Genetic Studies and Improvement of *Pinus caribaea* Morelet

Yongqi ZHENG

PhD

The University of Edinburgh

1996

To my Mother

PREFACE

This thesis actually consists of two parts of work done during two different periods. The work reported in Chapter 5 is basically the dissertation work for a MSc course I attended. I initially came to the University of Edinburgh for a MSc degree in Tree Improvement. After finishing the MSc dissertation work, I realised that there was much work to be done for the genetic improvement of *Pinus caribaea* in China, and this is suitable for me to pursue a PhD degree in Forest Genetics. There was almost no genetic studies had been done with the Cuban and Bahaman varieties particularly in the aspect of population genetic studies. Although breeding programmes are underway, neither information on genetic structure nor information on breeding system of the varieties is available. With support from Dr Richard Ennos, who supervised my MSc dissertation, Dr Douglas Malcolm, who is the co-ordinator of the UK-China forestry project, and Professor Huoran Wang who is the Chinese project co-ordinator, I finally convinced the Overseas Development Administration (ODA) to fund me for such studies. The MSc work has been strengthened and continued with population genetic studies of the two varieties. All these together leads to this PhD thesis.

The thesis consists of 7 Chapters. Chapter 1 presents a general description of *P. caribaea* and a comprehensive review of introduction, testing and improvement of the species around the world with particular reference to China. The purpose of the thesis is also outlined in this Chapter. The second Chapter reports a study of genetic structure and patterns of variation of two varieties of the species by use of isozyme markers. The impact of introduction and domestication of the species on its genetic structure and diversity is investigated. In Chapter 3 breeding system of the var. *caribaea* is studied. Single- and multi-locus outcrossing rates are estimated for 3 natural populations and 1 seed orchard population. Chapter 4 explores the application of DNA techniques in identification of an unknown seed source of *P. caribaea* in China. In Chapter 5 the use of integrated multiple function field test is explored and quantitative genetic analysis of such a trial is conducted. Based on the results obtained from these Chapters, Chapter 6 reports a proposed breeding strategy for *P. caribaea* in China. Finally, the thesis is ended with Chapter 7 in which a general summarisation and concluding remarks are presented. Critical appraisal on materials and methods used in the studies are discussed and priorities for future studies are recommended in this Chapter.

ACKNOWLEDGEMENTS

Thanks are given to the ODA of British government for the financial support for my studies and to the Division of Exotic Forestry, Chinese Academy of Forestry (CAF) for giving me the opportunity to study at the University of Edinburgh. I specially thank Dr. Richard Ennos for his invaluable supervision throughout my studies, from theory to practical points and to the usage of English language. I could not have made such achievements without the help of his wide and deep knowledge of genetics, biology and forestry. The thesis could not have been turned into present shape without many times of his careful read-through and important comments from structuring to spelling. Professor Huoran Wang at the CAF is greatly thanked for his genial support and encouragement for my studies. Particular thank is given for his help on providing the Chinese samples, data collected from the field trial and Chinese literature. Without these, my studies would not have been possible. I thank Dr Adrian Newton for his invaluable comments on the drafts of the thesis. I thank Mr. George Gibson for his important advice on my study during the MSc course and for his invitation and help for my travel from China to UK. Thanks are also given to Dr. Douglus Malcolm for his support and encouragement for my PhD study.

I appreciate the Forestry Commission at South Alice Holt, England for providing me with the samples of Cuban variety and Bahaman variety. I thank Dr. Kanowski, Oxford Forestry Institute (OFI) for his advice on the quantitative analysis of the field experiment. Advice on DNA studies from Dr. Steven Harris, OFI and help on arranging seed samples of Bahaman variety from Mr. Alan Pottinger, OFI are acknowledged. Help from Dr Xiaoru Wang, Swedish Agricultural University on using BIOSYS computer programme is appreciated. Help from Mr Carolyn Argentati, Natural Resources Library, North Canolina State University in providing useful information is acknowledged. I am very grateful to John Morman, Billy Sinclair, Anthony Langton, Eric Easton and Xingsheng Hu for their kind helps during the laboratory work. Establishment and management work of the experiment by the staff of Forestry Institute of Hepu County, China is acknowledged. My colleagues Mr. Lu Pengxin and Mr. Zang Daoqun are also acknowledged for their help in the experimental design and early establishment. I am grateful to many staffs and friends at IERM, ITE and Forestry Commission who provided help and support to my studies.

Finally but not the least, thanks are given to my family and relatives. I am very indebted to my wife Liu Yan and my daughter Katie for their continuous encouragement and support to my studies. I have not been able to spend enough time, to stay with them particularly with my new born daughter. I feel deeply indebted to my parents, I have not visited them for more than 7 years during my stay in UK and Beijing. Without their understanding and support, I could not have been strong enough to finish these studies.

LIST OF ABBREVIATION

ANOVA: Analysis of variance

ANCOVA: Analysis of Covariance

BSO: Breeding Seedling Orchard

CAF: Chinese Academy of Forestry

CBSO: Composite Breeding seedling Orchard

CFI: Commonwealth Forestry Institute

CPU: Central Processing Unit

CSO: Clonal Seed Orchard

DANIDA: Danish International Development Agency

d.f.(m.v.): Degree of Freedom (Missing Values)

EMP: Expected Mean Product

EMS: Expected Mean Square

ECTF: Edinburgh Centre of Tropical Forestry

FAO: Food and Agriculture Organisation

F pr.: F test probability

G × E: Genotype by Environment interaction

IERM: Institute of Ecology and Resource Management

ITE: Institute of Terrestrial Ecology

LATTSQ: Lattice Square

MAT: Mean Annual Temperature

MAR: Mean Annual Rainfall

MPBS: Multiple Population Breeding System

ODA: Overseas Development Administration

OFI: Oxford Forestry Institute

Pcb: *Pinus caribaea* var. *bahamensis*

Pcc: *Pinus caribaea* var. *caribaea*

Pch: *Pinus caribaea* var. *hondurensis*

RCB: Randomised Complete Blocks

REML: Restricted (or residual) Maximum likelihood

VCOMP: Variance Component

v.r.: Variance Ratio

Abstract

Pinus caribaea is a tropical pine that is widely planted as an exotic. It is an important plantation species for industrial purposes in China. The objective of this thesis was to collect basic information on genetic variation, breeding system and performance of a range of seed sources which have potential for incorporation into the Chinese breeding programme. Isozyme markers were used to look at genetic diversity, population structure and breeding system of the species. DNA techniques together with isozyme markers were used to distinguish an unknown seed source. Quantitative analysis of morphological traits was also conducted for a field provenance/progeny trial.

