GENETIC DIVERSITY AND GAIN

THE CONCEPT OF A STATUS NUMBER

A thesis submitted in partial fulfilment of the requirements for the Degree of

Doctor of Philosophy in Forestry

by

Luis Daniel Gea

School of Forestry
University of Canterbury
August 1997

ABSTRACT

A trade-off always tends to exist involving genetic gains and selection intensity, on the one hand, and the remaining effective population size (usually known as N_e), on the other. A new approach is presented and analysed for different breeding situations, using stochastic simulations, in terms of mating designs and subline sizes, guiding breeders through a new concept of *status number* (N_s) and its trade-off with gain.

Status number is defined as half the inverse of the average coancestry and depicts the current state of the population. The status number concept can easily be applied to deployment of different genotypes with unequal representation.

Breeding schemes with small breeding groups are slightly more efficient in preserving status number through multiple generations than breeding schemes with large groups. Medium- to large-size breeding groups showed a comparatively small reduction in aggregated status number over generations but showed greater increases in gain compared with small groups. Inbreeding in small elites becomes so great that it is likely to cause fertility problems and disturb selection considerably. Small breeding groups will probably not be useful for a sustainable long-term breeding strategy. Substantial benefits on status number for subdividing the population into small breeding groups will only be seen after numerous generations.

Selection schemes that maximise gain by unrestricted combined index selection will result in rapid inbreeding, and may not be sustainable in the long term. Selection procedures that place less emphasis on family information would best meet long-term diversity targets. However, gains may be too low for mating systems and selection procedures that do not include a between-family component, especially with low heritabilities. This is a good reason for using a large number of families as founders of the breeding population.

Going from selection within only 0.5 or 1 available cross per parent per generation (made equivalent to within-family selection) to 2.5 crosses per parent (restricting the number of individuals chosen per full-sib family) resulted in substantial increases in genetic gain, depending on heritability. However, increasing the number of crosses per parent up to 2.5 does carry a modest penalty of increased coefficient of inbreeding and reduced status number.

Higher levels of gain per unit of status number loss are obtained with a conservative within-family selection strategy but to reach the same level of gain more cycles of breeding will be required.

Effects of departures from assumptions (zero inbreeding coefficient and coancestry for the founders, genes being independently assorted, no mutation and interactions, or combinations from departures of the neutrality assumption) , singly and in various combinations will occur, meaning that calculations and predictions based on pedigrees will be biased. Future work will require modelling the effects for departures from the idealised assumptions and laboratory-based quantification of departures from some key assumptions.

Key words: *breeding strategies, effective population size, elite populations, genetic diversity, mating designs, simulation, status number, subline size*

A mi Madre y la memoria de mi Padre

Come epigrafi in un alfabeto indecifrabile, di cui metà delle lettere siano estate cancellate dallo smeriglio del vento carico di sabbia, così, voi resterete, profumerie, per l'uomo senza naso.

El nome, el naso
Sotto il sole giaguaro
Italo Calvino

Epigraphs in an undecipherable language, half their letters rubbed away by the sand-laden wind: this is what you will be, O *parfumeries*, for the noseless man of the future.

Translated by William Weaver

PREFACE

The structure of this thesis reflects several distinct phases of this study. During Dag Lindgren's sabbatical in New Zealand, he suggested and developed the idea of a new measure of diversity based on coancestry. Theoretical and practical uses of the new tool (status number) and the algorithm for its calculation are presented in the second chapter and Appendix 2. The logical follow-up steps (influences of population size, selection strategy and mating design on genetic gain, coefficient of inbreeding and status number) were accomplished using stochastic simulations and discussed in the third chapter. The reader will notice some repetition between Chapters 2 and 3 and also Appendix 2, which reflects the fact that these Chapters represent stand-alone papers.

Some important matters and assumptions of Chapters 2 and 3 were pulled together and are deeply discussed in Chapters 4 and 5.

There are two Appendices. The first analyses the advantages and disadvantages of creating an additional level (an elite stratum within the breeding population) in the population hierarchy on the actual breeding strategy, and points to the need for using a new tool for measuring diversity. The second reviews a little-known SAS Procedure that can calculate inbreeding coefficients for a pedigree.

LIST OF PAPERS

Chapter 2

Lindgren D., **Gea L.D.**, Jefferson P. 1996: Effective number and coancestry in breeding populations following within family selection. Silvae Genetica **45**: 52-59.

Lindgren D., **Gea L.D.**, Jefferson P. 1995: Status Number a measure of genetic diversity. In Proc. Joint Meeting of the IUFRO Working Parties S2.02.05; 06; 012 and 14. Evolution of Breeding Strategies for Conifers from the Pacific North West. Limoges, France 28 July-4th August. (Also in Forest Genetics **4**(2):69-76 1997)

Chapter 3

Gea L.D., Jefferson P., Lindgren D., Mullin T., Shelbourne C.J.A. 1995: Optimizing subline size for breeding populations. P.58 In Proc. Evolution and Tree Breeding Conference, Canadian Tree Improvement Association / Western Forest Genetic Association, Victoria, British Columbia, Canada, August 28 - September 1, 1995. (Ed. J. Lavereau).

Gea L.D., Lindgren D., Shelbourne C.J.A., Mullin T. (in prep.). The effect of population size, mating design and selection strategy on status number, coefficient of inbreeding and genetic gain.

Appendix I

Gea L.D., Shelbourne C.J.A. 1995: Gain expectations from control-pollinated main and elite breeding population and corresponding production populations. P. 876 In Proc. CRCTHF-IUFRO Conference. Eucalypt Plantations: Improving Fibre Yield and Quality, Hobart, Australia. (Awarded Third Prize in Poster Competition)

Appendix II

Gea L. D., Low C., Lindgren D. 1995: Proc Inbreed: The Shadow of forgotten ancestors. SUNZ'95. SAS users of New Zealand. 13th Annual Conference. Wellington, 18-19 September.

ACKNOWLEDGMENTS

This thesis could not have been possible without the support of a large number of people to whom I feel greatly indebted. A formal statement is scarcely indicative of my gratitude for their suggestions, assistance, guidance and thorough friendship.

I thank Prof. Geoff Sweet, who introduced me into the Biotechnology Division at the NZ FRI, guided the first steps of this research and found sufficient grants for the duration of the study. I also thank Prof. Roger Sands who took over Geoff Sweet's responsibility for supervision from PhD and since questioned continuously about the progress of my research.

I feel greatly indebted to Tony Shelbourne for his unconditional support. He has showed me the way to see beyond the numbers and has criticised and reviewed my work again and again. His patience with my written English deserves a special mention and his always-open office has been of invaluable support for the completion of this thesis. Thanks Tony!

I am very grateful to Paul Jefferson for his encouragement to complete this study and his valuable opinion, discussion and patience. Not sure that I have returned all his books and papers, this task might take longer than this thesis. He also introduced me into the dangerous field of 'Politically Correct Breeding' an area not easily covered by traditional textbooks.

Sue Carson submerged me into the obscure domains of timetables, deadlines, presentations, reports, etc. all of them as unknown as essential. She always shocked me with her red-pen comments, formulating the problems from a different perspective and visualising the challenges for the future.

Mike Carson has been balancing University demands and NZ FRI requirements until the last minute of this study and has supported and keenly criticised my research. He involved me with the GTI crew (Tony, Charles, Gerry, Toby, Danielle, Debbie, Charlie, Mark, Dean, Silvia, Susan, Jody, Robin, Melanie, Rod, Henry, Ruth, Satish, Sean, Simon) and to all of them I am very thankful.

The continuous challenge of answering Rowland Burdon's comments carried me through tons of reading material. He showed me that work is never finished because there are always many assumptions that can be violated and situations that can not be explained.

Finally, but not last, I like to express my gratitude to Dag Lindgren. He came to New Zealand for a sabbatical, running away from demanding students and he never expected to find another one here! I am greatly indebted to him and his imperative way of teaching science. I hope the friendship that started with a 'status number' will continue beyond this thesis.

The outstanding support from the University of Canterbury Doctoral Scholarship, T.W. Adams Scholarship in Forestry, and NZ FRI Doctoral Scholarship deserve distinguished mentions.

A special thanks for Victoria Mackisack and Jeanette Allen for their continuous enquires about my research and their help with every single University form.

A huge thanks to the Payn family from this virtual uncle.

Many friends have often questioned my sanity during the duration of this study, I hope to prove to them that they were wrong.

RESUMEN

Es lógico que tienda a existir un equilibrio entre ganancias genéticas e intensidad de selección por un lado y una reducción en el número efectivo de población por el otro (comunnmente conocido como Ne). Esta tésis presenta un nuevo método y analiza sus implicaciones en distintas situaciones de mejoramiento, diseños de cruzamiento y tamaño de sublineas, introduciendo al mejorador en un nuevo concepto de número estado (N_s) y los valores buscados de ganancias genéticas.

El número estado ha sido definido como la inversa de la mitad del promedio de endocría de la población y representa el estado actual de la población. El concepto de número estado puede ser facilmente aplicado a la distribución de genotipos en representación desigual. Es posible que algunas violaciones ocurran en los supuestos asumidos en el desarrollo del número estado (coeficiente de endocría y de coascendencia cero para la población inicial, mutaciones nulas e interacciones o combinaciones del supuesto de neutralidad), lo que indicaría que algunas de las predicciones y cálculos basados en los esquemas propuestos estén posiblemente sobrevaluados.

Esquemas de mejoramiento con pequeñas elites son significativamente mas eficientes en preservar el número estado a traves de un largo número de generaciones, que aquellos sitemas de mejoramiento con grupos grandes. Poblaciones elite de tamaño mediano a grande presentan una reducción en el número estado pero presentan ganancias genéticas mayores que los grupos pequeños. Poblaciones pequeñas posiblemente no constituyan una alternativa atractiva en el largo plazo. Los efectos de acumulación de mayor numero estado solo se hacen visible después de un largo número de generaciones.

Los esquemas de selección que maximizan ganancias genéticas a traves de índices de selección sin restricciones dan por resultado un coeficiente de endocría muy alto y posiblemente sea poblaciones insostenibles en el largo plazo. Métodos de selección que ponen menos énfasis en la información familiar logran mejores resultados en diversidad (medidos a traves del número estado). Las ganancias genéticas para sistemas de selección que no incluyen selección entre familias son muy bajas, particularmente cuando se tienen en consideración bajas heredabilidades. Esto constituye un buen argumento para justificar el uso de un gran número de familias en los inicios de un programa de mejoramiento.

Un incremento de 1 a 2,5 cruzamientos por padre asumiendo un sistema con restricciones en los índices de selección, resulta en un incremento substancial en los valores de ganancias genéticas (dependiendo de la heredabilidad). No obstante, el incremento en el número de cruzamientos tiene como contrapartida un incremento en el coeficiente de endocría y una reducción en el número estado.

Una estrategia conservadora, de selección dentro de las familias y nó entre familias, presenta mayores ganancias genéticas por unidad perdida de número estado, pero para alcanzar los niveles de ganancia genética de estrategias no tan conservadoras se necesitan muchos más ciclos de selección.

Palabras claves: *estrategias de mejoramiento, número efectivo de población, poblaciones elite, diversidad genética, simulaciones montecarlo, número estado, coeficiente de endocría, sublineas.*

TABLE OF CONTENTS

CHAPTER 1
BREEDING STRATEGIES AND GENETIC DIVERSITY IN FORESTRY.

CHAPTER 2
EFFECTIVE NUMBER AND COANCESTRY IN BREEDING POPULATIONS FOLLOWING WITHIN-FAMILY SELECTION.

LIST OF TABLES AND FIGURES

CHAPTER 3
THE EFFECT OF POPULATION STRUCTURE, MATING DESIGN AND SELECTION STRATEGY ON STATUS NUMBER, COEFFICIENT OF INBREEDING AND GENETIC GAIN.

LIST OF TABLES AND FIGURES

CHAPTER 4
MODEL AND REALITY.

CHAPTER 5
COMPLEMENTING INBREEDING COEFFICIENT INFORMATION WITH STATUS NUMBER: IMPLICATIONS FOR STRUCTURING BREEDING POPULATIONS.

APPENDIX I

DETERMINISTIC PREDICTIONS OF GAINS FROM CONTROL-POLLINATED MAIN AND ELITE BREEDING POPULATIONS AND CORRESPONDING PRODUCTION POPULATIONS.

RESULTS

LIST OF FIGURES AND TABLES

APPENDIX II
PROC INBREED: THE SHADOW OF FORGOTTEN ANCESTORS.

LIST OF TABLES AND FIGURES

CHAPTER 1

BREEDING STRATEGIES AND GENETIC DIVERSITY IN FORESTRY

INTRODUCTION

Tree breeding is now an accepted part of forest management not only in New Zealand but in approximately half the countries of the world. Gains in yield have come essentially through manipulating quantitative traits. Most strategies currently used in forestry have adopted the breeding method of "Recurrent Selection for General Combining Ability" (RS/GCA) to generate successive generations of a breeding population (Allard 1960; Burdon and Shelbourne 1971; Zobel and Talbert 1984; Mahalovich 1990; McKeand *et al* 1992; Shelbourne 1992; King and Johnson 1993; Loo-Dinkins 1993), and the strategies are thus designed to progressively increase in the breeding population the frequencies of genes that exhibit desirable additive effects, by means of cycles of selection and crossing.

Tree improvement has been widely organised within a framework of a population hierarchy: the gene resource population, the breeding population and the production population. This hierarchy involves increasing levels of selection and genetic improvement and decreasing levels of genetic variability and numbers of individuals (Shelbourne *et al* 1986).

Predicting genetic gains has always been an important mechanism of comparing different breeding strategies before committing resources to expensive breeding operations. Simulation is a powerful tool for efficiently comparing alternative strategies, and many of the most significant simulation studies of artificial selection systems have been carried out for domestic animals.

Deterministic simulation produces a single result, where a series of algebraic functions is used to predict the outcome for a given scenario of events. Although one might expect the term simulation to be reserved for predictive techniques where the outcome is driven by *stochastic* events, this term is still frequently used as a "buzz-word to describe any technique to represent real life events through an artificial medium" (Mullin and Park 1995). Deterministic models are useful for the rapid evaluation of selection responses from alternative strategies. The cumulative effects of stochastic variations about underlying parameter values may lead to substantial directional deviations from deterministic expectations. Stochastic simulations can also give sampling distributions about expected outcomes.

Early attempts to compare genetic gains in tree breeding were made using deterministic simulations and were mostly focused on different types of seed-orchard strategies (Namkoong *et al* 1966; Shelbourne 1969). Since then several authors have compared gains from different types of breeding and production populations using either deterministic simulations (Matheson and Lindgren 1986; Shaw and Hood 1985; Carson 1986; Cotterill 1986; van Buijtenen and Burdon 1990; Shelbourne 1992) or stochastic simulations (Mahalovich 1990; King and Johnson 1993; White *et al* 1992; Mullin and Park 1995).

Genetic gain and genetic diversity

This thesis addresses linked questions on genetic gains and measures of genetic diversity. The main issue, therefore, were to determine which procedures for selection and inter-mating generate greater gain while maintaining genetic variation and which 'tools' were the most of the appropriate to use. Genetic gains are accumulated through selection in each generation of the breeding population. When creating a new generation for recurrent selection the parents with the highest breeding values are likely to offer the greatest potential genetic gains and it is therefore worth concentrating the investment in genetic improvement on these genotypes, at least for the near future.

However, the strong artificial selection practised in forest tree breeding, which is enhanced by methods to use information from relatives, requires an evaluation of the consequences for genetic diversity. Moreover, finite populations suffer allele losses, with generation turnover; genes from some ancestors become abundant, while genes from some other ancestors become rare or lost, and with small populations such losses are potentially much greater. The process may be described as a loss of genetic diversity. The gains resulting from intensive selection are expected to balance with a reduction in the effective population size.

Effective population size has been suggested as an index for monitoring genetic diversity, but several objectionable features (described later) of this parameter for characterising genetic diversity makes it difficult to use in practice. Dr. Dag Lindgren's idea of a 'status number' as an operational measure of genetic diversity evolved during his sabbatical in New Zealand. Concepts and multiple-generation calculations have been described in **Chapter 2**. An example with a useful algorithm to calculate status number is also presented in **Appendix II** which reviews a little-known algorithm in SAS and the necessary steps for calculating status number.

Breeding programmes have moved to advanced generations of selection, and breeding uncertainties have arisen about the impact of inbreeding and narrowing of the genetic base. To guarantee the later production of unrelated genotypes, some breeding populations have been organised into sublines, and patterns for intermating individuals (called mating designs) have been developed to provide a population within which to perform the next generation of selection.

Population size within the subline has the widest range of consequences, since it has influences in both the short-term (involving selection differential, inbreeding and the reduction of genetic variation due to genetic drift) and the long term (affecting selection limits and the availability of new variation arising from mutation).

The increase in coancestry during successive generations of recurrent selection is perceived to be a major problem in long term breeding programmes. It can be delayed by restricting selection to within families, and by equal and symmetric representation of all parents and founders in mating schemes. Balanced mating designs are more advantageous for maximising effective population size and minimising allele loss than for immediate efficiency in realising genetic gains (Kang and Namkoong 1988; Burdon and van Buijtenen 1990; Dempfle 1975). Selecting predominantly within families is necessary for maintaining effective population size in the breeding population but has the disadvantages of disregarding family information, with the result that each generation it is expected to operate on only half the original additive variance (Burdon 1988).

Predicted gains, inbreeding coefficients and status numbers were compared after a number of generations with the help of stochastic simulations for different population sizes, selection strategies and a number of mating designs in **Chapter 3**. Changes in effective population size were evaluated by the new parameter (status number); this innovation was justified because conventional effective numbers express the average rate of change of the genetic base, and are not fully appropriate for studying the maintenance of effective number.

In recent years there has been interest in establishing another level in the breeding population hierarchy. In its simplest form, a subset of the large breeding population is selected and a new breeding population is established. This is what is called 'special-purpose breeds' (Carson 1986), 'nucleus breeding' (Cotterill 1986) or 'elite population' (Mahalovich 1990; White 1992). This idea came from animal breeding and has been proposed as 'nucleus breeding' for the breeding population of trees. Mahalovich (1990) found that an elite nucleus offered significant extra gain when associated with faster generation turnover for the elite breeding population. During the development of this thesis main attention is being given to the breeding population, although gains simultaneously available in the production population are also of interest. Deterministic simulations were used in **Appendix I** to quantify advantages for the elite breeding population strategy (cloned an uncloned) over the main breeding population linked with their respective seed orchards.

REFERENCES

Allard R.W. 1960: Principles of plant breeding. John Wiley & Sons Inc. New York.

Burdon R.D., Shelbourne C.J.A. 1971: Breeding populations for recurrent selection: conflicts and possible solutions. New Zealand J. For. Sci. **1**: 174-193.

Burdon R.D. 1988: Recruitment for breeding populations: Objectives, genetics and implementation. Pp. 555-572 *in* Proc. 2nd International Conference on Quantitative Genetics, (Eds Weir B.S. *et al*). Sinauer Associates.

Burdon R.D., van Buijtenen J.P. 1990: Expected efficiencies of mating designs for reselection of parents. Can. J. For. Res. **20**: 1664-1671.

Carson M.J. 1986: Control-pollinated seed orchards of best general combiners - a new strategy for radiata pine improvement. Agronomy Society of New Zealand. Special Publication No. **5**.

Cotterill P.P. 1986: Genetic gains from alternative breeding strategies including simple low cost options. Silvae Genetica **35**: 212-223.

Dempfle L. 1975: A note on increasing the limit of selection through selection within families. Genet. Res. Camb. **24**: 127-135.

Kang H., Namkoong G. 1988: Inbreeding effective population size under some artificial selection schemes. I. Linear distribution of breeding values. Theor. Appl. Genet. **75**: 333-339.

King J.N., Johnson G.R. 1993: Monte carlo simulation models of breeding population advancement. Silvae Genetica **42**: 68-78.

Loo-Dinkins J. 1993: Simulated advanced generation breeding populations: Options for coastal Douglas-Fir. Quantgen Forest Consulting. Canada.

Mahalovich M.R. 1990: Modelling positive assortative mating and elite populations in recurrent selection programs for general combining ability. PhD Thesis. North Carolina State University. 129 pp.

Matheson A.C., Lindgren D. 1986: Gains from the clonal and clonal seed orchard options compared for tree breeding programs. Theor. Appl. Genet. **71**: 242-249.

McKeand S., Bower R., Bridgwater F., Hendrickson J., McCutchan B., Mueller W., Purnell R., Weir R., Zoerb M. 1992: A Strategy for the third breeding cycle of Loblolly pine in the North Caroline State University Industry Cooperative Tree Improvement Program.

Mullin T.J., Park Y.S. 1995: Stochastic simulation of population management strategies for tree breeding: a new decision-support tool for personal computers. Silvae Genetica **44**: 132-141.

Namkoong G., Snyder E.B., Stonecypher R.W. 1966: Heritability and gain concepts for evaluating breeding systems such as seedling orchards. Silvae Genetica **15**: 76-84.

Shaw D.V., Hood J.V. 1985: Maximising gain per effort by using clonal replicates in genetics tests. Theor. Appl. Genet. **71**: 392-399.

Shelbourne C.J.A. 1969: Tree breeding methods. Forest Research Institute. New Zealand Forest Service. Technical Paper No. 55.

Shelbourne C.J.A., Burdon R.D., Carson S.D., Firth A., Vincent T.G. 1986: Development plan for radiata pine breeding. Forest Research Institute Rotorua, pp 142.

Shelbourne C.J.A. 1992: Genetic gains from different kinds of breeding population and seed or plant production population. S. Afr. For. J. **160**: 49-65.

Van Buijtenen J.P., Burdon R.D. 1990: Expected efficiencies of mating designs for advanced-generation selection. Can. J. For Res. **20**: 1648-1663.

White T., Matheson A.C., Boomsma D., Rout A.F. 1992: Logistics, costs and genetic gains of five options of nucleus breeding strategies. Technical report TR92-01 STBA Mount Gambier, South Australia.

Zobel B.J., Talbert J.T. 1984: Applied forest tree improvement. John Wiley & Sons, 505 pp.

CHAPTER 2

EFFECTIVE NUMBER AND COANCESTRY IN BREEDING POPULATIONS FOLLOWING WITHIN-FAMILY SELECTION*.

(*) A portion of this chapter was submitted to Silvae Genetica and accepted for publication in February 1996 (Silvae Genetica 45: 52-59, 1996). Status number concept and algorithm, were presented at the Evolution of Breeding Strategies for Conifers from the Pacific North West, Joint Meeting of the IUFRO Working Parties S2.02.05; .06; .12 and .14, Limoges, France 1 - 4th August 1995 and accepted for publication in Forest Genetics in May 1996. The status number concept has also been published as Progress Report 53 by the Swedish University of Agricultural Sciences, Department of Forest Genetics and Plant Physiology. ISBN 0280-7998

INTRODUCTION

Small populations suffer allele losses with generation turnover (e.g. Remmert 1994). This is a manifestation of genetic drift, which represents the stochastic changes in gene frequencies that arise when population sizes become small. The changes in population sizes and allele frequencies will be such that alleles may be lost while inbreeding and coancestry increases in small populations over successive breeding generations. Genes from some founders become abundant while genes from other founders became rare or are even lost. These developments are all part of the dynamic process of changing genetic diversity. The strong artificial selection practised in forest tree breeding, which is enhanced by methods that use information from relatives, makes it more important to evaluate the consequences for genetic diversity. The census number of the breeding population evidently gives insufficient information, so better measures are needed.

Average coancestry seems a useful measure for the status of a breeding population. It is suggested here that half the inverse of the average coancestry, which will be called status effective number (status number), is a valuable measure, and that for many purposes it is more informative than other definitions of effective population size. The concept is similar, but not identical, to one used by Burrows (1984b). Properties of the status number will be discussed and compared with other measures below. The measure is well-suited for the needs of forest-tree breeding, which is still only in the very early generations. Forest-tree breeders have seldom kept to the same mating designs over time or over the whole breeding population, and the breeding population is now often structured into sublines and elites.

Selection that allows the offspring of different parents to be differentially selected may erode diversity drastically (e.g. Wei and Lindgren 1995). Methods proposed to delay such loss of genetic diversity are aggressive use of inbreeding (Lindgren and Gregorius 1976), small breeding groups and equal representation for different families (Wei and Lindgren 1995). Choosing population replacements so that families (or rather parents) are equally represented reduces diversity as little as possible (Dempfle 1975). It is therefore of great interest to study how well diversity is preserved over generations in which each parent makes an equal contribution to the following generation, as the most efficient schemes for long-term preservation of genetic diversity can be expected to be found among such schemes.

The objectives of this study are to introduce a new tool to measure genetic diversity that will be called status number; to develop the mathematics needed for deriving status number, and to investigate how status number is influenced by choice of breeding schemes, in particular those that are expected to preserve diversity. A general algorithm and numerical examples to calculate the status number are included.

Theory

Definitions, symbols and equations are based as far as possible on those of Falconer (1989). Populations are viewed from the forest tree breeder's point of view, but much of the algebra should to be applicable to other situations.

Definitions

The breeding population is the set of genotypes that are intermated so that some of their genes are transmitted to the next generation. The concept of population is ambiguous, sometimes it refers to certain individuals and sometimes it is regarded as a variable. In this study the term generally refers to certain coexisting individuals in the same generation. In a breeding programme, membership of the breeding population for existing individuals can be considered as a matter of definition.

Coefficient of inbreeding is the probability that any pair of alleles in an individual are identical by descent.

Coancestry (or coefficient of kinship) is the probability that genes sampled from different (or the same) individuals are identical by descent. An equivalent definition is the coefficient of inbreeding that would exist after intermating or self-fertilization.

Average inbreeding is the average coefficient of inbreeding over a population.

Average coancestry is the average coancestry over all pairs of population members (including individuals with themselves). Another formulation is the probability that any two genes at a locus sampled with replacement from the breeding population are identical by descent.

Note that the concepts inbreeding and coancestry require that a reference base population is defined, in which (by definition) no inbreeding or relatives occur.

Diploid genotypes are assumed, thus each individual carries two homologous genes at a particular locus.

Notation

The following designations are used (following Falconer 1989):

N	Size of the breeding population (population size, census number).
N_e	Effective population size. Variations of this concept, according to different definitions are denoted by different characters appended to the subscript.
F_X	Coefficient of inbreeding for individual X.
F	Coefficient of inbreeding (average coefficient of inbreeding).
f_{PQ}	Coancestry between individuals P and Q.
f	Coancestry (average coancestry).
t	Time in number of generations. If t is used as a subscript it gives the generation of the breeding population. Note that sometimes entities in the same equation may refer to different generations, in which case the respective generations are specified in separate subscripts. For the base population t = 1. Note that whereas the founders are regarded here as comprising generation 1, in some other studies they have been regarded as generation 0.
m	Number of individuals that form a sib group.

The concept of status effective number

Status effective number, N_s, is defined as:

$$N_s = 0.5/f \tag{2.1}$$

where f is the average coancestry of the population (considering also the coancestry of individuals with themselves). It may be abbreviated as "status number".
In words, N_s may therefore be defined as "half the inverse of average coancestry".

The relative status number, N_r, is defined by

$$N_r = N_s/N. \tag{2.2}$$

It can thus provide a useful measure of how status number has been eroded during the history of the population.

A more pragmatic definition of N_s may be formulated: "the status number of the population is the number of unrelated and not inbred genotypes in an ideal panmictic population, which would produce progeny with the same coefficient of inbreeding as the offspring of the genotypes of the population following random mating". (Note that selfing and mating with relatives are allowed).

Other uses of status number

Status number may be used to describe the genetic diversity for other populations than the breeding population. In forestry the status number concept could be applied to a seed orchard crop, the forest holding of a company, the forests in a nation planted in a certain year, or clonal plantations. In principle, it can be always computable and represent very meaningful information.

Consider a simple example of the calculation of status number for a population of three unrelated individuals (Table 2.1). The coancestry of an individual with the other individuals is zero. The coancestry of each individual with itself is 0.5. The average coancestry in Table 2.1 is 1.5/9 = 0.16667. Therefore the status number of this population is :

$$\frac{0.5}{0.166667} = 3$$

If a seed orchard composed of equal numbers of clones, with each clone of equal fecundity, was in perfect panmixis, the status number of the population of orchard clones would be the number of unrelated individuals which under the same assumption of perfect panmixis, would produce offspring with the same average coefficient of inbreeding as the seed orchard progeny.

Table 2.1: Coancestry matrix for a population of three unrelated individuals, with equal frequencies of all nine mating.

Individuals	**1**	**2**	**3**
1	0.5	0.0	0.0
2	0.0	0.5	0.0
3	0.0	0.0	0.5

Equations for coancestry and inbreeding - general case

Consider the pedigree in Fig. 2.1.

