especially, the Halla population, because of its unique genetic structure. Finally, we can anticipate a restoration program for populations destroyed by natural and/or man-made disturbances. Some populations, such as Dukyoo in this study have been severely disturbed recently, by the construction of a ski resort and other activities. In this case, we can collect seeds, grow seedlings and transplant them back to the disturbed sites. If the seedlings survive and mature, this would help maintain an effective population size.

Literature

ALLARD, R. W., JAIN, S. K. and WORKMAN, P. L.: The genetics of inbreeding populations. Adv. Genet. **14**, 55–131 (1968). — BARTON, N. H. and SLATKIN, M.: A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. Heredity **56**, 409–415 (1986). — CHAKRABORATY, R., FUERST, P. A. and NEI, M.: Statistical studies on protein polymorphism in natural populations. III. Distribution of allele frequencies and the number of alleles per locus. Genetics **94**, 1039–1063 (1980). — CONKLE, M. T., HODGSKISS, P. D., NUNALLY, L. B. and HUNTER, S. C.: Starch gel electrophoresis of pine seed: a laboratory manual. U.S. Forest Service General Technical Report PSW-64 (1982). — CORNUET, J. M. and LUIKART, G.: Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics **144**, 2001–2014 (1996). — EL-KASSABY, Y. A. and YANCHUK, A. D.: Genetic diversity, differentiation and inbreeding in Pacific yew from British Columbia. J. Hered. **85**, 112–117 (1994). — GOVINDARAJU, D. R.: Relationship between dispersal ability and levels of gene flow in plants. Oikos **52**, 31–35 (1988). — GURIES, R. P. and LEDIG, F. T.: Genetic diversity and poulation structure in pitch pine (*Pinus rigida* MILL.). Evolution **36**, 387–402 (1982). — GÖÇMEN, B., JERMSTAD, K. D., NEALE, D. B. and KAYA, Z.: Development of random amplified polymorphic DNA markers for genetic mapping in Pacific yew (*Taxus brevifolia*). Can. J. For. Res. **26**, 497–503 (1996). — HAMRICK, J. L., GODT, M. J. W. and SHERMAN-BROYLES, S. L.: Factors influencing levels of genetic diversity in woody plant species. New Forests **6**, 95–124 (1992). — HERTEL, H.: Vererbung von Isoenzymmarkern bei Eibe (*Taxus baccata* L.). Silvae Genetica **45**, 284–290 (1996). — HERTEL, H. and KOHLSTOCK, N.: Genetische Variation und geographische Struktur von Eibenvorkommen (*Taxus baccata* L.) in Mecklenburg–Vorpommern. Silvae Genetica **45**, 290–294 (1996). — KIM, Z. S., LEE, S. W., LIM, J. H., HWANG, J. W. and KWON, K. W.: Genetic diversity and structure of natural populations of *Pinus koraiensis* (SIEB. et ZUCC.) in Korea. Forest Genetics **1**, 41–49 (1994). — LEDIG, F. T.: Heterozygosity, heterosis, and fitness in outbreeding plants. In: SOULE, M. E. (Ed.): Conservation Biology: The Science of Scarcity and Diversity. Sinauer, Sunderland, MA. pp. 77–104 (1986). — LEDIG, F. T.: Genetic variation in *Pinus,*. In: RICHARDSON, D.M. (Ed.): Ecology and Biogeography of *Pinus*. Cambridge University, pp. 251–280 (1998). — LEDIG, F. T., CONKLE, M. T., BERMEJO-VELQUEZ, B., EGUILUZ–PIEDRA, T., HODSKISS, P. D., JOHNSON, D. R. and DVORAK, W. S.: Evidence for an extreme bottleneck in a rare Mexican pinyon: genetic diversity, disequilibrium, and the mating system in *Pinus maximartinezii*. Evolution **53**, 91–99 (1999). — LEDIG, F. T., JACOB-CERVANTES, V., HODGSKISS, P. D. and EGUILUZ-PIEDRA, T.: Recent evolution and divergence among populations of a rare Mexican endemic, Chihuahua spruce, following Holocene climatic warming. Evolution **51**, 1815–1827 (1997). — LEE, S. W. and KIM, Z. S.: Computer program for hierarchical analysis of gene diversity. Korean J. Genet. **15**, 379–384 (1993). — LEE, T. B.: Dendrology. Hyang Moon Sa, Seoul. 331 pp. (in Korean) (1987). —LEWANDOWSKI, A., BURCZYK, J. and MEJNARTOWICZ, L.: Genetic structure of English yew (*Taxus baccata* L.) in the Wierzchlas Reserve: implications for genetic conservation. Forest Ecology and Management **73**, 221–227 (1995). — LUIKART, G. and CORNUET, J. M.: Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology **12**, 228–237 (1998). — NEI, M.: Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA **70**, 3321–3323 (1973). — NEI, M.: Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**, 583–590 (1978). —RAYMOND, M. and ROUSSET, F.: An exact test for population differentiation. Evolution **49**, 1280–1283 (1995a). — RAYMOND, M. and ROUSSET, F.: GENEPOP (ver. 1.2) : A population genetics software for exact test and ecumenicism. J. Hered. **86**, 248–249 (1995b). — ROUSSET, F. and RAYMOND, M.: Testing heterozygote excess and dificiency. Genetics **140**, 1413–1419 (1995). — SLATKIN, M.: Rare alleles as indicators of gene flow. Evolution **39**, 53–65 (1985). — SLATKIN, M.: Isolation by distance in equilibrium and non–equilibium populations. Evolution **47**, 264–279 (1993). — SNEATH, P. H. A. and SOKAL, R. R.: Numerical Taxonomy. Freeman, San Francisco, CA. 573 pp. (1973). — SWOFFORD, P. L. and SELANDER, R. B.: BIOSYS–1: A computer program for the analysis of allelic variation in population genetic and biochemical systematics. Release 1.7. User's Manual. Illinois Natural History Survey, IL. USA (1989). — WEIR, B. S. and COCKERHAM, C. C.: Estimating *F*-statistics for the analysis of population structure. Evolution **38**, 1358–1370 (1984). — WHEELER, N. C., JECH, K. S., MASTERS, S. A., O'BRIEN, C. J., TIMMONS, D. W., STONECYPHER, R. W. and LUPKES, A.: Genetic variation and parameter estimates in *Taxus brevifolia* (Pacific yew). Can. J. For. Res. **25**, 1913–1927 (1995). — WRIGHT, S.: The genetical structure of populations. Ann. Eugen. **15**, 323–354 (1951). — WRIGHT, S.: The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution **19**, 395–420 (1965).

Analysis of Half-diallel Mating Design with Missing Crosses: Theory and SAS Program for Testing and Estimating GCA and SCA Fixed Effects

By H. X. WU and A. C. MATHESON[1]

Abstract

The half-diallel mating design, particularly a series of disconnected half-diallels has been widely adopted as a mating design for estimating genetic parameters and for future selection in many commercially important tree species. Standard commercial statistical packages do not allow direct specification of the linear model associated with the half-diallel design and therefore are not capable of analysing diallel mating designs, even for balanced diallel matings (no missing crosses). Published special computer programs for diallel analyses do not provide an adequate solution for GCA and SCA fixed effects in diallels with missing crosses. This paper presents the least squares theory for analysing half-diallel mating designs with missing crosses, and a SAS computer program (DIAFIXED.SAS), developed to test the significance of GCA and SCA effects and estimate the GCA and SCA fixed effects. The program is flexible enough to accommodate different number of parents, multiple environments and missing individual trees as well as missing whole plots. The DIAFIXED.SAS output includes (1) hypothesis testing for GCA and

[1] CSIRO, Division of Forestry and Forest Products, PO Box E4008, Kingston, ACT 2604, Australia

SCA fixed effects and environmental effects, (2) estimates of GCA and SCA fixed effects, (3) estimates of standard errors of GCA and SCA fixed effects. Results from a 6 by 6 half-diallel for radiata pine planted in two sites are also presented.