Isozyme variation was studied in natural populations of two geographically separated varieties, *P. caribaea* var. *caribaea* and *P. caribaea* var. *bahamensis*, and in exotic populations of the varieties from China and Australia respectively. There was significant genetic differentiation between the two varieties. Within the varieties, populations of var. *bahamensis* were more differentiated than those of var. *caribaea*. Exotic populations of both varieties experienced higher inbreeding than natural populations. With var. *bahamensis*, the Australian population was genetically similar to natural populations. However there were large genetic differences between the Chinese population and the natural populations of both var. *caribaea* and var. *bahamensis*. The Chinese material was identified as a distinct species or hybrid by the use of chloroplast DNA variation. Mating systems in natural and seed orchard populations of var. *caribaea* were analysed using isozyme markers. Both single and multilocus estimates of outcrossing rate were significantly smaller than 1.0 (complete outcrossing) in the island population, but were not significantly less than 1.0 in mainland populations and the seed orchard population, indicating that stronger inbreeding exists in the island population. The small differences between single and multilocus estimates suggest that the inbreeding detected within the variety is caused by true selfing rather than consanguineous mating. The selfing rates ranged from 10.6 % in the island population to 1.5 % in the seed orchard based on 5 loci assayed. Estimates of female outcrossing rates were found to be variable among individual trees though this was probably caused by sampling error.

Quantitative analysis of three growth traits, height, diameter and crown width were conducted in a field trial of *P. caribaea* var. *bahamensis* located in China. Significant genetic variation was found in growth traits among regions, provenances within region and families within provenance, with the Abaco Island and New Providence provenances superior and those from Andros Island inferior. Growth performance was significantly correlated with latitude and longitude of provenance origin. Family heritabilities for all traits were much higher than individual heritabilities. Results suggest that provenance selection would be the most appropriate first step for the genetic improvement of the species and substantial genetic gain could be achieved by family selection within provenance. Genetic correlations among traits were found to be significant. A breeding strategy of multiple population breeding system (MPBS) featuring simple and cheap breeding methods for the first generation and intensive breeding for advanced generations is proposed. Finally, several problems pertaining to the breeding programme in China are discussed.

Key words: *Pinus caribaea,* Diversity, Genetic structure and variation, Isozyme, Chloroplast DNA, Mating system, Outcrossing rate, Breeding strategy, Provenance variation, Genetic parameter

1. Introduction

1.1 Exotic forestry and *Pinus caribaea* in China

Exotic forestry refers to the methodology of species transfer and introduction and domestication of exotic tree species. Exotic species refer to those grown in an area where they do not naturally occur (Zobel and Talbert, 1984). Exotic species are utilised to supplement or replace the local indigenous forests that cannot, or do not, produce the desired quantity and quality of forest products when the native forests have been destroyed or when suitable trees are not present (Zobel *et al.*, 1987). The introduction and domestication of exotic species and the development of exotic forestry have always been closely associated with human civilisation and economic development. Some temperate countries in which indigenous tree species are very few or cannot meet the demand for wood and wood products have a long history of exotic forestry. Nowadays, the most successful examples of exotic forestry are in the tropics and subtropics (Zobel *et al.*, 1987; Evans, 1992; Eldridge *et al.*, 1993). Although arguments have been made against exotic species by environmentalists who fear site degradation, environment upset and monoculture, exotic forestry continues to be very successful and plays an important role in the forest economy in many parts of the world. The reasons for success of exotic forestry have been discussed by Zobel and Talbert (1984) and Zobel *et al.* (1987). The most widely planted exotic species over the world are members of the genera *Eucalyptus* and *Pinus*.

Despite the fact that China has a very rich flora and many species are of important economic value, plantation forestry using exotic species plays an important role in the nation's forest economy. About 700 exotic species belonging to 85 families have been grown in China and about 20 of them, in genera such as *Pinus*, *Larix*, *Populus*, *Eucalyptus*, *Acacia*, *Casuarina* and *Robinia*, are of key importance for wood production and other economic uses. The area of plantations using exotic species is estimated at more than 8 million ha which accounts for a quarter of the total plantations in the country (Wang, 1994). Among the exotic pine species grown in China, the Caribbean pines are the most important plantation species in the tropics while the Slash pine and the Loblolly pine are important in the subtropics. Faster growth, better wood quality and higher resistance to pests make the Caribbean pine a favourable and promising species for industrial purpose in the tropics.

1.2 Description of *Pinus caribaea* Morelet

1.2.1 Botanical nomenclature

The taxonomic position of *Pinus caribaea* Morelet is in subsection *Australes*, section *Pinus*, subgenus *Pinus*, genus *Pinus*, family Pinaceae (Little and Dorman, 1954). Other species in the same subsection with *P. caribaea*, are *P. elliottii* Engelm, *P. cubensis* Griseb. and *P. occidentalis* SW. Other close relatives of *P. caribaea* Morelet are *P. oocarpa* Schiede and *P. tropicalis*. *P. oocarpa* Schiede is associated with *P. caribaea* in natural pine forests in the central America continent (including Honduras, Nicaragua) and belongs to subsection *Oocarpae*. *P. occidentalis* is associated with the Cuban variety of *P. caribaea* and belongs to subsection *Sylvestres*. *P. merkusii* Jungh and de Vriese and *P. kesiya* Royle ex Gordon are also in subsection *Sylvestres*.

The name *P. caribaea* was used before 1950 to cover not only the pines of the Bahamas, the Pino Macho of western Cuba and the continent provenances, but also pines, which are now named as *Pinus elliottii* var. *elliottii* and *P. elliottii* var. *densa* Engelm, in the Southern USA (Little and Dorman, 1954). Since the early 1950s the pine confined to the Southern USA has been referred to as a separate species *P. elliottii* Engelm. The Slash pine in South Florida was ranked as a variety named *P. elliottii* var. *densa*, while the name *P. caribaea* was retained for the Cuban, Bahaman and Central American populations. Since 1957 more intensive studies were carried out on the variation within the species (Luckhoff, 1964; Barrett and Golfari, 1962) and consequently *P. caribaea* was divided into three varieties: *P. caribaea* Morelet var. *caribaea* (typical), *P. caribaea* Morelet var. *hondurensis* Barr. and Golf., and *P. caribaea* Morelet var. *bahamensis* Barr. and Golf.

1.2.2 Botanical properties

The trees of the species are generally straight-boled and knot-free, reaching heights of up to 45 m and a DBH (Diameter at Breast Height) of over 100 cm. In all varieties the needles are between 15 and 25 cm long, 1.5 mm wide at the most, stiff and finely serrate, dark to yellowish-green in colour, and bearing stomata in whitish lines on all sides. Bark is thick and deeply fissured, reddish-brown to pale brownish-grey in colour. Seeds are ovoid, about twice as long as they are wide (Lamprecht, 1989). The major morphological differences among its three varieties are described as follows by Barrett and Golfari (1962):

var. *caribaea*: Needles in fascicles of 3 (rarely 4); cones 5-10 cm long; seeds with adnate wings (remaining attached).

var. *hondurensis*: Needles in fascicles of 3, occasionally of 4, 5 and 6 in young trees; cones 6-14 cm long; most of the seeds with articulate wings (wings becoming detached).

var. *bahamensis*: Needles in fascicles of 2 and 3; cones 4-12 cm long; seeds mostly with articulate wings, rarely adnate.