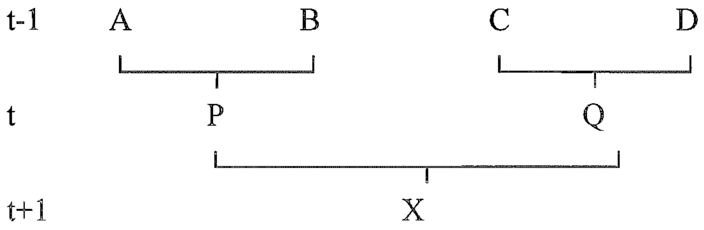


Figure 2.1: Pedigree showing the relationships between the individuals P and Q in generation t and their parents and progeny.

Basic relationships and assumptions

Note that coancestry values are commutative, thus $f_{PQ} = f_{QP}$.
The following basic relationships apply:

$$\begin{aligned} f_{PQ} &= 0.25(f_{AC}+f_{AD}+f_{BC}+f_{BD}) = F_X; \\ f_{AA} &= 0.5(1+F_A); \\ f_{PP} &= 0.5(1+F_P) = 0.5(1+f_{AB}) \end{aligned} \qquad \textbf{(2.3)}$$

In the founding population (base population, generation 1), A, B, C and D are not inbred and not related; and the following values apply:

$$F_1 = F_A = F_B = F_C = F_D = 0;$$

$f_{AB} = f_{CD} = 0;$
$f_{AA} = 0.5$ and
$f_t = 0.5/N_S$ **(2.4)**

Structure of coancestry and inbreeding

Knowing the pedigree, the coancestry for any pair of members of the breeding population and their average can be calculated. There are no theoretical problems with retrospective calculations of average coancestry, and the calculation problems are manageable with modern computers and algorithms. However, it is also desirable to make forecasts and predictions. In the following, predictive formulae are derived. Sometimes these are dependent on assumptions, which will be discussed here.

After some generations predictions of coancestry usually become complicated, because asymmetries appear between the members of the breeding population, in that they tend to be differently related to each other and inbred to varying degrees. All pairs of individuals in a generation, when needed, are assumed to be equally related to their full-sibs and to their half-sibs and to be equally inbred with coefficient of inbreeding F_t. This is correct for the founding first generation (t = 1). It is also correct for simple symmetric breeding systems like repeated selfing and repeated full-sib mating. There may be some other breeding structures for which it is correct, but generally it is not. There are, however, good reasons to believe that the lack of symmetry in real populations will not cause large errors in the numerical calculations as:

- Many of the equations involved are linear, such that the effects of symmetrically distributed departures from the simplifying assumptions will tend to cancel out.
- The average measures concerned are based on many individual values, and thus individual variations will not be important.
- The applied calculations will deal with pure within-family selection applied in a highly symmetric fashion.
- In a real multi-generation breeding program probably other considerations would be made when assigning breeding mates.

For calculations in non-symmetric cases more elaborate algebra than developed in this study would be required (cf. Burrows 1984b, Chesser *et al.* 1993). Note that the assumption of symmetry will sometimes impose a requirement that N is a multiple of 2.

In cases addressed here it is assumed that the breeding population is cycled from one generation to the next in discrete steps, and no distinction is made between males and females. As mentioned above, a base population comprising N genotypes is assumed, these genotypes being not inbred ($F_1 = 0$) and not related ($f_1 = 0$), and in genetic equilibrium (Hardy-Weinberg equilibrium), unless stated. (These assumptions are convenient, but may often be relaxed).

The members of the base population may be called founders. All genes are assumed to be strictly neutral with respect to selection and not to affect the probability that their carriers will belong to the breeding population. Diploid genotypes are assumed , such that each individual carries two homologous genes at a particular locus.

Although sexual propagation is considered in this study, the concept is easily extended to addressing the genetic diversity of populations comprising clones, which is an advantage with the concept.

Derivations of coancestry and inbreeding for specific mating systems

There are seven distinct possible types of relationships applicable to the pedigree in Fig. 2.1. Four of these comprise pedigrees including selfing (B = C = D = A; B = A & D = C; B = C = A; B = A). Of these only the first type (self-sibs) will be considered. The three more conventional types of relationships, which do not involve selfing in their pedigrees, are full-sibs (C = A & D = B), half-sibs (D = A) and non-sibs (i.e., unrelated).

Consider a breeding population of size N, where P belongs to a family with *m* members. The average coancestry for an individual (P) is the sum of the contributions from this individual and all individuals in the breeding population. The breeding population can for this purpose be seen as composed of three parts, the individual itself, its sibs and its non-sibs relatives (cousins, etc). The contribution from these components will be:

$$Nf_t = f_{PP} + (m-1)*(\text{coancestry for sibs to P}) + (N-m)*(\text{coancestry for other than sibs to p}) \qquad \textbf{(2.5)}$$

If all individuals in the breeding population have equivalent pedigrees the average coancestry over the whole breeding population will be equal to that for P. Note that one prerequisite for equivalent pedigrees is that families are of equal size. As the population is symmetric all members of a certain generation of the breeding population have equal inbreeding $F_B = F_A = F_{t-1}$ and all pairs that are not sibs have the same coancestry. In the following it is important to consider the generation. In general, average coancestry for one generation is derived as a function of particular coancestry relationships in the previous generation.

Coancestry for self-sibs

For many plant species, including many commercially important conifers, selfing is possible and usually results in viable plants, which have a potential value in breeding. Sibs from selfed parents, which are called self-sibs, form a uniparental progeny and are more closely related than ordinary full-sibs (Fig. 2.2).

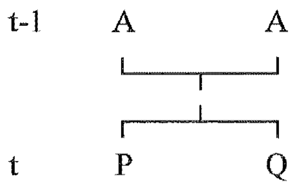


Figure 2.2: Pedigree showing the relationships between self-sibs.

The coancestry for a pair of self-sibs (cf Fig. 2.2) in a monoecious species will be:

$$f_{PQ} = f_{AA} = F_X \quad \textbf{(2.6)}$$

and the inbreeding following mating of self-sibs:

$$F_P = f_{AA} = 0.5(1+F_A) \quad \textbf{(2.7)}$$

A recursive relationship can be developed:

$$\begin{aligned} &\text{let } F_A = F_{t-1} \text{ and } F_P = F_t \\ &\text{then } F_t = 0.5(1+F_{t-1}) \end{aligned} \quad \textbf{(2.8)}$$

Consider a breeding population in generation t comprising a family with m self-sibs. Average f_t considering the individual P as a member a family of m self full-sibs will be:

$$\begin{aligned} mf_t &= f_{PP} + (m-1)f_{AA} \\ &= 0.5(1+0.5(1+F_{t-1})) + 0.5(m-1)(1+F_{t-1}) \\ &= 0.25 + 0.5m + F_{t-1}(0.5m-0.25) \end{aligned} \quad \textbf{(2.9)}$$

For the first progeny generation ($t = 2$):

$$\begin{aligned} mf_t &= 0.25 + 0.5m \\ N_s &= m/(m+0.5) \end{aligned} \quad \textbf{(2.10)}$$

N_s depends more on the number of unrelated families than it does on the number of sibs in a family.

$$N_r = m/(m+0.5) \qquad \textbf{(2.11)}$$

$$f_{PQ} = F_X = 0.5 \text{ and } F_P = 0.5.$$

Repeated selfing

Consider N selfing sublines each with $m = 1$ family members. The coancestry between individuals belonging to different sublines is 0. The average coancestry and inbreeding will be:

$$f_t = 0.75+0.25F_{t-1} \text{ and} \qquad \textbf{(2.12)}$$

$$F_t = 0.5(1+F_{t-1}) \text{ for } t>1. \qquad \textbf{(2.13)}$$

Constraints on matings

Mating is controlled such that selfing does not occur in the breeding population. The members of the breeding population form pairs at random independently of relationship between them. Note that $F_t = f_{AB}$ if A and B are mating partners (cf. Fig.2.1), thus the average coancestry over all pairs in the previous breeding generation (excluding selfing) may be denoted F_t.

Coancestry for full-sibs

Consider a pair of full-sibs (Fig. 2.3).

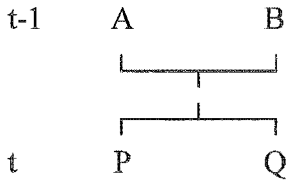


Figure 2.3: Pedigree showing the relationships between full-sibs.

The coancestry for a pair of full-sibs will be:

$$f_{PQ} = 0.25(2f_{AB} + f_{AA} + f_{BB}) = F_X;$$

$$\text{if } F_B = F_A = F_{t-1} \text{ and } F_X = F_{t+1} \text{ then}$$

$$f_{PQ} = F_{t+1} = 0.25(1 + 2F_t + F_{t-1}) \qquad \textbf{(2.14)}$$

Consider a breeding population comprising a single family with m full-sibs:

$$mf_t = 0.5(1 + f_{AB}) + 0.25(m-1)(2f_{AB} + f_{AA} + f_{BB})$$

$$mf_t = 0.5(1+F_t) + 0.25\,(m-1)(2F_t + 2*0.5(1 + F_{t-1})) \qquad \textbf{(2.15)}$$

$$\begin{aligned} &\text{If } F_{t-1} = F_t = 0 \text{ then } mf_t = 0.25 + 0.25m; \\ &N_s = 2m/(m+1) \text{ and} \qquad \textbf{(2.16)} \\ &N_r = 2/(m+1) \qquad \textbf{(2.17)} \end{aligned}$$

Repeated full-sib mating

Consider a breeding population of size N composed of unrelated sublines with repeated full-sib mating (Fig. 2.4)

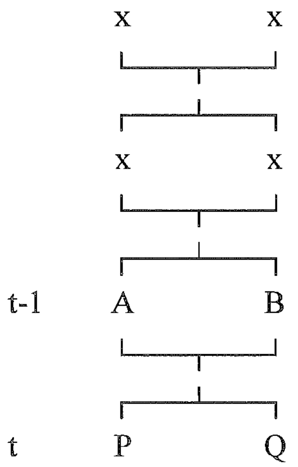


Figure 2.4: Pedigree for repeated full-sib mating.

Use equation **2.15** with $m = 2$.

$$\begin{aligned} mf_t &= Nf_t = 0.5(1 + F_t) + 0.25\,(2F_t + (1 + F_{t-1})) \\ &= 0.75 + F_t + 0.25F_{t-1} \end{aligned} \qquad \textbf{(2.18)}$$

Single-pair mating with random pairing of parents

Consider a breeding system where in each generation N genotypes are arranged in pairs. Each pair becomes parents to a full-sib family. From each full-sib family two individuals form the next generation of the breeding population (Fig. 2.1). The average f_t is calculated using equation 2.5 with $m = 2$.

$$\begin{aligned} Nf_t &= 0.5(1 + f_{AB}) + 0.25(2f_{AB} + f_{AA} + f_{BB}) + (N-2)0.25(f_{AC} + f_{AD} + f_{BC} + f_{BD}) \\ &= 0.5(1 + F_t) + 0.25(2F_t + 1 + F_{t-1}) + (N-2)F_t \\ &= 0.75 + (N-1)\,F_t + 0.25F_{t-1} \end{aligned} \qquad \textbf{(2.19)}$$

A recursive formula for F in the offspring of the parent pairs is needed. This is dependent on restrictions. It can be derived using $F_{t+1} = F_X = f_{PQ}$.

Consider a situation allowing for sib-mating but not selfing, when the inbreeding in the next generation can be derived from the average of coancestry considering the terms of equation 2.19 corresponding to when mating actually may occur.

$$\begin{aligned} F_{t+1}(N-1) &= 0.25(2F_t + 1 + F_{t-1}) + (N-2)F_t \\ &= 0.25 + (N-1.5)F_t + 0.25F_{t-1} \end{aligned} \qquad \textbf{(2.20)}$$

Coancestry for half-sibs

Consider a pair of half-sibs (Fig. 2.5).

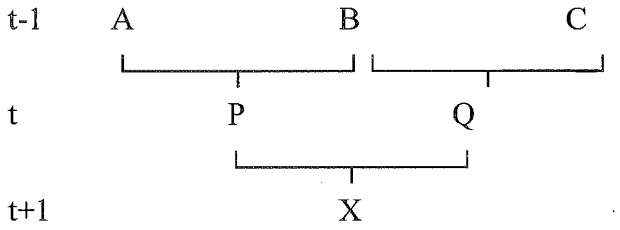


Figure 2.5: Pedigree showing the relationships between half-sibs.

The coancestry can be expressed (B is the common parent):

$$f_{PQ} = 0.25(f_{AC} + f_{BC} + f_{BA}) = F_X$$

Consider a breeding population comprising a single family with m half-sibs

$$\begin{aligned} mf_t &= 0.5(1+f_{AB}) + 0.25(m-1)(f_{AC} + f_{BC} + f_{BA}) \\ mf_t &= 0.5(1+F_t) + 0.25(m-1)(3F_t + 0.5(1+F_{t-1})) \end{aligned} \qquad \textbf{(2.21)}$$

If $F_{t-1} = F_t = 0$ $(t = 2)$ then $f_{PQ} = F_X = 0.125$;

$mf_t = 0.375 + 0.125m$

$N_s = m/(0.25m + 0.75)$, and **(2.22)**

$N_r = 1/(0.25m + 0.75)$ **(2.23)**

Note that in the present circumstances half-sib families introduce one more existing founder than a full-sib family, and the number of founders is dependent on family size, each member of a half-sib family is associated with an increase of the number of founders.

Double-pair mating with random pairing of parents

Consider a breeding system where in each generation N genotypes are arranged so each genotype is mated to exactly two different mates. From each full-sib family one individual is accepted as a member of the next generation of the breeding population, in that way a new breeding population generation of size N is recruited. The average f_t is calculated using equation 2.3 for half-sibs noting that m = 3 (P with two half-sibs).

$$\begin{aligned} Nf_t &= 0.5(1+f_{AB}) + 2*0.25(f_{AC}+f_{AA}+f_{BC}+f_{BA}) + (N-3)0.25(f_{AC}+f_{AD}+f_{BC}+f_{BD}) \\ &= 0.5(1+F_t) + 0.5(3F_t + 0.5(1+F_{t-1})) + (N-3)F_t \\ &= 0.75 + (N-1)F_t + 0.25F_{t-1} \qquad \textbf{(2.24)} \end{aligned}$$

A recursive formula for F in the offspring of the parent pairs is required. This is dependent on restrictions. It can be derived using $F_{t+1} = F_X = f_{PQ}$. If the parent pairs are formed at random.

$$\begin{aligned} F_{t+1}(N-1) &= 0.5(3F_t + 0.5(1+F_{t-1})) + (N-3)F_t \\ &= 0.25(1+F_{t-1}) + (N-1.5)F_t \qquad \textbf{(2.25)} \end{aligned}$$

Note that there is no difference between single-pair mating and double-pair mating as far as the accumulation of coancestry and inbreeding is concerned.

An ideal random mating population

An "ideal" random mating population is useful as a benchmark. Such a population is different from the populations discussed above, in which mating was also at random, but selfing excluded and the progeny size kept constant. In an "ideal" random-mating situation (Falconer 1989), selfing is allowed and the progeny size is Poisson-distributed (if the population is large, else, binomial) among parents.

In an ideal population the average coancestry in one generation will be the average coefficient of inbreeding in the next, thus:

$$F_t = f_{t-1}.$$

After t generations of random mating where for the base population is t = 1 the forecast inbreeding and coancestry is

$$F_{t+1} = f_t = 1-(1-0.5/N)^t \qquad \textbf{(2.27)}$$

(cf. Falconer 1989, note that Falconer regards the founders as t = 0).

If t/N is small, then $F_{t+1} = f_t \approx 0.5t/N$ and

$N_r \approx 1/t,$

In particular after the first generation (t = 2) $N_r \approx 0.5$.

RESULTS

Values for inbreeding, average coancestry and status number have been calculated for a number of situations (Tables 2.4-2.8). Formulae given in Tables 2.4 to 2.8 are easily derived from the theory section above. Similar inbreeding calculations have often been presented (eg. Falconer 1989, Kang and Namkoong 1988), but similar calculations for coancestry have not been presented and status number is a new concept.

The influence of family size

The status number of a single family with parents that are not inbred or related is shown in relation to family size in Table 2.2. These values can give an indication of the number of offspring to be included in the breeding population from different families. If a single selection is made from each family, the status number will be lower for self-sibs than for other families. This is because inbred individuals may carry identical allele copies in both their homologous genomes. If several selections are made per family, half-sib families are able to provide a higher status number. This is because each selection from such families introduces new founders into the breeding population; however, this is only relevant for the first breeding cycle.

Systems with regular inbreeding in small sublines or breeding groups

The average inbreeding, average coancestry and relative status number were calculated over a number of generations for repeated selfing and full-sib mating (Table 2.3). Note that for repeated selfing $F_t = f_{t-1}$ and for repeated full-sib mating $F_t = f_{t-2}$. The status number decreases quickly in each of these small sublines towards the minimum 0.5. Note that the relative status number (N_R) is independent of N.

Systems that let each parent be represented by two progenies in large sublines or breeding groups

The development of status number over generations is calculated for a breeding scheme where each parent contribute two offspring to the next generation. This offspring are mated at random, excluding selfing (Table 2.4). This may be called restricted panmixis.

The relative status number (N_R)is almost independent of the population size down to N = 4, but it appears that at very low population sizes such as N = 2 or N = 1 (Table 2.3) that the small population size really influences the evolution of status number over generations, maintaining N_R relatively well.

The decrease in relative status number over generations is delayed if sublines are small (Table 2.4). Thus small sublines are a way of reducing the erosion of status number. The difference between small and large sublines increases with generations. However, the difference between sublines of size four and sublines of infinite size does not become important before generation 10, and the difference cannot be considered as a viable argument for sublines of larger size than four. Inbreeding becomes substantial after a few generations. Inbreeding will become a problem with small sublines even if the increase of inbreeding by generation is minimised.

Random mating populations

In random mating (panmictic) populations (Table 2.5) coancestry and inbreeding accumulate more quickly over generations and the relative status number is lower than for populations where parents contribute equally (Table 2.4).

Comparison among measures of effective number in simple populations

A comparison was made between status number and conventional effective number in two simple populations (Table 2.6). The size of the breeding population was kept constant at 100. Individuals for producing the next generation were chosen either completely at random (thus random number of offspring per parent) or by selecting exactly two offspring per parent (thus maximising N_e and minimising loss of N_S). These individuals were mated at random (selfing excluded) and a new offspring generation was generated. Two different approaches were simulated. The first one was Monte Carlo simulation using the breeding simulator POPSIM® (Mullin and Park, 1995) and let the programme calculate how status number and inbreeding effective number developed over time 100 times and presented the averages (marked "sim" in Table 2.6). The second method was using the deterministic equations presented above (marked "pred" in Table 2.6).

For conventional effective number Wright's 1931 equation was used;

$$Ne = \frac{(4N-2)}{2+\sigma_k^2} \qquad \textbf{(2.28)}$$

Note that σ_k^2 is the variance in offspring between parents, which is 0 for constant offspring and 2 for random size offspring (maintaining constant N).

There is a remarkably good agreement between predictions and average outcome in actual populations for N_S, but the conventional number disagreed drastically from the status number values. The traditional effective size is not defined for situations when the inbreeding does not increase.

DISCUSSION

Alternative concepts: Status number versus other effective number concepts

The concept "effective population size" was recently reviewed by Caballero (1994). It was introduced by Wright (1931). The concept of an ideal population was introduced. The ideal population is characterised by distinct generations, constant size, reproductive isolation and random mating (including selfing) (Falconer 1989, p 48). Traditional effective population size (effective number) was defined as the number of individuals that would give rise to either the calculated sampling variance, N_{eFV}, or rate of inbreeding, N_{eFI}, that they would have if they bred in the manner of the idealised population (Falconer 1989). The "variance" and "inbreeding" definitions often give the same but sometimes different results. These concepts will be called "traditional". Inbreeding definition is considered as most conventional, and use it for comparative purposes; most of the difficulties encountered with the "inbreeding" definition arise also for the "variance" definition. Note that "effective population size" was defined as a measure of rate but not of state.

There are several inherent problems in the traditional concepts. The concept of the population is ambiguous in itself. It is often used in a sense that its members can change over time, although sometimes it refers to a snapshot describing the situation at a certain moment. The traditional measures describe a process. But the term "the effective size of the breeding population" implies the state of the breeding population at a certain moment rather than the mating tactics used or the resulting dynamics. The same actual breeding population can be associated with widely different traditional effective numbers depending on its assumed future mating pattern. The traditional concepts use an ideal population as a reference. Random mating is a dubious benchmark, as progenies of a real population can be both more or less variable than progenies from a random mating population, and N_{eFV} and N_{eFI} may be both bigger and smaller than the real population size. It is not appealing to use an effective number that can be much bigger than the census number. One may well question whether a comparison with the behaviour of an "ideal", unnatural, very hypothetical and inherently very unrealistic population is the most important aspect of a mating system.

Some of the features of the two concepts are illustrated by numerical examples in Table 2.6. Shortcomings of traditional measures have been pointed out by several researchers, e.g. Chesser *et al.* (1993).

Desirable characteristics for a concept of an effective number are listed in Table 2.7. The status effective number, N_s, meets most of these requirements, and for symmetric situations all demands are met. As a comparison, inbreeding effective number is used, N_{eFI}, as it was defined by Falconer (1989). The judgments implied in some of the comments may reflect the specific interest of tree breeders.

Since traditional concepts of effective population size have some rather awkward properties when used for characterising the state of populations, status number would be useful for many purposes.

An effective population number concept should imply that its size gives a relevant intuitive message about the current state of the population. This aspect is important because the depletion of biodiversity is, apart from the technical significance, a very political issue, and the effective number may have implications far beyond science.

Direct comparisons between status number and the inbreeding effective population size are now possible using the stochastic simulator POPSIM (Mullin & Park 1995) which can calculate both numbers for simulated breeding operations.

Breeders, geneticists and ecologists want to discuss distinct populations and changes over generations. To use a single number, which must be some sort of average, means that details are lost, and its nature of an approximation means that it can go badly wrong in some situations. The traditional measures are dependent on what specific interval of generations are considered. The comparison with an ideal population means that problems arise when overlapping generations of variable length are handled with traditional effective number.

The concept N_{eFI} is based on the rate at which inbreeding accumulates, thus change between generations. What is important in that connection is the relative change of inbreeding, F (Falconer 1989), which is a function of the mating system. Thus, what is called "inbreeding effective population size" actually has very little to do with the actual state of the population and much more with the mating systems. As long the mating system is constant, the traditional effective population size does not change much with time, even if the population has become almost completely homozygous. When $F = 0$, inbreeding effective population size is infinite (or undefined). As long as a single pair of unrelated individuals exist in a breeding population, however, the rate of inbreeding can always be put to zero and thus N_{eFI} can be made arbitrarily high.

There can be situations where N_s is lower in one generation than in the subsequent, although this happens only in highly asymmetric situations with sublines of different size. Situations in which inbreeding effective number is lower in one generation than in the following are much more frequent.

There is a close link between status effective number and previous definitions of effective population size considering a single generation turnover. Burrows (1984a) made a definition of effective size that is similar to the status number (denoted here as N_{eB}); he defined effective number as:

$$N_{eB} = 1/rQ \qquad \textbf{(2.29)}$$

where 0.5 r Q is the average pairwise coancestry of the selected group (excluding self-pairings). It is therefore the average inbreeding coefficient of progeny obtained by random mating among selected individuals (excluding self-matings). The key difference is that status number considered the coancestry of the individual with itself for calculation of average coancestry. If the coancestry of the individual with itself is excluded, N_{eB} will be infinite (or undefined) if there are no relatives, and N_{eB} can easily exceed the census number. Calculations by Wei and Lindgren (1995b) (*cf.* Table 2.1) demonstrate some undesirable features with N_{eB} like, for example, that the effective number may sometimes be increased by decreasing census number.

Chesser *et al* (1993 eq. 53) made a complicated definition of a "coancestral effective size" with limited scope to be used for intra-group gene correlations, which can be derived from the above formulation.

Robertson (1961) assuming an ideal population, defined effective population size as:

$$N_{eR} \;=\; N^2/\,m_j^2 \qquad \textbf{(2.30)}$$

where m_j is the number of individuals selected from family j (or expected contribution) and N = $\Sigma\, m_j$. A better interpretation may be that m_j is the number of contributions of parent *i*. N_{eR} as specified is independent of the type of families and whether families are related and inbred or not, and thus is not suitable as a general measure.

Robertson's definition has been quoted by Wei and Lindgren ($1995_{a,b}$), Kang and Namkoong (1988) and many others for analysing effects of selection schemes that generate differences in the expected size of progeny groups. When applied to unrelated full-sib families N_{eR} is equivalent to status number, except for a scaling factor. For this case it can also be interpreted in a restricted sense as equivalent to both variance effective number and inbreeding effective number (Kang and Namkoong 1988), and is thus an appropriate measure for unrelated full-sib families that are not affected by inbreeding.

In many situations the inbreeding coefficient may be sufficient, but when the breeding population is divided into unrelated sublines, which have zero gene exchange, the inbreeding coefficient alone is an insufficient descriptor for population changes in gene frequency and status.

Burrows ($1984_{a,b}$) formulated average coancestry values for a single cycle of selection for phenotypic and combined index selection for different regular mating plans, which have been extended here to multiple generations.

Conservation of rare alleles

A feature of interest in the long-term management of a population is the ability to preserve rare alleles. The coancestry measures whether two alleles are identical by descent as a result of the pedigrees linking the founders with the current situation. Thus **N_s carries information on the probability that alleles from the founders still remain in the population**. There should be an approximately 50% probability that an allele with initial frequency $0.5/N_s$ still remains, and the chance that an allele with initial frequency 1/N still remains ought to be approximately N_r.

Decline of the status effective number

The decline of status number by generation is illustrated in Figure 2.6. There is a conceptual problem with status number in that it declines quickly in early generations. This may be an artifact created by the concept of an initial population with unrelated founders, whereas in a way members of a species must, by definition, be considered to be related. A numerical example may be illuminating.
The status number drops by approximately half following a single generation of random mating (Table 2.5), yet it takes approximately 3 or more further generations to reduce the status effective number to half of the value it has at generation two, and five or six further generations to halve again. This indicates that the rate of decay is very dependent on the definition of the founding population. This restriction may make the concept much less suitable for situations where no well-defined starting point exists.

From this point of view, the measure may, however, be very suitable for forest tree breeding, where only a few well-pedigreed generations have passed since domestication started by composing breeding populations from the whole range of often large natural populations. A general goal of gene conservation can be formulated as keeping the accumulation of coancestry at a low level, and thus to keep the status number high. Three tools for this purpose are demonstrated in the tables provided. The first one is to make the contributions to the next generation as equal as possible among parents, the second is to keep sublines of the breeding population small, and the third one is to encourage early inbreeding.

Random contributions to the next generation with an average of two offspring per parent, to mantain constant N, (Table 2.5) resulted in considerably lower relative status numbers (N_r) than if there are two offspring per parent (Table 2.4) regardless of whether the offspring are full-sibs or half-sibs. The breeding population needs to be approximately 75% larger if the same status number is to be obtained after 10 generations under random mating compared with a population in which each parent contributes two offspring to the next generation (Table 2.6).

It seems possible therefore, to conserve status number considerably better in artificial breeding populations than in random mating populations.

The relative status number (N_r) is increased by decreasing the size of the sublines, but this increase is minor unless the subline size is small (Tables 2.6 and 2.7). Thus, structuring a population in large sublines (N>10) is not an efficient strategy for preserving status number. The dependence of status number on population size increases by generations, and is smaller following equal contributions rather than random mating.

Extremely small subline sizes associated with strong inbreeding (N = 1 or 2) are, in principle, able to preserve a high relative status number (Table 2.3). Sib-mating is more efficient for gene conservation than mating with less related individuals, and by a program of repeated full-sib mating a breeding population of less than half the size is required compared with random mating. To avoid high coancestry without expanding the size of the breeding population repeated selfing is, theoretically, by far the most efficient scheme. However, inbreeding depression would very often make schemes with close inbreeding impracticable (William and Savolainen 1996).

Status number is reduced even if there is a perfectly balanced mating design, whereby the members of one generation contribute equally to the next. Thus even though within-family selection (Dempfle, 1975) is the most conservative strategy, it does not prevent a rapid erosion of status number. Two-thirds of the status number will be lost in most programmes within a few breeding generations. This emphasises the large cost in increased coancestry when a closed population is cycled.