Key words: Diallel mating, least squares, GCA, SCA, radiata pine.

Introduction

Diallel mating designs are frequently used in plant and tree breeding research to obtain genetic information, such as genetic effects for a fixed set of parent (fixed effects or Model 1) or variance components for general (GCA) and specific (SCA) combining abilities and heritability for a population based on randomly chosen parents (random effects or Model 2) (VAN BUIJTENEN *et al.,* 1976; ZHANG and KANG, 1997). Diallel matings are also deployed to provide the maximum opportunity to manage coancestry in a breeding population and to maximise the selection differential for further selection (BRIDGWATER, 1992; JOHNSON and KING, 1998).

GRIFFING (1956) divided diallel matings into 4 categories (methods). Among the four diallel mating methods, Method 4 is the most popular with tree species because it only includes half of crosses in one direction, no selfs involved. Half-diallel, particularly a series of disconnected half-diallels was widely adopted as the mating design for many commercially important tree species, such as *Pseudotsuga menziesii* (YEH and HEAMAN, 1987; STONECYPHER *et al.,* 1996), *Pinus radiata* (WILCOX *et al.,* 1975; SHEPHERD, 1976; DEAN, 1990; CARSON, 1991; MATHESON *et al.,* 1994), *Pinus palustris* (SNYDER and NAMKOONG, 1978), *Pinus patula* (BURLEY *et al.,* 1966), *Pinus elliottii* (SQUILLACE, 1973; HUBER *et al.,* 1992), *Pinus taeda* (WEIR and GODDARD, 1986), *Pinus sylvestris* (JONSSON *et al.,* 1992), *Picea abies* (SKROPPA, 1996), to name a few. Among common mating designs in plant and tree breeding (open-pollinated, polycross, single-pair, nested, factorial, diallel mating), the diallel mating design is the most difficult to analyse. Standard commercial statistical packages do not allow direct specification of the diallel mating model and therefore are not capable of analysing the diallel mating design from existing models or procedures, for balanced and unbalanced diallel matings.

Systematic formulae were developed by GRIFFING (1956) to analyse diallel mating designs for balanced mating structures (no missing crosses). In tree breeding literature, a Fortran computer program called DIALL was developed by SCHAFFER

and USANIS (1969) to analyse diallel matings for unbalanced data, but the program estimates GCA and SCA variances only, in a completely random model. The enhancement by SNYDER (1975) could estimate GCA effect, and SCA effects, but could not test the statistical significance of GCA and SCA fixed effects. Estimates of reliability (standard errors) for GCA and SCA effects are also desirable. In this paper, a SAS program with a least squares analysis was developed to analyse and estimate GCA and SCA fixed effects for a half-diallel mating design with missing and non-missing crosses, along with statistical testing of GCA and SCA fixed effects. First, the theory of analysing half-diallel mating designs with missing crosses is presented in order to facilitate the understanding of the mechanism, assumptions and algorithm underlying the analysis. Through understanding of these matrix operations, readers could implement the algorithm into any other commercially statistical packages. A SAS program was developed because it is not only a powerful and versatile statistical analysis system with a flexible data management system, but is also increasingly popular in tree breeding.

Theory of Least Squares Estimates for Half-diallel Mating Design With Missing Crosses

The SAS program for testing and estimating GCA and SCA fixed effect is based on a least squares analysis of the model appropriate for a half-diallel mating design. The linear model for this mating design is usually specified as (model 1)

$$Y_{ijk} = \mu + g_i + g_j + s_{ij} + e_{ijk} \quad (1)$$

where μ is the grand mean, g_i and g_j are the general specific combining abilities for i^{th} and j^{th} parents, respectively, s_{ij} is the specific combining ability between i^{th} and j^{th} parent and e_{ijk} is the residual. The theory of general least square indicates that the best estimates of the parameter g_i, g_j and s_{ij} could be derived by minimising the sum of squares of residuals ($\Sigma e_{ijk}^2 = \Sigma (Y - \mu - g_i - g_j - s_{ij})^2$), if e_{ijk} is assumed independently, identically distributed with same variance. An ordinary least squares estimate for the parameter vector $\mathbf{b}$ (= μ g_i ... g_n s_{ij} ... $s_{(n-1)n}$) in equation 1 could be found by solving the following equation

$$\mathbf{b = (X'X)^{-} X'Y}$$

where $\mathbf{X}$ is the design matrix for model 1. The difficulty with standard statistical packages is that they can not build the design matrix $\mathbf{X}$ directly from the linear model 1. The design matrix $\mathbf{X}$ for the half-diallel mating design is of the form

		μ	g_1	g_2	g_3	$\cdots$	$g_{(n-1)}$	g_n	s_{12}	s_{13}	$\cdots$	s_{1n}	s_{23}	s_{24}	$\cdots$	s_{2n}	$\cdots$	$s_{(n-2)(n-1)}$	$s_{(n-2)n}$	$s_{(n-1)n}$
y_{12}		1	1	1	0	$\cdots$	0	0	1	0	$\cdots$	0	0	0	$\cdots$	0	$\cdots$	0	0	0
y_{13}		1	1	0	1	$\cdots$	0	0	0	1	$\cdots$	0	0	0	$\cdots$	0	$\cdots$	0	0	0
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_{1n}		1	1	0	0	$\cdots$	0	1	0	0	$\cdots$	1	0	0	$\cdots$	0	$\cdots$	0	0	0
y_{23}		1	0	1	1	$\cdots$	0	0	0	0	$\cdots$	0	1	0	$\cdots$	0	$\cdots$	0	0	0
y_{24}	$\mathbf{X} =$	1	0	1	0	$\cdots$	0	0	0	0	$\cdots$	0	0	1	$\cdots$	0	$\cdots$	0	0	0
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_{2n}		1	0	1	0	$\cdots$	0	1	0	0	$\cdots$	0	0	0	$\cdots$	1	$\cdots$	0	0	0
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$y_{(n-2)(n-1)}$		1	0	0	0	$\cdots$	1	0	0	0	$\cdots$	0	0	0	$\cdots$	0	$\cdots$	1	0	0
$y_{(n-2)n}$		1	0	0	0	$\cdots$	0	1	0	0	$\cdots$	0	0	0	$\cdots$	0	$\cdots$	0	1	0
$y_{(n-1)n}$		1	0	0	0	$\cdots$	0	1	0	0	$\cdots$	0	0	0	$\cdots$	0	$\cdots$	0	0	1