1.2.3 Natural Distribution

The natural distribution of *P. caribaea* is wide and often discontinuous (Figure 1.1). It covers an area with around 15 degrees of latitudinal range, from 12°13'N in Nicaragua to 27°00'N at Little Abaco island in the Bahamas (Lamprecht, 1989). The longitudinal range is nearly 17 degrees, from 74°40'W in the Caicos Islands to 89°25'W at Poptun in Guatemala (Lamb, 1973). The altitudinal range is from sea level to 12 m in the Bahamas and Caicos Islands, from sea level to 280 m in Cuba and from sea level to 1,000 m in Honduras. The range covers four large islands in the Bahamas, three small islands in the Caicos group, parts of the western province of Pinar del Rio in Cuba, north-western parts of the Isle of Pines off Cuba, parts of the northern and coastal plain and hills of Honduras, isolated localities in eastern Guatemala, along the northern coast and in upland valleys of Honduras, on Guanaja Island and as an extensive provenance down the eastern sea board of Honduras and Nicaragua. The natural distribution of *P. caribaea* including its three varieties is mapped in Figure 1.1.

1.2.4 Ecological requirements

1.2.4.1 Climate

Throughout its range, *P. caribaea* experiences a summer rainfall, winter dry season regime. The climate is frost-free and warm, and humid for most of the year (Greaves, 1978). It naturally occurs in permanently moist to seasonally moist tropical lowlands in areas with predominantly high relative humidity. Differences exist among the three varieties. The distribution ranges and main climatic data of the three varieties are given in Table 1.1.

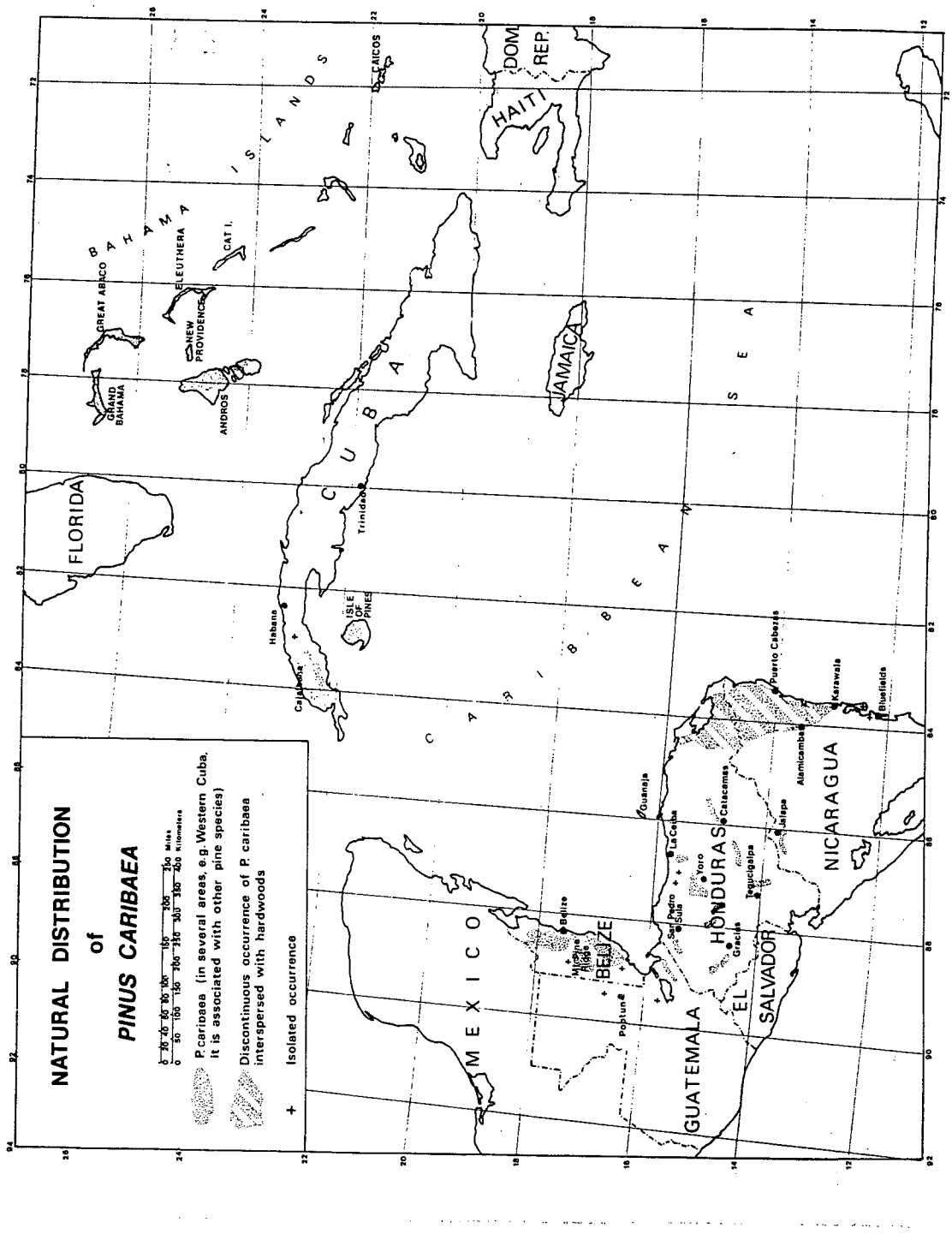

Figure 1.1 Map of natural distribution of Pinus caribaea Morelet

Table 1.1 Natural distribution and climatic data in the natural range of the varieties of *P. caribaea*.

Variety	Latitudinal range	Altitudinal range (m)	Geographic range
bahamensis	21°45'-27°N	0-50	Bahamas, including Caicos
caribaea	21°35'-22°50'N	40-350	Western Cuba, Isle of Pines
hondurensis	12°13'-18°00'N	0-800	East Nicaragua, South-eastern Mexico

(Continued)

Variety	MAR (mm)	Dry Season (month)	MAT (C°)	Mean Monthly Temperature (°C)		
				Warmest	Coldest	Extremes
bahamensis	800-1400	5-7	-	28	22	15-31
caribaea	1000-1800	4-6	24-27	28	22	12-34
hondurensis	900-4000	2-6	20-17	29	21	5-37

MAR: Mean annual rainfall; MAT: Mean annual temperature

- var. *bahamensis*

  Summer type rainfall decreases from north to south. Grand Bahama has a mean annual rainfall of 1,469 mm. On Andros three degrees of latitude farther south it is 1,055 mm. The mean annual temperatures in the Bahamas vary slightly (24.2°C to 25.4°C), but the southern islands are generally warmer in the winter months than the northern. The greatest difference between minimum and maximum temperatures is 30°C at Grand Bahama (min. 5°C, max. 35°C). The whole range is subject to light north-east winds from October to January and stronger east and south-east winds from February to September. Occasionally when cool north winds from North America occur during the winter, the temperature reaches its coldest level (Greaves, 1978).