Use of half-sib families can be an efficient tool for representing more founders in the breeding population at an early stage (Table 2.2); however, this works only as long as new, fully unrelated recruits are brought in, and it is not applicable, for example, if a polycross design is used (as is typical of some tree breeding programmes).

Overlapping generations

It is a less than ideal feature of status number that it frequently becomes lower if descendants are pooled with their ancestors. This is because the ancestors genes tend to be unequally represented in later generations. However, if it is possible to make substantial progress, ancestors cannot be regarded as members of the breeding population for very long and it is not meaningful to create more progeny from them for long term breeding. Thus this somewhat paradoxical behaviour of status number led us to consider a philosophical question rather than a real problem: For how long should the ancestors be consider as part of an active breeding population?

The neutrality assumption

One may ask about the evolutionary significance of this rapid decrease in status number. Counter-balancing forces must be important in affecting maintenance of genetic diversity, such as mutation and disruptive selection. In the real world it is likely that the departure from the assumption that genes are neutral would lead to many calculations and predictions based on pedigrees giving misleading predictions. This is a disadvantage for traditional effective numbers as well as for status number.

If selective forces are at work some genes may make it more likely that the genotypes carrying them are selected, and other genes may be linked with those. Chromosome fragments from ancestors carrying such genes may be more common than indicated by the pedigree, and that would increase coancestry. Such effects may be of importance even for within-family selection.

It may be noted that natural tree populations seem to have $F = 0$ as some sort of natural base-line for the breeding population. It suggests that in one sense the founder concept is adequate. It does not, however, ensure that $f=0$ in a set of founder parents. The excess of homozygotes formed by selfing and mating between relatives seem to die off between fertilisation and maturation. This indicates that the assumption of neutrality may often be badly fulfilled.

The introduction of molecular markers makes it possible to actually measure coancestry and compare the predictions and measurements (Bernardo, 1993). Such comparisons will probably help to identify the situations when predictions are unreliable.

REFERENCES

Bernardo R. 1993: Estimation of coefficient of coancestry using molecular markers in maize. Theor. Appl. Genet. **85**: 1055-1062.

Burrows P.M. 1984a: Inbreeding under selection from unrelated families. Biometrics **40**: 357-366.

Burrows P.M. 1984b: Inbreeding under selection from related families. Biometrics **40**: 895-906.

Caballero A. 1994: Development in the prediction of effective population size. Heredity **73**: 657-679.

Chesser R.K., Rhodes E.O., Sugg D.W., Schnabel A. 1993: Effective sizes for subdivided populations. Genetics **135**: 1221-1232.

Dempfle L. 1975: A note on increasing the limit of selection through selection within families. Genet. Res. Camb. **24**: 127-135.

Falconer D.S. 1989: Introduction to Quantitative Genetics. 3rd edition. Longman Scientific & Technical, England.

Golden B.L., Brinks J.S., Bourdon R.M. 1991: A performance programmed method for computing inbreeding coefficients from large data sets for use in mixed-model analysis. J. Anim. Sci. **69**: 3564-3573

Kang H., Namkoong G. 1988: Inbreeding effective population size under some artificial selection schemes, 1. Linear distribution of breeding values. Theor. Appl. Genet. **75**: 333-339.

Lindgren D., Gregorius 1976: Inbreeding and coancestry. Pp. 49-79 *in* Proc. IUFRO Meeting on Advanced Generation Breeding, June 14-18, Bordeaux, INRA, France.

Mullin T.J., Park Y.S. 1995: Stochastic simulation of population management strategies for tree breeding: a new decision-support tool for personal computers. Silvae Genetica **44**: 132-141.

Remmert H. (ed.) 1994: Minimum Animal Populations. Ecological studies 106. Springer-Verlag Heidelberg.

Robertson A. 1961: Inbreeding in artificial selection programs. Genetical Research. **2**: 189-194.

Tier B. 1990 Computing inbreeding coefficients quickly. Genetic Selection and Evolution **22**: 419-430

Wei R.P., Lindgren D. 1995a: Optimal family contributions and a linear approximation. Theoretical Population Biology. **48**: 318-332

Wei R.P. and Lindgren D. 1995b: Gain and effective population size following index selection with variable weights. Forest Genetics **1**: 147-155.

William C. and Savolainen O. 1996 Inbreeding depression in conifers: implications for breeding strategy. Forest Science **42**: 102-117

Wright S. 1931: Evolution in mendelian populations. Genetics **16**: 97-159.

Table 2.2: The status number of a family as a function of family size related to number of families (= 1); parents in the first generation (N_1); and offspring in the second generation ($N_2 = m$).

	Self sib, $N_1 = 1$		**Full-sib, $N_1 = 2$**			**Half-sib, $N_1 = 1 + m$**		
Family size	**Per family = per parent**	**Per offspring**	**Per family**	**Per parent**	**Per offspring**	**Per family**	**Per parent**	**Per offspring**
m	$N_s = m/(m + .5)$	$N_r = 1/(m + .5)$	$N_s = 2m/(m + 1)$	$m/(m + 1)$	$N_r = 2/(m + 1)$	$N_s = mN_r = m/(.75 + .25m)$	$mN_r/(1 + m)$	$N_r = 1/(.25m + .75)$
1	0.667	0.667	1	0.5	1	1	.5	1.
2	0.800	0.400	1.333	0.667	0.667	1.600	0.533	0.800
3	0.857	0.286	1.500	0.750	0.500	2.000	0.500	0.667
4	0.889	0.222	1.600	0.800	0.400	2.286	0.457	0.571
5	0.909	0.182	1.667	0.833	0.333	2.500	0.417	0.500
10	0.952	0.095	1.818	0.909	0.182	3.077	0.280	0.308
$\to\infty$	$\to 1$	$\to 0$	$\to 2$	$\to 1$	$\to 0$	$\to 4$	$\to 0$	$\to 0$

Table 2.3: Average inbreeding, average coancestry and relative status number for some mating system with close inbreeding. Note that the relative status number is independent of N and thus valid for a breeding population of any size structured in unrelated sublines.

	Repeated selfing, m = N = 1			**Repeated full-sib mating, m = N = 2**		
Entity	**F_t**	f_t	$N_s = N_r$	**F_t**	f_t	$N_r = N_s/N$
Formula	**$0.5(1 + F_{t-1})$**	**$0.75 + 0.25F_{t-1}$**		**$0.25(1 + 2F_{t-1} + F_{t-2})$**	$\frac{0.75 + F_{t-1} + F_{t-2}}{2}$	$\frac{1}{2fN}$
Generation						
t = 1	0	0.5	1	0	0.25	1
2	0.500	0.750	0.667	0	0.375	0.667
3	0.750	0.875	0.571	0.250	0.500	0.500
4	0.875	0.938	0.533	0.375	0.594	0.421
5	0.938	0.969	0.516	0.500	0.672	0.372
10	0.998	0.999	0.500	0.826	0.886	0.282
→ ∞	→ 1	→ 1	→ 0.5	→ 1	→ 1	→ 0.25

Table 2.4: Relative status number and inbreeding in a population composed by two progenies, from each of the members in the previous generation as a function of generation and population size.

	N = 4		**N = 10**		**N = 50**		**N = 200**		
Entity	**F_t**	**N_r**	**F_t**	**N_r**	**F_t**	**N_r**	**F_t**	**N_r**	**$N_r = 0.5/(1 + (t-3)/4)$**
Generation									
t = 1	0	1	0	1	0	1	0	1	1
2	0	0.667	0	0.667	0	0.667	0	0.667	0.667
3	0.083	0.500	0.028	0.500	0.0051	0.500	0.0013	0.500	0.500
4	0.153	0.407	0.054	0.402	0.0102	0.400	0.0025	0.400	0.400
5	0.218	0.347	0.080	0.338	0.0152	0.334	0.0038	0.334	0.333
10	0.474	0.219	0.197	0.195	0.0399	0.184	0.0100	0.182	0.182
$\rightarrow\infty$	$\rightarrow 1$	$\rightarrow .125$	$\rightarrow 1$	$\rightarrow .05$	$\rightarrow 1$	$\rightarrow 0.01$	$\rightarrow 1$	$\rightarrow 0.0025$	$\rightarrow 0$

Formulae used:

$Nf_t = 0.75 + (N-1)\,F_t + 0.25F_{t-1}$

$F_{t+1}(N-1) = 0.25(1 + F_{t-1}) + (N-1.5)F_t$ (selfing excluded)

$N_r = 0.5/(Nf_t)$

Table 2.5: Inbreeding and relative status number in a random mating population (selfing occurs and progeny size varies at random) as a function of generation and population size.

	N = 10		**N = 50**		**N = 200**		**N**
Entity	**F_t**	**N_r**	**F_t**	**N_r**	**F_t**	**N_r**	**$N_r = 1/t$**
Generation							
t = 1	0	1	0	1	0	1	1
2	0.050	0.513	0.010	0.503	0.002	0.501	0.5
3	0.098	0.351	0.020	0.337	0.005	0.334	0.333
4	0.143	0.270	0.030	0.254	0.007	0.251	0.250
5	0.185	0.221	0.039	0.204	0.010	0.201	0.200
10	0.370	0.125	0.086	0.105	0.022	0.101	0.100
$\to\infty$	→1	→ .05	→1	→0.01	→ 1	→0.0025	→0.5/N

Formulas used:
$F_{t+1} = f_t = 1-(1-0.5/N)^t$
$N_r = 0.5/(Nf_t)$

Table 2.6: Status effective number over generations (N_s). Comparison with traditional effective number (N_e) for a breeding population of size 100 as a function of generation.

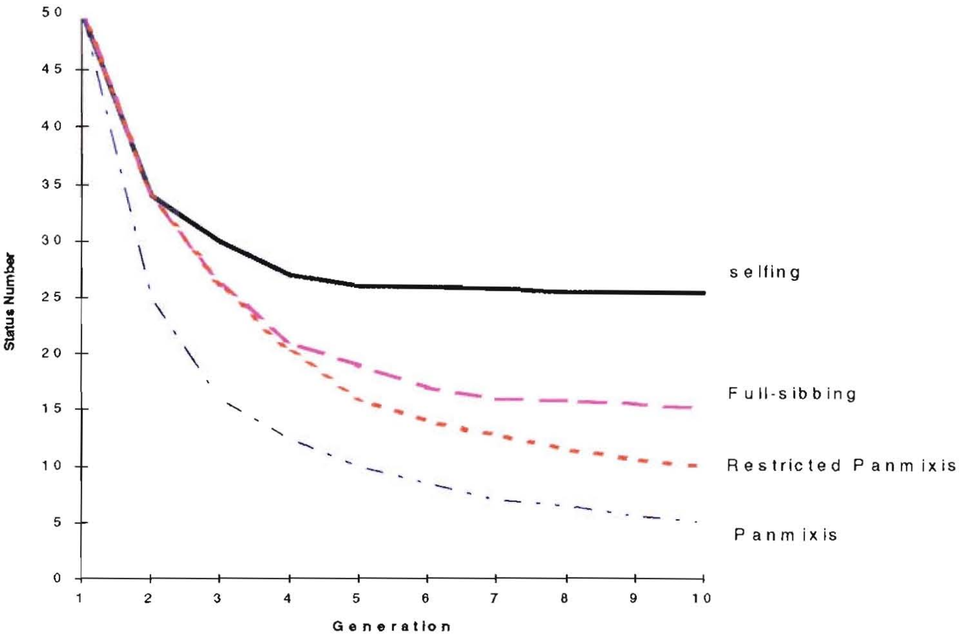
Generation	Two offspring per parent				Random offspring per parent			
	Ns (pred)	N_S (sim)	N_e (pred)	N_e (sim)	N_S (pred)	N_S (sim)	N_e (pred)	N_e (sim)
0	100	100	199	∞	100	100	99.5	∞
1	66.67	66.67	199	∞	50.12	50.42	99.5	∞
2	50.00	50.00	199	$\rightarrow\infty$	33.50	33.60	99.5	$\rightarrow\infty$
3	40.02	40.02	199	358.69	25.19	25.45	99.5	212.4
4	33.37	33.38	199	297.64	20.20	20.41	99.5	147.57
5	28.63	28.63	199	254.29	16.88	17.07	99.5	133.31
6	25.08	25.08	199	244.23	14.50	14.64	99.5	126.22
7	22.32	22.32	199	238.80	12.72	12.84	99.5	122.73
8	20.11	20.11	199	231.13	11.33	11.43	99.5	116.53
9	18.30	18.29	199	223.29	10.23	10.32	99.5	116.30

Table 2.7: Desirable characteristics of an effective number and how well they are met by a suggested and a traditional effective number concept. "Yes" means that the characteristic is reasonable well met.

Desirable characteristics of an effective number	N_S	N_{eFI}
A breeding population with a certain effective number should for some important aspect behave as an ideal population with that number.	Yes	Yes
It should be a characteristic of the current state of a population, not of assumptions of its history, its mating tactics or a scenario about its future. [5]	Yes	No
It should not be dependent on an assumption of discrete generations.	Yes	No
It should never be higher than the real number.	Yes	No
It should be equal to the census number if there is no inbreeding or relatedness.	Yes	No
It should give an idea about how related individuals are.	Yes	A little
The effective number when merging two unrelated populations should be the sum of the effective numbers of the two subpopulations.	Yes [3]	No
It should never get smaller than 0.5 (corresponding to a homozygous line or a gamete).	Yes	No
It should never be larger for the filial generation than for the parental generation for a closed population and equal only as a limiting case when numbers become large.	Yes [2]	No
It should be derivable from the pedigree. There should not exist populations with undefined or infinite effective number.	Yes	No
Expected sampling variation should be possible to specify	Yes	Yes
The definition should be simple and not ambiguous.	Yes	No
It should be theoretically possible to measure experimentally.	Yes	Yes
It should carry information of evident importance to biologists or breeders.	Yes	Yes
It should complement average coefficient of inbreeding and census number.	Yes	Yes
It should give some information about likelihood of losing rare alleles.	Yes	Yes
The difference between census and effective number should be intuitively acceptable.	Yes	No
It should increase if more members per progeny are included in the breeding population.	Yes [1]	Yes

1. Statement is true for symmetric populations, thus when all individuals have the same type and number of relatives. Asymmetric populations may be constructed where the statement is false.
2. Statement is true for repeated regular mating designs, cases where it is not true may be constructed.
3. Statement is true if the merged populations are symmetric, in many other realistic cases it is not far from the truth.
5. Actually this is border-line case, and it may be desirable to have both types of measures.

Figure 2.6. Preservation of status number in a population of size 50 by different breeding systems.



"Selfing" refers to repeated selfing (50 sublines of size 1). "Full-sibbing" refers to repeated mating of full-sibs (subline size 25, sublines of size 2). "Restricted panmixis" refers to a system where each parent give rise to two progenies, which mate at random (no selfing, subline size 50). "Panmixis" refers to random mating, where the number of progenies per parent is binomially distributed (selfing allowed, subline size 50).

CHAPTER 3

THE EFFECT OF POPULATION STRUCTURE, MATING DESIGN AND SELECTION STRATEGY ON STATUS NUMBER, COEFFICIENT OF INBREEDING AND GENETIC GAIN

A summary of this chapter was presented at the Evolution and Tree Breeding Conference, Canadian Tree Improvement Association/Western Forest Genetic Association, Victoria, British Columbia, Canada, August 28 - September 1, 1995

INTRODUCTION

As breeding programmes move to advanced generations of selection and breeding, uncertainties arise about inbreeding and narrowing of genetic base.
To guarantee the continued existence of unrelated genotypes, breeding populations have been organised into sublines. Sublines as defined by Burdon and Namkoong (1983) are:

'subunits of a breeding population (or a gene resource) which represent parallel or replicate populations that:
(i) differ only due to genetic sampling error or random drift, and
(ii) unless a qualification is stated, remain strictly unrelated (or disconnected), except due to common ancestry effects arising from random sampling in large base populations'.

If subunits of a breeding population are deliberately differentiated they amount to multiple populations (Namkoong, 1988).

A subline might be constituted with many breeding groups, the subdivision into breeding groups going beyond the need purely to control inbreeding at any time in the future.
Breeding groups, providing that they are kept strictly unrelated, can also be considered as sublines.

For managing sublines/breeding groups two concepts are essential:

- **size** of the breeding population required for selection programmes, and;

- coefficient of **inbreeding** and **coancestry**.

Population size has a wide range of consequences since it has impact in both the short term (influencing selection differential, inbreeding depression and the reduction of genetic variance due to genetic drift) and the long-term (affecting selection limits and the utilisation of new variation arising from mutation -outside the scope of this study-). Nicholas (1980) considered the coefficient of variation of response as a criterion for the calculation of the minimum effective population size required for selection experiments. The coefficient of variation of response is the ratio of the standard deviation of the response to its expectation; this parameter can be used as a criterion for determining the appropriate size of a breeding population required for recurrent selection approach and for comparing the relative results achieved from various selection schemes. Agreey *et al.* (1995) reviewed and adjusted the formulae to include variance due to drift and measurement error.

The coefficient of **inbreeding** and coefficient of **coancestry** have been used by breeders for a number of years (White *et al.* 1993; Mahalovich and Bridgwater 1989; Cotterill 1989; Shelbourne *et al.* 1986). The coefficient of inbreeding has been defined as the probability that both of any pair of alleles in a diploid individual are identical by descent. The coefficient of coancestry (which measures the relatedness of the prospective parents) has been defined as the probability that genes sampled from the same loci in different (or the same) individuals are identical by descent; this is equivalent to the coefficient of inbreeding which would occur after mating (Caballero 1994; Crow and Kimura 1970).

The patterns of intermating of individuals are called mating designs. Toro and Perez-Enciso (1990) have suggested the use of mating designs that minimise the average coefficient of inbreeding in the following generation of selection. One important purpose of a mating design is to provide a population within which to perform the next generation of selection.

Van Buijtenen and Burdon (1990) studied the expected efficiencies of alternative mating designs for obtaining genetic gains from 'forwards' selection in advanced generations, and Burdon and van Buijtenen (1990) also explored the differences in gain between mating designs for estimating breeding values for reselection of parents. They found that numerous crosses per parent, while increasing costs, added little to expected gain from forwards selection, unless specific combining ability (SCA) was substantial. On the other hand, with small numbers of crosses, selection of sufficient unrelated individuals could be expected to entail much reduced gains. Burdon (1988) emphasised the risk of losing or wasting the genes of the parent if each parent was committed to only a single pair-cross. This argument favoured some sort of double-pair mating despite the fact that it showed no clear advantage in expected gains (from forwards selection) from increased numbers of crosses (van Buijtenen and Burdon 1990).

The increase in coancestry during successive generations of recurrent selection is potentially a major problem in long-term breeding programmes. It can be delayed by restricting selection to within families, and by equal and symmetric representation of all parents and founders in mating schemes. Balanced mating designs are more advantageous than unbalanced for maximising effective population size and minimising allele loss than for immediate efficiency in realising genetic gains (Kang and Namkoong 1979; van Buijtenen and Burdon 1990; Dempfle 1975). Selecting predominantly within families is necessary for maintaining effective population size in the breeding population even if this disregards family information and in each generation operates on only half the original additive variance (Burdon 1988).

The objective of this chapter was to investigate the effects that different population sizes, mating designs and selection strategies have on status number, inbreeding coefficient and genetic gain. Changes in effective population size were evaluated by a new parameter, status number, N_S (Lindgren *et al.* 1996); this innovation was used because conventional effective numbers express the average rate of change of genetic base, and are inappropriate for studying the maintenance of effective number to a given point in time.

METHODS

Status number (Status effective number)

Assuming complete unrelatedness of parents and infinite number of alleles, status number is defined as half the inverse of the average coancestry. Average coancestry for an individual member of a population is the average of the coancestry between that individual and all individuals in the population including the individual itself (Lindgren *et al.* 1996).

Status number is expressed mathematically as:

$$N_s = 0.5/f \qquad \textbf{(3.1)}$$

where N_s is the status number;
f is the average coancestry of the population

The relative status number, N_r, is defined by:

$$N_r = N_s/N \tag{3.2}$$

where N is the census number.

For the calculations, use the stochastic simulation of Tier's (1990) algorithms (see also Appendix 2).

Genetic gain

Genetic gain and inbreeding coefficient were calculated using POPSIM (version 2.3.00), a computer programme that stochastically simulates changes in the genetic structure of managed tree populations under various tree breeding scenarios (Mullin and Park 1995).

In POPSIM the additive genetic effect calculated is the average additive effect of the female (*f*) and the male (*m*) parents, and the additive effect for each offspring (A_i') may be generated around a mean with a variance equivalent to the within-family portion of the additive genetic variance, σ_A^2 ; this variance having been adjusted for the expected reduction due to increasing inbreeding of the parents. In the infinitesimal gene-effects model, the additive variance arises from variation among an infinite number of loci, each controlling the expression of the trait by an infinitely small amount. As described by Dempfle (1975), the variation within families is due to Mendelian segregation at segregating loci. When alleles are identical by descent, the variation due to segregation is reduced. Dempfle suggested that the reduction in within-family variance due to inbreeding of the parents could be accounted for with the general expression:

$$\begin{aligned} Var(A_w) &= \left[\frac{1-F_P}{4} + \frac{1-F_Q}{4}\right] * \sigma_A^2 \\ &= \tfrac{1}{4}\left[4-\left(1+F_P\right)-(1+F_Q)\right] * \sigma_A^2 \\ &= \tfrac{1}{2}\left[1-\tfrac{1}{2}\left(F_P+F_Q\right)\right] * \sigma_A^2 \end{aligned} \tag{3.3}$$

where Var(A_W) is the within-family additive genetic variance of the full-sib cross between parents *P* and *Q*, whose inbreeding coefficients are F_P and F_Q, respectively. F and σ_A^2 refers to a reference population and F_1=0.

Inbreeding coefficient

The inbreeding coefficient for each tree in the breeding population is determined by analysing the pedigree of parent trees to determine the degree of common ancestry. The calculations use Tier's (1990) algorithms to make the calculation of inbreeding straightforward. The average inbreeding of the population is calculated simply as the average inbreeding coefficient for all trees.

POPSIM estimates the inbreeding effective population size (N_e), from the rate at which inbreeding accumulates (ΔF). Falconer (1981: equation 3.12) gives the inbreeding coefficient in any generation t, assuming an infinite randomly mated population and referring to the base population at t_0 as:

$$F_t = 1 - (1 - \Delta F)^t \qquad \textbf{(3.4)}$$

By rearranging Falconer (1981: equation 4.1) the calculation of inbreeding effective size is:

$$N_e = \frac{1}{2\left(1-(1-F_t)^{-t}\right)} \qquad \textbf{(3.5)}$$

note that the inbreeding effective size of the base population is undefined, if no inbreeding has accumulated, it will be infinitely large in the ensuing generation.

Details of model and settings of the simulations

The initial plus-tree selection is referred to as generation 1, and the expected breeding value of these trees is set to zero (with stochastic variation). These trees, which comprise the breeding population of initial plus trees, are mated according to a specified design and generate progenies, whose genetic structure depends on the mating design. From these progenies, new trees are selected to form the parents of the next cycle of the breeding population through the following 10 simulated generations.

To observe the effect that population size has on genetic gain, inbreeding coefficient and status number after ten generations of breeding, a simple double-pair-cross mating design was chosen for population sizes of 12, 32, 64 and 128 parents and the selection of trees was made according to a combined index selection (CI). Combined index selection involves predicting the breeding value of each tree and using these predictions as the basis of selection. The predicted breeding value is based on the phenotype of the tree and the mean performance of its full-sibs, weighted according to heritability of within-family effects and heritability (or repeatability) of sib-family effects respectively. Thus, the weighting will depend not only on individual heritability but also on family size.

To investigate the effect that mating design and selection strategies have on gain, inbreeding coefficient and status number, a small breeding group of 12 trees was selected. The selection of trees was made according to two methods: (i) phenotypic selection or mass selection (P) and; (ii) combined index selection (CI). Superimposed to these two main options, simulations were done either without any restriction on the numbers of trees selected per full-sib family or with the restriction of one or two selections per full-sib family (but no restriction on the number of selections per half-sib family).

All simulations were repeated with 100 replicate runs per set of conditions to obtain a reliable prediction.

In this study the breeding population represents a single breeding group and is characterised by the following parameters, each generation:

- Average breeding value (G) or genetic gain. The reference point, designated 0, is the expected breeding value of the initial parents.
- Average coefficient of inbreeding (F), assuming that there was no inbreeding or coancestry in the initial parents.
- Status number (N_S). The status number is half of the inverse of the average coancestry in the breeding population.

Summary of Assumed Population Parameters

• Breeding group size	12 to 128
• Total offspring	100 per full-sib cross
• Number of generations	10
• Additive genetic variance σ_A^2	1
• Dominance and epistatic variance σ_D^2; σ_I^2	0
• Environmental variance σ_E^2	19 and 4
• Narrow sense heritability $h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_E^2}$	0.05 and 0.2

Expected gain from phenotypic selection

The gain from selecting the best proportion of phenotypes can be predicted as (Falconer, 1981):

$$\Delta G = ih^2\sigma_P \tag{3.6}$$

where:

i = selection intensity

h^2 = heritability, and

σ_P = phenotypic variance.

Noting that $h = \sigma_A / \sigma_P$, we may rewrite this equation in the form:

$$G = ih\sigma_A \tag{3.7}$$

If $\sigma_A^2 = 1$ then:

$$G = ih = \frac{i}{\sigma_P} \tag{3.8}$$

Mating designs

Eight different balanced-symmetric designs (with the same number of crosses for each parent) were considered and are presented in Figure 3.1. Family size was kept constant at 100 individuals per full-sib cross, assuming a fixed resource scenario in terms of numbers of individuals per offspring.

Status number for regular, symmetrical mating designs

As the mating designs and family sizes are balanced, the expected average coancestry will be the same for all members of the breeding population (e.g. N=12) after the first round of selection. In other words, a single row or column in the symmetric coancestry matrix will suffice to give the expected average coancestry for the population.
Assuming that the founders are completely unrelated, and not inbred, during the first generation the average coancestry (Falconer, 1981) ($\bar{f}_t$) is:

$$\bar{f}_t = [(0.5 \text{ -if selfed-}) + (\text{no. of full-sibs}) \times 0.25 + (\text{no. of half-sibs}) \times 0.125\ ]/N \tag{3.9}$$

and

$$N_S = 0.5/f_t\text{, or}$$

$$= 0.5/[(0.5) + (\text{no. of full-sibs}) \times 0.25 + (\text{no. of half-sibs}) \times 0.125]/N \qquad \textbf{(3.10)}$$

For the first generation or for fully symmetric cases calculation of status number is easy but after several generations the relationships become complex and asymmetric (Lindgren *et al.* 1996).

RESULTS

The results of the first simulation study (in terms of gain, coefficient of inbreeding and status number) for a trait of heritability 0.2, ten generations, double pair-cross mating design and different population sizes are shown in Table 3.1 (A, B and C) and Figures 3.2, 3.3, 3.4 and 3.5.

Results for the second simulation study which involve heritabilities of 0.2 and 0.05 for all the one-cross-per-parent designs (DPC, DPF, PFM, DD3) and the other designs after ten generations, are shown in Table 3.2 (A, B and C). The results from the four mating designs with one cross per parent, shown in Figures 3.6, 3.7, 3.8 and 3.9 were pooled for ease of interpretation, because there were no evident differences amongst them (see below).

Status Number

For different breeding group sizes

Status number under conservative strategies (with restrictions to one individual selected per pair-cross), is better preserved by small groups (Figure. 3.2). However this difference is initially marginal and becomes recognisable only in later generations. After 10 generations the smallest breeding group (12) has a relative status number of 0.18 while a population of 128 parents has a relative status number of 0.17. If more individuals in any one pair-cross can be selected, the overall status number declined faster and is then preserved much better with small breeding groups (Table 3.1 and Figure 3.2). When no restrictions were applied on the number of individuals selected per cross almost 90% of the status number of the small breeding group was lost in one generation (and nearly 80% for a large group). It was under this extreme scenario that the advantage of small group size was greatest. Status number levels off at around generation 3 when no restriction in the number of individuals selected per cross were applied, and around generation 8 for two individuals selected per pair-cross, while with a conservative strategy status number is expected to level off even later (Figure. 3.2).

Higher levels of gain per unit of status number loss are obtained with a conservative within-family selection strategy, but to reach the same level of gain, this strategy will take more generations to achieve (Figure. 3.3).

For different mating designs

Within-family selection, made so as to save exactly one individual per full-sib cross (Table 3.2C and Figure. 3.2), maintained the highest status number. Even so, the status number dropped to approximately one-quarter of the original value in three generations, with the largest decrease occurring in the first generation. Status number values also fell faster for mating designs with more crosses per parent (Figure. 3.4). The four double-pair mating systems (DPC, DPF, PFM, DD3) are very similar for status number (Figure. 3.5).