where y_{ij} represents the cross between the ith and j[th] parents, the parameters displayed above the design matrix label the appropriate column for each parameter and 1's and 0's denote the presence or absence of a parameter in the model for the observation (data vector, Y). The design matrix **X** is composed of two submatrices: the GCA submatrix (columns g_1 to g_n) and the SCA submatrix (columns s_{12} to $s_{(n-1)n}$). To estimate the parameter vector **b** we need to invert matrix **X'X**. Because **X** is not of full-rank, there is no unique estimate for the vector **b**. In order to have an invariant estimate for the parameter vector **b,** linear constraints for the parameter vector **b** are set in such a way that the remaining columns in **X,** after deleting dependent columns according to the parameter constraints, become linearly independent. Thus, the design matrix **X** with these constraints becomes non-singular (full-rank). To define these linear constraints, the number of linearly dependent columns in **X** should be determined first. For a half-diallel mating design of p parents, there should be p+1 linear dependencies, whether there are missing crosses or not. Therefore there should be 6 linearly dependent columns for a p=5 half-diallel mating design. One of common constraints for a half-diallel mating design is to set $\Sigma g_i = 0$ and $\Sigma s_{ij} = 0$ for each i or j with $s_{ij} = s_{ji}$. Thus for the p = 5 half-diallel design without missing crosses, we could estimate ten independent and 6 dependent parameters of the vector **b**. One of the more convenient choices for the ten independent parameters is μ, g_1, g_2, g_3, g_4, s_{12}, s_{13}, s_{14}, s_{23}, s_{24}. With such a choice and the parameter constraints, the six dependent parameters can be estimated as $g_5 = -g_1 -g_2 -g_3 -g_4$, $s_{15} = -s_{12} -s_{13} -s_{14}$, $s_{25} = -s_{12} -s_{23} -s_{24}$, $s_{34} = -s_{12} -s_{13} -s_{14} -s_{23} -s_{24}$, $s_{35} = s_{12} + s_{14} + s_{24}$, $s_{45} = s_{12} + s_{13} + s_{23}$. These relationships are illustrated in the following **X** matrix for the case of p = 5 without missing crosses.

		μ	g_1	g_2	g_3	g_4	s_{12}	s_{13}	s_{14}	s_{23}	s_{24}
y_{12}		1	1	1	0	0	1	0	0	0	0
y_{13}		1	1	0	1	0	0	1	0	0	0
y_{14}		1	1	0	0	1	0	0	1	0	0
y_{15}		1	0	−1	−1	−1	−1	−1	−1	0	0
y_{23}	**X** =	1	0	1	1	0	0	0	0	1	0
y_{24}		1	0	1	0	1	0	0	0	0	1
y_{25}		1	−1	0	−1	−1	−1	0	0	−1	−1
y_{34}		1	0	0	1	1	−1	−1	−1	−1	−1
y_{35}		1	−1	−1	0	−1	1	0	1	0	1
y_{45}		1	−1	−1	−1	0	1	1	0	1	0

Similarly, the design matrix could be derived for all other sizes of half-diallel mating design. Working out these relations is relatively easy for no missing crosses, and indeed, these relationships for most common half-diallel matings have been reported (ZHANG and KANG, 1997). But with missing crosses, working out these relationships become tedious due to variation in the number and distribution pattern of missing crosses. For example, for p = 5 with two missing crosses (say y_{14} and y_{34}), there are only eight independent parameters left with the same six dependent parameters. These 8 independent parameters could be μ, g_1, g_2, g_3, g_4, s_{12}, s_{13}, s_{23}. The remaining six dependent parameters are estimated as $g_5 = -g_1 -g_2 -g_3 -g_4$, $s_{15} = -s_{12} -s_{13}$, $s_{25} = s_{13}$, $s_{24} = -s_{12} -s_{13} -s_{23}$, $s_{35} = -s_{13} -s_{23}$, $s_{45} = +s_{12} +s_{13} +s_{23}$, with the **X** matrix defined as

		μ	g_1	g_2	g_3	g_4	s_{12}	s_{13}	s_{23}
y_{12}		1	1	1	0	0	1	0	0
y_{13}		1	1	0	1	0	0	1	0
y_{15}		1	0	−1	−1	−1	−1	−1	0
y_{23}		1	0	1	1	0	0	0	1
y_{24}	**X** =	1	0	1	0	1	−1	−1	−1
y_{25}		1	−1	0	−1	−1	0	1	0
y_{35}		1	−1	−1	0	−1	0	−1	−1
y_{45}		1	−1	−1	−1	0	1	1	1

In order for a computer to derive the design matrix **X** for a half-diallel mating with missing crosses, certain rules should be used for generating **X**. Fortunately, these rates are relatively simple for half-diallels with no missing crosses and with some patterns of missing crosses, as described below.

First, determine which parents have no missing crosses. If there is at least one parent with no missing crosses, then following four steps apply to form the design matrix X: 1. Construct a GCA submatrix of p parents according to the structure of linear model 1, then delete a single column corresponding to any one parent with no missing crosses in the GCA submatrix. While deleting the column in the GCA submatrix, add minus one ('-1') to the elements on the other columns along each row containing '1' for the deleted column (SEARLE, 1987; HUBER *et al.,* 1992). This is equivalent to setting the parameter in the deleted column as the linearly dependent column according to the constraint $\Sigma g_i = 0$. 2. The SCA submatrix can be formulated from the horizontal direct product: e.g. the s_{ij} column is formed by element-wise multiplication between the g_i and the g_j columns. 3. The next step is to delete columns in the SCA submatrix corresponding to the missing crosses. Since the remaining SCA submatrix after the deletion of columns would always have one more linear dependent columns, this linearly dependent column should be deleted in the remaining SCA submatrix in order to make the design matrix full-rank (usually the last column in the remaining SCA submatrix is deleted). 4. When deleting this last linearly dependent column in the SCA submatrix, elements in other columns of the SCA submatrix in the row containing '1' in the deleted column will also be altered: e.g. if the element is one in the deleted column, add minus one on all the elements of other columns in the same row.

For example, if cross y_{45} was missing in a 5 x 5 half-diallel mating experiment, then parents 1, 2, and 3 have no missing crosses. Any column corresponding to parent 1, or 2, or 3 could be selected for deletion. If parent 3 (g_3) was selected as the linearly dependent parameter, the column corresponding to g_3 in the GCA submatrix is deleted and minus one is added to all elements on the same row if the element is one in the g_3 column (such as row y_{13}, y_{23}, y_{34}, and y_{35}). This produces following GCA submatrix

		μ	g_1	g_2	g_4	g_5
y_{12}		1	1	1	0	0
y_{13}		1	0	−1	−1	−1
y_{14}		1	1	0	1	0
y_{15}		1	1	0	0	1
y_{23}	**X** =	1	−1	0	−1	−1
y_{24}		1	0	1	1	0
y_{25}		1	0	1	0	1
y_{34}		1	−1	−1	0	−1
y_{35}		1	−1	−1	−1	0

Next, the SCA submatrix is formed by horizontal direct product between the g_i and the g_j columns and with this, a new GCA and SCA matrix is formed as

	μ	g_1	g_2	g_4	g_5	s_{12}	s_{14}	s_{15}	s_{24}	s_{25}
y_{12}	1	1	1	0	0	1	0	0	0	0
y_{13}	1	0	-1	-1	-1	0	0	0	1	1
y_{14}	1	1	0	1	0	0	1	0	0	0
y_{15}	1	1	0	0	1	0	0	1	0	0
y_{23}	1	-1	0	-1	-1	0	1	1	0	0
y_{24}	1	0	1	1	0	0	0	0	1	0
y_{25}	1	0	1	0	1	0	0	0	0	1
y_{34}	1	-1	-1	0	-1	1	0	1	0	1
y_{35}	1	-1	-1	-1	0	1	1	0	1	0

with $\mathbf{X} =$ shown for the left block.

