investigated in eleven isoenzymic loci in *Salix eriocephala* Michxaux progenies that conformed to a 2 × 2 balanced factorial mating design. Joint segregation was tested in 49 two-locus combinations over different families and in a total of 107 two-locus tests. Linkage investigations involved contingency table χ^2 tests, while recombination fractions (θ) and associated standard errors were estimated with maximum likelihood methods. One possible linkage group involving *Aco-2, Ppo-1* and *Cto-1* was found. Positive interference was inferred. When results were compared to those reported in other willow species, it was suggested that linkage conservation exists above the species level. The merits and shortcomings of isoenzyme markers in gene mapping are briefly discussed.

Key words: *Salix eriocephala*, linkage, interference, linkage conservation, isoenzymes

INTRODUCTION

Information of linkage relationships among genetic markers is very important in plant genetics and breeding. Such information assists in answering questions pertaining to the organisation and establishment of genetic variation in natural populations (ZOUROS *et al.* 1974), is essential for population models dealing with the maintenance of balanced polymorphism through epistatic selection (LEWONTIN 1974), permits the assessment of the distribution of available loci in the genome, and provides references to locate unmapped genes on chromosomes. Studies on the genetic structure of any population should include linkage analysis (RUDIN 1986). The usefulness of identifying instances of joint segregation is also paramount in breeding. Knowledge of the positions of marker loci can permit the study and manipulation of important genes either single or of quantitative nature, within or among species (TANKSLEY 1993). Linkage maps are increasingly being used in evolutionary biology (MITCHELL-OLDS 1995), and have become important in taxonomy and phylogenetic analyses through the study of linkage conservation above the species level.

However, information regarding linkage relationships in forest angiosperms is still scarce. Linkage studies in forest trees were focused almost exclusively on conifers. Today the genomes of *Pinus* and *Picea* are becoming relatively adequately mapped (BUCCI *et al.* 1997, SEWELL *et al.* 1999, YAZDANI *et al.* 1995). In forest woody angiosperms, few linkage groups have been identified, for example in *Betula pendula* (HATTEMER *et al.* 1990), in *Prunus avium* (SANTI & LEMOINE 1990), in *Liriodendron* spp. (PARKS *et al.* 1990), and in *Liquidambar* spp. (HOEY & PARKS 1990). *Eucalyptus* spp. present a notable exception (GRATTAPAGLIA & SEDEROFF 1994). Limited information exists in *Salicaceae*. ARAVANOPOULOS (1998) and THORSEN *et al.* (1997) recently reported the presence of two linkage groups in *Salix exigua* and four possible linkage groups in *S. viminalis* respectively. In poplars only the hybrid *Populus trichocarpa* × *deltoides* has been adequately mapped (BRADSHAW *et al.* 1994). There is one report on linkage, regarding poplars of the *Tacamahaca* section, where three linkage groups were detected (MÜLLER-STARK 1992). Analyses of joint segregation by MALVOLTI *et al.* (1991) in *Populus deltoides*, RAJORA (1990) in *P. nigra* and *P. maximoviczii*, and HYUN *et al.* (1987) in *P. tremuloides*, did not lead in the identification of linkage groups.

This study investigates the linkage relationships among isoenzymic loci in *Salix eriocephala* Michx. (section *Cordatae*, subgenus *Vetrix*) a species important in short rotation intensive culture biomass plantations (ZSUFFA & ARAVANOPOULOS 1989). The genetic basis of enzymatic variation in this species has already been reported (ARAVANOPOULOS *et al.* 1994).