- var. *caribaea*

  The Cajalbana pine forest of Cuba receives about 1,780 mm of rainfall due to its extra elevation (330m). Further to the south-west in the piedmont country at Guane the annual rainfall is 1,295 mm and on the plain of the south coast it falls to 1,060 mm. On the Isle of Pines rainfall at Nueva Gerona is 1,500 mm. Atmospheric humidity never falls to a low level in these islands (Lamb, 1973).

- var. *hondurensis*

  Rainfall in northern coastal savannas of Honduras and on the western edge of the Mountain Pine Ridge is about 1,500 mm. The climatic conditions are generally rather similar throughout the coastal range of *P. caribaea,* with mean annual rainfall increasing progressively southwards, so that on the Island of Guanaja it is 2,300 mm and reaches about 3,900 mm at the extreme southern limit of the range at 12°13'N. Rainfall on the inland sites is generally lower than on the coast. The mean annual rainfall is about 1,690 mm with four months having rainfall less than 75 mm (Haufe, 1969) at the most northerly stands, near Poptun in Guatemala. Most inland stands of *P. caribaea* in Honduras and Nicaragua have an annual rainfall from 1,200 to 1,500 mm. However, some stands in the upper valley of the Choluteca River in Honduras have an annual rainfall as low as 660 mm with six consecutive months of rainfall less than 40 mm (Lamb, 1973).

The natural range of *P. caribaea* in the Bahamas and Cuba has a warm equable climate with no frost. Mean temperature varies from 22.1 °C during the coldest month (January) to 28.3 °C during the hottest month (August). In the mainland natural range, the climate is more variable and less equable except in the coast area. In upland valleys of the Mountain Pine Ridge the temperature falls to 5 °C in January. In the hot season in May it may rise to 37 °C in inland areas. In coast districts temperatures rarely fall below 15.6 °C or rise above 32.3 °C. Further south to the east coast of Nicaragua temperature fluctuates slightly and ranges from a mean minimum of 23.9 °C in January-February to about 29.4 °C in May-June (Lamb, 1973).

The climate in the natural range has an important influence on its potential in exotic environments. Generally, the Honduran variety is more suitable for warmer tropical temperatures and the other two varieties particularly var. *bahamensis,* are more suitable for cooler tropical conditions.

1.2.4.2 Soil

P. caribaea grows under a wide variety of soil conditions. However, soil moisture is generally more important than nutrient availability for good growth. *P. caribaea* can typically be found growing in acid soils with a pH from 5 to 5.5, but the var. *bahamensis,* which can thrive in

soils with a pH of up to 8.5, is an exception to this. The soil types typically preferred by the 3 varieties can be characterised as follows:

- var. *caribaea*

  Flat but well drained oxisols and shallow sandy soils; waterlogged during the wet season, very dry in the arid months.

- var. *hondurensis*

  Belize: In the coastal plain in leached soils of alluvial origin and sandy loams with a clay horizon impermeable to water at a depth of 1.0-1.5 m. In hilly regions in weathered, well drained sandy soils. Not on calcareous sites.

  Guatemala: On well drained, strongly weathered calcareous soils with manganese concretions between 0.5 and 1 m below the surface.

- var. *bahamensis*

  In poor, porous, to some extent shallow and to some extent well drained soils on coral reefs. The pH of the soil for var. *bahamensis* can be as high as 8.5 and water table is at a depth of about 60 cm (Lamprecht, 1989). The reason for this tolerance to high pH is not known yet, but is probably due to endogenous or mycorrhizal factors (Greaves, 1978).

1.2.5 Physiology

1.2.5.1 Drought and the root system

Since *P. caribaea* grows naturally in a climate free from frost and it appears to be little influenced by length of day, the predominant environmental factor controlling the growth, apart from a sufficient supply of nutrients, is soil moisture. Moisture stress has an obvious influence on the root system as it reduces not only root extension but also mycorrhizal development (Lamb, 1973). The root system is variable from site to site depending on the soil condition. It can reach the source of moisture if the subsoil permits root penetration, and therefore soil texture is of vital importance. Haigh (1966) studied root systems of plantation trees of var. *hondurensis* in South Africa in conditions where water stress in the soil was absent and found that the trees produced a very small surface root system, from which were

sent down one or several vertical tap roots to the vicinity of the water table. Water absorbing roots were abundant where the perennial water occurred. Lamb (1973) reported that the root system of var. *bahamensis* growing in the Bahamas is a flat sheet about one foot thick on top of the coral limestone platform with no taproot. This effect is caused by the soil texture which does not permit the roots to penetrate.

There is no reported study on the root system of var. *caribaea.*

1.2.5.2 Stem form

Stem form is of key importance in commercial plantations whatever their end use will be. There are great variations in stem form between the varieties and between the sites. The form known as foxtail or branchless stem is extremely rare under natural conditions in var. *hondurensis* and has not been observed at all in natural stands of the other 2 varieties (Lamb, 1973). However, outside its natural habitat, var. *hondurensis* has extremely variable stem form and the percentage of foxtail trees can be as high as 30%. This phenomenon of foxtailing was explained by Lamb (1973) as the rapid early growth due to the adequate supply of nutrients and moisture under the plantation condition, whereas in natural habitats the nutrients and moisture are often in stress and competition with weeds is strong. He also demonstrated that the variation in foxtailing is not only strongly controlled by genetic inheritance but also influenced by the growing conditions. In windy conditions foxtailed trees are bent over by the wind because the timber of these trees is very weak due to the rapid growth. Consequently, foxtailing produces curled stems (Lamb, 1973). However, in non-windy conditions, foxtailing tends to have a beneficial rather than detrimental effect on wood properties at early life of the trees (Plumptre, 1984). Foxtailed trees are taller and slimmer than normal trees but do not differ significantly in mean wood density (Plumptre, 1984; Whyte *et al.*, 1981; Woods *et al.*, 1979).

The abnormal growth of var. *hondurensis* can change to normal and *vice versa* in plantations outside its natural habitat (Kozlowski and Greathouse, 1970). Individual trees can change from normal to foxtail form and from foxtail to normal. Foxtailed trees grow continuously whereas normal trees grow periodically and foxtailing varies with the quality of the sites.

1.2.5.3 Growth pattern

P. caribaea can normally keep growing throughout a year under a tropical climate. However its growth is correlated to the rainfall distribution in a year. Rapid growth usually happens during the high humidity months (Greaves, 1978).