Now, deleting the last linearly dependent column (S_{25}) and adding minus 1 to the remaining columns of same row of the SCA submatrix if elements in the S_{25} column are one (such as row y_{13}, y_{25} and y_{34}). This creates the final full-rank design matrix

	μ	g_1	g_2	g_4	g_5	s_{12}	s_{14}	s_{15}	s_{24}
y_{12}	1	1	1	0	0	1	0	0	0
y_{13}	1	0	-1	-1	-1	-1	-1	-1	0
y_{14}	1	1	0	1	0	0	1	0	0
y_{15}	1	1	0	0	1	0	0	1	0
y_{23}	1	-1	0	-1	-1	0	1	1	0
y_{24}	1	0	1	1	0	0	0	0	1
y_{25}	1	0	1	0	1	-1	-1	-1	-1
y_{34}	1	-1	-1	0	-1	0	-1	0	-1
y_{35}	1	-1	-1	-1	0	1	1	0	1

with $\mathbf{X} =$ for the left block.

If all parents have missing crosses, the above four steps do not apply, and the exact SCA submatrix needs to be worked out differently. There are two different situations with all parents having missing crosses:

A. All crosses for one or more parent are missing, but there is no more missing cross in any one of the remaining parents after the deletion of parents with all crosses missing. In this situation, the mating structure can be rearranged so that parents with all missing crosses are excluded in the rearranged mating structure. The design matrix for GCA and SCA can be derived from this rearranged mating structure as outlined above. For example, in our five-parent half-diallel mating case, if all crosses for parent 4 (y_{14}, y_{24}, y_{34}, y_{45}) plus y_{23} were missing, the mating structure could be rearranged as

		Parent J		
		2	3	5
	1	x	x	x
Parent I	2			x
	3			x

and this could be analysed as a four parent half-diallel mating design with one missing cross.

B. With other types of missing crosses when all parents having missing crosses, there seems to be no simple rule for deriving the SCA submatrix. However, it can be derived manually using the constraint that $\Sigma s_{ij} = 0$ for each i or j.

For example, in our five-parent half-diallel mating example, if crosses y_{12}, y_{34}, and y_{45} were missing, then all parents have missing crosses. The design matrix cannot be derived by the above method. But according to the same constraints that $\Sigma s_{ij} = 0$ for each i or j, the seven independent parameters and six linear dependent parameters can be derived for this pattern of missing crosses. If μ, g_1, g_2, g_3, g_4, s_{13}, s_{14} are considered as the independent parameters, the six dependent parameters can be worked out as: $g_5 = -g_1 -g_2 -g_3 -g_4$, $s_{23} = -s_{13}$, $s_{24} = -s_{14}$, $s_{15} = -s_{13} -s_{14}$, $s_{25} = s_{13} +s_{14}$, $s_{35} = 0$ with the corresponding design matrix specified as

	μ	g_1	g_2	g_3	g_4	s_{13}	s_{14}
y_{13}	1	1	0	1	0	1	0
y_{14}	1	1	0	0	1	0	1
y_{15}	1	0	-1	-1	-1	-1	-1
y_{23}	1	0	1	1	0	-1	0
y_{24}	1	0	1	0	1	0	-1
y_{25}	1	-1	0	-1	-1	1	1
y_{35}	1	-1	-1	0	-1	0	0

with $\mathbf{X} =$ for the left block.

Therefore, in any case of missing crosses in the half-diallel mating design, a full-rank design matrix can be always derived and inverted to obtain the least-squares estimator for the parameter vector $\mathbf{b}$.

Besides the design matrix for the half-diallel mating structure, progenies are usually planted at multiple sites with replications in a designed experiment. Therefore, when the design matrix for the mating design is obtained properly, it should be incorporated into such an environmental design matrix to establish the entire design matrix for all effects in the analysis. Thus, when these design matrices are set up, any statistical packages which can invert them can be used to estimate the fixed effects for any diallel mating design.

The SAS GLM Procedure makes it easy to incorporate all experimental factors (such as site and replication) into the design matrix. When environmental factors are properly specified in the model statement, SAS GLM automatically builds the entire design matrix combining the mating design matrix with the environmental design matrix. Thus, the estimation and hypothesis testing for fixed effects of GCA and SCA can be conducted in a single step. In the following, a SAS program was developed to illustrate use of these design matrices for testing and estimating fixed GCA and SCA effects in half-diallel mating designs.

Detail of the SAS Program

A SAS program DIAFIXED.SAS was developed to estimate and test fixed GCA and SCA effects for a half-diallel mating design with missing crosses. Diallels with no missing crosses are special, simpler cases for the program. DIAFIXED.SAS can handle any number of parents (it is not necessary to change the code for different numbers of parents, such as is the case in DIALLEL.SAS). The program was divided into four steps and is available upon request from the senior author. The first step was to generate a proper design matrix, a proper linear contrast matrix for each genetic component and a matrix of estimable functions for each GCA and SCA effect according to the

linear model for the diallel mating design. The SAS Macro facility was used for this purpose (SAS Institute Inc., 1987). The contrast and estimable matrices were set up based on the generated design matrix. Before running the program, four macro variables must be specified for properly generating the design matrix. The four macro variables are:

(1). N_PARENT = Number of parents in the half-diallel (e.g. N_PARENT = 5 for a 5 × 5 half-diallel design);

(2). NMCP = Any parent having no missing crosses (e.g. NMCP = 3 for parent 3 which has no missing crosses);

(3). MC = List of all missing crosses (e.g. MC = S13 S15 for missing crosses of y13 and y15);

(4). ALLPMC = Whether or not all parents have missing crosses (e.g. ALLPMC = NO for not all parents having missing crosses).

Specification of one parent (any one parent) with no missing crosses is essential to generate proper design matrix, since for half-diallel designs with all parents having missing crosses, there is no simple rule for generating the design matrix. Therefore, if all parents have missing crosses (ALLPMC = YES), the design matrix for SCA should be worked out manually and input as an external file (named as SCA_DN in the program). A macro program (D_SCA) was constructed to indicate whether computer will generate the SCA submatrix or whether it should be input as an external file.