MATERIALS AND METHODS

Inheritance analysis in *S. eriocephala* revealed 11 segregating loci: *Acp–2, Aco–2, Adh–2, Alp–1, Cto–1, Idh–2,*

Per–1, Pgd–1, Pgi–2, Ppo–1, and *Sdh–2,* in 11 enzyme systems: acid phosphatase, aconitase, alcohol dehydrogenase, alkaline phosphatase, cytochrome oxidase, isocitrate dehydrogenase, peroxidase, 6-phosphogluconate dehydrogenase, phosphoglucose isomerase, polyphenol oxidase and shikimate dehydrogenase (ARAVANOPOULOS *et al.* 1994). This study employed the parental clones and full-sib progeny of a 2 × 2 balanced factorial mating design, involving female clones E16 and E269, and male clones E263 and E292. The sample size of each progeny was 33 and the total sample size including the parental clones was 136.

Non-random joint segregation of locus pairs was estimated by the LINKAGE–1 software of SUITER *et al.* (1983) version 3.50. This program assesses linkage employing χ^2 contingency tables to compare observed two-locus segregation data with those expected from actual single-locus segregation ratios. In this study the use of the χ^2 test was preferred, since sample sizes were relatively small and the closeness of small-sample distribution to the asymptotic χ^2 approximation is better achieved by the χ^2 than its alternative the G-test (LARNTZ 1978). Additionally, Monte Carlo studies suggest that at least for tests conducted at a nominal 0.05 level of significance, the χ^2 goodness-of-fit statistics accomplish the desired level of approximation, even when minimum expected cell values approximate 1.0 (FIENBERG 1985). When significant deviation from independent assortment was observed, the recombination fraction (θ) and its standard error (SE), were estimated using the maximum likelihood formulae developed by ALLARD (1956).

Locus pairs that could be tested in one family only, were not considered. When a pair of loci segregated in more than one family and produced additive offspring genotypic classes across families, then the test for linkage and the estimation of recombination fractions was conducted from individual families, as well as jointly from all available families. Linkage tests in these cases were based both on the total χ^2 value across families, and the χ^2 estimate of the pooled data set. If it was not possible to pool the data, a weighted average recombination frequency from individually informative families, was computed according to COLQUHOUN (1971), with a standard error estimated according to BAILEY (1961). Pooling data from different families produces more robust tests when compared to estimates of the total χ^2 value across families due to the reduction in the degrees of freedom, however since parental phase is unknown there is a chance that pooling may mask true association among locus pairs. Furthermore, the absence of data on parental phase prompted the investigation of the stability of linkage groups found. The stability of such linkages was examined by using bootstrap data sets. Replicate data sets were created by resampling individuals in a linked locus pair with replacement from the original data set. At the end of 1000 bootstraps the percentage of recombination fractions equal to 0.5 was examined. The program LINKAGE developed by Dr. K. Ritland[1] was used in this investigation. Homogeneity of the recombination frequencies from different families was estimated according to BAILEY (1961).

The mapping function suggested by KOSAMBI (1944) was employed. This function has been used in many organisms including its almost exclusive application in plant genetics, since general levels of known interference (SWANSON *et al.* 1981) support its application. Distance between linked loci was calculated according to KOSAMBI (1944), with a standard error according to OWEN (1950). The impact of chiasma interference was estimated by the coefficient of coincidence (C) and its variance according to BAILEY (1961).

RESULTS AND DISCUSSION

From the 55 possible two locus combinations of the 11 variable loci, 49 could be investigated. The number of families employed and the total number of genotypes used in linkage analysis are presented in Table 1. It should be noted that 88% of the two locus combinations could be tested in two or more families (with the number of individuals employed ranging from 62 to 131). A total of 107 two-locus tests were initially investigated over the four full-sib families of *S. eriocephala.* In a number of cases, recombination results varied among families. Parallel observations have been demonstrated in many taxa and have been attributed mainly to sampling error, but also to genetic (inversion polymorphisms, pre or post-zygotic selection), or environmental factors (GRELL 1966, RUDIN and EKBERG 1978). The original data of two groups that showed possible linkage are presented in Table 2, while the linkage statistics of these groups are shown in Table 3 and discussed below.