1.2.5.4 Flowering and fruiting

Female flowers are produced over many months and are generally receptive in the winter dry season. Studies in Queensland for all three varieties (Slee and Nikles, 1968) gave the following results:

Table 1.2 Flowering and cone maturing time of the varieties of *P. caribaea* in Queensland

Variety	Months when female flowers receptive	Pollen ripen	Cone maturity	
			Month	Months after flowering
caribaea	June-July	July	January	18
bahamensis	Mid. Apr.-Mid June	Apr.-June	February	21
hondurensis	March-June	Mid Apr.-June	February	21

Young plantations usually start bearing female cones when they are 3-4 years old but these do not produce fertile seed, probably due to the inadequate supply of pollen at this age. In natural regeneration in Honduras, cones are usually produced after 10 years (Lamb, 1973). Latitude has an effect on the time of cone ripening.

Cones start to shed their seeds 2-3 weeks after turning brown on the tree depending on the atmospheric humidity at the time. All seeds are shed soon after the cones ripen (Lamb, 1973).

1.2.6 Silviculture

1.2.6.1 Nursery

All varieties of *P. caribaea* can normally germinate easily within 8-21 days under tropical conditions if the quality of seeds is high. There is little difference between soaking in water at room temperature for 24 to 60 hours and moist cold storage for 14 days or both combined (Lamb, 1973). The number of seeds per kg is different for each variety: 59,000 seeds/kg for var. *caribaea*, 52,000-72,000 seeds/kg for var. *hondurensis* and 81,000 seeds/kg for var.

bahamensis (Lamprecht, 1989). The germination rate of fresh seeds is 70-80%. This rate is maintained for at least a year. The viability can be conserved for 8-9 years by reducing the humidity to 8% and storing the seed at 0-10 °C (Lamb, 1973)

For the establishment of a plantation, direct seeding is uncommon. North of the equator, it is usual to sow in the months of October and November, to transfer the seedlings to nursery beds in December or January, and to transplant them to the field in May or June (Lamprecht, 1989). A wide range of different nursery techniques can be used for the species. Inoculation with mycorrhizal fungi is practised with container raised seedlings.

1.2.6.2 Planting

Plants should not be moved to the field until they are 20-30 cm tall. Planting of container raised seedlings is recommended. The survival rate and growth rate of young plants were much higher for container raised seedlings than for bare-root planting stock. Application of phosphorus and nitrogen fertiliser at planting will increase growth.

1.2.6.3 Spacing and rotation

Spacing can be variable from 2 x 2 m to 5 x 5 m, depending on production aims. Rotation could be from 12-15 years for pulp wood and from 15-30 years for saw timber production.

1.2.6.4 Pest and disease

The main pest in the native habitat of this species is the bark beetle *Dendroctonus frontalis*. Considerable growth losses are occasionally caused by leaf-cutter ants of the genus *Atta* in Latin American plantations, and by various termite species such as *Coptotermis currignatus* in Southeast Asia and by tip moth in China (Lamprecht, 1989; Pan and You 1991).

1.2.7 Timber production and utilisation

Within the native habitat, *P. caribaea* stands have low standing volumes and growth increments. In Guatemala, the standing volume was estimated as 46.3 m^3 ha^{-1} (±26%) on average, and the annual volume increment was 2.6 m^3 ha^{-1} (Lamb, 1973). Annual volume increments were measured between 3 and 37 m^3 ha^{-1} on various plantations in Queensland

indicating that *P. caribaea* plantations are by no means economically justified on all sites (Lamprecht, 1989).

The wood is hard and generally durable, with yellowish sapwood and reddish-brown heartwood. There are large differences between slowly grown wood from natural stands and quickly grown plantation wood (Lamb, 1973). Wood quality is variable among varieties and usually decreases with the increase of growth rate (Table 1.3).

Table 1.3 Difference of wood properties between natural stands and plantations

Wood property	Natural stands	Plantations
Specific gravity	0.65-0.8 (kg m^{-3})	0.4-0.5 (kg m^{-3})
Resin content	lower	greater
Susceptibility to blue stain	more or less resistant to fungal attack	fungal attack
Sawability	good	poor

Wood from natural stands is suitable for all outdoor and indoor applications. Plantation wood is principally used for paper production. The species is important as a source of resin in Central America, playing a role similar to that of *P. pinaster* in Europe (Lamprecht, 1989).

1.3 Introduction, domestication and genetic improvement of *P. caribaea* with particular reference to China

1.3.1 Plantation forestry of *P. caribaea* and its potential

P. caribaea is a well-known fast-growing species native to South and Central America. The first introductions of *P. caribaea* in the world took place in as early as 1929, but it was not until 1960, when its potential for softwood plantation development in tropical countries was recognised, that a general interest in the species by these tropical countries developed (Greaves, 1978).

As a widely planted exotic tropical plantation species, *P. caribaea* can be found throughout the tropics and subtropics over the world: From 14°N (Nigeria) to 35°S (South Africa) in Africa; from 18°N (Puerto Rico) to 33°S (Brazil) in Latin America; from 1°N (Malaysia) to 33°N (India) in Asia; and from 1°S (New Guinea) to 28°S (Australia) in Oceania (Lamb, 1973; Lamprecht, 1989). Commercial plantations have been established in probably more

than 40 countries in Africa, South America, the Caribbean Islands, Asia, the Pacific Islands and in Australia.

In recent years, *P. caribaea* has been increasingly planted in Southern China on a large scale. It has been selected as a major fast-growing plantation species for industrial purposes and given top priority in the national afforestation programme. The species is increasingly playing an important role in forestry industries in China.

However, interest in var. *caribaea* and var. *bahamensis* has always been overshadowed by var. *hondurensis* which generally out-performs them in growth except in a narrow range of environments where conditions are too temperate for var. *hondurensis* and too tropical for *Pinus elliottii.* Nevertheless, their excellent stem form and branch form and good growth, and also high resistance to shoot tip moth (*Rhyaciona* and *Dioryctria* spp.) which has devastated the var. *hondurensis* in Southeast Asian countries, means that both var. *caribaea* and var. *bahamensis* may have great potential over a huge area, because they are the next most productive softwood after *P. caribaea* var. *hondurensis* (Baylis and Barnes, 1989).

Before the 1980s, a number of projects had been carried out with *P. caribaea* in many different countries and regions; these studies concentrated on tree improvement, breeding and silviculture (Nikles *et al.,* 1978; Wang *et al.,* 1995). Numerous studies of quantitative genetic variation have been conducted on the species, especially the var. *hondurensis* (Nikles *et al.,* 1978; Barnes and Gibson, 1984; Gibson *et al.,* 1989). The majority of the literature available is on var. *hondurensis* while relatively few results are available on the other two varieties.