The first step of the program produces a contrast statement for the effects of GCA, SCA, GCA by environment and SCA by environment interactions, respectively. These contrasts are used to produce sums of squares, mean squares and hypothesis testing for these effects. The following is an example of the contrast statements generated for a 6 by 6 half-diallel mating design with no missing crosses:

C_GCA = CONTRAST 'GCA' G1 1, G2 1, G3 1, G4 1, G5 1;

C_SCA = CONTRAST 'SCA' S12 1, S13 1, S14 1, S15 1, S23 1, S24 1, S25 1, S34 1, S35 1;

C_GCAENV = CONTRAST 'GCA*ENV' G1*ENV 1 -1, G2*ENV 1 -1, G3*ENV 1 -1, G4*ENV 1 -1, G5*ENV 1 -1;

C_SCAENV = CONTRAST 'SCA*ENV' S12*ENV 1 -1, S13*ENV 1 -1, S14*ENV 1 -1, S15*ENV 1 -1, S23*ENV 1 -1, S24*ENV 1 -1, S25*ENV 1 -1, S34*ENV 1 -1, S35*ENV 1 -1;

The first step also generates statements of estimable functions for estimating GCA and SCA effects. Below is a series of sample statements generated for a 6 by 6 half-diallel mating design with no missing crosses:

```
ESTIMATE "G1" G1 1; ESTIMATE "G2" G2 1; ESTIMATE "G3" G3 1; ESTIMATE "G4" G4 1; ESTIMATE "G5" G5 1;
ESTIMATE "G6" G1 –1 G2 –1 G3 –1 G4 –1 G5 –1;
ESTIMATE "S12" S12  1 S13  0 S14  0 S15  0 S23  0 S24  0 S25  0 S34  0 S35  0;
ESTIMATE "S13" S12  0 S13  1 S14  0 S15  0 S23  0 S24  0 S25  0 S34  0 S35  0;
ESTIMATE "S14" S12  0 S13  0 S14  1 S15  0 S23  0 S24  0 S25  0 S34  0 S35  0;
ESTIMATE "S15" S12  0 S13  0 S14  0 S15  1 S23  0 S24  0 S25  0 S34  0 S35  0;
ESTIMATE "S16" S12 –1 S13 –1 S14 –1 S15 –1 S23  0 S24  0 S25  0 S34  0 S35  0;
ESTIMATE "S23" S12  0 S13  0 S14  0 S15  0 S23  1 S24  0 S25  0 S34  0 S35  0;
ESTIMATE "S24" S12  0 S13  0 S14  0 S15  0 S23  0 S24  1 S25  0 S34  0 S35  0;
ESTIMATE "S25" S12  0 S13  0 S14  0 S15  0 S23  0 S24  0 S25  1 S34  0 S35  0;
ESTIMATE "S26" S12 –1 S13  0 S14  0 S15  0 S23 –1 S24 –1 S25 –1 S34  0 S35  0;
ESTIMATE "S34" S12  0 S13  0 S14  0 S15  0 S23  0 S24  0 S25  0 S34  1 S35  0;
ESTIMATE "S35" S12  0 S13  0 S14  0 S15  0 S23  0 S24  0 S25  0 S34  0 S35  1;
ESTIMATE "S36" S12  0 S13 –1 S14  0 S15  0 S23 –1 S24  0 S25  0 S34 –1 S35 –1;
ESTIMATE "S45" S12 –1 S13 –1 S14 –1 S15 –1 S23 –1 S24 –1 S25 –1 S34 –1 S35 –1;
ESTIMATE "S46" S12  1 S13  1 S14  0 S15  0 S23  1 S24  0 S25  1 S34  0 S35  1;
ESTIMATE "S56" S12  1 S13  1 S14  1 S15  1 S23  0 S24  1 S25  0 S34  1 S35  0;
```

The default program includes hypothesis testing only for environment, replication, GCA, SCA, GCA by environment and SCA by environment interaction for the following linear model

$$Y_{ijmkn} = \mu + l_k + r_{m(k)} + g_i + g_j + s_{ij} + lg_{ik} + lg_{jk} + ls_{ijk} + e_{ijmkn}$$

where l_k is the kth environment (location) effect, $r_{m(k)}$ is the

m^{th} replication effect in the kth environment, lg_{ik} and lg_{jk} are k^{th} environment with i^{th} and j^{th} GCA interactions respectively, and ls_{ijk} is the lth environment with ij^{th} SCA interaction. The user can accommodate other effects such as interaction between GCA and replication and between SCA and replication.

The second step is to input and process raw data. There is virtually no limitation on the format for raw data since SAS can process most raw data structures to match the format of the design matrix (SAS Institute Inc., 1990). The raw data should be sorted by SAS into the order of two variables: I (first parent, such as 1, 2, 3 for 4 parent half-diallel) and J (second parent, such as 2, 3, 4 for 4 parent half-diallel) so as to merge with the design matrix.

The third step is to test the effects of family (f_i), and environment, replication and interactions by the model

$$Y_{imkn} = \mu + l_k + r_{m(k)} + f_i + lf_{jk} + rf_{im(k)} + e_{imkn}$$

where f_i is i^{th} family effect. Again, users can modify the model according to the requirements of their experiments.

The last step in the program is to test the significance of GCA and SCA effects, interaction effects between environment and GCA and between environment and SCA, and to estimate GCA and SCA fixed effects using the SAS GLM Procedure (SAS Institute Inc., 1989). Which components should be included in the model, contrast, and estimate statements in the GLM model is entirely determined by the design of the actual experiments and the form of analysis.

For a typical multiple site experiment with many replications and multiple tree plots, there are two types of data imbalance. The first type is due to the mating design itself such as missing crosses and the second type is due to other factors such as unequal numbers of trees in plots or missing plots. The design matrix with missing crosses generated by DIAFIXED. SAS accounts only for the imbalance of mating design. Therefore there will be some unwanted consequences if there is data imbalance from missing trees or plots when estimating GCA and SCA effects if this imbalance is not accounted for in the program. Fortunately, the SAS GLM Procedure does account for imbalance such as missing plots (cells) or unequal observation numbers among cells in the construction process of the design matrix. However, with missing plots, GCA and SCA effects may be not entirely estimable under certain linear models. For example, in a single site experiment with multiple replications, the model used is usually

$$Y_{ijmn} = \mu + r_m + g_i + g_j + s_{ij} + rg_{im} + rg_{jm} + rs_{ijm} + e_{ijmn}.$$

If there is a missing plot, then GCA and SCA will not be estimable with the above linear model due to inclusion of replication by GCA and SCA. To make GCA and SCA estimable, the cross interactions in the missing plot should be deleted in the data set if the replication effect is significant or the interactions between replication and GCA and between replication and SCA should be deleted and the model be altered to

$$Y_{ijmn} = \mu + r_m + g_i + g_j + s_{ij} + e_{ijmn}.$$

In this model there are no interaction terms involving replication and so GCA and SCA would be estimable. Similarly, when multiple sites are used, estimability of GCA and SCA effects depends on the pattern of data imbalance and the linear model. Therefore selection of the proper model and restructuring of data are essential for estimating GCA and SCA effects when missing plots are involved.

Radiata Pine Example

Data for diameter at breast height (DBH) measured at age 11 in two radiata pine trials were used to demonstrate use of the SAS program DIAFIXED.SAS. The mating design of the

Table 1. – Radiata pine example of missing crosses in a half-diallel mating scheme, illustrating effect of missing crosses on GCA, SCA estimates and their hypothesis testing using the SAS Program DIAFIXED.SAS.