Inbreeding Coefficient

For different breeding group sizes

Breeding group size is an important factor for delaying inbreeding. A selection strategy based on keeping no more than two individuals per cross under a large group will have a significantly smaller inbreeding coefficient than the conservative strategy of one individual per pair-cross (Figure 3.6).

The coefficient of inbreeding becomes extremely high for those selection strategies with no restrictions in the number of individuals per family. If such strategy is applied the level of inbreeding exceeds 0.2 after only four generations of breeding, for all group sizes. Inbreeding coefficient quickly reached 0.2 in the smallest breeding group. However, for breeding groups larger than 64, the coefficient of inbreeding was still negligible after 10 generations (Table 3.1B) even under a strategy of two individuals kept per cross.

For different mating designs

The three mating systems (DPC, DPF, DD3) are rather similar for inbreeding coefficient (Figure 3.5). Status number and inbreeding coefficient for the single-pair mating design (SPM) are similar to those for the schemes with one cross per parent. Inbreeding coefficients were also higher for mating designs with more crosses per parent (Figure 3.7). With one cross per parent, an approximate formula for calculating the inbreeding coefficients in a random-mating population is $F_{t+1}= t/2N$ (Lindgren *et al.* 1996). For a population of size 12, this is $F_{t+1}= t/24 = 0.042$ per generation. The inbreeding coefficient in the simulated populations seems to build up slower, which can be expected, since selection is more even between families than for random selection, especially with the restrictions applied. Inbreeding and N_S are quite similar for the same selection scenarios (Table 3.2B and C) at different heritabilities.

Genetic gain

For different breeding group sizes

Breeding-population gain (Figure 3.8) increased with group size, especially in later generations. There is little difference in gain for the first two generations under the restrictions. Gain for the 12-parent breeding group with no restriction in the number of individuals kept per cross slows down around generation 5. This 'plateauing' in gain for the smaller population under an aggressive selection strategy coincides with high inbreeding coefficient values, without actually invoking inbreeding depression.

For different mating designs

The predicted gains for each mating design with one cross per parent were similar (Figure 3.5), though gain for the disconnected diallel design (DD3) was slightly (but not significantly - $p > 0.05$) higher than for the partial factorial mating (PFM), double-pair factorial (DPF) and double-pair cyclic (DPC) designs; these gains are averaged in Table 3.2A.

Substantial extra gain results from restricted combined index selection (but less for unrestricted selection) if more crosses per parent are made (Figure 3.8 and 3.9); for instance, during the first four generations of breeding gain almost doubles with the increase from 0.5 crosses per parent to 1.5 crosses per parent (DD4). The extra gain is accompanied by more inbreeding and lower N_S, and the improvement in gain with increasing number of crosses per parent occurs up to 2.5 crosses (DD6) but not beyond (DIA).

Almost exactly the same status numbers, coefficients of inbreeding and marginally higher gains can be obtained with a strategy based on the same number of crosses as parents (DPM) and a maximum of two individuals selected per family as those values obtained with a mating design of 2.5 crosses per parent (DD6) and a maximum of 1 individual per family.

SPM is expected to produce slightly higher gain from within-family selection (as the selection intensity for the same selected proportion is higher) but less gain from among-family selection (as there are fewer families to select from), leaving almost no net difference (Figure 3.10).

Combined index selection yields higher expected gains than phenotypic selection at both heritabilities, essentially at low heritability (Table 3.2A).

Unrestricted combined index selection did not result in more gain, after 10 generations, than combined index selection with restriction, unless the number of crosses per parent was one or less, or the selection was restricted to no more than one individual per full-sib family. Thus with these exceptions it is possible to choose a strategy that after ten generations of breeding will produce more expected gain than unrestricted combined index selection (Table 3.2A). The high genetic gains of unrestricted selection during the first generations (Table 3.1A, B and C) were made at the cost of a reduced status number and an increased coefficient of inbreeding. When no restrictions were applied, the status number dropped in two generations to less than a fifth of its initial value. This loss of status number was also accompanied by a rapid increase in the coefficient of inbreeding (to 0.2 in 2 generations for N=12).

Constraints on the number of individuals selected per family severely curtailed gain, especially at low heritability and with one or fewer crosses per parent (Table 3.2A). For the low-heritability case and mating designs with the same number of crosses as parents, a 50 % increase in gain after ten generations was obtained by relaxing the constraint of equal representation of all families (one tree per full-sib family) to two individuals from the best full-sib families. This scenario (relaxing the constraint) also carried a corresponding increase in the inbreeding level and a reduction in the N_S (Table 3.1). With mating designs that involved increasing crosses per parent (DD4, DD6 and DIA), the increase in gain progressively decreased. Because of the "fixed resources" model used, gains from among-family selection increase but gains from within-family selection decrease as number of crosses per parent increases and number of individuals per family decreases, especially when the crosses per parent number more than 2.5.

Unrestricted phenotypic selection showed relatively higher gains (if expressed in percentages) at high heritability, but these are still inferior to those from combined index selection (Table 3.2A). Using family information thus becomes less important at higher heritabilities. Increasing the number of crosses per parent at both heritabilities did not increase gain proportionately, with gains appearing to plateau at 1.5 to 2.5 crosses per parent, for the fixed-resource case (Table 3.2A).

Phenotypic selection, and combined index selection restricted to a maximum of two selections per cross, resulted in similar values for status number and inbreeding coefficient; only when more crosses per parent are used does phenotypic selection give higher status number and lower inbreeding levels and gains (Table 3.2). Phenotypic selection resulted in less gain at the same status number when heritability was low, but at the higher heritability the difference was small.

DISCUSSION

Size of the breeding group

Genetic Gain

If we consider the equation for gain prediction (equation 3.6), gain can be affected by changes in the selection intensity (*i*), heritability (h^2) and additive variance (σ_A) and group size may influence the contribution of all three of these factors.

The genetic variance is smaller in a sample than in the full population. The variance in a finite population sample is approximately (1-1/n) of that in a large population. If the offspring generated are numerous, this sampling effect is unlikely to be of significance in a reasonable-sized breeding programme.

Reduction of the additive variance due to selection is expected. Bulmer (1971) showed how to anticipate changes in genetic variance and heritability during the first two to three generations of selection; he used models of infinitely many genes, so the reduction in variance was necessarily due to gametic-phase disequilibrium. Sorensen and Hill (1982) showed that even with a small number of loci, linkage disequilibrium still accounted for the major part of the changes in variance.

Lindgren and Mullin (1996) noted that the within-family variance for a cross between any two parents arises from the fraction of genes that are not identical by descent and therefore able to contribute to family genetic variance. Small populations after a couple of generations would carry more individuals identical by descent than larger populations, and their within-family variance would be smaller and with less opportunity for selection.

Bridgwater (1995) compared population variances for 20 generations, for subline sizes from 4 to 12 and showed that additive variance for subline size 4 increases for 7 to 8 generations before slowly declining. He also showed that the additive genetic variance is reduced more rapidly in larger sublines.

With the model assumptions (additive variance influenced by inbreeding coefficient, equation 3.2 used here), the additive variance within the group for a population size of 12 will start declining after generation 7 with a conservative selection strategy, and levels-off after generation 3 if no restrictions are applied in the number of individual selected per family (Figure 3.8).

Lindgren (personal communication) suggested that parental variance might decay at $(1\text{-}0.5/N_S)t$ (instead of $(1\text{-}0.5/N)t$ as it is usually known, Bulmer 1970), if so the effects will be rather substantial for group sizes below 20. Since the effects are cumulative they will become larger as N_S drops differently for different group sizes with time. If the above assumption is correct, the decay of additive variance during the first generation of breeding could be as illustrated in Table 3.3.

Thus, to prevent the decay of additive variance per generation, a conservative selection strategy (i.e. restricting the degree of among-family selection) is recommended more than an emphasis on population size. Bridgwater (1995) argued that genetic variances averaged over all groups decrease much more rapidly in larger groups, this will only be true when selection strategies are relaxed so that status number is better preserved by small groups.

Table 3.3. Decay of within-groups parental additive variance per generation with different group sizes

Parents per group	Additive variance per group (1-0.5/Ns)
128	0.996
64	0.992
32	0.984
16	0.968

Inbreeding Coefficient

A desirable goal in tree breeding is to guarantee access to unrelated genotypes that can be intercrossed within the production population. This goal can be achieved by arranging the breeding population in unrelated groups. Large groups can provide unrelated individuals for seed orchards over many generations, while considerably delaying the build-up of inbreeding, conferring an advantage for large groups over small ones.

Theoretically (Falconer 1981), inbreeding will reduce the within-family variance but unmask some total additive variance, which will counterbalance the within-group reduction. Based on this principle, Bridgwater (1995) advocated the use of small sublines (that in this case will apply also for small breeding groups) for managing breeding populations. If inbred groups are merged the additive variance will increase compared to the average for the within inbred groups, and opportunities for selection gains will be restored.

Breeding groups with less than 10 parents will be impossible to maintain for longer than a few generations without inbreeding, which may become so severe that it is likely to cause fertility problems and to hamper selection, depending on the level of genetic load carried by the species or, rather, the particular subset of parents. Possible exceptions are species that have few lethal alleles and undergo inbreeding with no symptoms of depression (Russell *et al.*. 1995). If inbreeding within the groups became severe, full-sib testing, clonal testing and parental ranking will become less efficient for estimation of breeding values and less efficient for long-term breeding. Also, as the families within groups become related there will be less usable variation among them. Inbred breeding populations will also require to produce large numbers of progeny per replicate per group because the probability of extinction for each group is expected to be high. Advantages of highly inbred groups can include perfect assortative mating, increased selection efficiency among groups and increased uniformity within groups (Williams and Savolainen 1996).

The use of inbreeding in forest trees is controversial. For most conifers, high levels of inbreeding will not be the best method for reducing inbreeding depression in small groups. Durel (1990) proposed a strategy where inbreeding was promoted every two generations of breeding, replacing with selfing the traditional polycross. Williams and Savolainen (1996) following the same line, claimed that sib- or random mating within small breeding groups is a better option than selfing in the early generations of conifer domestication because it can be expected to prevent losses to low offspring survival and adult fecundity, and thus affords the best way purging deleterious alleles. If inbreeding depression is based on deleterious mutations then it should theoretically decline with a strong purging of deleterious alleles. Based in a partial dominance model such that inbreeding depression can be largely eliminated through strong selection, selfing and high levels of inbreeding in early generations of breeding does not seem to be the best option for most conifers, with the possible exception of conifers with a low lethal load (Russell *et al.* 1995; Williams and Savolainen 1996).

Even small groups can be maintained without inbreeding for the first generation, and no need to make decisions about group size before the second generation. Small groups can always be merged, but large groups will become impossible to later split into smaller units (Burdon 1986, McKeand and Bridgwater 1992). Decisions might, therefore, be postponed until needed.

In conclusion, small groups of trees will presumably fail if, as Meuwissen and Woolliams (1994) described, they enter a "down-ward spiral of ever-decreasing fitness" as they reach high levels of inbreeding and low offspring survival and adult fecundity.

In the absence of mutation, directional selection (and to some extent stabilizing selection) usually erodes mainly additive genetic variance, and partially affecting dominance variance. Detrimental effects of inbreeding depression are closely associated with non-additive effects. Consequently, inbreeding depression should reflect high levels of dominance variance not modelled here (pseudo over-dominance and linkage blocks). Meuwissen and Woolliams (1994) showed the important effect of very rare recessive deleterious (or even lethal) genes on fitness. The additive genetic model used here was not directed in particular towards such genes, because the problems caused by these genes may be more efficiently tackled by genetic markers (Williams and Savolainen 1996), which may be used to help purge them from breeding groups. However, some genes producing less detrimental inbreeding depression are more likely to increase in frequency and collectively lead to a substantial effect.

Status Number

The status number is a useful quantitative measure of the current state of genetic diversity in a breeding population and extends information given just by inbreeding and coancestry coefficients. Although smaller groups raise inbreeding coefficients and lower gain within the groups, they allow higher aggregate status numbers to be maintained.. Substantial cumulative effects on status number, however, will only be seen after a large number of generations even with very small groups. Hence, breeding schemes with small groups are slightly more efficient in preserving status number through a large number of generations than breeding systems with large groups, thus, reducing the chance that rare alleles included in the population will be lost (Bridgwater 1995). Medium-to-large-size groups showed a comparatively small reduction in aggregate status number over generations but showed larger increases in gain compared with small groups (Figure 3.5).

Group size may also affect genetic gains by changes in selection intensity. The most important source of gain is the exploitation of effective number by intensive among-family selection. The larger the status number in the breeding population, the larger the gain that can be achieved in the production population. Thus, a lower immediate gain in the breeding population by keeping a higher status number does not necessarily mean a reduced gain in the production population. A higher status number in the breeding population means that it can be exploited harder when transferring the gain to the production population.

Mating designs and selection strategy

From the results of these simulations, unrestricted combined family and individual index selection always resulted in higher expected genetic gain than phenotypic selection and combined index selection with restriction on the number of individuals selected per family, agreeing with other findings in the literature e.g. Falconer 1981; Namkoong *et al.* 1988; Pswarayi and Barnes 1994. However, combined index selection with no restrictions was **less** effective than phenotypic selection in conserving genetic diversity -status number- (Table 3.2C), though combined index selection with restrictions was **more** effective than phenotypic selection in this function.

Choice of a selection strategy for a breeding population, will normally be strongly influenced by its ability to maintain genetic diversity in the long term as well as its capacity for realising high gains. The drop in status number following unrestricted combined index selection is likely to be unacceptable, even after just a single generation, and using any mating design. Verrier *et al.* (1993), concluded that selection procedures that place less emphasis on family information would best meet long-term objectives of diversity (Burdon 1988). Selection schemes that maximise gain by unrestricted combined index selection will result in rapid inbreeding, which is likely to cause fertility problems and bias selection (Williams and Savolainen 1996).

However, gains may be unacceptably low for mating systems and selection procedures (SPM, DPC, DPF, PFM, DD3) that do not include a between-family selection component (Figure 3.1), especially in low-heritability situations. It also seems to be impossible to avoid having the status effective number decreasing to something like one-third of its initial value after three rounds of selection, if the breeding programme is going to be reasonably effective in achieving gains. This is a good reason for using a large number of families as founders of the breeding population.

Van Buijtenen and Burdon (1990) addressing a single generation of breeding, asserted that SPM was an optimal mating design if all parents were equally represented, though some sort of double-pair mating design was seen to be preferable to avoid the risk of losing or wasting parents in the current round of crossing. I have shown, however, that SPM can be an efficient way to preserve status number but is sub-optimal for multi-generation gain. Van Buijtenen and Burdon (1990) showed that under a fixed-resource scenario, in a single generation and without special restrictions on relatedness of selections, numerous crosses per parent (more than 5) both raise costs and added little to expected genetic gain from forwards selection, unless SCA is substantial. However, in this study with multiple generations, an additive genetic model and "fixed resources", an increase from 0.5 or 1 cross per parent (equivalent to within-family selection) to 2.5 crosses per parent (DD6) with use of restricted combined-index selection resulted in substantial increases in expected genetic gain. This accords with the stochastic simulation results of King and Johnson (1993) and deterministic simulation results of Gea and Shelbourne (1995). However, under the given selection constraints, increasing the number of crosses per parent up to 2.5 does carry a penalty of increased coefficient of inbreeding and reduced status number.

If a minimal reduction of genetic diversity is required, however, families should be equally represented. Any selection scheme that is based exclusively on within-family selection will produce low gains, but after some generations will result in a higher status effective number than selection schemes that maximise immediate gain. Wei and Lindgren (1995), showed that genetic diversity was eroded quickly, even in a single generation, if selection allowed unequal representation of the progeny of different parents. In our study, gain increased substantially when restrictions were relaxed from no more than one individual per full-sib family to up to two individuals per full-sib family while the status number decreased from 6 to about 4 (for one-cross-per-parent designs after 10 generations). It seems desirable to apply constraints but to make them flexible enough to permit some unequal representation of parents.

With conservative selection strategies smaller group sizes raise inbreeding coefficients and lower gain but they lead to *slightly higher* aggregate status numbers. However, *substantial* cumulative effects on status number will only be seen with a strategy of unrestricted selection (Lindgren *et al.* 1996).

The use of effective population size (N_e) as an index for monitoring genetic diversity has been proposed (King and Johnson 1993), but the disadvantages and unattractive features of this parameter for characterising genetic diversity (discussed by Lindgren *et al.* 1996) makes it difficult to use in practice. Status number appears to be a more practical and efficient tool although its properties and limitations need to be understood.

IMPLICATIONS OF THE RESULTS

Some breeding programmes (Bridgwater 1995; White 1993; Borralho 1992) have been revised by a shift to smaller breeding groups.

In an advanced breeding population strategy, small breeding groups can be the vehicles of delivering rapid gains through a concentration of breeding resources (testing, selection and crosses and maybe cloning) combined with a reasonably short breeding cycle. Analysis of mating designs have shown that optimal gains are obtained with up to 2.5 crosses per parent and a more relaxed selection strategy. Thus, small breeding groups can take full advantage of this strategy, and provide individuals with high levels of gain, while a large breeding population can be managed with a more conservative and less expensive approach and fulfil the objectives of a long-term breeding strategy. Furthermore, small breeding groups are slightly more efficient in preserving status number through a large number of generations than large groups but small groups by themselves will probably not constitute a sustainable long-term breeding strategy.

If small breeding groups are kept closed (<20), minimising inbreeding depression within the group will be a guiding consideration. If inbreeding within the group becomes important, full-sib family information, clonal testing and parental ranking will become less efficient for estimating breeding values for a long-term breeding population. As the families within the group become related there will be less useful variation among them. One way of overcoming inbreeding will be by adopting an enrichment strategy where the breeding group is enriched every generation with genotypes from a large main breeding population, to periodically reduce inbreeding levels and thus maintain gain (Mahalovich and Bridgwater 1989).

Critical levels of status number need yet to be determined; however the simulations presented here have shown that unrestricted selection will lose almost 90 % of the genetic diversity (expressed as status number) in only one generation. On the other hand, within-family selection have shown the highest levels of gain per unit loss of status number but it will take longer to achieve it.

Small breeding groups can also be used as experimental populations in which high levels of inbreeding associated with high genetic gain can be managed and accelerated breeding techniques implemented without risking the success of long-term breeding population management.

Almost exactly the same status numbers, coefficients of inbreeding and gains can be obtained with a strategy based on the same number of crosses as parents (DPM) and a maximum of two individuals selected per family as those values obtained with a mating design of 2.5 crosses per parent (DD6) and a maximum of one individual per family; however, I would favour the second option because selection will be made on a multi-trait selection index and more crosses per parent will provide more opportunities for recombination, which could allow more options in selecting for the production population.

The most important source of the immediate gain that is needed for a production population is the exploitation of effective number by intensive among family selection. The larger the status number in the breeding population, the larger the gain that can be achieved in the production population. Thus, a lower immediate gain in the breeding population by keeping a higher status number does not necessarily mean a reduced gain in production population. A larger status number in the breeding population means that it can be exploited harder when transferring the gain to the production population.

REFERENCES

Aggrey S.E., Lin C.Y., Chen K.M. 1995: Size of breeding populations required for selection programs. Theor. Appl. Genet. **91**: 553-556.

Borralho N.M.G., 1992. Gains expected from clonal forestry under various selection and propagation strategies. Pp.328-338 *in* Proc. IUFRO Meeting on Mass Production Technology for Genetically Improved Fast Growing Forest Tree Species, Vol.1. Bordeaux, AFOCEL, France.

Bridgwater, F. 1995. A locus based model for advanced generation breeding. Pp. 43 *in* Proc. Evolution and Tree Breeding Conference Vol. 2. Canadian Tree Improvement Association / Western Forest Genetic Association, Victoria, British Columbia, Canada, August 28 - September 1, 1995. (Ed. J. Lavereau).

Bulmer M.G. 1971: The effect of selection on genetic variability. American Naturalist **105**: 201-211.

Burdon R.D. 1988: Recruitment for breeding populations: objectives, genetics and implementation. Pp. 554-572 *in* Proc. 2nd International Quantitative Genetics Conference, Raleigh NC, USA, 1-5 June 1987. (Eds. B.S. Weir *et al.*) Sinauer, Sunderland, MS, USA.

Burdon R.D., van Buijtenen J.P. 1990: Expected efficiencies of mating designs for reselection of parents. Can. J. For. Res. **20**: 1664-1671.

Burdon R.D., Namkoong G. 1983: Short Note: Multiple populations and sublines. Silvae Genetica **32:** 221-222.

Caballero, A., 1994. Development in the prediction of effective population size. Heredity **73**: 657-679

Cotterill P.P., 1989 The Nucleus breeding system. Pp. 36-44 *in* Proc. 20th South Forest Tree Improvement Conference. June 26-30, 1989, Charleston, SC, USA.

Crow J. F., Kimura M. 1970: An introduction to population genetics theory. Harper and Row, New York.

Dempfle L. 1975: A note on increasing the limit of selection through selection within families. Genetic Research Cambridge. **24**: 127-135.

Durel C. 1990: Parametres génétiques et selection en deuxieme generation d'amélioration du pin maritime *Pinus pinaster* ait. PhD. Thesis. Institut National Agronomique Paris-Grignon.

Falconer D.S. 1981: Introduction to quantitative genetics. 3rd edition. Longman 340 pp.

Gea L.D., Shelbourne C.J.A. 1995: Gain expectations from control pollinated main and elite breeding population and corresponding production populations. Pp. 876 *in* Proc. CRCTHF-IUFRO Conference. Eucalypt Plantations: Improving Fibre Yield and Quality, Hobart, Australia.

Kang H., Namkoong G. 1979: Limits of artificial selection under balanced mating systems. Silvae Genetica **28**: 53-60.

King J.N., Johnson G.R. 1993: Monte Carlo simulation models of breeding-population advancement. Silvae Genetica **42:** 2-3.

Lindgren D., Nilsson J. 1985: Calculations concerning selection intensity. Swedish University of Agricultural Sciences. Department of Forest Genetics and Plant Physiology. Umeå, Report 5.

Lindgren D., Gea L.D., Jefferson P. 1995: Status Number a measure of genetic diversity. Pp. 120-126 *in* Proc. Joint Meeting of the IUFRO Working Parties S2.02.05; 06; 012 and 14. Evolution of Breeding Strategies for Conifers from the Pacific North West. Limoges, France 28 July-4th August.

Lindgren D., Gea L.D., Jefferson P. 1996: Effective number and coancestry in breeding populations following within family selection. Silvae Genetica **45**: 52-59.

McKeand S.E., Bridgwater F.E. 1992: Third-generation breeding strategy for the North Carolina State University-Industry Cooperative Tree Improvement Programme. Pp. 234-240 *in* Proc. IUFRO Section S2.02-08, Breeding Tropical Trees. 'Resolving Gene Conservation and Domestication for New Species', Cartagena, Colombia.

Mahalovich M.R, Bridgwater F.E., 1989. Modelling elite populations and positive assortative mating in recurrent selection programs for general combining ability. Pp. 43-49 *in* Proc. 20th Southern Forest Tree Improvement Conference, Charleston, South Carolina. Springfield Virginia: National Technical Information Service.

Meuwissen T.H.E., Woolliams J.A. 1994: Effective sizes of livestock populations to prevent a decline in fitness. Theor. Appl. Genet **89**: 1019-1026.

Mullin T.J., Park Y.S. 1995: Stochastic simulation of population management strategies for tree breeding: a new decision-support tool for personal computers. Silvae Genetica **44**: 132-141.

Namkoong G., Roberds J. 1974: Choosing mating designs to efficiently estimate genetic variance components for trees. Silvae Genetica **23**: 43-54.

Namkoong G., Kang H.C., Brouard J.S. 1988: Tree breeding: principles and strategies. Springer Verlag. 174 p.

Nicholas F.W. 1980: Size of population required for artificial selection. Genetic. Research Cambridge **2**: 189-194.

Pswarayi I.Z., Barnes R.D. 1994: Genetic gains expected from phenotypic and combined index selection in *Pinus elliottii*. Zimbabwe Forestry Commission Research Paper No. 5. 11 Pp.

Russell J.H., Burdon R.D., Yanchuk A. 1996: Inbreeding depression and experimental selfed populations in western redcedar (*Thuja plicata*). Pp. 58 *in* Proc. Evolution and Tree Breeding Conference Vol. 2. Canadian Tree Improvement Association / Western Forest Genetic Association, Victoria, British Columbia, Canada, August 28 - September 1, 1995. (Ed. J. Laverau).

Shelbourne C.J.A., Burdon R.D., Carson S.D., Firth A., Vincent T.G. 1986: Development Plan for Radiata Pine Breeding. Forest Research Institute, Rotorua, New Zealand. Pp.142.

Sorensen D.A., Hill G.H. 1982: Effects of disruptive selection on genetic variance. Theor. Appl. Genet. **65**: 173-180.

Tier B. 1990: Computing inbreeding coefficients quickly. Genet. Sel. Evol. **22**: 419-430.

Toro M.A., Perez-Enciso M. 1990: Optimizing selection under restricted inbreeding. Gen. Sel. Evol. **22**: 93-107.

Verrier E., Colleau J.J., Foulley J.L. 1993: Long term effects of selection based on the animal model BLUP in a finite population. Theor. Appl. Genet. **87**: 446-454.

van Buijtenen J.P., Burdon R.D. 1990: Expected efficiencies of mating designs for advanced-generation selection. Can. J. For. Res. **20**: 1648-1663.

Wei R.P. 1995: Predicting genetic diversity and optimizing selection in breeding programmes. PhD. dissertation. Swedish University of Agricultural Sciences, Umeå. Department of Forest Genetics and Plant Physiology.

Wei R.P., Lindgren D. 1995: Optimal family contributions and a linear approximation. Theor. Pop. Biol. **48**.

White T., Hodge G.R., Powell G.L., 1993. An advanced-generation tree improvement plan for slash pine in the southeastern United States. Silvae Genetica

Williams C., Savolainen O. 1996: Inbreeding depression in conifers: implications for breeding strategy. Forest Science **42**: 102-117

Table 3.1: Cumulative Gains, Coefficients of Inbreeding and Relative status Numbers for ten generations of breeding, different population sizes and selection strategies.

A) Cumulative Gains

Generation	1 individual per family				2 individuals per family				no restrictions			
	12	32	64	128	12	32	64	128	12	32	64	128
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.59	0.62	0.60	0.58	1.02	1.10	1.10	1.09	1.43	1.62	1.79	1.84
2	1.18	1.21	1.19	1.17	1.94	2.06	2.07	2.06	2.52	2.82	3.10	3.14
3	1.74	1.77	1.78	1.76	2.80	2.99	3.02	3.00	3.48	3.92	4.39	4.46
4	2.31	2.36	2.38	2.36	3.64	3.89	3.97	3.96	4.38	5.00	5.57	5.67
5	2.85	2.93	2.95	2.96	4.44	4.80	4.90	4.91	5.14	5.98	6.56	6.82
6	3.41	3.49	3.55	3.55	5.24	5.68	5.85	5.85	5.83	6.90	7.51	7.92
7	3.94	4.06	4.12	4.13	5.99	6.54	6.76	6.79	6.45	7.67	8.36	8.80
8	4.48	4.64	4.70	4.72	6.72	7.43	7.68	7.73	6.97	8.35	9.12	9.60
9	5.02	5.20	5.27	5.31	7.42	8.29	8.59	8.65	7.43	8.97	9.83	10.40
10	5.51	5.77	5.85	5.89	8.08	9.12	9.49	9.58	7.81	9.50	10.49	11.00

B) Coefficients of Inbreeding

Generation	1 individual per family				2 individuals per family				no restrictions			
	12	32	64	128	12	32	64	128	12	32	64	128
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.023	0.008	0.003	0.002	0.053	0.018	0.008	0.006	0.203	0.146	0.129	0.099
3	0.047	0.016	0.008	0.004	0.099	0.040	0.018	0.010	0.334	0.272	0.246	0.194
4	0.067	0.026	0.012	0.006	0.150	0.060	0.030	0.016	0.450	0.365	0.352	0.278
5	0.089	0.031	0.015	0.008	0.190	0.075	0.038	0.020	0.540	0.464	0.435	0.359
6	0.106	0.039	0.020	0.010	0.230	0.096	0.050	0.026	0.620	0.549	0.510	0.441
7	0.127	0.047	0.023	0.012	0.270	0.111	0.057	0.030	0.686	0.619	0.581	0.509
8	0.144	0.055	0.027	0.014	0.300	0.128	0.067	0.035	0.741	0.675	0.637	0.556
9	0.163	0.062	0.031	0.016	0.330	0.147	0.076	0.040	0.783	0.726	0.685	0.608
10	0.182	0.069	0.035	0.018	0.370	0.163	0.086	0.044	0.823	0.769	0.725	0.648

C) Relative Status Numbers

Generation	1 individual per family				2 individual per family				no restrictions			
	12	32	64	128	12	32	64	128	12	32	64	128
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	0.67	0.67	0.67	0.67	0.49	0.48	0.47	0.47	0.20	0.11	0.07	0.05
2	0.50	0.50	0.50	0.50	0.32	0.30	0.30	0.30	0.12	0.06	0.03	0.02
3	0.40	0.40	0.40	0.40	0.24	0.22	0.22	0.22	0.09	0.04	0.02	0.02
4	0.34	0.33	0.33	0.33	0.19	0.17	0.17	0.17	0.08	0.03	0.02	0.01
5	0.29	0.29	0.29	0.29	0.16	0.14	0.14	0.14	0.07	0.03	0.02	0.01
6	0.26	0.25	0.25	0.25	0.14	0.12	0.12	0.12	0.06	0.02	0.01	0.01
7	0.23	0.23	0.22	0.22	0.13	0.11	0.10	0.10	0.06	0.02	0.01	0.01
8	0.21	0.20	0.20	0.20	0.12	0.10	0.09	0.09	0.05	0.02	0.01	0.01
9	0.19	0.19	0.18	0.18	0.11	0.09	0.08	0.08	0.05	0.02	0.01	0.01
10	0.18	0.17	0.17	0.17	0.10	0.08	0.08	0.07	0.05	0.02	0.01	0.01

Table 3.2: Cumulative Gains, Coefficient of Inbreeding and Status Number after ten generations of breeding for different mating designs and selection strategies for a population size of twelve and two heritabilities.