1.3.2 Breeding history and current status

1.3.2.1 World

The first export of seed of *P. caribaea* happened in 1927. Seeds from the coastal plain at Old Man's Rest in British Honduras were introduced to Placerville, California and South Africa (Luckhoff, 1964). The seed sent to South Africa was raised and planted out in 1929 (Lamb, 1973). Australia started to grow the var. *caribaea* and var. *bahamensis* in Queensland in 1930 (Nikles, 1967; Lamb, 1973). In the early 1940s, British Honduras (1943 onwards) and Cuba (1945 onwards) established plantations with their own variety. Since then, the

Honduran variety was introduced to Surinam, Guyana and Western Malaysia (1953). East and Central African countries started trials from 1957. Thereafter the FAO seminar on Tropical pines in Mexico in 1960 probably did more than anything else to stimulate interest in the species by tropical countries (FAO, 1961). Replicated provenance trials were established in Fiji, the Sudan, Uganda, Kenya, Tanzania, Zambia, Rhodesia, Malawi, South Africa, and Queensland, followed by Gambia, Sierra Leone, Nigeria, The French speaking countries in West and equatorial Africa, Madagascar, Brazil, Argentina, Venezuela, Jamaica, Colombia, Costa Rica, Sabah, Sarawak, India, Sri Lanka, Northern Territory of Australia, the Solomon islands, and Hawaii also started trials. Ghana, Dahomey, Togo and New Guinea started to grow the species in the early 1970s (Lamb, 1973). Most of these early introductions were of the var. *hondurensis* until the Tropical Silviculture Unit at the Commonwealth Forestry Institute (CFI) was set up in 1963. From 1964 to the 1970s it was difficult to supply seed of var. *caribaea* from Cuba (Lamb, 1973).

P. caribaea has shown great differences among its 3 varieties in growth rates and stem form. Experiments in South Africa showed that var. *caribaea* had the slowest growth rate, regular stiff ascending multinodal branches and a very straight stem; the var. *bahamensis* grew faster than var. *caribaea* and had excellent form with longer, slender, multinodal branches; the var. *hondurensis* had extremely variable form, containing 39% of foxtail trees, together with many straight-stemmed trees, and it grew faster than either of the other two varieties in height (Lamprecht, 1989). Similar results have also been observed from many other studies (Nikles *et al.*, 1978).

Much genetic improvement work has been done with the var. *hondurensis* (Nikles, 1979; Nikles, 1989; Nikles and Robinson, 1989). However, work on var. *bahamensis* and var. *caribaea* started later due to the difficulty of their seed supply and fewer results are available. Most breeding work on var. *bahamensis* as an exotic species has been carried out since the 1980s after the intensive seed collection by OFI (Baylis and Barnes, 1989).

1.3.2.2 China

1.3.2.2.1 Environments

The area of planting of *P. caribaea* in China geographically ranges in latitude from 18°55' to 23°30' N with an extension along the eastern coast to Fuzhou and a separate area in southern Yunnan (Figure 1.2).

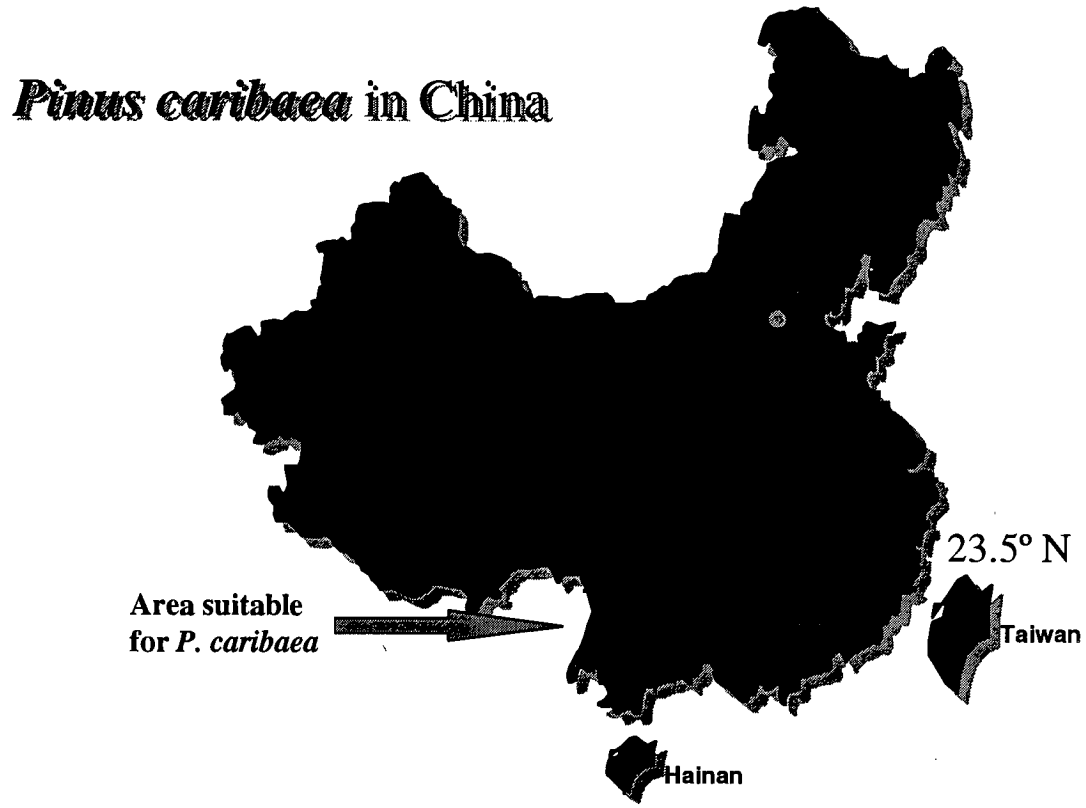

Figure 1.2 Planting area for *P. caribaea* in Southern China

The climate in this range is subtropical to tropical. Most of the trial plantations were established at elevations below 100 m on coastal tableland or undulating hills, but in Yunnan the altitude is 1,060 m above sea level (Wang *et al.*, 1995). Soils are lateritic or yellowish-red soils with a pH value from 4.5-5.5, derived from shallow marine sediment, granite, basalt or sandstone. Planting sites for *P. caribaea* in southern China are normally exposed with skeletal, seriously eroded and iron-pan soils with an organic matter content of less than 1%.

Mean annual temperatures in the region of planting vary from 19 to 24 °C. Absolute maximum temperature ranges from 36.8 to 38.1 °C and the minimum is about -1 °C. Mean annual rainfall varies from 1,527 to 2,253 mm with the peak occurring in the summer from

environments of the species but are influenced by the more continental climate, which causes the temperature dramatically to drop to below zero, sometimes for several days in winter. This is a major limitation to growing *P. caribaea* in China (Wang *et al.*, 1995).