Sources of Variation[e]	None missing[a]			Missing 3 crosses[b]			Missing 6 crosses[c]			Missing 4 crosses[d]		
	DF	MS	Pr > F	DF	MS	Pr > F	DF	MS	Pr >	DF	MS	Pr > F
ENV	1	55559	0.001	1	46970	0.001	1	41234	0.001	1	47771	0.001
REP(ENV)	4	549	0.586	4	1168	0.215	4	1647	0.113	4	1365	0.153
GCA	5	1432	0.103	5	820	0.394	4	1748	0.094	5	1044	0.267
SCA	9	1733	0.020	6	2408	0.006	4	2600	0.020	5	2611	0.007
GCA*ENV	5	1173	0.186	5	730	0.464	4	2059	0.054	5	1597	0.083
SCA*ENV	9	1165	0.146	6	1724	0.045	4	872	0.406	5	889	0.359
ERROR	271	775		215	788		162	868		194	805	
Parameter[f]	Estimate	S.E.[g]		Estimate	S.E.		Estimate	S.E.		Estimate	S.E.	
G1	−1.553	2.884		0.771	3.566		1.939	3.991		−1.700	4.224	
G2	3.001	2.924		2.446	3.374		2.889	3.669		1.939	3.426	
G3	−6.110	2.877		−3.897	3.502		−7.747	3.837		−6.242	4.012	
G4	3.289	2.872		3.662	3.174		6.505	3.702		5.319	3.423	
G5	4.338	2.939		1.665	3.926					5.482	5.290	
G6	−2.964	2.861		−4.647	3.321		−3.587	3.638		−4.799	3.373	
S12	9.823	5.218		8.493	4.829		6.737	4.627		9.568	4.718	
S13	−3.172	4.693										
S14	−9.974	4.744		−12.26	4.375		−15.01	4.376		−12.44	4.413	
S15	−6.078	4.946		−5.377	4.584					−7.884	4.221	
S16	9.400	4.863		9.143	4.559		8.273	4.462		10.76	4.505	
S23	−6.209	4.702		−7.514	4.528		−2.853	4.343		−5.685	4.428	
S24	5.162	5.078		5.556	4.783		3.141	4.538		3.174	4.373	
S25	0.432	5.072										
S26	−9.209	4.802		−6.536	4.197		−7.025	4.390		−7.056	4.229	
S34	7.956	4.862		5.724	4.365		7.985	4.348		5.388	4.419	
S35	9.927	5.393		10.68	4.776					7.884	4.221	
S36	−8.502	4.728		−8.891	4.508		−5.132	4.340		−7.587	4.395	
S45	−7.868	4.701		−5.304	4.242							
S46	4.725	4.897		6.284	4.641		3.884	4.484		3.882	4.321	
S56	3.586	4.924										

[a] No missing crosses.

[b] Missing y_{13} y_{25} y_{56}.

[c] Missing y_{13} y_{15} y_{25} y_{35} y_{45} y_{56} (all crosses for parent 5 plus cross 13).

[d] Missing y_{13} y_{25} y_{45} y_{56} (all parent having missing crosses).

[e] ENV (environment), REP (replication).

[f] $G_1 \ldots G_6$ – GCA of parent 1 ... 6; $S_{12} \ldots S_{56}$ – SCA for pair of parent 1 and 2 ... 5 and 6.

[g] S.E. standard error of the estimates.

trials is a 6 by 6 half-diallel structure in which each family was replicated three times at each site using 4-tree row plots. The analyses were carried out for no missing crosses and 3 scenarios of missing crosses as follows:

(1). 3 missing crosses (y_{13}, y_{25}, y_{56});

(2). 6 missing crosses (y_{15}, y_{25}, y_{35}, y_{45}, y_{56}, y_{13}, the case that all crosses involving parent 5 are missing);

(3) 4 missing crosses (y_{13}, y_{25}, y_{45}, y_{56}, the case that all parents have missing crosses). The three missing cross scenarios were derived by deleting corresponding crosses from the original complete data set.

In scenario (1), all parents except parent 4 have missing crosses. Therefore, the column for parent 4 in the GCA sub-matrix was used as the linearly dependent column. In scenario (2), all crosses involving parent 5 were missing, therefore parent 5 was deleted first. After deleting parent 5, there were no more missing crosses for parents 2, 4 and 6 in the newly organised 5 × 5 half-diallel. Thus GCA for any one of the three parents (2, 4, 6) could be selected as a dependent parameter and the corresponding column deleted in the GCA submatrix. In scenario (3), all parent have missing crosses and so the design matrix is derived by manual calculation. The design matrices for the complete data set and scenarios (1) and (2) are generated by computer. In the last scenario (missing crosses y_{13}, y_{25}, y_{45}, y_{56}), there should be 5 independent SCA parameters and 6 dependent parameters. If S_{12}, S_{14}, S_{15}, S_{23}, and S_{24} are taken as independent parameters, the other parameter could be estimated as: $S_{35} = - S_{15}$; $S_{34} = - S_{12} - S_{14} - S_{23} - S_{24}$; S_{35} = $- S_{15}$; $S_{16} = -S_{12} - S_{14} - S_{15}$; $S_{26} = - S_{12} - S_{23} - S_{24}$; $S_{36} = S_{12} + S_{14} + S_{15} + S_{24}$; $S_{46} = S_{12} + S_{23}$. These relationships were input as the design matrix SCA_DN for the last scenario.

A sample output from DIAFIXED.SAS is listed in Appendix A and the raw data for the example is listed in Appendix B for readers to verify their results. The results of hypothesis testing for GCA, SCA effects and interaction effects between GCA, SCA and site effects are listed in table 1 together with the estimates of GCA and SCA effects and their standard errors for the complete data set and the three missing data scenarios. It was observed that missing a few crosses relative to the complete half-diallel design did not significantly alter the hypothesis testing for the effect of environment (site), replication, GCA, SCA, and their interactions. Site and SCA effects were significant for the complete data and for all three missing data scenarios. Although deleting crosses did alter the estimated values of GCA and SCA effects, it did not significantly affect the relative GCA ranking of parents or SCA ranking for parent combinations.

Acknowledgements

References

BRIDGWATER, F. E.: Mating designs. In: Handbook of quantitative forest genetics. Edited by FINS, L., FRIEDMAN, S. T. and BROTSCHOL, J. V., pp. 69–95. Kluwer Academic Publishers. (1992). — BURLEY, J., BURROWS, P.