Aco–2:Ppo–1. This group could be investigated in two families (Tables 2, 3). Both χ^2 tests were significant (p values were 0.008 and 0.020 respectively). Homogeneity of recombination fractions among families was observed (Table 3). The total χ^2 test was also highly significant ($p < 0.005$, 2 d.f.), nevertheless the pooled test (1 d.f.) was not significant. Resampling indicated that in 1000 bootstrapped samples a recombination fraction of 0.5 was observed only in 4.5 % of cases in family E16 × E263 and 8.9 % of cases in family E269 × E263. The crosses that could be tested involved three parental clones which originated from unrelated populations (ARAVANOPOULOS *et al.* 1994). Therefore a difference in the parental phase

[1] Dr. K. Ritland, Department of Forest Sciences, University of British Columbia

Table 1. The number of *S. eriocephala* families (below the diagonal) and the total number of trees (above the diagonal) examined in each particular pair of 11 polymorphic loci.

Locus	*Aco-2*	*Acp-2*	*Adh-2*	*Alp-1*	*Cto-1*	*Idh-2*	*Per-1*	*Pgd-2*	*Pgi-2*	*Ppo-2*	*Skdh-2*
Aco-2		-	33	-	66	65	62	32	-	64	66
Acp-2	-		33	65	-	65	65	65	65	64	65
Adh-2	1	1		33	33	66	64	66	33	65	66
Alp-1	-	2	1		-	65	65	65	65	64	131
Cto-1	2	-	1	-		66	62	98	-	64	66
Idh-2	2	2	2	2	2		127	98	65	127	131
Per-1	2	2	2	2	2	4		98	65	127	131
Pdg-1	1	2	2	2	1	3	3		65	97	98
Pgi-2	-	2	1	2	-	2	2	2		64	65
Ppo-1	2	2	2	2	2	4	4	3	2		131
Sdh-2	2	2	2	2	2	4	4	3	2	4	

Table 2. Potential linkage groups, parental genotypes (first locus denoted with capital letters, second locus with small letters), individual family progeny classes and pooled progeny classes.

Locus pair	Cross	Parental genotypes	Progeny classes			
			AA//aa	AA//ab	AB//aa	AB//ab
Aco-2:Ppo-1	E16 × E263	AA//aa × AB//ab	2	13	10	7
	E269 × E292	AA//aa × AB//ab	12	4	5	10
	Pooled		14	17	15	17
Cto-1:Ppo-1	E16 × E263	AA//aa × AB//ab	7	10	5	10
	E269 × E292	AA//aa × AB//ab	9	1	8	13
	Pooled		16	11	13	23

among these clones, may produce different categories of 'parental' and 'recombinant' progenies in the two crosses and lead in an insignificant result when the pooled data are considered. This can be appreciated from a comparison between original family data and pooled data depicted in Table 2. It was decided to regard this pair as possibly linked.

Cto–1:Ppo–1. This group could also be investigated in two families (Tables 2, 3). In one family (E269 × E263) the χ^2 test was highly significant ($p = 0.007$), but in the other the test was insignificant. In the former case the percentage of the bootstrapped samples that resulted in $\theta = 0.5$ was only 8 %, but in the latter it was 56.5 %. There was homogeneity of recombination fractions among families. The total χ^2 test was significant ($p < 0.025$), while the pooled χ^2 test weakly significant ($p = 0.070$). It was decided to consider this pair as possibly linked.

The loci presumed as linked should form a single linkage group in the order *Aco–2:Ppo–1:Cto–1*, and the mapping distances would then be about 34 cM and 48 cM respectively (Table 2). The same linkage group was also detected in *Salix exigua* Nutt.(section *Longifoliae*, subgenus *Salix*; ARAVANOPOULOS 1998). Three consecutively linked loci permit the estimation of the coefficient of coincidence. The estimate $C = 0.970$ ($s_C = 0.322$) reflected a positive interference, which would mean that the occurrence of one exchange between homologus homologous chromosomes in *S. eriocephala* would reduce the likelihood of another in its vicinity. Similar interference levels have been reported in *Salix exigua* (ARAVANOPOULOS 1998), in conifers (RUDIN & EKBERG 1978, STRAUSS & CONKLE 1986) and in various other plant and animal species (SWANSON *et al.* 1981).