A) Cumulative Gains

	$h^2 = 0.2$						$h^2=0.05$				
Mating Design	**SPM**	**DP[1]**	**DD4**	**DD6**	**DIA**		**SPM**	**DP[1]**	**DD4**	**DD6**	**DIA**
No. crosses per parent	0.5	1	1.5	2.5	5.5		0.5	1	1.5	2.5	5.5
Family size	200	100	67	40	18		200	100	67	40	18
Phenotypic selection	7.20	7.73	7.98	7.95	7.96		3.84	4.40	4.47	4.39	4.30
Comb. Index (unrestricted)	7.22	7.91	8.08	8.00	7.93		4.74	5.57	5.23	5.13	4.90
Comb. Index 2 per Full-Sib	5.29	8.08	8.73	8.87	8.76		2.51	5.51	5.83	5.84	5.57
Comb. Index 1 per Full-Sib		5.51	7.51	8.23	8.64			2.68	4.52	5.28	5.45

B) Coefficients of Inbreeding

Phenotypic selection	0.430	0.420	0.400	0.400	0.380		0.350	0.350	0.340	0.340	0.330
Comb. Index (unrestricted)	0.820	0.820	0.800	0.770	0.700		0.840	0.830	0.810	0.780	0.690
Comb. Index 2 per Full-Sib	0.180	0.370	0.440	0.500	0.520		0.180	0.370	0.420	0.490	0.460
Comb. Index 1 per Full-Sib		0.180	0.270	0.350	0.420			0.180	0.260	0.330	0.370

C) Status Number

Phenotypic selection	1.07	1.08	1.13	1.11	1.16		1.27	1.24	1.27	1.28	1.29
Comb. Index (unrestricted)	0.58	0.58	0.59	0.61	0.67		0.57	0.57	0.58	0.60	0.68
Comb. Index 2 per Full-Sib	2.14	1.19	1.02	0.91	0.88		2.14	1.19	1.06	0.99	0.97
Comb. Index 1 per Full-Sib		2.14	1.57	1.24	1.06			2.14	1.60	1.32	1.18

1. DP represents an average for the one-cross-per-parent mating designs, DPC, DPF, PFM and DD3.

Figure 3.1: Mating designs (for a population size of 12 parents)

SINGLE PAIR MATING (SPM)

	1	2	3	4	5	6	7	8	9	10	11	12
1		X										
2												
3				X								
4												
5						X						
6												
7								X				
8												
9										X		
10												
11												X
12												

Cross per parent=0.5, family size 200

DOUBLE PAIR CYCLIC MATING (DPC)

	1	2	3	4	5	6	7	8	9	10	11	12
1		X										
2			X									
3				X								
4					X							
5						X						
6							X					
7								X				
8									X			
9										X		
10											X	
11												X
12	X											

Cross per parent:1, family size 100

DOUBLE PAIR FACTORIAL (DPF)

	1	2	3	4	5	6	7	8	9	10	11	12
1			X	X								
2			X	X								
3												
4												
5							X	X				
6							X	X				
7												
8												
9											X	X
10											X	X
11												
12												

Cross per parent:1, family size 100

PARTIAL FACTORIAL MATING (PFM)

	1	2	3	4	5	6	7	8	9	10	11	12
1							X					
2								X				
3									X			
4										X		
5											X	
6												X
7												
8	X											
9		X										
10			X									
11				X								
12					X							

Cross per parent:1, family size 100

DISCONNECTED DIALLEL (DD3)

	1	2	3	4	5	6	7	8	9	10	11	12
1		X	X									
2			X									
3												
4					X	X						
5						X						
6												
7								X	X			
8									X			
9												
10											X	X
11												X
12												

Cross per parent:1, family size 100

DISCONNECTED DIALLEL (DD4)

	1	2	3	4	5	6	7	8	9	10	11	12
1		X	X	X								
2			X	X								
3				X								
4												
5						X	X	X				
6							X	X				
7								X				
8												
9										X	X	X
10											X	X
11												X
12												

Cross per parent:1.5, family size 67

DISCONNECTED DIALLEL (DD6)

	1	2	3	4	5	6	7	8	9	10	11	12
1		X	X	X	X	X						
2			X	X	X	X						
3				X	X	X						
4					X	X						
5						X						
6												
7								X	X	X	X	X
8									X	X	X	X
9										X	X	X
10											X	X
11												X
12												

Cross per parent:2.5, family size 40

HALF DIALLEL (DIA)

	1	2	3	4	5	6	7	8	9	10	11	12
1		X	X	X	X	X	X	X	X	X	X	X
2			X	X	X	X	X	X	X	X	X	X
3				X	X	X	X	X	X	X	X	X
4					X	X	X	X	X	X	X	X
5						X	X	X	X	X	X	X
6							X	X	X	X	X	X
7								X	X	X	X	X
8									X	X	X	X
9										X	X	X
10											X	X
11												X
12												

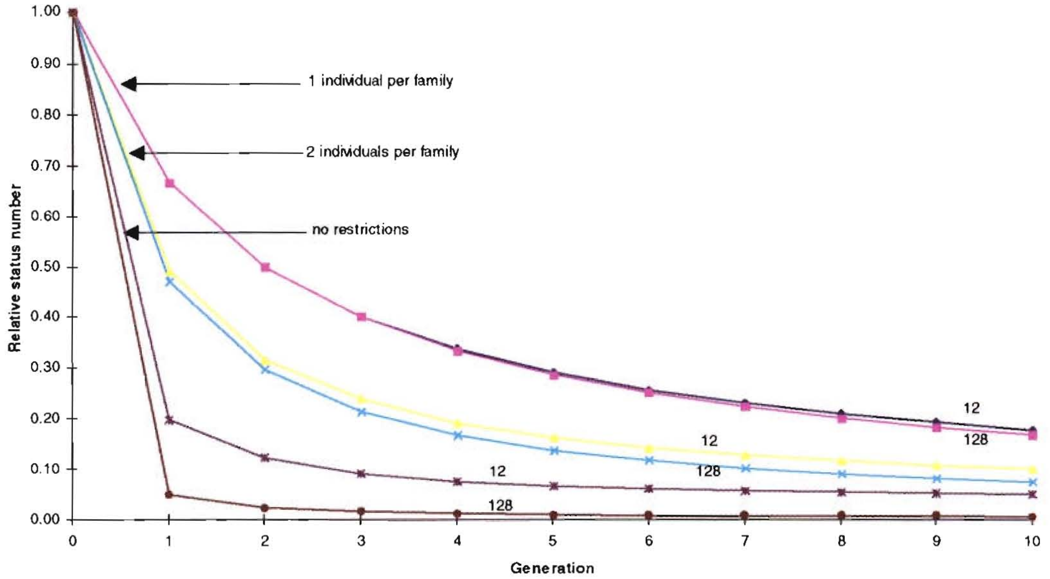
Cross per parent:5.5, family size 18

The mating designs are characterised by the number of crosses per parent:

Single-Pair Mating (**SPM**)	0.5	crosses per parent,
Double-Pair Mating (**DPM**), Double-Pair Factorial (**DPF**), Partial Factorial (**PFM**) and Disconnected Diallel with parents in groups of 3 (**DD3**)	1	cross per parent,
Disconnected Diallel with parents in groups of 4 (**DD4**)	1.5	crosses per parent,
Disconnected Diallel with parents in groups of 6 (**DD6**)	2.5	crosses per parent,
Half Diallel (**DIA**)	5.5	crosses per parent

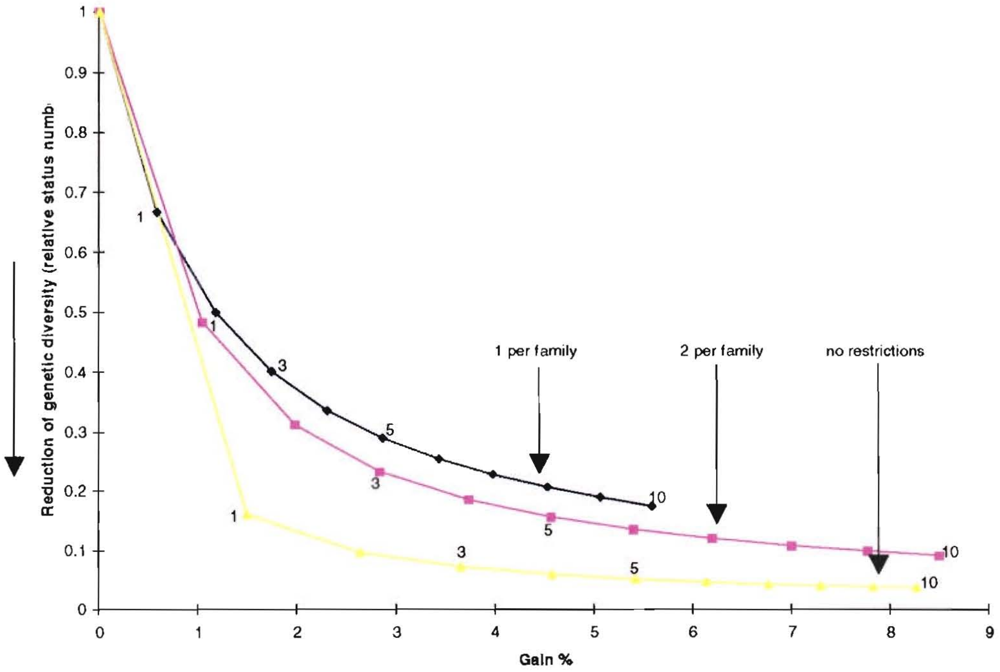
Family size was kept constant at 100 individuals per full-sib cross, assuming an expanding resource scenario.

Figure 3.2. Relative status number per generation for two population sizes (12 and 128), three selection strategies and a double-pair mating design



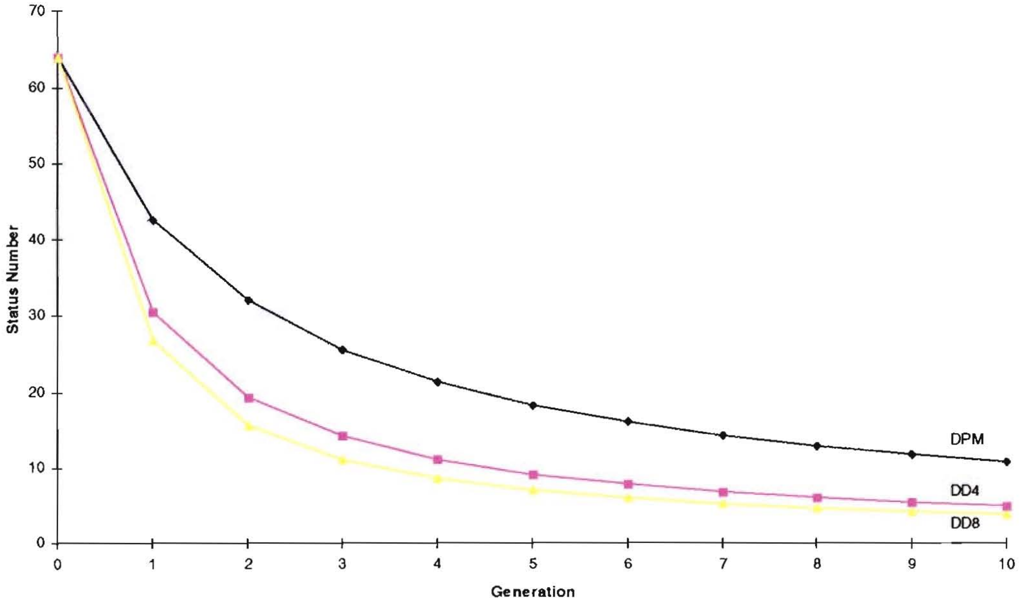
Relative status number under conservative strategies is marginally better preserved by small groups. If more individuals per cross are kept, the overall status number decayed faster and status number is then preserved much better in small groups. When no restrictions in the number of individuals selected per cross were applied almost 90% of the status number of the small group was lost in one generation (and nearly 80% for a large group).

Figure 3.3. Relative status number and genetic gain per generation for different selection strategies with DPM



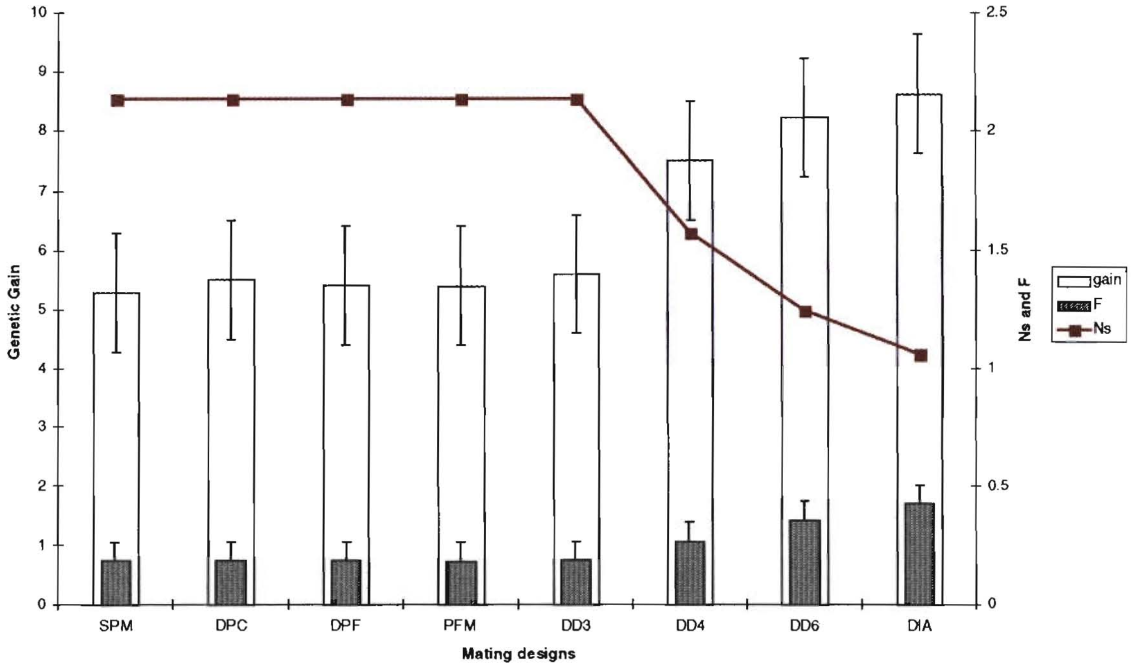
Higher levels of gain per unit of status number loss are obtained with a conservative within-family selection strategy but to reach the same level of gain, this strategy will take more generations to achieved it

Figure 3.4. Status number per generation for different mating designs and one individual selected per family.



Within-family selection with one individual chosen per full-sib cross maintained the highest status number (see Figure 3.1). Even so, the status number dropped to approximately one fourth of the original value in three generations, with the largest decrease occuring in the first generation. Status number values were also smaller for mating designs with more crosses per parent.

Figure 3.5. Genetic gain, coefficient of inbreeding and aggregated status number at generation ten for different mating designs; one individual selected per family (when feasible) and a population size of twelve.

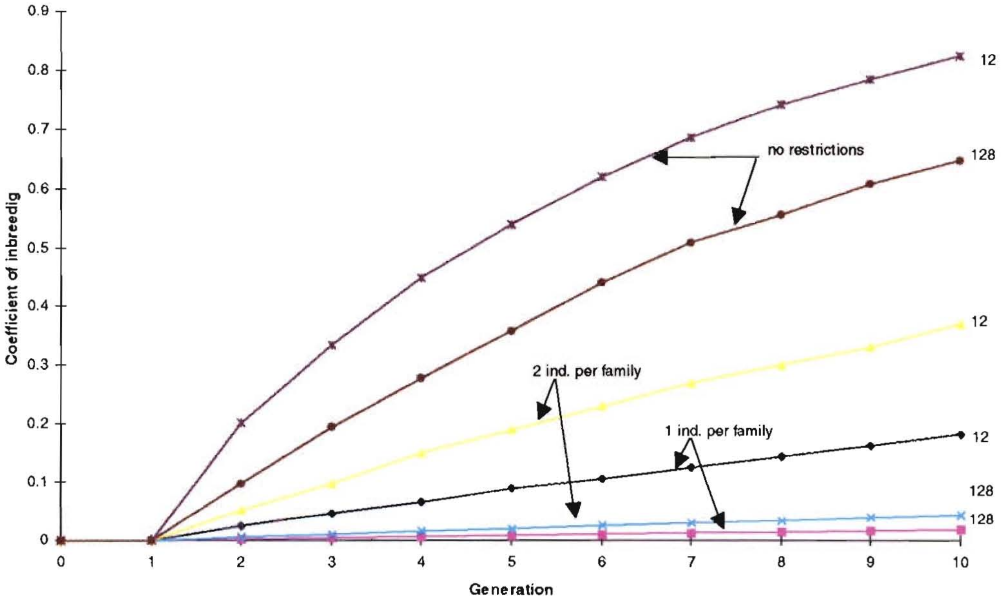


The predicted gains, inbreeding coefficients and status numbers for each mating design with one cross per parent were similar. However, gain for the disconnected diallel design (DD3) was slightly (but not significantly) higher than for the partial factorial mating (PFM), double-pair-factorial (DPF) and double-pair-cyclic (DPC) designs.

Predicted gains and inbreeding coefficents are higher for mating designs with more crosses per parent while status number values were smaller.

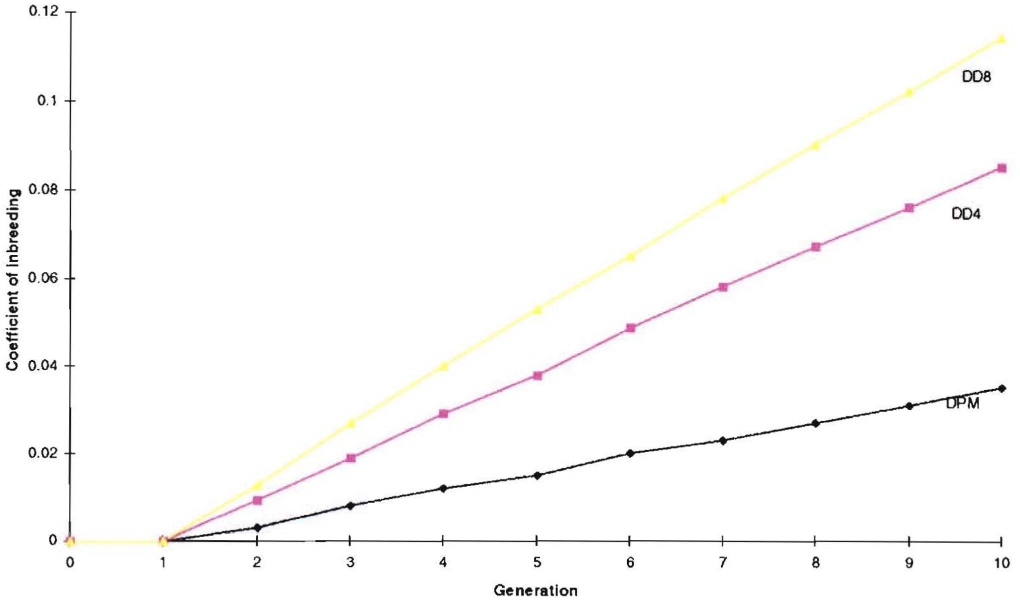
Predicted gains, status number and inbreeding coefficients for the single-pair mating design (SPM) are similar to the schemes with one cross per parent.

Figure 3.6. Coefficient of inbreeding per generation for two population sizes (12 and 128); three selection strategies and double-pair mating design



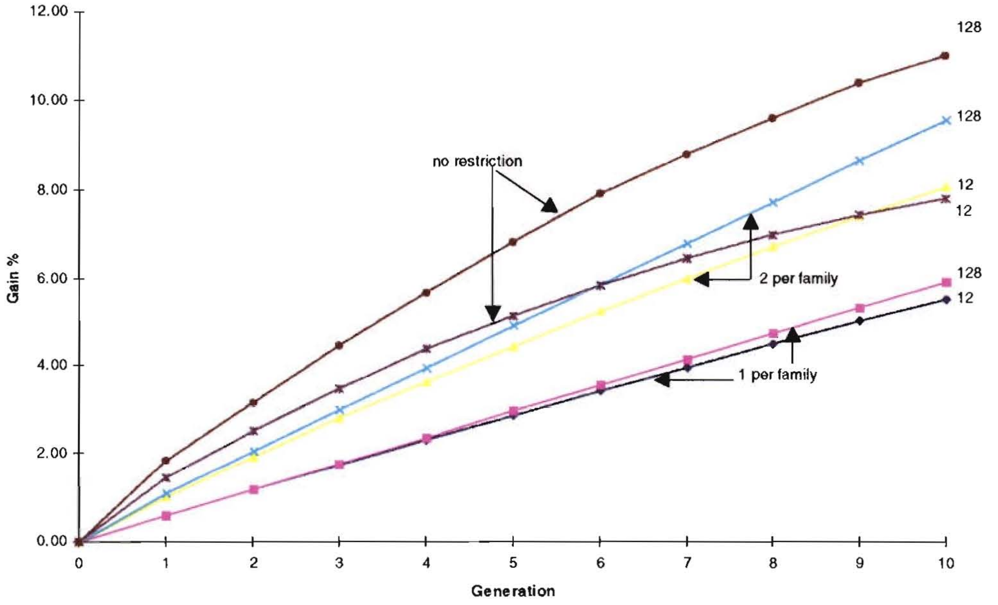
Population size is an important factor for delaying inbreeding. A selection strategy based on keeping two individuals per cross under a large group will have a significantly smaller inbreeding coefficient than a conservative strategy of one individual per family in a small population. The coefficient of inbreeding becomes extremely high for those selection strategies with no restrictions in the number of individuals per family.

Figure 3.7. Coefficient of inbreeding per generation for different mating designs and one individual selected per family



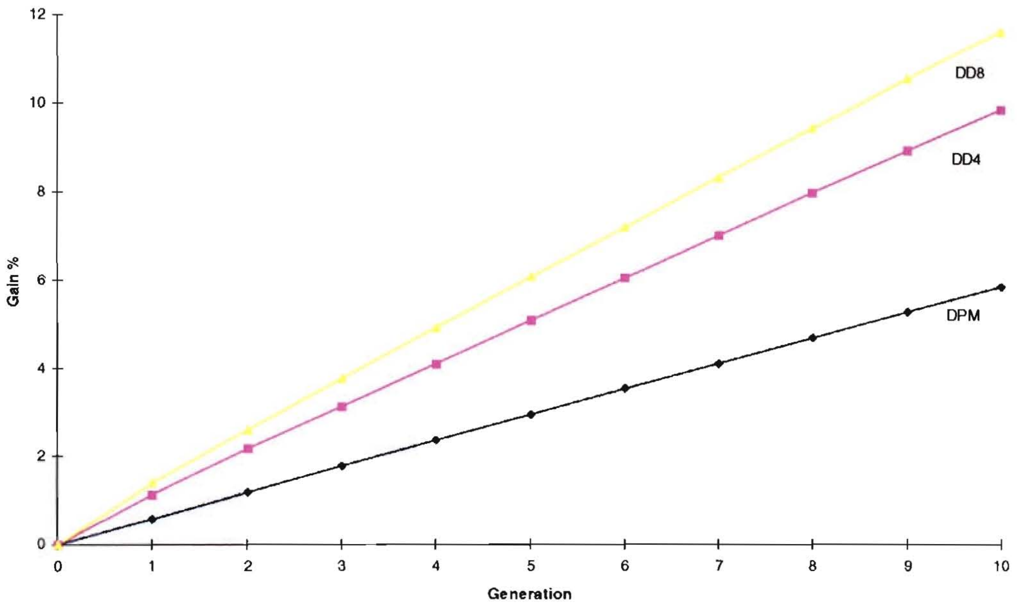
Inbreeding coefficients are higher for mating designs with more crosses per parent.

Figure 3.8. Genetic gain per generation for two population sizes (12 and 128); three different selection strategies and a double-pair mating design



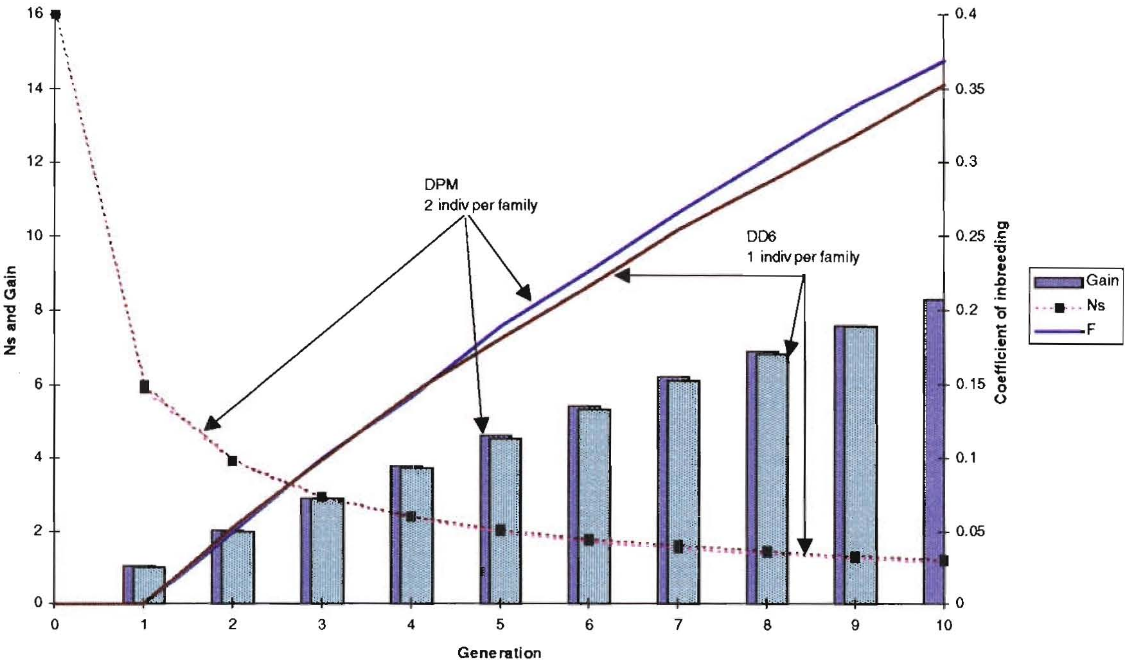
Breeding population gain increased with group size and is more influenced in later generations. There is little difference in gain for the first two generations under the restrictions. Gain for the 12 parent breeding group with no restriction in the number of individuals kept per cross slow its growth around generation 5. This 'plateauing' in gain for the smaller population under an aggressive selection strategy coincide with high inbreeding coefficient values.

Figure 3.9. Genetic gain per generation for different mating designs and 1 individual selected per family



The predicted gains for each mating design with one cross per parent were similar (see Figure 3.4). Substantial extra gain results from restricted combined index selection (and less so for unrestricted selection) if more crosses per parent are made; for instance, during the first four generations of breeding gain almost doubles from 0.5 cross per parent to 1.5-crosses per parent (DD4).