135

M., Armitage, F. B. and Barnes, R. D.: Progeny test designs for *Pinus patula* in Rhodesia. Silvae Genet. **15**, 166–173 (1966). — Carson, S. D.: Genotype X environment interaction and optimal number of progeny test sites for improving *Pinus radiata* in New Zealand. N. Z. J. For. Sci. **21**, 32–49 (1991). — Dean, C. A.: Genetics of growth and wood density in radiata pine. Ph. D. Dissertation, University of Queensland (1990). — Griffing, B.: Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. Biol. Sci. **9**, 463–493 (1956). — Hallauer, A. R. and Miranda, J. B.: Quantitative genetics in maize breeding. Iowa State University Press, Ames (1981). — Huber, D. A., White, R. L., Little, R. C. and Hodge, G. R.: Ordinary least square estimation of general and specific combining abilities from half-diallel mating designs. Silvae Genet. **41**, 263–273 (1992). — Johnson, G. R. and King, J. N.: Analysis of half-diallel mating designs: I – A practical analysis procedure for ANOVA approximation. Silvae Genet. **47**, 74–79 (1998). — Jonsson, A., Dormling, I., Eriksson, G. and Norell, L.: GCA variance components in 36 *Pinus sylvestris* L. full-sib families cultivated at five nutrient levels in a growth chamber. For. Sci. **38**, 575–593 (1992). — Matheson, A. C., Spencer, D. J. and Magnussen, D.: Optimum age for selection in *Pinus radiata* basal area under bark for age:age correlations. Silvae Genet. **43**, 352–357 (1994). — SAS Institute Inc.: SAS Guide to Macro processing, Version 6. First Edition. Cary, NC. SAS Institute INC., 233 pp. (1987). — SAS Institute Inc.: SAS/STAT User's Guide. Version 6. Fourth Edition, Volume 1 and 2, Cary, NC. SAS Institute INC., 943 pp., 846 pp. (1989). — SAS Institute Inc.: SAS/LANGUAGE: Reference. Version 6. First Edition. Cary, NC. SAS institute INC., 1042 pp. (1990). — Schaffer, H. E. and Usanis, R. A.: General least squares analysis of diallel experiments. North Carolina State University, Genetics Department Research Report, Number 1, 61 pp. (1969). — Searle, S. R.: Linear models for unbalanced data. John Wiley and Sons, New York, NY. 536 pp. (1987). — Shepherd, K. R.: Report of Research Working Group 1 (Forest Genetics) to Standing Committee of the Australian Forestry Council. CSIRO Division of Forest Research, Canberra (1976). — Skroppa, T. and Tho, T.: Diallel crosses in *Picea abies*. I. Variation in seed yield and seed weight. Scand. J. For. Res. **5**, 355–367 (1990). — Snyder, E. B.: Combining-ability determinations for incomplete mating designs. USDA Forest Service, Southern Forest Experiment Station, General Technical Report SO-9 (1975). — Snyder, E. B. and Namkoong, G.: Inheritance in a diallel crossing experiment with longleaf pine. *In:* USDA For. Serv. Res. Pap. SO-140. South. For. Exp. Stn., New Orleans, LA. 31 pp. (1978). — Squillace, A. E.: Comparison of some alternative second-generation breeding plans for slash pine. *In:* South. For. Tree Improve. Conf. June 12 to 13, 1973, Baton Rouge, LA. pp. 2–13 (1973). — Stonecypher, R. W., Piesch, R. F., Helland, G. G., Chapman, J. G. and Reno, H. J.: Results from genetic tests of selected parents of Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) in an applied tree improvement program. For. Sci. **42**(2), Monograph 32 (1996). — van Buijtenen, J. P.: Mating Designs. *In:* Proc. of the IUFRO Joint Meeting of Genetic Working Parties on Advanced Generation Breeding. Bordeaux, pp. 11–27 (1976). — Weir, R. J. and Goddard, R. E.: Advanced generation operational breeding programs for loblolly and slash pine *In:* South. Coop. Series Bull. 309. Lousiana Agric. Exp. Stn. Baton Rouge, LA. pp. 21–26 (1986). — Wilcox, M. D., Shelbourne, C. J. A. and Firth, A.: General and specific combining ability in eight selected clones of radiata pine N. Z. J. For Sci. **5**, 219–225 (1975). — Yeh, F. C. and Heaman, J. C.: Estimating genetic parameters of height growth in seven-year old coastal Douglas-fir from disconnected diallels. Forest Science **33**: 946–957 (1987). — Zhang, Y. and Kang, M. S.: Diallel-SAS: A SAS program for Griffing's diallel analyses. Agron. J. **89**, 176–182 (1997). — Zobel, B and Talbert, J.: Applied Forest Tree Improvement. John Wiley & Sons, New York (1984).

Appendix A

SAS output from radiata pine example of missing three crosses (y_{13} y_{25} y_{56})

SAS Program 'DIAFIXED.SAS' To Estimate Fixed Effects For Diallel Mating Design

With Missing Crosses By Least Squares Method For Method 4

Testing And Estimating Fixed GCA And SCA Effects

General Linear Models Procedure

Class Level Information

Class Levels Values

ENV 2 M T

REP 3 1 2 3

Number of observations in data set = 281

Note: Due to missing values, only 243 observations can be used in this analysis.

General Linear Models Procedure

Dependent Variable: DBH

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	27	83200.79338778	3081.51086621	3.91	0.0001
Error	215	169418.42471922	787.99267311		
Corrected Total	242	252619.21810700			

	R-Square	C.V.	Root MSE	DBH Mean
	0.329353	14.76600	28.07120719	190.10699588

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ENV	1	46970.46793345	46970.46793345	59.61	0.0001
REP(ENV)	4	4671.07138468	1167.76784617	1.48	0.2087
G1	1	36.81051988	36.81051988	0.05	0.8291
G2	1	414.08042793	414.08042793	0.53	0.4693
G3	1	975.54051166	975.54051166	1.24	0.2671
G5	1	141.76239716	141.76239716	0.18	0.6719
G6	1	1542.30015991	1542.30015991	1.96	0.1632
S12	1	2437.59219508	2437.59219508	3.09	0.0800
S15	1	1083.78227875	1083.78227875	1.38	0.2422
S16	1	3169.43407398	3169.43407398	4.02	0.0462
S23	1	2169.30729214	2169.30729214	2.75	0.0985
S26	1	1910.60273424	1910.60273424	2.42	0.1209
S35	1	3940.56378816	3940.56378816	5.00	0.0264
G1*ENV	1	1309.49110397	1309.49110397	1.66	0.1987
G2*ENV	1	979.67332685	979.67332685	1.24	0.2661
G3*ENV	1	684.84129718	684.84129718	0.87	0.3523
G5*ENV	1	55.85787684	55.85787684	0.07	0.7903
G6*ENV	1	24.91326380	24.91326380	0.03	0.8590
S12*ENV	1	451.53185662	451.53185662	0.57	0.4499
S15*ENV	1	31.97388312	31.97388312	0.04	0.8405
S16*ENV	1	32.69507824	32.69507824	0.04	0.8388
S23*ENV	1	545.75875703	545.75875703	0.69	0.4062
S26*ENV	1	707.08720302	707.08720302	0.90	0.3446
S35*ENV	1	3895.44655514	3895.44655514	4.94	0.0272

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
GCA	5	4100.90940137	820.18188027	1.04	0.3945
SCA	6	14450.09987482	2408.34997914	3.06	0.0069
GCA*ENV	5	3652.62126081	730.52425216	0.93	0.4643
SCA*ENV	6	10345.06222451	1724.17703742	2.19	0.0453

Parameter	Estimate	T for H0: Parameter=0	Pr > |T|	Std Error of Estimate
G1	0.7708561	0.22	0.8291	3.56655077
G2	2.4463017	0.72	0.4693	3.37465014
G3	–3.8971417	–1.11	0.2671	3.50255231
G5	1.6652829	0.42	0.6719	3.92616544
G6	–4.6473583	–1.40	0.1632	3.32187123
G4	3.6620593	1.15	0.2500	3.17479969
S12	8.4935304	1.76	0.0800	4.82912901
S13	0.0000000	99999.99	0.0001	0.00000000
S14	–12.2600826	–2.80	0.0055	4.37593900
S15	–5.3766914	–1.17	0.2422	4.58463933
S16	9.1432436	2.01	0.0462	4.55900972
S23	–7.5136072	–1.66	0.0985	4.52844436
S24	5.5562412	1.16	0.2467	4.78315919
S25	0.0000000	99999.99	0.0001	0.00000000
S26	–6.5361645	–1.56	0.1209	4.19757885
S34	5.7239181	1.31	0.1912	4.36569727
S35	10.6806990	2.24	0.0264	4.77618989
S36	–8.8910100	–1.97	0.0499	4.50880651
S45	–5.3040076	–1.25	0.2126	4.24234910
S46	6.2839308	1.35	0.1772	4.64132776
S56	0.0000000	99999.99	0.0001	0.00000000

Appendix B

Appendix B. – 11 year old DBH data for a 6 x 6 half-diallel mating in radiata pine at two sites (E1 and E2).