1.3.2.2.2 Tree introduction

The introduction of *P. caribaea* started at the beginning of the 1960s. In 1961, seeds of var. *caribaea* were introduced from a seed source in Cuba and the seedlings were planted in Fuzhou Arboretum in 1963 (Li *et al.*, 1991; Liang and Zhou, 1991). A second introduction of var. *caribaea* was made in 1964, when seeds were directly imported from Cuba to Guangdong and Guangxi provinces and species trials were established at 10 sites in the two provinces. In Guangdong province, trials were located in Leizhou Peninsula, Shantou and Zhaoqing districts and the Arboretum of Guangdong Forestry Institute in Guangzhou and covered a total area of 1.6 ha. Of these earliest introductions of var. *caribaea*, about 3,000 individuals (nearly half), still exist today (Wang *et al.*, 1995). In Guangxi province 6 trials were set up by Forestry Institutes and State Forest Farms of Nanning, Hepu and Yulin areas (Anonymous, 1985; Zhu *et al.*, 1986; Chen *et al.*, 1989; Wang and Lai, 1990).

An increase in the planting of var. *caribaea* was made in 1975, when the seeds produced from the second introduction in Guangxi Forestry Institute were transferred to Xishuangbanna area in Yunnan province. Trees were planted out in 1978 (Ma, 1986). The area is at the northern edge of the tropics and experiences a lowland humid tropical climate. There are no indigenous pine species growing naturally in the same area. The native pine species *P. yunnanensis* occurs naturally at higher altitudes.

The var. *hondurensis* and the var. *bahamensis* were first introduced into Guangdong and Guangxi provinces in 1973 (Pan and You, 1991; Wang *et al.*, 1995). The var. *hondurensis* was introduced into Hainan province in the following year (Wang *et al.*, 1995). The origin of var. *hondurensis* was Guatemala and seeds of var. *bahamensis* were provided by the CFI, now the OFI. After these two introductions, the total area of experimental plantings was estimated of 440 ha for var. *caribaea* and 5 ha for var. *bahamensis* (Zhu *et al.*, 1986). However, no estimates were available for var. *hondurensis*.

These early introductions of *P. caribaea* started to produce seeds in 1975. A very good seed production was obtained in 1982 and 1983, which enabled larger scale plantations to be established in Leizhou Peninsula. The area of the species extended to more than 3,500 ha, of which 90% was var. *caribaea* (Wang *et al.*, 1995).

As the potential of the species became more widely recognised, interest in the species was constantly increasing and resulted in a large demand for seeds for planting. The seed production from these early introductions was insufficient to meet such a big requirement. Large amount of seeds of *P. caribaea* were imported from commercial sources abroad yearly. No records of source identity for these imports are available today. The total quantity of imported seeds of *P. caribaea* was estimated to be up to 6,000 kg during the period between 1973 and 1988 (Pan and You, 1991). The total area of *P. caribaea* plantations approached 19,700 ha by the end of 1988 (Wu and Ye, 1992). The present plantation area of *P. caribaea* is estimated to be more than 40,000 ha, mainly in Guangdong, Guangxi and Hainan provinces.

1.3.2.2.3 Genetic tests

Generally speaking, the introduction of *P. caribaea* in China has followed the basic procedures common in forest tree introduction (Burley and Wood, 1976; Zobel *et al.*, 1987). Firstly, species or variety trials were established at the beginning of the species introduction in different environments. Then provenance trials were established in a range-wide test followed by a limited-range test. Following this progeny trials were established. Results from these tests have been applied to the planting programme of the species. However, it should be noted that due to the limitation of seed availability, some plantings were still using untested sources and those from commercial sources were also not well documented in local planting programmes.

Provenance trials have been established for each of the three varieties since 1973. In these trials only limited sources were used. Until 1983, as a participant of the international provenance trials on *P. caribaea*, China has carried out a series of provenance trials of these three varieties of *P. caribaea.* Species(variety)/provenance trials were set up at various sites within Guangdong (Zhu *et al.*, 1986; Wang *et al.*, 1989), Guangxi (Ling and Huang, 1991), Hainan (Liang and Zhou, 1991) and Yunnan (Luo *et al.*, 1992) provinces. A wider range of provenances was used in these tests. Some of these trials have been reported by local institutes

in local newsletters or journals. However, there were almost always problems in tracking down the origins of their materials due to poor project management.

From 1983, the introduction and domestication of *P. caribaea* have been included in the national priority research project and extensive research on tree improvement of the species has been carried out in the 7th and 8th five-year national plans. All this work was co-ordinated by the CAF, Beijing. Lu *et al.*(1990) reported comprehensive results on the genetic variation and provenance selection of *P. caribaea.* This is the first publication of the results from the national research project on *P. caribaea.* They found that there were significant differences among varieties in growth rates and pest resistance, and large genetic variation within var. *bahamensis* and var. *hondurensis.* In the following year, Pan (1991) and Pan and You (1991) edited two books "Growing exotic trees in China" and "Genetic improvement and silviculture of the American Southern Pines and the Caribbean Pines", in both of which many details of introduction, domestication and genetic improvement of *P. caribaea* were covered in regard to the whole country. In these publications, they concluded that the introduction of the species was very successful. *P. caribaea* grows faster than the native pine *P. massoniana* and the American pine *P. elliottii* in the tropical and southern subtropical regions in China. All the varieties of the species can be normally reproduced by seeds. *P. caribaea* was recommended as best suited for the area south to 22°20' N including Hainan island and Leizhou Peninsula, south of Yangjiang, Dianbai, Lianjiang and Hepu (Figure 1.2). The area between 23°30' and 22°20' is also suitable for *P. caribaea* but in this region it does not flower. The growth of *P. caribaea* in this area is still greater than of *P. elliottii.* With respect to the varieties, the var. *caribaea* is suitable for the lowland tropics and south tropics. Inland mountain areas are suitable for the var. *hondurensis.* However, in Leizou Peninsula, where the tip moth is severe, var. *bahamensis* and var. *caribaea* are better than var. *hondurensis.*

The species shows great potential for industrial purpose in the tropical area in southern China. However, it appears that there are obvious differences among the three varieties in growth rates, stem form and pest and disease resistance. Overall results from earlier tests showed that var. *hondurensis* has fast growth rates in the early stages, but highest susceptibility to shoot tip moth and highest percentage of foxtailed forms. The var. *caribaea* and the var. *bahamensis* have excellent stem form and are more resistant to the shoot tip moth. Genetic variation and genotype by environment interaction were also demonstrated within the varieties.

The large genetic variation found within the species provides opportunities for genetic improvement.

However, the genetic base of the material used in the above provenance trials was relatively narrow. This is particularly true for the var. *bahamensis* which was introduced on a significant scale only in the 1980s. Nonetheless, significant provenance variation was shown in the trials. In 1991 a range-wide provenance trial combined with open-pollinated progeny tests of the var. *bahamensis* was established at Hepu County, Guangxi Province with seeds provided by the OFI (Oxford Forestry Institute). The trial consists of 121 families collected from natural forests. This was the initial step of genetic improvement of var. *bahamensis*. The quantitative genetic analysis of this trial is a main part of this PhD study. In co-operation with British institutions and scientists, a comprehensive genetic improvement programme for *P. caribaea* has been underway since 1992. As a part of the project, one of the major activities of the programme is an extensive range-wide seed collection of the var. *caribaea*, which was made during the Summer in 1994. In the following year a combined provenance/progeny trial was set up in Guangdong province using these seeds. The trial consists of 195 open-pollinated families from natural stands and one seed orchard, and 25 families from land races developed in Brazil and China. This seed collection also provides the majority of the materials for this PhD study.