Figure 3.10. Genetic gain, inbreeding coefficient and status number for two different mating designs and selection strategies and a population size of twelve.



Almost exactly the same status numbers, coefficients of inbreeding and marginally higher gains can be obtained with a strategy based on the same number of crosses as parents (DPM) and a maximum of two individuals selected per family as those values obtained with a mating design of 2.5 crosses per parent (DD6) and a maximum of 1 individual per family.

CHAPTER 4
MODEL AND REALITY

PROBLEMS ASSOCIATED WITH THEORETICAL COANCESTRY

The application of the principle of identity by descent to genetic models has allowed the formulation of general expressions for the covariance of relatives belonging to an ideal population and, for instance, the general expression of traditional effective population size numbers and status number. However, departures from the basic assumptions can be mentioned and the effects of such departures on the current theory (mostly for artificially constructed breeding populations) have not been fully investigated. Traditional effective population size (effective number) was defined as the number of individuals that would give rise to the same increase in the level of inbreeding, if they bred in the manner of the idealised population (Chapter 2 Pp. 8). The new concept proposed here "Status Effective Number" (N_S), is a departure from traditional effective numbers, and is more like a snapshot of the population at a given stage (rather than a movie, depicting how the population would behave if left to follow a particular course).

Askew and Burrows (1983) and Burrows (1984) developed an average coancestry-based measure (N_{eB}), which is similar to N_s and under some circumstances behaves as traditional effective population size for a single cycle. Calculated average coancestries; however, differ from those using N_s, as the self-coancestry is not considered. Some undesirable features, for example, that the effective number may sometimes be increased by decreasing census number, have already been addressed by Wei and Lindgren 1995 (see also Chapter 2).

UNCERTAINTIES OF THE MODEL

The basic concepts of quantitative genetics and the accuracy of the genetic parameters used in this thesis, including additive genetic variation and heritability depend on a set of assumptions defined in the model. The likely impacts of possible violations of the most important assumptions of the models used in this thesis; infinitesimal model, selective neutrality, no linkage, no mutation, initial F=0 and simultaneous departures from several assumptions will be reviewed below.

Infinitesimal model

The model known as 'Fisher's infinitesimal model' assumes that inherited differences among individuals for any quantitative trait are due to many independent genes, each of small effect on the phenotype, and that these effects are cumulative. Contrary to Fisher's model, for forest trees genes with major effects rather than small and equal effects seem to be involved in the control of a significant proportion of the variation in a number of quantitative traits (as described by Grattapaglia, 1995 for *Eucalyptus*). Barker (1995) suggested that while short-term responses can be predicted using the infinitesimal model, models for long-term predictions should include genes with large effects.

Neutrality assumption

While great advances have been made in tree breeding programmes covering a diversity of breeding goals, the appropriate stance towards inbreeding is still debatable (Williams and Savolainen 1996, Burdon 1995). Inbreeding has strong theoretical attractions as a breeding tool, since (in the absence of 'hard' genetic load) it offers the benefits of exposing a greater proportion of the additive variance to recurrent selection, enabling more rapid fixation of favourable genes under recombination (Falconer 1981). The rapid fixation of any major genes under directional selection, is in principle particularly especial under inbreeding. Thus, quantitative trait loci (QTL) can be more easily tracked through pedigrees. However, estimated levels of inbreeding depression are relatively high in many forest tree species, and this, combined with the conservative approach appropriate to genetic management of long-rotation crops, has generally inhibited the aggressive use of inbreeding. A promising exception, however, is *Thuja plicata* (Russell *et al* 1996).

The neutrality assumption will be violated in some ways by the very fact of selective breeding. This assumption is especially important to whether true inbreeding as theoretically expected approach towards homozygosity proceed. The excess of homozygotes formed by selfing and mating between relatives seems to disappear through natural selection between fertilisation and reproduction (Williams and Savolainen 1996), due to genetic load, providing a concrete indication that the assumption of neutrality may often be strongly violated.

No linkage

A problem with N_s is that natural tree populations seem to have inbreeding coefficient zero as some sort of natural base-line for the breeding population, reflecting non-neutrality. Under natural selection, genes linked with genes for increased fitness will be maintained in higher frequencies. Chromosome fragments or linkage groups from ancestors containing such genes will be more common than indicated by the pedigree, leading to increased coancestry. Similar effects will occur under artificial selection, where very rare alleles in the original population will become common (Libby, 1991) if they are favourable. Thus, status number can only be a strictly valid indicator for the diversity of genes that are neutral under selection.

No mutation

Hill (1987) showed that good approximations are possible with simple models if we assume symmetric distribution of mutant effects, which seems most unlikely. However, for long-term predictions used in designing and executing breeding programs the aim must be, as far as possible, to obtain parameters for mutation appropriate for insertion into models.

Initial f_1=0

A rapid erosion of status number occurred even with conservative strategies like within-family selection (Chapter 3 Table 3.2). As much as two-thirds of the status number will be estimated as lost in most programs within a few breeding generations if the assumption of f_1=0 for the founders is correct.

Table 4.1 summarises the effects on N_s and coefficient of inbreeding of a population of 100 individuals under random mating, assuming inbreeding and coancestry in the initial population of 0; 0.01; 0.02 or 0.05. The effect of different coancestry coefficient levels is shown by the ratio among the status number at generation 9 and the status number at generation 0. It is clear that the initial run-down of N_s is very dependent on the initial assumptions about the level of relatedness of the founders. Perhaps N_s will be a more useful concept if the initial coancestry is set at 1 or 2 percent, which may be more realistic than f = zero (Figure 4.1). Relationships between relatives will usually be unchanged and the picture will be less alarming.

Simultaneous departures from several assumptions

So far, the likely impacts of departures from the assumptions have addressed one assumption at a time. However, it is probably when two or more assumptions are violated simultaneously that one gets the biggest impact. For instance, departures from neutrality, in conjunction with linkage and the presence of major genes (which could often reflect recent mutation) are likely to drag along whole linkage groups and calculated N_s could underestimate the remaining diversity or genetic base, particularly with selection for balanced heterozygotes which does appear to occur in nature.

INDICATIONS FOR FUTURE WORK

Barker (1995) suggested that while short-term responses can be predicted using the infinitesimal model, models for long-term predictions should include genes with large effects. Hill (1987) warned about models that include no mutation and no linkage (even though he also proved that the influence of linkage is likely to be small even for genes on the same chromosome; only when the distribution of mutant effects was asymmetric were linkage effects important). Wherever there is reliable information to go on, and it can be incorporated, more realistic parameters should definitely be inserted into models.

Recent information from molecular biology challenges the current understanding of the nature of at least some of the quantitative genetic variation (Lande and Thompson, 1990). Genes with major effects seem to be involved in the control of a significant proportion of the variation in a number of quantitative traits (as described by Grattapaglia, 1995 for *Eucalyptus*).

To investigate the departure from the assumptions listed above and thereby make tree breeding more precise, techniques from molecular biology offer some powerful tools. The introduction of molecular markers will make it potentially feasible to measure coancestry and compare predictions with measurements (Bernardo 1993). Such comparisons will probably help to identify the situations when there are discrepancies between true and theoretical coancestries, also it will reveal levels of departure from the neutrality assumptions and will give us real information about linkage and possibly even about the contribution of mutation to genetic variation. With this knowledge, adjusting the models will be easier.

CONCLUSIONS

- Including the coancestry of the individual with itself in the computation of the status number, resulted in a number that avoids the instability of traditional measures of population sizes for situations where the traditional N_e is undefined. This scenario was obscure under Burrows average coancestry-based measured.

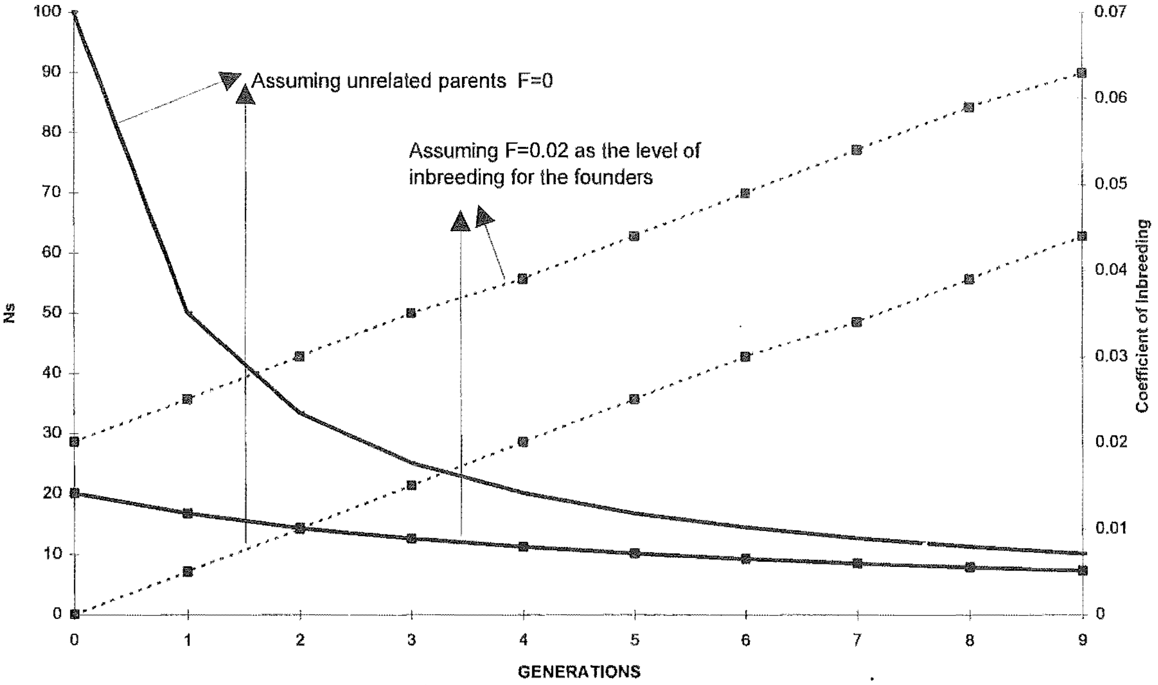
- It is likely that the assumptions used for the development of status number (zero coefficient of inbreeding for the founders, in conjunction with genes independently assorted, no mutation and interactions or combinations from departures of the neutrality assumption), means that calculations and predictions based on pedigrees will be biased.

- Status number is more applicable to breeding populations rather than for natural populations where it is more likely that f exceeds zero. Even so, N_s looks to be a more useful concept if the initial inbreeding in breeding populations is set at 1 or 2 percent, which may be far more realistic than f_0 = zero.

Table 4.1 Status Number and Inbreeding coefficient under random mating for a population of 100 individuals assuming different levels of relatedness for the founders.

Gen.	**F**	**N_s**	**F**	**N_s**	**F**	**N_s**	**F**	**N_s**
0	**0.000**	100.00	**0.010**	33.45	**0.020**	20.08	**0.05**	9.13
1	0.005	50.12	0.015	25.15	0.025	16.79	0.055	8.40
2	0.010	3350	0.020	20.18	0.030	14.44	0.059	7.79
3	0.015	25.18	0.025	16.86	0.035	12.67	0.064	7.26
4	0.020	20.20	0.030	14.49	0.039	11.29	0.069	6.80
5	0.025	16.78	0.035	12.71	0.044	10.19	0.074	6.39
6	0.030	14.50	0.039	11.32	0.049	9.29	0.078	6.04
7	0.034	12.72	0.044	10.22	0.054	8.54	0.083	5.72
8	0.039	11.33	0.049	9.32	0.059	7.91	0.087	5.44
9	0.044	10.22	0.054	8.56	0.063	7.36	0.092	5.18
RATIO N_{s0}/N_{s9}		9.77		3.90		2.72		1.76

Figure 4.1 Status Number (N_s)(—) and Inbreeding coefficient (F)(- -) under random mating for a population of 100 individuals assuming different levels of relatedness for the founders.



REFERENCES

Askew G.R., Burrows P.M. 1983: Minimum coancestry selection I. A *Pinus taeda* population and its simulation. Silvae Genetica **32**: 125-131.

Barker J.S.F., 1995: Quantitative genetic models: past, present and future challenges. Pp. 3-16 *in* Proc. Canadian Tree Improvement Association, Part 2. CTIA/WFGA Conf." Evolution and Tree Breeding", Victoria, BC, 28 Aug.-1 Sept. 1995 (Ed. J. Lavereau).

Burdon R.D. 1995: Future directions in tree breeding: Some questions of what we should seek, and how to manage the genetic resource. Pp. 24-39 *in*: Proc. Canadian Tree Improvement Association, Part 2. CTIA/WFGA Conf." Evolution and Tree Breeding", Victoria, BC, 28 Aug.-1 Sept. 1995 (Ed. J. Lavereau).

Burrows P.M. 1984: Inbreeding under selection from related families. Biometrics **40**: 895-906.

Falconer D.S. 1981: Introduction to Quantitative Genetics. Longman, Inc. 340 pp

Grattapaglia D., 1994: Genetic mapping of quantitatively inherited economically important traits in *Eucalyptus*. PhD. Thesis, North Carolina State University.

Hill W.G. 1987: Increasing the rate of genetic gain in animal production. University of New England, Armidale, NSW, Australia.

Libby W.J. 1991: Use of genetic variation for breeding forest trees. Pp.101-117 *in* Proceedings Plant Breeding in the 1990s (Ed. Stalker and Murphy). North Caroline State University, Raleigh, March 1991.

Russell J.H., Burdon R.D., Yanchuk A. 1996: Inbreeding depression and experimental selfed populations in western redcedar (*Thuja plicata*). Pp. 58 *in* Proc. Canadian Tree Improvement Association, Part 2. CTIA/WFGA Conf." Evolution and Tree Breeding", Victoria, BC, 28 Aug.-1 Sept. 1995 (Ed. J. Lavereau).

Wei R.P., Lindgren D. 1995: Optimal family contributions and a linear approximation. Theoretical Population Biology **48**: 318-332

William C. and Savolainen O. 1996: Inbreeding depression in conifers: Implications for breeding strategy. Forest Science **42**: 102-117

CHAPTER 5

COMPLEMENTING INBREEDING COEFFICIENT INFORMATION WITH STATUS NUMBER: IMPLICATIONS FOR STRUCTURING BREEDING POPULATIONS.

Both ongoing genetic improvement and maintenance of genetic variability for longer-term genetic gain are generally achieved in most breeding populations, reflecting the relatively high emphasis placed on starting with a broad genetic base. A desirable effective population size of upwards of 200 is a widely accepted guideline (White 1992). A major development has been put forward by Namkoong (1976), see also Eriksson *et al* (1994) with the multiple-populations concept, which addresses a diverse array of breeding goals by subdividing the breeding population. This system enhances also the probability of retaining alleles if they have selective advantages within any sub-unit while the overall frequency of neutral alleles may remain essentially unchanged. Other aspects of breeding-population management have also been addressed and reviewed, like choice of mating designs, elite sub-units, variable representation of parents etc. (Lindgren *et al* 1997).

An important objective in tree breeding is to guarantee future access to combinations of unrelated genotypes for actual plantation establishment; in order to avoid effects of inbreeding depression in the production population. This goal can be achieved by arranging the breeding population in unrelated sublines as defined by Burdon and Namkoong (1983). The most important source of gain in low-heritability situations is the exploitation of genetic variance by intensive among-family selection. As discussed in Chapter 3 (Table 3.2), the larger the genetic variance (which is assumed to be tracked by status number) in the breeding population, the larger the gain that can be achieved in the production population. Thus, accepting a lower immediate gain in the breeding population by keeping a higher status number does not necessarily mean a reduced gain in the production populations. Optimisation of gain can come from achieving the appropriate balance of selection and diversity across both the breeding and production populations.

Breeding schemes with small groups are slightly more efficient in preserving status number through a large number of generations than breeding schemes with large groups, but medium- to large-size breeding groups showed larger increases in gain (Chapter 3 Table 3.2). Inbreeding in small breeding groups may become so severe as to cause fertility problems and considerably reduce the efficiency of selection for additive gene effects. Thus, using very small breeding groups will probably not provide a sustainable long-term breeding strategy. Nevertheless, small groups may form a critical component of breeding strategies that employ marker-assisted selection, since the maintenance of marker-QTL associations will be facilitated in these small populations.

When inbred breeding groups are merged, the variance within the new composite group will increase compared to within the pre-existing variance and if all the genetic variance is not additive the absolute and relative amounts of additive and non-additive genetic variance will change (Falconer, 1981).

Combining small groups will also change allelic frequencies and there will (possibly) be a short-term period of gametic-phase disequilibrium and (possibly) a longer-term period of linkage disequilibrium as population-specific chromosomes are recombined (Libby, 1991). For instance, total genetic variances will be initially larger or smaller than if unlinked, and they will approach the unlinked values as the disequilibria decay. Although smaller breeding groups raise the inbreeding coefficient and lower gain, they do lead to slightly higher aggregate status number. Only by using very small groups however will substantial cumulative benefits accrue in maintenance of status number after many generations.

Large groups will considerably delay the build-up of inbreeding level, conferring an advantage in this respect over small groups. Breeding groups smaller than 10 parents will be impossible to maintain for longer than a few generations without inbreeding. However, even small groups can be maintained without inbreeding for the first two generations, and there are no really compelling reasons to make a decision about size before significant inbreeding becomes unavoidable. Small groups can always be merged, but big groups will become impossible to split into smaller, fully disconnected units.

A breeding strategy with small breeding groups will differ from a multiple-elite strategy primarily in respect of breeding goals addressed. While breeding groups can have the same overall breeding objectives, a multiple-elite strategy might be designed to attend different breeding goals on each elite.

Several small elites might be an effective way to select for special-purpose traits as well as provide extra gain for a general-purpose breed (Appendix 1). Increasing the number of crosses in the small elites appears to be a cost-effective strategy for increasing gains in the short term. However, under a closed strategy, the coefficient of inbreeding will increase to quickly with less than 12 parents per elite.

Restricted selection

Often, a tree breeding strategy is based on mating and field-test designs in which randomised, replicated progenies are used to solidly estimate General Combining Ability (GCA) values of parents for seed production, and at the same time provide material in which to make forward selections for the next generation of breeding (Carson *et al* 1990). With selection so critical to the success of every breeding programme it is important to improve the selection efficiency and effectively use all the information available in breeding decisions. Choice of a selection strategy within a breeding population will normally be determined mainly by its ability to maintain genetic diversity in the long term.

Crow and Kimura (1970) showed that truncation selection is the most efficient form of directional selection in changing gene frequencies. But although truncation selection without restrictions will maximise genetic gain (in low-heritability situations), it will be expected to have a high cost in terms of reduced effective population size (by contrast, within-family selection will maximise effective population size but genetic gain will be small). Even for the large elite simulated in this thesis, the drop in status number following unrestricted combined index selection may be considered much too rapid to be acceptable with any mating design. Furthermore, selection schemes that maximise gain by unrestricted combined index selection will result in rapid inbreeding, and may not be sustainable in the long term, but could still be a feasible approach for short-term gains and for species that can cope with high levels of inbreeding.

Unbalanced mating designs

Unbalanced breeding schemes, which contain elements of assortative mating (best with best), superimposed upon random mating components, appear to offer a means of increasing gain over random mating for a given selection intensity (Falconer, 1981). Lindgren and Gregorius (1976) proposed promoting some of the features of assortative mating by the use of unbalanced family contributions to a breeding population. Kang and Namkoong (1988) showed that, in some circumstances, unbalanced designs where mating frequency is in proportion to an individual's breeding value could even generate larger effective population sizes (N_e) for the same level of gain than truncation selection followed by balanced mating. Lindgren (1993) showed the superiority that "linear deployment" has when compared with truncation selection followed by non-assortative mating. The simulations presented in this thesis showed that gains may be unacceptably low, especially in low-heritability situations, for mating systems and selection procedures that do not include a between-family component, which promotes some kind of unbalanced family contribution. In any case, it seems impossible to avoid having the status number dropping to something like one quarter of its initial value after three rounds of selection (cf. Chapter 3 Figure 3.2).

Single-pair mating

In a multi-generation situation, single-pair mating is an efficient way to preserve status number but is sub-optimal for gain. An increase from 0.5 or 1 cross per parent (applied as within-family selection) to 2.5 crosses per parent with use of restricted combined index selection resulted in substantial increases in genetic gain (of two- or three-fold), depending on heritability (cf. Chapter 3 Table 3.2). However, increasing the number of crosses per parent up to 2.5 does carry a modest penalty under the prescribed selection constraints of increased coefficient of inbreeding and reduced status number.

Any selection scheme that is based exclusively on within-family selection will produce low short-term gains, but after some generations may result in both a higher cumulative gain and a higher status effective number than selection schemes that maximise immediate gain (cf. Chapter 3 Table 3.2).

CONCLUSIONS

- Status number is not intended to replace any other measures but rather it is complementary to existing measures and provides information for situations when the inbreeding coefficient and traditional measures of effective population size are less useful.

- The larger the genetic variance (as estimated by status number) in the breeding population, the larger the gain that can be expected to be achieved in the production population. Optimisation of gain can come from achieving the appropriate balance of selection and diversity across both the breeding and production populations.

- Breeding schemes with small breeding groups are slightly more efficient in preserving status number through a large number of generations than breeding schemes with large groups, but medium- to large-size breeding groups showed larger increases in gain.

- The drop in status number following unrestricted combined index selection may be considered much too rapid to be acceptable with any group size and mating design. Selection schemes that maximise gain by unrestricted combined index selection will also result in rapid inbreeding, and will not be sustainable in the long-term, but could still be a feasible approach for short-term gains (eg. for production populations) and for species that can cope with high levels of inbreeding.

- Gains may be unacceptably low, especially in low-heritability situations, for mating systems and selection procedures that do not include a between-family component of selection. Any selection scheme that is based exclusively on within-family selection will produce low short-term gains, but after some generations may result in both a higher cumulative gain and a higher status effective number than selection schemes that maximise immediate gain.

- Single-pair mating is an efficient way to preserve status number but is sub-optimal for gain. Increasing the number of crosses per parent up to 2.5 does carry a modest penalty of increased coefficient of inbreeding and reduced status number.

- How aggressively the breeder can use inbreeding as a tool will depend on the tolerance of the species to different levels of inbreeding.

REFERENCES

Barker J.S.F., 1995: Quantitative genetic models: past, present and future challenges. Pp. 3-16. *in* Proc. Canadian Tree Improvement Association, Part 2. CTIA/WFGA Conference." Evolution and Tree Breeding", Victoria, BC, 28 Aug.-1 Sept. 1995 (Ed. J. Lavereau).

Bernardo R. 1993: Estimation of coefficient of coancestry using molecular markers in maize. Theor. Appl. Genet. **85**: 1055-1062

Burdon R.D., Namkoong G. 1983: Multiple populations and sublines. Silvae Genetica **32**: 221-222.

Carson M., Burdon R., Carson S.D., Firth A., Shelbourne C.J.A., Vincent T.G., 1990: Realising genetic gains in production forests. *in* Proc. IUFRO Working Parties on Douglas-Fir, Lodgepole pine, Stika and *Abies* spp. Breeding and Genetic Resources Session.

Crow J.F., Kimura M. 1970: An Introduction to Population Genetics Theory. Harper and Row Publishers, New York.

Falconer D.S. 1981: Introduction to Quantitative Genetics. Longman, Inc. 340 pp

Eriksson G., Namkoong G., Roberds J.H. 1994: Dynamic gene conservation for uncertain futures. For. Ecol. and Manage. **62**: 15-37

Grattapaglia D., 1994: Genetic mapping of quantitatively inherited economically important traits in *Eucalyptus*. PhD. Thesis, North Carolina State University.

Hill W.G. 1987: Increasing the rate of genetic gain in animal production. University of New England, Armidale, NSW, Australia.

Kang H., Namkoong G. 1988: Inbreeding effective population size under some artificial selection schemes. I. Linear distribution of breeding values. Theor. Appl. Genet. **75**: 333-339.

Lande R., Thompson R. 1990 : Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics **124**: 734-756.

Libby W.J. 1991: Use of genetic variation for breeding forest trees. Pp.101-117. *in* Proc. Plant Breeding in the 1990s. North Carolina State University, Raleigh, March 1991 (Ed. H.T. Stalker and J.P. Murphy).

Lindgren D. 1993: Quantitative comparison between truncation selection and a better procedure. Hereditas **118**: 289-292.

Lindgren D., Wei R.P, Lee S.J. *In Press*: How to Calculate Optimum Family Number when Starting a Breeding Program. Forest Genetics

Lindgren D., Gregorius H.R. 1976: Inbreeding and Coancestry. Pp 49-72. *in* Proc. IUFRO Joint Meeting on Advanced Generation Breeding , Bordeaux, France.

Namkoong G. 1976: A multiple index selection strategy. Silvae Genetica **25**:199:201.

White T.L. 1992: Advanced-generation breeding populations: Size and structure. Pp 208-22. In Proc. IUFRO Meeting on Breeding Tropical Trees, Cartagena & Cali, Colombia, Oct 1992. CAMCORE, Raleigh, NC, USA.

Williams C. and Savolainen O. 1996: Inbreeding depression in conifers: Implications for breeding strategy. Forest Science **42**: 102-117

APPENDIX I

DETERMINISTIC PREDICTIONS OF GAINS FROM CONTROL-POLLINATED MAIN AND ELITE BREEDING POPULATIONS AND CORRESPONDING PRODUCTION POPULATIONS

Presented at the Poster Session of the CRCTHF-IUFRO Conference. Eucalypt Plantations: Improving Fibre Yield and Quality. IUFRO Working Parties S2.02-09 (Eucalypt provenances and breeding) P2.02-01 (Productivity of eucalypts). Hobart, Australia. 19-24 February 1995

INTRODUCTION

Genetic gains are accumulated each generation in the breeding population through selection and intermating. Progeny of the selected individuals are planted to become the next generation breeding population for application of future selection (Allard 1960; Cotterill 1989; Shelbourne *et al* 1986). Delivery of gains is made by propagating the best individuals as parents in clonal seed orchards, and amplification of improved seed through vegetative propagation (which may include micropropagation. Control-pollinated orchards have resulted in genetic gains that are larger than those obtained from wind-pollinated orchards (Carson *et al* 1992).

The question of how to integrate selection, breeding and deployment in commercial crops to generate greater gains has been tackled by several authors and their predictions have been an important mechanism of assessing different breeding strategies without committing resources to expensive breeding operations (Namkoong *et al* 1966; Libby 1969; Shelbourne 1969; Matheson and Lindgren 1985; Shaw and Hood 1985; Carson 1986; Cotterill 1986; Cotterill and Jackson 1989; Lindgren and Werner 1989; van Buijtenen and Burdon 1990; Mahalovich 1990; Shelbourne 1992; King and Johnson 1993; White *et al* 1993; Mahalovich 1992; Mullin 1993).

In general, combining higher gains with faster generation turnover to maximise genetic gain per unit of time is the aim. However, the increase in relatedness during successive generations of recurrent selection is perceived to be a major problem in long-term breeding. It can be delayed by restricting selection to within families, and by equal and symmetric representation of all parents in mating schemes (Kang and Namkoong, 1979; Burdon 1988; Dempfle 1975). Selecting predominantly within families will minimise reductions in effective population size in the breeding population but has the disadvantages of disregarding family information, resulting in selection on each generation utilising only half the original additive variance (Burdon 1988).

Within-family selection is potentially inefficient if each genotype is unreplicated. The potential for clonal replicates to increase test efficiency, however, is influenced by the heritability for the trait (Shaw and Hood 1985; Mullin 1993). When all the genetic variance is additive, using clonal replicates in test plantations will often increase the average genetic gain (even for a fixed resource), and will be especially effective at lower heritabilities.

The concept of stratify the breeding population into a small nucleus (or elite) and a larger 'main' component was introduced into tree breeding from animal breeding by Cotterill *et al* (1986) and was further adapted by Mahalovich (1990); McKeand *et al* (1992) and White *et al* (1992). The purpose is to focus efforts and resources of the programme (in terms of breeding and testing) on the best selections. This development has been driven by the need to achieve high recurrent gains in the short term. Adding this level to the hierarchy of breeding populations was reported to increase predicted gains per year over those from the main-line population (Mahalovich 1990), but mainly because a shorter breeding generation was assumed for small elite populations versus a large main.

Deterministic simulation produces a single result because a series of algebraic functions are used to predict the outcome for a given scenario of events. Deterministic models provided accurate predictions for evaluating expected genetic response from alternative breeding structures. They are useful for the rapid evaluation of selection responses expected from alternative strategies. In one study, predictions of response to selection, genetic variance and inbreeding levels from a deterministic model have been shown to agree closely with results from stochastic simulation (Keller *et al* 1991).