Iᵃ	J	R	T	E1	E2	I	J	R	T	E1	E2	I	J	R	T	E1	E2	I	J	R	T	E1	E2
1	2	1	1	213	.	1	5	3	2	154	164	2	5	2	3	154	242	3	6	1	4	214	.
1	2	1	2	185	.	1	5	3	3	184	203	2	5	2	4	177	191	3	6	2	1	163	214
1	2	1	3	197	.	1	5	3	4	.	203	2	5	3	1	202	200	3	6	2	2	187	150
1	2	1	4	211	.	1	6	1	1	195	201	2	5	3	2	168	.	3	6	2	3	167	144
1	2	2	1	186	220	1	6	1	2	187	223	2	5	3	3	191	.	3	6	2	4	171	220
1	2	2	2	175	207	1	6	1	3	.	.	2	5	3	4	178	262	3	6	3	1	.	179
1	2	2	3	172	235	1	6	1	4	194	246	2	6	1	1	173	133	3	6	3	2	161	60
1	2	2	4	173	237	1	6	2	1	229	206	2	6	1	2	165	171	3	6	3	3	134	231
1	2	3	1	186	218	1	6	2	2	205	.	2	6	1	3	183	197	3	6	3	4	164	240
1	2	3	2	187	215	1	6	2	3	175	210	2	6	1	4	170	.	4	5	1	1	174	183
1	2	3	3	167	.	1	6	2	4	182	196	2	6	2	1	.	185	4	5	1	2	183	247
1	2	3	4	188	.	1	6	3	1	139	235	2	6	2	2	165	192	4	5	1	3	187	227
1	3	1	1	163	212	1	6	3	2	180	118	2	6	2	3	188	212	4	5	1	4	191	225
1	3	1	2	177	164	1	6	3	3	196	254	2	6	2	4	185	205	4	5	2	1	188	.
1	3	1	3	165	205	1	6	3	4	159	.	2	6	3	1	157	248	4	5	2	2	182	159
1	3	1	4	184	170	2	3	1	1	202	203	2	6	3	2	156	213	4	5	2	3	195	154
1	3	2	1	169	176	2	3	1	2	151	142	2	6	3	3	165	236	4	5	2	4	199	203
1	3	2	2	172	.	2	3	1	3	183	198	2	6	3	4	147	.	4	5	3	1	179	207
1	3	2	3	166	224	2	3	1	4	184	220	3	4	1	1	172	255	4	5	3	2	212	209
1	3	2	4	174	186	2	3	2	1	178	224	3	4	1	2	216	179	4	5	3	3	188	159
1	3	3	1	188	212	2	3	2	2	160	216	3	4	1	3	188	225	4	5	3	4	152	225
1	3	3	2	188	61	2	3	2	3	175	204	3	4	1	4	.	224	4	6	1	1	179	.
1	3	3	3	181	218	2	3	2	4	154	236	3	4	2	1	.	147	4	6	1	2	158	217
1	3	3	4	196	235	2	3	3	1	169	60	3	4	2	2	183	170	4	6	1	3	163	223
1	4	1	1	169	199	2	3	3	2	177	286	3	4	2	3	190	224	4	6	1	4	186	223
1	4	1	2	170	239	2	3	3	3	158	159	3	4	2	4	187	252	4	6	2	1	178	245
1	4	1	3	199	123	2	3	3	4	174	.	3	4	3	1	.	241	4	6	2	2	148	271
1	4	1	4	219	159	2	4	1	1	192	.	3	4	3	2	158	216	4	6	2	3	185	.
1	4	2	1	179	192	2	4	1	2	179	260	3	4	3	3	141	.	4	6	2	4	146	.
1	4	2	2	128	195	2	4	1	3	159	216	3	4	3	4	189	241	4	6	3	1	154	227
1	4	2	3	198	204	2	4	1	4	181	.	3	5	1	1	174	.	4	6	3	2	153	217
1	4	2	4	170	205	2	4	2	1	187	203	3	5	1	2	185	.	4	6	3	3	163	215
1	4	3	1	186	227	2	4	2	2	183	219	3	5	1	3	168	.	4	6	3	4	185	.
1	4	3	2	158	196	2	4	2	3	144	.	3	5	1	4	190	225	5	6	1	1	192	.
1	4	3	3	139	.	2	4	2	4	215	242	3	5	2	1	170	.	5	6	1	2	170	182
1	4	3	4	191	.	2	4	3	1	.	254	3	5	2	2	201	.	5	6	1	3	176	262
1	5	1	1	.	178	2	4	3	2	157	.	3	5	2	3	150	175	5	6	1	4	171	.
1	5	1	2	189	.	2	4	3	3	179	.	3	5	2	4	175	216	5	6	2	1	160	201
1	5	1	3	196	215	2	4	3	4	179	227	3	5	3	1	191	.	5	6	2	2	185	238
1	5	1	4	152	228	2	5	1	1	192	214	3	5	3	2	182	214	5	6	2	3	177	214
1	5	2	1	.	192	2	5	1	2	196	222	3	5	3	3	176	.	5	6	2	4	175	.
1	5	2	2	201	187	2	5	1	3	168	200	3	5	3	4	197	283	5	6	3	1	211	.
1	5	2	3	178	216	2	5	1	4	195	.	3	6	1	1	182	128	5	6	3	2	193	210
1	5	2	4	198	213	2	5	2	1	176	.	3	6	1	2	172	196	5	6	3	3	193	177
1	5	3	1	164	.	2	5	2	2	197	.	3	6	1	3	178	210	5	6	3	4	185	226

ᵃ) I and J are first and second parent for diallel crosses, reciprocal effect was not considered, therefore I or J could be male or female; R-replication; T-tree number within plot; E1 and E2 represent DBH (mm) data for two sites and . indicating missing value.

Genetic Subdivisions of the Range of Scots Pine (*Pinus sylvestris* L.) Based on a Transcontinental Provenance Experiment

By A. M. SHUTYAEV[1]) and M. GIERTYCH[2])[3])

Summary

Studies were continued on the variability of 113 Scots pine provenances based on an experiment established at 33 locations in the former USSR in 1974 to 1976. Following on the analysis presented earlier for height measurements (SHUTYAEV and GIERTYCH, 1997) now an analysis is made of data on survival, stem diameter and stem straightness. A synthetic volume estimate (based on height, diameter and survival) was evaluated for phenotypic stability. On the basis of growth performance in various environments the range of Scots pine in the former USSR is divided into 10 regions (A- to J) and these divisions are compared with earlier attempts at subdividing this vast area. There is agreement in the opinions about

[1]) Research Institute of Forest Genetics and Breeding, Lomonosova 105, 394043 Voronezh, Russia

[2]) Institute of Dendrology, PL-62-035 Kórnik, Poland

[3]) M. GIERTYCH is the corresponding author