The detection of this possible linkage group warrants some discussion with regards to the corresponding linkage group detected in *S. exigua.* When compared to *S. eriocephala*, *Cto–1* and *Ppo–1* appear in reverse positions in relation to *Aco–2* in the *S. exigua* sequence (ARAVANOPOULOS 1998). The estimates of the map distances between these loci in *S. exigua* are somewhat smaller, however it is important to note that they are analogous in the two species (the estimate of distance between *Cto–1* and *Ppo–1* is greater than the estimate of distance between *Aco–2* and its adjacent locus in these species). This finding suggests that although these genes are probably common to both species, one of the species may contain a chromosomal paracentric inversion in the chromosome marked by this linkage group. Chromosomal inversions can have profound implications

Table 3. Potential linkage groups, their recombination fractions (θ) and standard errors (s) within and across informative families. χ^2-tests within families, total and pooled and associated probabilities (p), χ^2-tests for the homogeneity of recombination fractions χ^2_r and associated probabilities (p_r), mapping distances (D) and associated standard errors (s_D). F: percent frequency of $\theta = 0.5$ is 1000 bootstrapped samples.

Locus pair	Cross	χ^2 (df)	p	θ	s	χ^2_r (df)	p	D	s_D	F (%)
Aco-2:Ppo-1	E16 × E263	7.036 (1)	0.008	0.281	0.079					04.5
	E269 ×	5.427 (1)	0.020	0.290	0.081					08.9
	E292	12.463	<0.005	0.293	0.057	0.017	>0.900	33.6	8.7	
	Total	(2)	0.859			(1)				
	Pooled	0.031								
		(1)								
Cto-1:Ppo-1	E16 × E263	0.209 (1)	0.647	0.468	0.088					56.5
	E269 ×	7.369 (1)	0.007	0.290	0.082					08.0
	E292	7.578 (2)	<0.025	0.373	0.058	2.174	>0.100	48.2	13.1	
	Total	3.330 (1)	0.070			(1)				
	Pooled									

on evolution, speciation and phylogenetic reconstruction (DOBZANSKY 1970). Such an inversion would represent a notable event in the evolution and speciation of *Salix* and offers a challenge for further genetic and cytological studies.

As it has been reported for many organisms including higher plants, in general linkage conservation appears to exist above the species level (WEEDEN & WENDEL 1989). The presence of the group identified above in *S. exigua*, was already discussed. Another locus pair that was found linked in *S. exigua Acp–1:Alp–1* (ARAVANOPOULOS 1998), could be tested in two families of *S. eriocephala*. This pair was not found linked (total $\chi^2 = 3.040$, $0.20 < p < 0.25$). In agreement with the results of this study, THORSEN *et al.* (1997) found no linkage between *Acp–2:Pgm–1* and *Acp–2:Sdh–2* in *Salix viminalis*. Furthermore, in concordance with the above results, PER and 6PGD loci were also found not to be linked in *Populus* spp. (HYUN *et al.* 1987, RAJORA 1990). On the other hand, MÜLLER-STARCK (1992) reported the possible linkage of *Pgi–2* and *Idh–2* in poplars of the *Tacamahaca* section, but this group was not found to be linked in *S. eriocephala* as well as in *Salix viminalis* (THORSEN *et al.* 1997).

This study has shown the limitations of linkage investigations with isoenzyme markers. Despite the fact that the original set of markers developed included 47 isoenzyme loci (ARAVANOPOULOS *et al.* 1994), only 11 were variable in the full-sib families studied (the breeding material needed in woody angiosperms in order to study linkage) and only three of them form one possible linkage group. These markers are highly stable and completely penetrant offering very good anchor points in any genome, but are unfortunately very few. The development of molecular marker maps will cover a major part of *S. eriocephala* genome, but at a greater expense including extensive tests in order to verify marker stability and repeatability.