This chapter compares deterministically predicted genetic gains, for a single trait, from control-pollinated main and elite breeding populations after two generations and gains from their respective production-population seed orchards. Within-family selection was used for the main breeding population to delay the effects of inbreeding. The effects of cloning and increasing the number of crosses in the elite breeding population are also simulated.

MATERIAL AND METHODS

Gain expectation equations used here are based on Shelbourne (1969, 1992) and assumed a univariate additive genetic model, therefore with a single trait or composite index considered. Formulae considered the effect of reduced variance associated with comparisons within finite groups (Burdon 1982) but were simplified by omitting the effects on truncation of additive variance through selection and the consequent gametic-phase disequilibrium.

Gains were predicted on the basis of two-stage, among and within-family selection to allow direct control of selection intensities:

$$\Delta G = i\sigma_p b \qquad \textbf{(AI.1)}$$

where i = selection intensity
where b = heritability for the particular selection system (Namkoong *et al* 1966)

$$= k\frac{\sigma_A^2}{\sigma_P^2} \qquad \textbf{(AI.2)}$$

where k = fraction of the total additive genetic variance in the covariance of additive values for the particular relative in question, eg.
k = ¼ for selection among half-sib families
k = ¾ for selection within half-sib families,
σ_P^2 = phenotypic variance
σ_A^2 = additive genetic variance

This basic gain formula has been adapted to a number of breeding and production populations options by Shelbourne (1992); where:

i = selection intensities between (i_1)and within families (i_2) ,
c = number of clones per family,
r = number of ramets per clone,
n = number of individuals per family (c∗r),
σ_e^2 = environmental variance,
σ_A^2 = additive genetic variance,
$\sigma_{wf}^2 = \frac{1}{2}\sigma_A^2 + \sigma_e^2$ Within full-sib family variance,
$\sigma_{fmf}^2 = \frac{1}{2}\sigma_A^2 + \sigma_{wf/n}^2$ Full-sib family means variance for seedling breeding population,
$\sigma_{fmf}^2 = \frac{1}{2}\sigma_A^2 + \left(\frac{\frac{1}{2}\sigma_A^2}{c} + \frac{\sigma_e^2}{c*r}\right)$ Full-sib family means variance for cloned breeding population, and
$\sigma_{\bar{c}(ff)}^2 = \left(\frac{c-1}{c}\right)\left[\frac{\frac{1}{2}\sigma_A^2}{c} + \frac{\sigma_e^2}{c*r}\right]$ variance of clonal mean within full-sib variance

Thus for a seedling breeding population of full-sib families and a clonal orchard formed from forwards selection in a full-sib breeding population

$$\Delta G = i_1 \tfrac{1}{2}\sigma^2_A / \sigma_{fmf} + i_2 \tfrac{1}{2}\sigma^2_A / \sigma_{wf} \qquad \textbf{(AI.3)}$$

between full-sib family + within full-sib family

For a *cloned breeding population* of full-sib families and a clonal orchard formed from forwards selection on clone means in a cloned full-sib breeding population;

$$\Delta G = i_1 \tfrac{1}{2}\sigma^2_A / \sigma_{fmf} + i_2\left(\frac{n-1}{n}\right)\tfrac{1}{2}\sigma^2_A / \sigma_{\bar{c}(ff)} \qquad \textbf{(AI.4)}$$

between cloned full-sib + within cloned full-sib family

Breeding-population gains were calculated as a percentage increase over the previous generation's mean. No adjustments were made for truncated additive variance. Also the impact of reduced effective population size was not simulated. It was assumed that additive genetic variance was the only source of genetic variance and that dominance and epistatic variance were zero. Different heritability levels were obtained by adjusting the environmental variance while keeping the additive variance constant.

Optimal number of ramets per clone was determined by keeping the total number of plants constant at 100 while varying the number of ramets per clone from 2 to 10. Ageing effects from vegetative propagation, maternal and paternal effects and common environment effects ('C' effects), which could inflate among-clone differences or differences among ramets of a clone were assumed to be absent. Rooting characteristics are assumed not to be correlated with the trait under selection. Clone-site interaction were assumed to be zero. Selection intensity was obtained from the tables for finite population sizes of Becker (1984) using the appropriate selection ratios and were incorporated into a computer programme. Selection was made between families and within families, with selection ratios assigned independently and not through combined selection.

The broad-based main breeding population after one generation of breeding and selection, provided the founder material for the first cycle of breeding and selection in the elite (Fig.AI.1). It should be noted that the selected elite parents are also part of the second-generation main population, and will be included in crosses made in the main population. Second-generation main and elite populations were forwards-selected from the previous generation, with the best individual per family usually selected. However, where the number of full-sib families was greater than the number of parents of the elite and production population, selection between families ensured that a new population of parents of the same size as the previous one was selected. Seed-orchard clones from corresponding main and/or elite populations were forwards-selected from the previous generation main or elite breeding population as the best individuals of the best 10 families per subline.

Variances used in the simulations were taken from a *P. radiata* progeny test aged 7.5 years, for a trait of 0.2 heritability. Variances for traits of heritability 0.1 and 0.4 were derived from these by adjusting the environmental variance.

Five different breeding-population options, sixteen production-population and seven cloned production-population (clonal forestry) options have been revised in different gain expectation equations and they are listed at the end of the chapter in Table AI.3.

The various assumed values for population parameters are shown below.

Population parameters

No. of Parents main Breeding Population	150 per subline
No. of Parents elite Breeding Population	20 and 40 per subline
Orchard clones (production population parents)	10 per subline
Number of sublines	2
Number of crosses per parent Main	2
Elite	2 and 4
Number of clones per family	100, 50, 20, 10
Number of ramets per clone	1, 2, 5, 10
Total number of plants per family	100
Heritabilities	0.1, 0.2, 0.4
Additive variance	560
Environmental variance	5040, 2235, 840

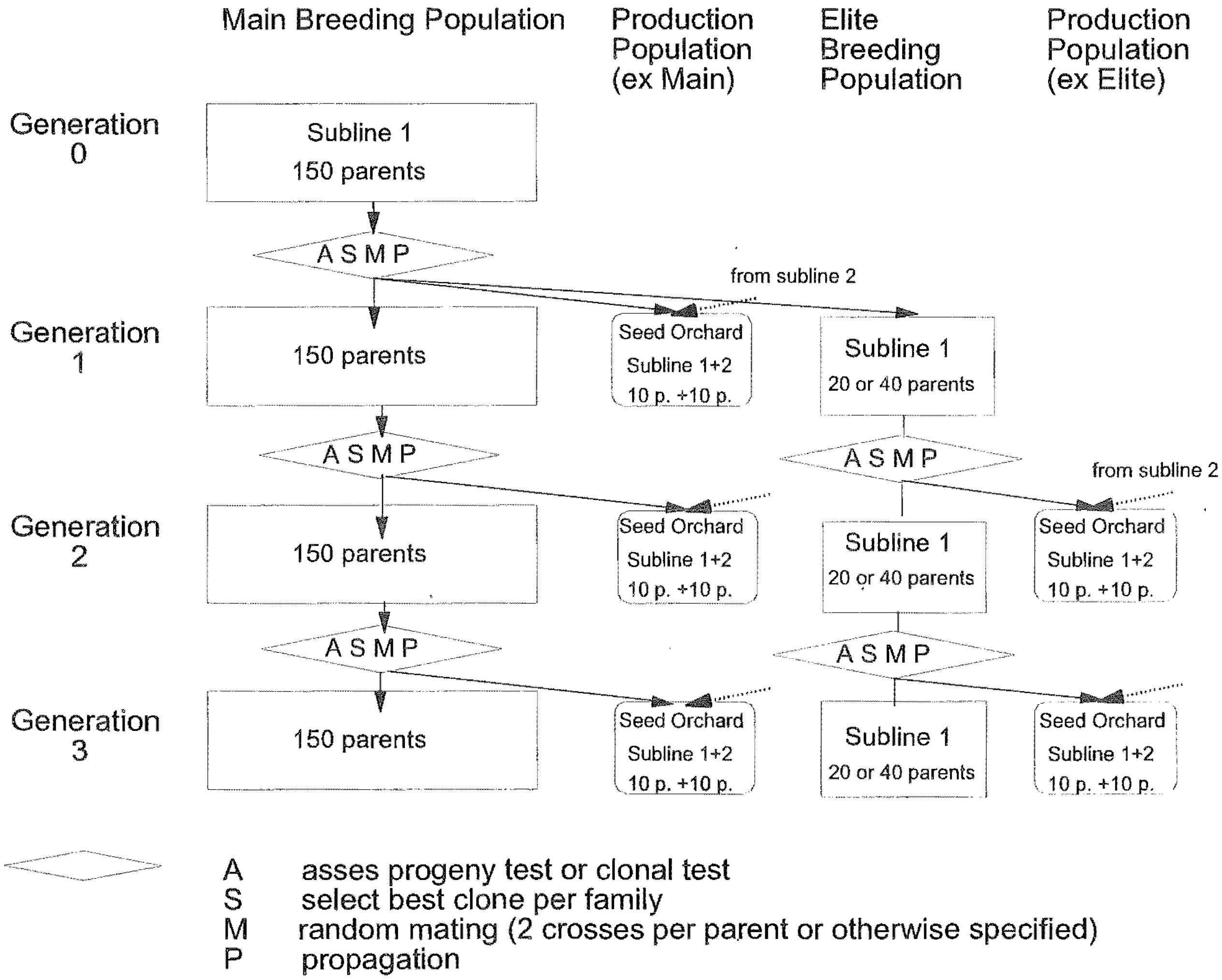
RESULTS

Predicted gains from main and elite breeding populations and corresponding orchards

Second-generation cumulative gains, for main and elite breeding populations (seedling and cloned) and corresponding production populations derived from these, at heritabilities 0.1, 0.2 and 0.4 are summarised in Table AI.1 and Fig. AI.2.

Selection to form the elite breeding populations resulted in almost doubling the gain shown by the main breeding population (Table AI.1). This was a direct result of increasing the proportion of families culled from 0 % for the main to 13 and 33 % for the 20-and 50-parent elites respectively. Thereafter and for a given heritability, recurrent gains for elite and main breeding population were similar where family selection ratios were the same. However, seed orchards originating from elite populations showed consistently larger gains over seed orchards originating from the main for all heritabilities (Table AI.1).

Figure AI.1. Simulated Breeding Strategy



Under an expanding-resource situation and doubling the number of crosses in the elite (from 20 to 40 for 20-parent elite and 50 to 100 for 50-parent elite) increased the cumulative gain by nearly 4 % independent of the size of the elite (and the number of crosses involved). When selection for the orchards was made, approximately 8 % extra cumulative gain was obtained over a seed orchard derived from the main. The 20-parent elite with 20 crosses (20/20) gave almost the same seed orchard gains as the much more costly 50-parent elite with 50 crosses; the orchard from the 20-parent elite with 40 crosses showed about 4% more gain and this was about the same as for the orchard from the 50-parent elite with 100 crosses.

The potential advantage of the elite breeding strategy is based on getting extra gains from an intensively-managed small elite population. Establishing the best selections from smaller elite breeding populations in cloned orchards resulted in increasing the gain over the orchard originating from the large main by around 4 or 5 %. However, the extra costs of running an elite need to be carefully examined. Comparing the seedling-based main population orchard with one from a cloned elite, this advantage was about doubled. Unless more resources (in the form of more crosses) are devoted to the elite, gains from this strategy can only be marginally worthwhile. However, with a cloned elite of 20 parents and 40 crosses the resulting 10-clone orchards gave 12 % more gain than an orchard derived from the seedling main population, a realistic comparison of the two strategies. The extra cost of the extra 40-family cloned elite over an 150-family main population would be about one-third of the basic cost.

Trade-offs between number of clones per family and ramets per clone

Trade-offs in predicted gains between number of clones and ramets per clone in a fixed-resource situation changed at different heritability levels. At low heritability and with a fixed resource, gains for cloning are relatively greater and suggested the use of more ramets and therefore fewer clones; but with higher heritability the opposite occurred (Table AI.2 and Fig. AI.3). Cumulative gain differences for main and elite breeding populations were due to the use of differing selection intensities. Set-up gain for the smaller elite (20) was highest while the only gain for the main was from within-family selection.

Seedling versus cloned population gains

With a cloned breeding population (20 clones, 5 ramets per clone) an extra 2 % recurrent gain per generation was obtained for heritability 0.1 and 0.2, and 1.5 % extra recurrent gain per generation for heritability 0.4 (Table AI.1).

GENERAL DISCUSSION

Several small elites might be an effective way to select for special-purpose traits as well as provide extra gain for a general-purpose breed. King and Johnson (1993) remarked that an elite strategy may be best suited in the production populations, where gain should be maximised and there is less concern about diversity. Increasing the number of crosses in the small elites appears to be a cost-effective strategy for increasing gains in the short term, although effective population sizes (not simulated here) would decline rapidly.

This simulation exercise with a closed elite has not considered the consequences of increasing inbreeding and reduction of diversity associated with small populations. Inbreeding depression and reduction in gain with decreasing population sizes for the breeding population may cause problems even though derived orchards are outcrossed.

Cloning has in most circumstances improved gains from within-family selection by effectively increasing heritability; owing to the constraints imposed in this study, only a small proportion of the increased gain derived from enhanced among-family selection. Results for estimating an optimal number of ramets per clone agreed with those found by Shaw and Hood (1985) and Russell and Libby (1986) who found that numerous clones should be tested with relatively few (i.e. 2-6) ramets per clone (Table AI.2 and Fig. AI.2). Using the average of 5 ramets per clone to rank the seedlings from a given family increased the precision of the ranking and therefore the gain from selection sufficiently to offset reduction in selection intensity.

IMPLICATIONS OF ASSUMPTIONS AND RESULTS

Certain results achieved using the assumptions of this simulation require a further analysis.

- The use of deterministic simulations for predicting gains over many generations can be questioned. Deterministic models are useful for the rapid evaluation of selection responses expected from alternative strategies. However, complicated multi-generational and multi-trait scenarios often cannot be simulated either easily or satisfactorily with deterministic simulations. In some situations, contrary to the findings of Keller et al. (1991), the cumulative effects of stochastic variations about underlying parameter values may lead to substantial directional deviations from deterministic expectations. Stochastic simulations can give sampling distributions about specific outcomes.

- There was no truncation of variance due to genetic fixation after recombination. The model assumed that σ_A^2 did not change from generation to generation. Shelbourne (1992) reported that gains can be overestimated in one generation by almost 3% for a trait of heritability 0.4 .

- The model used only a single trait, while breeding is often been for multiple traits. If genetic correlations amongst traits are strong, this simulation will require severe adjustments and the use of a multiple-trait index. If between-trait genetic correlations are favourable the compound trait will have a good heritability, if the genetic correlations are adverse the compound (index) trait will have low heritability.

- This simulation does not take account of 'C-effects', and results may be biased if these are important. 'C-effects' can contribute to variance within clones and variance among clones; in the later case, estimates of genetic gain might be inflated.

- Cotterill (1986) remarked on the benefits of keeping a breeding strategy as simple as possible, but this simulation has demonstrated the utility of adding elite breeding populations to the hierarchy of populations currently employed in tree breeding. However, if elite breeding populations were created for each breeding goal the added complexity and costs associated with this strategy should be evaluated as well as the potential benefits. Some elite breeding populations may be turned through generations faster than others as Mahalovich (1990) assumed, creating a very complex strategy of overlapping generations and different levels of gains that could generate some difficulties in the analysis and comparisons of gains.

- It was assumed that inbreeding did not affect the within-family variation, and the additive variance did not change from generation to generation. If elite breeding populations are kept relatively small (20), minimising the impacts of inbreeding depression within the elites will be an important requirement. If inbreeding coefficients within elites became high, full-sib testing, clonal-testing and parental ranking will become less precise for estimation of breeding values and small elite will be not viable as long-term breeding populations.

- A closed breeding population was simulated. If the families within the elites become related there will be less variation among them. Since breeding populations are structured in hierarchies, and to counteract this problem, Mahalovich (1990) suggested an enrichment strategy where the elite is enriched every generation with genotypes from the main to periodically ameliorate inbreeding levels and thus maintain gain; at the same time, the main breeding population can be, in principle, enriched with selections from the gene resource population or even other breeding programmes. None of these options was included in the present simulations.

- Only forwards selection of offspring was simulated and it is likely that a breeding population will also contain an element of backwards selection as well. While most of seed orchards are backwards-selected, the orchards originating from elite and main in this paper are forwards-selected. Backwards-selected orchards offered a more reliable prediction of the performance of the progeny than forwards-selected orchards. However, if cloning is considered this alternative would increase the predictability of the performance of forwards orchards. Also with advanced generations forwards-derived orchards will become more reliable.

REFERENCES

Allard R.W. 1960: Principles of Plant Breeding. John Wiley & Sons Inc. New York.

Becker W.A. 1984: Manual of Quantitative Genetics. 4th ed. Academic Enterprises. Pullman, Washington.

Bishir J., Roberds J. 1996: Limit theorems and a general framework for risk analysis in clonal forestry. (in prep.).

Borralho N.M.G. 1992: Gains expected from clonal forestry under various selection and propagation strategies. Pp. 328-338. *in* Proc. IUFRO Meeting on Mass Production Technology for Genetically Improved Fast Growing Forest Tree Species. Bordeaux 14-18 September, AFOCEL, France.

Burdon R.D. 1982: Selection indices using information from multiple sources fro the single-trait case. Silvae Genetica **31**: 81-85.

Burdon R.D., Namkoong G. 1983: Multiple populations and sublines. Silvae Genetica **32**: 221-222.

Burdon R.D. 1988: Recruitment for breeding populations: objectives, genetics and implementation. Pp. 554-572 *in* Proc. 2nd International Quantitative Genetics Conference, , Raleigh NC, USA, 1-5 June 1987, Sinauer , Sunderland, MS, USA (Eds. B.S. Weir *et al.*)

Carson M., Vincent T.G., Firth A. 1992: Control-pollinated and meadow seed orchards of radiata pine. Pp 100-109 *in* Proc. Symposium on Mass Production Technology for Genetically Improved Growing Fast Growing Tree Species, Bordeaux, France.

Cotterill P.P. 1984: A plan for breeding radiata pine. Silvae Genetica **33**: 84-90.

Cotterill P.P. 1986: Genetic gains from alternative breeding strategies including simple low cost options. Silvae Genetica **35**: 212-223.

Cotterill P.P., Jackson N. 1989: Gains expected from clonal orchards under alternative breeding strategies. Forest Science **35**: 183-196.

Cotterill P.P. 1989: The nucleus breeding system. *in* Proc. 20th South Forest Tree Improvement Conference. June 26-30, Charleston, SC.

Dieters M.J., White T.L., Hodge G.R. 1995: Genetic parameter estimates for volume from full-sib tests of slash pine (*Pinus elliottii*). Can. J. For. Res. **25**: 1397-1408.

Falconer D.S. 1981: Introduction to Quantitative Genetics. Longman, Inc. 340 pp.

Gea L.D., Lindgren D., Mullin T.J., Shelbourne C.J.A. 1997: The effect of population size, mating design and selection strategy on status number, coefficient of inbreeding and genetic gain. (in ms.).

King J.N., Johnson G.R. 1993: Monte carlo simulation models of breeding population advancement. Silvae Genetica **42**: 68-78

Keller D.S., Quinton M. and Smith C. 1991: A comparison of deterministic versus stochastic modelling of closed nucleus breeding schemes for beef cattle. Pp 315: 319 *in* Proc. of the 4th World Congress Genet. Appl. Livest. Prod.

Libby W.J. 1969: Some possibilities of the clone in forest genetics research. Genetics Lectures Vol. 1. Ed. Bogart R. Oregon State University Press, Corvallis.

Libby W.J. 1982: What is a safe number of clones per plantation? in "Resistance to Disease and Pests in Forest Trees". Pp 342-360 *in* Proc. 3rd International Workshop on the Genetics of Host-Parasite Interactions in Forestry, Wageningen ,The Netherlands. (Eds. H.M. .Heybroek, B.R. Stephan, K. von Weissenberg).

Libby W.J. 1992: Use of genetic variation for breeding forest trees. in Plant Breeding in the 1990's. Pp.101-118 *in* Proc. of the Symposium on Plant Breeding in the 1990's. North Carolina State University (Ed. H.T. Stalker, J.P. Murphy).

Lindgren D. 1986: How should breeders respond to breeding values? Pp. 361-372 *in* Proc. A Joint Meeting of IUFRO Working Parties on Breeding Theory and Progeny Testing and Seed Orchards. Williamsburg, VA (Ed. R.J. Weir).

Lindgren D., Werner M. 1989: Gain efficiency of different Norway Spruce seed orchard designs. (including an appendix by Ö Danell). Pp 189-206 *in* Proc. Norway Spruce: provenances, breeding and genetic conservation. Institutet för Skogsforbättring (Eds. L.G. Stener and M. Werner).

Lindgren D. 1991: Optimal utilisation of genetic resources. Forest Tree Improvement **23**: 49-67.

Loo-Dinkins J. 1993: Simulated advanced generation breeding populations: Options, for coastal Douglas-Fir. Quantgen Forest Consulting. Canada.

Manly B.F.J. 1991: Randomization and Monte Carlo Methods in Biology. Chapman and Hall. 281 pp.

Mahalovich M.R., Bridgwater F.E. 1989: Modelling elite populations and positive assortative mating in recurrent selection programs for general combining ability. Pp. 43-49 *in* Proc. 20th Southern Forest Tree Improvement Conference, Charleston, South Carolina, National Technical Information Service. Springfield, Virginia.

Mahalovich M.R. 1990: Modelling positive assortative mating and elite populations in recurrent selection programs for general combining ability. PhD Thesis. North Carolina State University. 129 pp.

Matheson A.C., Lindgren D. 1985: Gains from clonal and the clonal seed orchard options compared for tree breeding programs. Theor. Appl. Genet. **79**: 841-848

McKeand S., Bower R., Bridgwater F.E., Hendrickson J., McCutchan B., Mueller W., Purnell R., Weir R., Zoerb M. 1992: A strategy for the third breeding cycle of loblolly pine in the North Carolina State University Industry Cooperative Tree Improvement Program.

Mullin T.J., Park Y.S. 1991: Estimating genetic gains from alternative breeding strategies for clonal forestry. Can. J. For. Res. Vol. **22**: 14-23

Mullin T.J. 1992: Making choices -Seedling and clonal options for breeding and reforestation in New Brunswick. Canada New Brunswick Cooperation Agreement on Forest Development under Program 5.1 Applied Research and Development.

Namkoong G., Snyder E.B., Stonecypher R.W. 1966: Heritability and gain concepts for evaluating breeding systems such us seedling orchards. Silvae Genetica **15**: 76-84.

New Zealand Forestry Statistics. 1993: Ministry of Forestry, Wellington, New Zealand.

Russell J.H., Libby W.J. 1986: Clonal testing efficiency: the trade-off between clones tested and ramets per clone. Can. J. For. Res. **16**: 925-930

Russell J.H., Loo-Dinkins J.A. 1992: Distribution of testing effort in cloned genetic tests. Silvae Genetica **42**: 98-104.

Shaw D.V., Hood J.V. 1985: Maximising gain per effort by using clonal replicates in genetic tests. Theor. Appl. Genet. **71**: 392-399.

Shelbourne C.J.A. 1969: Tree breeding methods. New Zealand Forest Service. Forest Research Institute, Technical Paper No. 55.

Shelbourne C.J.A., Burdon R.D., Carson S.D., Firth A., Vincent T.G. 1986: Development plan for radiata pine breeding. New Zealand Forest Service. Forest Research Institute, Rotorua, New Zealand. pp.142.

Shelbourne C.J.A. 1992: Genetic gains from different kinds of breeding population and seed or plant production population. S. A. For. J. **160**: 49-65

van Buijtenen J.P., Burdon R.D. 1990. Expected efficiencies of mating designs for advanced-generation selection. Can. J. For. Res. **20**: 1648-1663

White T., Hodge G.R., Powell G.L. 1993: An advanced-generation tree improvement plan for slash pine in the southeastern United States. Silvae Genetica **42**: 359-371.

White T., Matheson A.C., Boomsma D., Rout A.F. 1992: Logistics, costs and genetic gains of five options of nucleus breeding strategies. Technical Report TR92-01 STBA Mount Gambier, South Australia.

Table AI.1 Second-generation cumulative gains for main and elite breeding populations (seedlings and cloned) and corresponding production populations (10 clones per subline) derived from these, at h² 0.1, 0.2, 0.4

h^2=0.1

100 seedlings

	SUBLINE SIZE				
	Main	**Elite 20**		**Elite 50**	
Number of crosses	150	20	40	50	100
Breeding Population	9.00	20.30	26.60	16.70	22.80
Recurrent Gain	4.30	4.30	9.70	4.30	9.80
Production Population	22.60	26.60	30.30	27.00	30.60
Extra Gain		4.00	7.30	4.40	8.00

20 Clones / 5 ramets per clone

	SUBLINE SIZE				
	Main	**Elite 20**		**Elite 50**	
Number of crosses	150	20	40	50	100
Breeding Population	13.20	24.80	31.00	21.00	27.20
Recurrent Gain	6.40	6.40	11.70	6.40	11.80
Production Population	27.00	31.00	34.70	31.60	34.40
Extra Gain		4.00	7.70	4.60	7.40

h^2=0.2

100 Seedlings

	SUBLINE SIZE				
	Main	**Elite 20**		**Elite 50**	
Number of crosses	150	20	40	50	100
Breeding Population	13.10	25.20	31.90	21.20	22.90
Recurrent Gain	6.30	6.30	11.90	6.30	12.00
Production Population	27.90	32.10	36.10	32.40	35.60
Extra Gain		4.20	8.20	4.50	7.70

20 Clones / 5 ramets per clone

	SUBLINE SIZE				
	Main 150	**Elite 20**		**Elite 50**	
Number of crosses	150	20	40	50	100
Breeding Population	17.60	31.90	36.70	26.10	32.60
Recurrent Gain	8.50	8.50	14.00	8.50	14.10
Production Population	32.50	36.10	40.50	37.40	40.30
Extra Gain		3.60	8.00	4.90	7.80

h^2=0.4

100 Seedlings

	SUBLINE SIZE				
	Main 150	**Elite 20**		**Elite 50**	
Number of crosses	150	20	40	50	100
Breeding Population	19.80	32.90	39.60	28.80	35.60
Recurrent Gain	9.50	9.50	15.20	9.50	15.30
Production Population	35.10	40.00	43.70	40.50	43.70
Extra Gain		4.90	8.60	5.40	8.80

20 Clones / 5 ramets per clone

	SUBLINE SIZE				
	Main 150	**Elite 20**		**Elite 50**	
Number of crosses	150	20	40	50	100
Breeding Population	22.80	35.70	42.70	31.60	38.40
Recurrent Gain	10.90	10.90	16.50	10.90	16.60
Production Population	38.30	42.50	46.70	43.40	45.30
Extra Gain		4.20	8.40	5.10	7.00

Figure AI.2. Cumulative gains for main and elite breeding populations (20 parents) and corresponding production populations (10 clones per subline). Elite breeding populations with 20 and 40 crosses. $h^2=0.1$

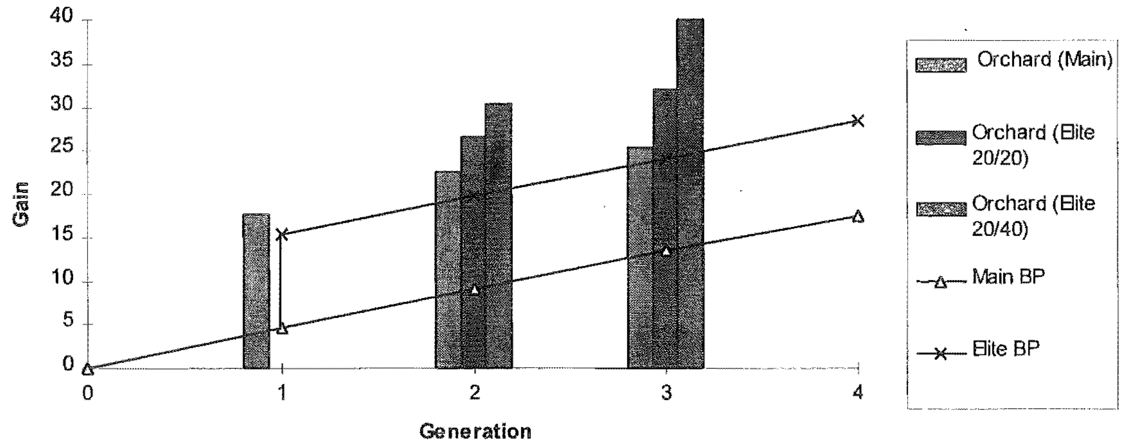


Table AI.2 Predicted gains from control-pollinated families with different number of ramets per clone

Plants per family	Heritabilities	ramets per clone			
		1	2	5	10
50	0.1	3.85	4.58	5.13	4.68
	0.2	5.59	6.49	6.83	5.84
	0.4	8.39	9.18	8.7	6.88
100	0.1	4.31	5.31	6.39	6.57
	0.2	6.27	7.51	8.51	8.2
	0.4	9.4	10.61	10.84	9.65
200	0.1	4.74	5.95	7.5	8.19
	0.2	6.88	8.42	9.98	10.23
	0.4	10.32	11.89	12.71	12.03
500	0.1	5.24	6.7	8.77	10.03
	0.2	7.69	9.49	11.67	12.51
	0.4	11.43	13.41	14.86	14.71
1000	0.1	5.6	7.23	9.63	11.24
	0.2	8.15	10.23	12.81	14.02
	0.4	12.2	14.46	16.32	16.49

Figure AI.3 Gain (%) from clone-within-family selection in cloned control-pollinated breeding population. Constant number of plants per family (100).

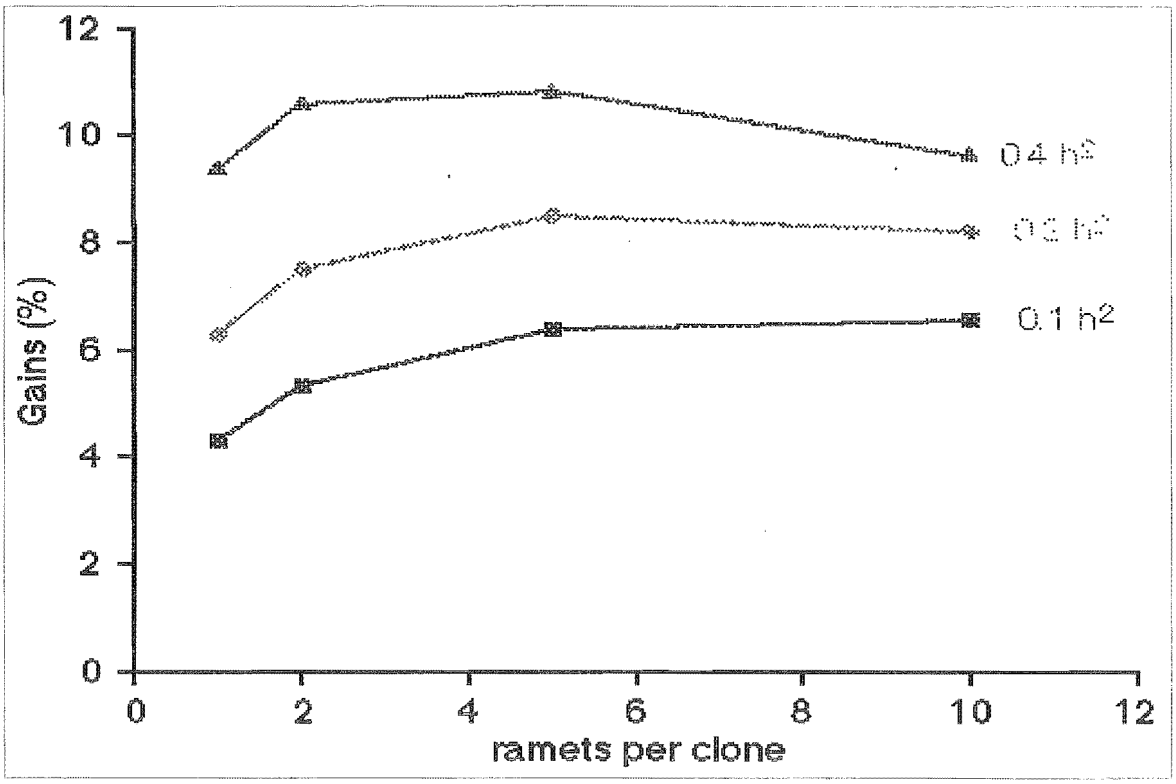


Table AI.3 Breeding population options

1 Phenotypic selection

$\Delta G_F = \frac{1}{2} i_1 \sigma_A^2 / \sigma_p$

$\Delta G_M = \frac{1}{2} i_2 \sigma_A^2 / \sigma_p$

$\Delta G_{Total} = \Delta G_F + \Delta G_M$ (and for other options)

2 Breeding population of half-sib (OP) families

$\Delta G_F = \frac{1}{2} i_1 \frac{1}{4} \sigma_A^2 / \sigma_{fm} + \frac{1}{2} i_2 \left(\frac{n-1}{n}\right) \frac{3}{4} \sigma_A^2 / \sigma_w$

$\Delta G_M = \frac{1}{2} i_3 \frac{1}{4} \sigma_A^2 / \sigma_{fm} + \frac{1}{2} i_4 \left(\frac{n-1}{n}\right) \frac{3}{4} \sigma_A^2 / \sigma_w$

3 Cloned breeding population of half-sib (OP) families

$\Delta G = i_1 \frac{1}{4} \sigma_A^2 / \sigma_{fm} + i_2 \left(\frac{n-1}{n}\right) \frac{3}{4} \sigma_A^2 / \sigma_{\bar{c}(f)}$

4 Breeding population of full-sib (CP) families

$\Delta G = i_1 \frac{1}{2} \sigma_A^2 / \sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) \frac{1}{2} \sigma_A^2 / \sigma_{wf}$

5 Cloned breeding population of full-sib (CP) families

$\Delta G = i_1 \frac{1}{2} \sigma_A^2 / \sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) \frac{1}{2} \sigma_A^2 / \sigma_{\bar{c}(ff)}$

Table AI.4. Plant production population options

1 Breeding population of half-sib families, thinned for seed production

$$\Delta G_F = ½\, i_1\, ¼\, \sigma_A^2/\sigma_{fm} + ½\, i_2 \left(\frac{n-1}{n}\right) ¾\, \sigma_A^2/\sigma_w$$

$$\Delta G_M = ½\, i_3\, ¼\, \sigma_A^2/\sigma_{fm} + ½\, i_4 \left(\frac{n-1}{n}\right) ¾\, \sigma_A^2/\sigma_w$$

$$\Delta G_T = \Delta G_F + \Delta G_M$$

2 Breeding population of full-sib families, thinned for seed production

$$\Delta G_F = i_1 ½\, \sigma_A^2/\sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{wf}$$

3 Cloned breeding population of half-sib (OP) families, thinned for seed production on clone means

$$\Delta G = i\, \sigma_A^2/\sigma_{\bar{c}}$$

$$\text{or } \Delta G = i_1\, ¼\, \sigma_A^2/\sigma_{fm} + i_2 \left(\frac{n-1}{n}\right) ¾\, \sigma_A^2/\sigma_{\bar{c}(ff)}$$

4 Clonal breeding population of full-sib (CP) families, thinned for seed production on clone means

$$\Delta G = i\, \sigma_A^2/\sigma_{\bar{c}}$$

$$\text{or } \Delta G = i_1\, ½\, \sigma_A^2/\sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{\bar{c}(ff)}$$

5 Clonal orchard from forwards selection in half-sib breeding population

$$\Delta G = i_1\, ¼\, \sigma_A^2/\sigma_{fm} + i_2 \left(\frac{n-1}{n}\right) ¾\, \sigma_A^2/\sigma_w$$

6 Clonal orchard from forwards selection in full-sib breeding population

$$\Delta G = i_1\, ½\, \sigma_A^2/\sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{wf}$$

7 Clonal orchard from forwards selection on clone means in cloned half-sib breeding population

$$\Delta G = i_1\, ¼\, \sigma_A^2/\sigma_{fm} + i_2 \left(\frac{n-1}{n}\right) ¾\, \sigma_A^2/\sigma_{\bar{c}(f)}$$

8 Clonal orchard from forwards selection on clone means in cloned full-sib breeding population

$$\Delta G = i_1\, ½\, \sigma_A^2/\sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{\bar{c}(ff)}$$

9 Clonal orchard from forwards selection and roguing in half-sib breeding population

$$\Delta G = i_1\, ¼\, \sigma_A^2/\sigma_{fm} + i_2 \left(\frac{n-1}{n}\right) ¾\, \sigma_A^2/\sigma_w + 2i_3 \left(\frac{n-1}{n}\right) ¼\, \sigma_A^2/\sigma_{fm}$$

10 Clonal orchard from forwards selection and roguing in full-sib breeding population

$$\Delta G = i_1\, ½\, \sigma_A^2/\sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{wf} + 2i_3 \left(\frac{n-1}{n}\right) ¼\, \sigma_A^2/\sigma_{fm}$$

11 Clonal orchard from forwards selection and roguing in cloned half-sib breeding population

$$\Delta G = i_1\, ¼\, \sigma_A^2/\sigma_{fm} + i_2 \left(\frac{n-1}{n}\right) ¾\, \sigma_A^2/\sigma_{\bar{c}(f)} + 2i_3 \left(\frac{n-1}{n}\right) ¼\, \sigma_A^2/\sigma_{fm}$$

12 Clonal orchard from forwards selection and roguing in cloned full-sib breeding population

$$\Delta G = i_1\, ½\, \sigma_A^2/\sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{\bar{c}(ff)} + 2i_3 \left(\frac{n-1}{n}\right) ¼\, \sigma_A^2/\sigma_{fm}$$

13 Clonal orchard from backwards selection in half-sib breeding population

$$\Delta G = i_1\, ¼\, \sigma_A^2/\sigma_{fm} + i_2 \left(\frac{n-1}{n}\right) ¾\, \sigma_A^2/\sigma_w + 2i_3 \left(\frac{n-1}{n}\right) ¼\, \sigma_A^2/\sigma_{fm}$$

14 Clonal orchard from backwards selection in full-sib breeding population

$$\Delta G = i_1\, ½\, \sigma_A^2/\sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{wf} + 2i_3 \left(\frac{n-1}{n}\right) ¼\, \sigma_A^2/\sigma_{fm}$$

15 Seedling orchard from forwards selection in half-sib breeding pop.and within-family thinning of the orchard

$$\Delta G_F = ½\, i_1\, ¼\sigma_A^2/\sigma_{fm} + ½\, i_2 \left(\frac{n-1}{n}\right) ¾\sigma_A^2/\sigma_w$$

$$\Delta G_M = ½\, i_3\, ¼\sigma_A^2/\sigma_{fm} + ½\, i_4 \left(\frac{n-1}{n}\right) ¾\sigma_A^2/\sigma_w$$

$$\Delta G_{(Thinning)} = i_5 \left(\frac{n-1}{n}\right) ¾\sigma_A^2/\sigma_w$$

16 Seedling orchard from forwards selection in full-sib breeding pop.and within-family thinning of the orchard

$$\Delta G = i_1\, ½\, \sigma_A^2/\sigma_{fmf} + i_2 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{wf}$$

$$\Delta G_{(Thinning)} = i_5 \left(\frac{n-1}{n}\right) ½\, \sigma_A^2/\sigma_{wf}$$

Table AI.5. Clonal forestry options

1 Clonal selection in cloned half-sib and full-sib breeding populations

$$\Delta G = i\,\sigma_A^2/\sigma_{\bar{c}}$$

2 Clonal selection in cloned half-sib breeding population with selection between and within families

$$\Delta G = i_1\,\tfrac{1}{4}\,\sigma_A^2/\sigma_{fm} + i_2\left(\tfrac{n-1}{n}\right)\tfrac{3}{4}\,\sigma_A^2/\sigma_{\bar{c}(f)}$$

3 Clonal selection in cloned full-sib breeding population in selection between and within families

$$\Delta G = i_1\,\tfrac{1}{2}\,\sigma_A^2/\sigma_{fmf} + i_2\left(\tfrac{n-1}{n}\right)\tfrac{1}{2}\,\sigma_A^2/\sigma_{\bar{c}(ff)}$$

4 Clonal selection in cloned production population (from cloned half-sib breeding population).

At the same time that new breeding population parents are selected and mated within sublines, an intensive selection of the best trees in each subline is made and these are crossed between sublines, to form a Production Population. A few seedlings from each family are cloned and clonal tests planted. The best clones are re-selected and then re-propagated for planting commercially

$$\Delta G_1 = i_1\,\tfrac{1}{4}\sigma_A^2/\sigma_{fm} + i_2\left(\tfrac{n-1}{n}\right)\tfrac{3}{4}\,\sigma_A^2/\sigma_{\bar{c}(f)}$$

$$\Delta G_2 = i_3\left(\tfrac{n-1}{n}\right)\sigma_A^2/\sigma_{\bar{c}}$$

5 Clonal selection in cloned production population (from cloned full-sib breeding population)

$$\Delta G_1 = i_1\,\tfrac{1}{2}\sigma_A^2/\sigma_{fmf} + i_2\left(\tfrac{n-1}{n}\right)\tfrac{1}{2}\,\sigma_A^2/\sigma_{\bar{c}(ff)}$$

$$\Delta G_2 = i_3\left(\tfrac{n-1}{n}\right)\sigma_A^2/\sigma_{\bar{c}}$$

6 Clonal selection in cloned production population (from seedling half-sib breeding population)

$$\Delta G_1 = i_1\,\tfrac{1}{4}\sigma_A^2/\sigma_{fm} + i_2\left(\tfrac{n-1}{n}\right)\tfrac{3}{4}\,\sigma_A^2/\sigma_w$$

$$\Delta G_2 = i_3\left(\tfrac{n-1}{n}\right)\sigma_A^2/\sigma_{\bar{c}}$$

7 Clonal selection in a cloned production population (from seedling full-sib breeding population

$$\Delta G_1 = i_1\,\tfrac{1}{2}\sigma_A^2/\sigma_{fmf} + i_2\left(\tfrac{n-1}{n}\right)\tfrac{1}{2}\,\sigma_A^2/\sigma_{wf}$$

$$\Delta G_2 = i_3\left(\tfrac{n-1}{n}\right)\sigma_A^2/\sigma_{\bar{c}}$$

APPENDIX II.

PROC INBREED : THE SHADOW OF FORGOTTEN ANCESTORS

L.D.Gea[1], C. Low[2] and D. Lindgren[3]

This paper was presented at the SUNZ'95 SAS Users of New Zealand. 13th Annual Conference. Wellington 18-19 September 1995

SUMMARY

Genetic diversity is a key issue. To monitor and predict genetic diversity, measures are urgently needed. Lindgren *et al* (1995) suggested that half the inverse of the average coancestry in the breeding population, called status effective number (status number), is a suitable measure for describing the status of a breeding population.

A little-known SAS procedure (PROC INBREED) can calculate the covariance or inbreeding coefficients for a pedigree, which permits the estimation of status number for overlapping generations and for a large number of individuals.

Key words: *genetic diversity, inbreeding coefficient, status number, coancestry.*

---

[1] and [2] NZ FRI Genetics and Tree Improvement, Private Bag 3020, Rotorua, New Zealand
[3] Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, S 901 83 UMEÅ, SWEDEN.

INTRODUCTION

Small populations undergo changes in allele frequencies with generation turnover. This phenomenon, if independent of the effects of selection, is called genetic drift. Thereby alleles may be lost, while inbreeding and coancestry increases in small populations over succesive breeding generations. Genes from some founders become abundant while genes from other founders became rare or lost. While mutation may counteract these effects, too little of its role is known for it to be a dependable factor in managing genetic diversity.

The strong artificial selection practised in forest tree breeding, which is amplified by methods to use information from relatives, makes it important to evaluate the consequences for genetic diversity. Management of genetic resources requires actions to monitor and conserve genetic diversity. To do this, an operational quantitative measure is needed, and surprisingly enough we could not find such a measure in the literature.

Problems of inbreeding in tree species

Most tree species produce abundant flowers and pollen on the same tree, so that they are capable of producing self-fertilised seed (selfing) in the absence of unrelated pollen. They have several mechanisms which favour pollination by unrelated pollen (outcrossing), but in a natural situation various levels of inbreeding occur.

It is also possible for many tree species to reproduce other than by seed (i.e. vegetatively), so that there may be many individual trees that have identical genes, referred to as members of the same clone. This does not often occur naturally, except in the case of species that sprout new trees from surface roots, but is increasingly common in many tree species that are propagated in tree nurseries.

There have been many studies of the growth and performance of selfed trees compared to outcrosses, Wilcox (1982) showed growth losses of up to 50% and great susceptibility to disease in selfed radiata pine trees. A study on the growth and performance of brother/sister crossing showed growth losses of up to 10% compared with outcrossing. Release from the effects of inbreeding can be a considerable part of the gains made in the first part of a tree breeding programme.

These days, study of DNA molecules can effectively identify individual genes in chromosome chains, and can show how many different gene combinations (genetic diversity) there are in a given tree species. In general the greater the genetic diversity for a species, the more successful that species will be in adapting to new environments or to the threat of new diseases.

The concept of status number (Status effective number)

Status number is proposed as a measure of genetic diversity and is expressed mathematically as:

$$N_s = 0.5/f \qquad \textbf{(AII.1)}$$

where N_s is the status number;
f is the average coancestry of the population

Average coancestry for an individual member of a population is the average of the coancestry between that individual and all individuals in the population including the individual itself. Average coancestry of a population is the average of the coancestry for all individuals in the population.

Alternative Concepts

Inbreeding

Inbreeding is caused by coancestry among parents, and it can be calculated as the covariance of the individual minus 1. Because coancestry is an earlier event and a cause, it seems more basic than inbreeding. A breeding population can usually be managed so it comprises pairs of unrelated individuals, so that by mating these inbreeding can always be forced back to zero. A cross between two homozygous lines has an inbreeding coefficient of zero, although it does not carry more genetic diversity than a single individual. Thus, the coefficient of inbreeding says very little about the long-term genetic potential of a population or about its genetic diversity.

The Traditional Concept of Effective Population Size

The traditional "effective population size" seems to have been developed for approximate calculations of inbreeding or variance in complex situations, but not as a diversity measure. **The traditional 'effective population size' (N_e) is defined as an implied measure of rate but not of current state, and is thus basically different from status number.**

$$Ne = \frac{1}{2\Delta F} \qquad \textbf{(AII.2)}$$

where :
N_e is the effective population size, and
ΔF is the rate of inbreeding.

To maintain a rate is very different from conserving a state, and it may be questioned whether the traditional measure of effective population size has a place in discussions about conservation of genetic diversity. The traditional effective population size is undefined or infinite for the initial population, as a rate requires a difference between several states. If related individuals do not mate there is no inbreeding, and that corresponds to an infinite ideal population.

In the first cycle of a breeding programme (where many forest tree improvement programmes still operate) the traditional inbreeding effective population size is infinite as related individuals are not mated. If trees are grouped into a number of units (sublines or breeding groups), by merging sublines when inbreeding occurs the coefficient of inbreeding can decrease, and this can be interpreted as an effective population size beyond infinity. A better interpretation is of course that the effective population size suddenly becomes undefined.

The effective population size is dependent on for which generation the rate of change is considered. The effective population size is approximately constant when the number of generations considered is expanded, although it may both decrease and increase. Thus, effective population size cannot be considered as a measure which can be used for studying how genetic diversity declines with time, and it is rather meaningless to give its value for breeding populations of forest trees.

We argue than an effective population number concept requires that **its size gives a relevant intuitive message about the current state of the population**. This is important because the depletion of biodiversity can be also a very political question, and the effective number may have implications far beyond science.

Characteristics of Status Number

- It can be derived for any population that is linked by pedigree to known initial genotypes with known coancestry and inbreeding.
- It can be predicted as a function of the mating and selection tactics of previous generations.
- It can be the same as the number of individuals (census number) for a population with unrelated non-inbred genotypes.
- It can never be higher than the census number.

- It can never be lower than 0.5.
- It generally declines with generation shifts.
- It is only dependent on the current state of a population, not the history or scenarios about the future. It does not measure a rate of change but a state.
- The sum of the status number of sublines is that of the whole population when genotypes from unrelated sublines are pooled for reasonably symmetric situations.

SEARCHING FOR AN EFFICIENT ALGORITHM FOR CALCULATION OF STATUS NUMBER

Any calculation method should be as simple as possible and should be able to calculate the statistic in large populations.

The PROC INBREED Procedure

This procedure explained in the SAS Technical Report P-229 Changes and Enhancements (1992); calculates the covariance or inbreeding coefficients for a pedigree. PROC INBREED is unique in that it handles very large populations and allows the division of the population into non-overlapping generations.
Memory requirement, for PROC INBREED, depends on the processing mode (SAS Technical Report P-229 pg. 214).

- Non-overlapping generations the memory requirement in bytes is

$$s = 64n + 2f(f+1) \qquad \textbf{(AII.3)}$$

where n is the maximum number of individuals in a generation and f is the maximum number of families in a generation .

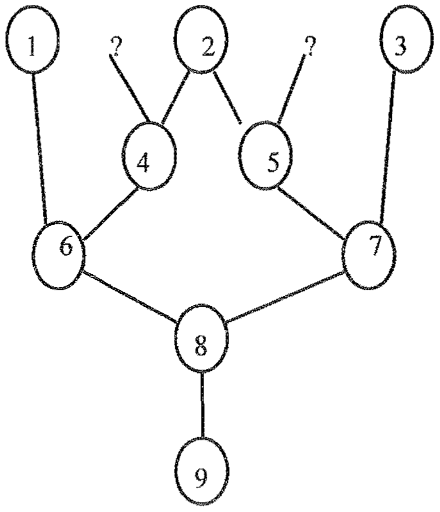
- If the population is not divided into non overlapping generations, the minimum memory requirement in bytes is

$$s = 32\text{n}$$

An example proposed by Tier (1990) has been adapted to show the use of PROC INBREED to calculate status number. The method is based on the calculation of the numerator relationship matrix from which coancestry and therefore status number can be calculated.

A rather complicated scheme is chosen to illustrate the complications and general nature of the algorithm. For calculation of the inverse relationship matrix the pedigrees have to be sorted as in Table AII. 1, so that the parents precede their progeny. The numerator relationship matrix is given in Table AII. 2.

Figure AII.1 and Table AII.1: Pedigree of sample population



Individual	Parent 1	Parent 2
1	unknown	unknown
2	unknown	unknown
3	unknown	unknown
4	2	unknown
5	2	unknown
6	4	1
7	5	3
8	6	7
9	8	8

The following SAS code was used to get the "status number" for the relationships shown in Table AII. 1 :-

```
data temp;          *this step makes the basic datafile  (temp) from the numbers in Table AII.1;
input clone mother father;
cards;
1   .   .
2   .   .
3   .   .
4   2   .
5   2   .
6   4   1
7   5   3
8   6   7
9   8   8
;

proc inbreed data=temp covar matrix outcov=mono;
var clone mother father;          *this step ouputs the covariance matrix to the datafile "mono";

proc summary nway data=mono;
var col1-col9;                    *this step outputs the means of each column as a row of numbers;
output out=new mean=;
```

```
data set2 (drop = _type_ _freq_) ;          *this step drops unwanted variables _type_ and _freq_ ;
set new;

proc transpose data=set2 out=transet2;
var col1-col9;          * this step turns the row of means into a column (or variable) called col1;

proc summary nway data=transet2;
var col1;                         * this step takes the mean of the previously created means ;
output out=acoan mean= average ;

data end;
set acoan;
avc=average/2;                    *this step computes the status number ;
status=1/(2*avc);

proc print data=end;              *this step prints the statistics calculated by the previous step' ;
var avc status;
title 'Average coancestry and Status number for Relatedness Example' ;

run;
```

The elements of the numerator relationship matrix are equivalent to twice the coancestry between individuals. For example the coancestry between individual 1 and 6 is half the value in the cell at row 1 and column 6 that is 0.25. Thus **average coancestry** is calculated by taking the cell average of the relationship matrix and dividing by 2. In the example above, the average coancestry (*f*) is 0.182. **The status number is half the inverse of this value, $N_s = ½f = 2.75$**

Tree breeding populations often consist of overlapping generations. The above method can be adapted to calculate status number in populations of overlapping generations. All the parents in the pedigree must be in the pedigree list. For example, consider the population in Table AII.2; if individuals 4 and 5 were not considered in the population they would still have to be included in the pedigree listing to allow the calculation of the correct coefficient of relationship. The average coancestry (*f*) is then calculated from a reduced matrix with rows 4 and 5 and columns 4 and 5 removed giving a status number of 2.45.

Table AII. 2: The numerator relationship matrix for the sample population in Table AII. 2. The elements of the corresponding coancestry matrix are half of the elements of this table.

Individual	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	cov(1,1) 1.000	cov(1,2) 0.000	cov(1,3) 0.000	cov(1,4) 0.000	cov(1,5) 0.000	cov(1,6) 0.500	cov(1,7) 0.000	cov(1,8) 0.250	cov(1,9) 0.250
2	0.000	1.000	0.000	0.500	0.500	0.250	0.250	0.250	0.250
3	0.000	0.000	1.000	0.000	0.000	0.000	0.500	0.250	0.250
4	0.000	0.500	0.000	1.000	0.250	0.500	0.125	0.312	0.312
5	0.000	0.500	0.000	0.250	1.000	0.125	0.500	0.312	0.312
6	0.500	0.250	0.000	0.500	0.125	1.000	0.062	0.531	0.531
7	0.000	0.250	0.500	0.125	0.500	0.062	1.000	0.531	0.531
8	0.250	0.250	0.250	0.312	0.312	0.531	0.531	1.031	1.031
9	0.250	0.250	0.250	0.312	0.312	0.531	0.531	1.031	1.516

CONCLUSIONS

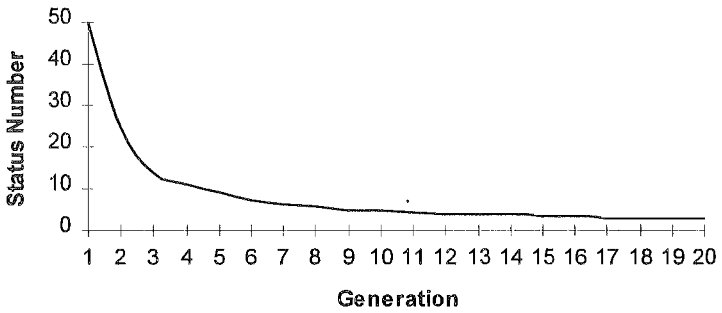
The graph below (cf. Lindgren *et al.* 1996) demonstrates that genetic diversity measured by status number can be rapidly lost when generations turn over.

1) We have a problem:
Avoiding the mating of relatives when managing breeding and production populations.

2) We know how to measure coancestry and genetic diversity.
Using the concept of status number, in forestry for example, a seed orchard crop, the forest holding of a company, the forest in a nation planted a certain year or clonal plantations yields, etc. can be measured .

3) We have an efficient algorithm:
Proc Inbreed, that allow us to calculate the covariance or inbreeding coefficient for a pedigree and handles very large populations. Thus, we can programme our crosses to avoid the shadow of forgotten ancestors.

Figure AII.2 Status number of a random population of size 50



REFERENCES

Burrows P.M. 1984a: Inbreeding under selection from unrelated families. Biometrics, **40**: 357-366.

Burrows P.M. 1984b: Inbreeding under selection from related families. Biometrics, **40**: 895-906.

Falconer D.S. 1989: Introduction to Quantitative Genetics. 3rd edition. Longman Scientific & Technical, England.

Golden B.L., Brinks J.S., Bourdon R.M. 1991: A performance programmed method for computing inbreeding coefficients from large data sets for use in mixed-model analysis. J. Anim. Sci. **69**: 3564-3573.

Johnson R., King J., Low C., Firth A., Vincent T. 1987: Age 7 Assessment of the 875 series diallel progeny test. New Zealand Radiata Pine Breeding Cooperative Report 8.

Lindgren D., Gea L.D., Jefferson P. 1995: Status Number a measure of genetic diversity. Proceedings Joint Meeting of the IUFRO Working Parties S2.02.05; 06; 012 and 14. Evolution of Breeding Strategies for Conifers from the Pacific North West. Limoges, France 28 July-4th August.

Robertson A. 1961: Inbreeding in artificial selection programs. Genetical Research, **2**: 189-194.

SAS Institute Inc.,1992: SAS Technical Report P-229 SAS/STAT Software *Changes and Enhancements*. Release 6:07, Cary, NC, SAS Institute Inc., 1992, 620 pp

Tier B. 1990: Computing inbreeding coefficients quickly. Genet. Sel. Evol. **22**: 419-430.

Wilcox M. D. 1982: Inbreeding depression and genetic variances estimated from self- and cross-pollinated families of *Pinus radiata*. IUFRO Meeting, Rotorua, New Zealand, session 4.11.

Wright S. 1931: Evolution in Mendelian Populations. Genetics **16**: 97-159.