ACKNOWLEDGEMENTS

The experimental part and the data of this research was conducted at the University of Toronto. Furhter data analysis was conducted and the manuscript was written while the author was associated with the Aristotle University of Thessaloniki. This research was supported by the Natural Science and Engineering Research Council of Canada (N.S.E.R.C.) through Grant OG0020542. Special thanks are due to Dr. K. Ritland for fruitful discussions on linkage analysis and for the provision of software. The author would also like to thank Drs. J. E. Eckenwalder, M. Hubbes, D. N. Roy and J. Anderson for carefully reviewing an earlier draft of this work. Dr. L. Zsuffa is especially thanked for enlightening and productive discussions. Financial assistance from the University of Toronto (Special Fellowships), the Faculty of Forestry (Jonson Scholarships) and the Government of Ontario (Ontario Graduate Scholarships) is gratefully acknowledged.

REFERENCES

ALLARD, R.W. 1956: Formulae and tables to facilitate the calculation of recombination values in heredity. *Hilgardia* **4**: 235–278.

ARAVANOPOULOS, F. A. 1998: Analysis of genetic linkage in *Salix exigua* Nutt. *Silv. Genet.* **47**: 127–131.

ARAVANOPOULOS, F. A, ZSUFFA, L. and CHONG, K. X. 1994: Inheritance analysis of isoenzyme genes in *Salix eriocephala* Michx. *J. Hered.* **85**: 381–388.

BAILEY, N. T. J. 1961: Introduction to the mathematical theory of genetic linkage. Clarendon Press, Oxford.

BRADSHAW, H. D., VILLAR M., WATSON B. D., OTTO K. G., STEWART S. & STETTLER, R. F. 1994: Molecular genetics of growth and development in *Populus*. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. *Theor. Appl. Genet.* **89**: 167–178.

BUCCI, G., KUBISIAK, T. L., NANCE, W. L. & MENOZZI, P.

1997: A population consensus partial linkage map of *Picea abies* Karst. based on RAPD markers. *Theor. Appl. Genet.* **95**: 643–654.

COLQUHOUN, D. 1971: Lectures on Biostatistics. Claredon Press, Oxford.

DOBZANSKY, T. 1970: Genetics and the Evolutionary Process. Columbia University Press, New York.

FIENBERG, S. E. 1985: The Analysis of Cross-classified Categorical Data. MIT Press, Cambridge, Massachusetts, USA.

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

GRELL, R. 1966: The meiotic origin of temperature induced crossovers in *Drosophila melanogaster* females. *Genetics* **54**: 411–421.

HATTEMER, H. H., STEINER, W. & KOWNATZKI, D.1990: Genetic markers in birch. *Silv. Genet.* **39**: 45–50.

HOEY, M. T. & PARKS, C. R. 1990: Isozyme inheritance in the genus *Liquidambar* L. *J. Hered.* **81**: 393–397.

HYUN, J. O., RAJORA, O. P. & ZSUFFA, L. 1987: Inheritance and linkage of isozymes in *Populus tremuloides* Michx. *Genome* **29**: 383–388.

KOSAMBI, D. D. 1944:. The estimation of map distances from recombination values. *Ann. Eugen. Lond.* **12**: 172–175.

LARNTZ, K. 1978:. Small sample comparisons of exact levels for chi-square goodness-of-fit statistics. *J. Am. Statist. Assoc.* **73**: 253–263.

LEWONTIN, R. C. 1974: The Genetic Basis of Evolutionary Change. Columbia University Press, New York, USA.

MALVOLTI, M.E., TEISSIER DU CROS, E., FINESCHI, S. & PACIUCCI, M. 1991: Biochemical markers in eastern cottonwood (*Populus deltoides* Bartr.): Enzymatic variation in a factorial mating design. *In:* Biochemical Markers in the Population Genetics of Forest Trees. (ed. Fineschi, S., Malvotti M. E., Cannata F. and Hattemer H. H.). pp. 31–40. SPB Academic Publ, The Hague.

MITCHELL-OLDS, T. 1995: The molecular basis of quantitative genetic variation in natural populations. *Trends Ecol. Evol.* **10**: 324–328.

MÜLLER-STARCK, G. 1992: Genetic control and inheritance of isoenzymes in poplars of the *Tacamahaca* section and hybrids. *Silv. Genet.* **41**: 289–292.

OWEN, A. R. G. 1950: The theory of genetic recombination. *Adv. Genet.* **3**: 117–157.

PARKS, C. R., WENDEL, J. F., SEWELL, M. M. & QIU, Y. L. 1990: Genetic control of isozyme variation in the genus *Liriodendron* L. (*Magnoliaceae*). *J. Hered.* **81**: 317–323.

RAJORA, O. P. 1990: Genetics of allozymes in *Populus deltoides* March., *P. nigra* L. and *P. maximoviczii* Henry. *J. Hered.* **81**: 301–308.

RUDIN, D. 1986: Developmental trends in the field of biochemical genetics of forest trees. *In:* Proc. 18th IUFRO World Congress. Div. 2. Forest Products and Forest Protection, Vol. 2, pp. 577–588, Ljubljana.

RUDIN, D. & EKBERG, I. 1978: Linkage studies in *Pinus silvestris* L., using macrogametophytic allozymes. *Silv. Genet.* **27**: 1–12.

SANTI, F. & LEMOINE, M. 1990: Genetic markers for *Prunus avium* L.: inheritance and linkage of isozyme loci. *Ann. Sci. For.* **47**: 131–139.

SEWELL, M. M., SHERMAN B. K. & NEALE D. B. 1999: A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. *Genetics* **151**: 321–330.

STRAUSS, S. H. & CONKLE, M. T. 1986: Segregation, linkage and diversity of allozymes in knobcone pine. *Theor. Appl. Genet.* **72**: 417–434.

SUITER, K. A., WENDEL J. F. & CASE, J. S. 1983: LINKAGE-1: A PASCAL computer program for the detection of analysis of genetic linkage. *J. Hered.* **74**: 142–153.

SWANSON, C.P., MERZ, T., and YOUNG, W.J. 1981: Cytogenetics. Prentice-Hall Inc., New Jersey.

TANKSLEY, S.D. 1993: Mapping polygenes. *Ann. Rev. Genet.* **27**: 205–233.

THORSEN, J., JORDE, P. E.,, ARAVANOPOULOS, F.A., GULLBERG, U. & ZSUFFA, L. 1997: Inheritance and linkage of isozyme loci in the basket willow *Salix viminalis* L. *J. Hered* **88**: 144–150.

WEEDEN, N.F. & WENDEL, J.F. 1989: Genetics of plant isozymes. *In:* Isozymes in Plant Biology. (Ed. Soltis, D. E. and Soltis P. S.). pp. 46–72. Dioscurides Press, Portland.

YAZDANI, R., YEH, F. C. & RIMSHA, J. 1995: Genomic mapping of *Pinus sylvestris* L. using random amplified polymorphic DNA markers. *Forest Genetics* **2**: 109–116.

ZOUROS, E., KRIBAS, C. B., TSAKAS, S. & LOUKAS, M. 1974: Genic versus chromosomal variation in natural populations of *Drosophila subobscura*. *Genetics* **78**: 1223–1244.

ZSUFFA, L. & ARAVANOPOULOS, F. A. 1989: Genetics and breeding of *Salicaceae*. *In:* Recent Developments in Poplar Selection and Propagation. Institute of Forest Tree Breeding. pp. 1–17, Hann. Münden.