1.3.2.2.4 Breeding and seed production

Breeding work on *P. caribaea* in China has concentrated on hybridisation. Experiments have been carried out on hybridisation between *P. elliottii* and *P. caribaea* var. *caribaea* successfully (Ye, 1990). The hybrid vigour in growth rates remains in the F_2 generation. However, the use of F_1 hybrids has not been implemented in large-scale planting programmes due to insufficient seed supply (Wang *et al.*, 1995).

11 seed orchards or seed production areas of var. *caribaea* were established in Guangdong with a total area of about 70 ha in the period from 1973 to 1978 (Zhu *et al.*, 1986). Plus trees were grafted to the rootstocks of *P. elliottii* or *P. massoniana* in seed orchards. It was reported that these seed orchards had a very narrow genetic base because all selections were made from some 3,000 individual trees from the second introduction established in 1964 with Cuban seed source (Zhong *et al.*, 1993).

Because the selection and design of the seed orchard were not well planned, the genetic gain obtained from the orchards was rather low (Wang *et al.*, 1995). The seed production varied from year to year ranging from 8 to 50 kg/ha (Wu and Ye, 1993). Present seed production of var. *caribaea* in China is around 1000 kg which is far less than the demands in planting programmes. About 500-1000 kg seeds is supplemented by imports from native countries (Wang *et al.*, 1995).

1.3.2.2.5 Silviculture in China

Tests showed variation in germination rates among the varieties and seed sources, and this was also related to quality of seeds and germination conditions (Zhang, 1994). A recent test showed that germination rates were strongly dependent upon length of seed storage. Fresh seeds revealed the highest germination rates (Zheng, unpublished data).

Nursery techniques are critical for the success of raising seedlings of *P. caribaea*. The nursery management in southern China is generally not intensive. Seeds are usually sown in trays or in soil beds. Seedlings are transplanted into containers before they are one month old. The containers are usually plastic bags filled with a mixture of subsoil, river sand and burned soil (burned surface-layer soil under forests) in a ratio of 1:1:1, with an addition of some superphosphate. Seedlings grow in the container for 4-5 months and are then outplanted to the field (Wang *et al.*, 1995). It is a common practice in China to spray liquid urea onto the seedlings. This helps to ease nutrition stress of the seedlings.

It is essential that seedlings of *P. caribaea* are inoculated with mycorrhizas to promote the seedlings' growth in the nursery and during early establishment in the field. A comparison showed that inoculated seedlings were 20-35 cm tall at age of 90 days, while those without inoculation reached only 14-19 cm in height at 120 days (Lin, 1992). The inoculation is usually done by taking soil from plantations of *P. caribaea* or *P. elliottii*. The major mycorrhiza fungi are *Pistolithus tinctorus*, *Tylopilus balloui*, *Hebeloma* sp. and *Cortinarius decipiens* (Zhu *et al.*, 1986).

Vegetative propagation by cutting has successfully been conducted with either shoot or fascicles in experiments (Zhu *et al.*, 1986; Peng, 1988). However, this has not yet been applied to operational use.

The best time for planting *P. caribaea* in China is from late March to early April. Site preparation is done by digging pits of 50 x 50 x 40 cm. Sometimes prescribed fire is practised for site clearance. Fertiliser is normally applied before planting. The site preparation is very intensive in most of the planting programmes in China. *P. caribaea* is usually planted as pure stands, although recently a few experimental mixed stands were established with *Acacia auriculiformis* in Hainan to improve the site conditions.

One of the major insect pests which seriously damages many pine species, including *P. thunbergii* Parl., *P. elliottii*, *P. merkusii*, *P. taeda* L. and *P. caribaea* in Guangdong province, is *Hemiberlesia pitysophyila*. Zhou (1991) reported that *P. caribaea* was more resistant to insects than any other pine species grown in Guangdong province. Within *P. caribaea*, var. *hondurensis* is most susceptible to *Dioryctria alternatus*. The insect reproduces rapidly and it is difficult to control. This is one of the reason for the poor stem form of var. *hondurensis* in China (Zeng and Yang, 1992). Both chemical and biological methods are used to control insects in practice.

1.3.2.2.6 Utilisation

The main purpose for growing *P. caribaea* in China is timber and pulp production. The species is also used for other industrial uses.

1.3.3 Prospects for *P. caribaea* in China

As a well-known fast-growing tropical tree species, *P. caribaea* will continue to be an important industrial species in tropical countries. In China, the total area of plantations of this species is now estimated to be over 40,000 ha. The role played by the species in plantation forestry in China is becoming more important as economic development requires more timber and pulp consumption. The area of *P. caribaea* in China is expected to reach 100,000 to 150,000 ha by the year 2010.

While the plantations of the species are rapidly expanding, genetic research and tree breeding lag far behind. Although priorities have been given to the genetic improvement for the species in the last two national five-year research plans, the applicable results obtained are still limited. In order to promote the productivity of the species in practical operations, a parallel genetic research and breeding programme on a national scale is needed, in which useful genetic information on population and quantitative genetics will be explored and a comprehensive breeding strategy will be developed.

Although numerous quantitative studies are available on *Pinus caribaea,* these are particularly with var. *hondurensis* and studies on the other two varieties are rarely available. Knowledge of evolutionary processes, genetic structure, breeding systems, genetic variation and patterns of distribution (for both selectively neutral marker genes and selectively adaptive quantitative traits) for var. *bahamensis* and var. *caribaea* remains unknown. Obtaining this genetic information is essential for the breeding programme. This needs to employ a variety of genetic methods including the use of isozyme markers, DNA techniques and quantitative traits.

1.4 Aims of this thesis

1.4.1 Background and importance of the project

P. caribaea is a well-known fast-growing species native to South and Central America. It is one of the most important plantation species and is widely planted as an exotic species throughout the tropics and subtropics over the world. The species plays an important role in plantation forestry around the world.

Quantitative genetics of this species have been extensively studied by field trials to assess quantitative traits (growth rates, stem form and wood properties), such as provenance trials and progeny tests (Nikles *et al.,* 1978; Matheson *et al.,* 1989; Zheng *et al.,* 1994). Among these studies, the most significant one is the international provenance trials co-ordinated by the CFI (now the OFI), in collaboration with the FAO (Greaves and Kemp, 1978). This study was conducted in many tropical countries and many consecutive research results from the trials were published, mostly in three proceedings of the IUFRO workshops (Nikles *et al.,* 1978; Barnes and Gibson, 1984; Gibson *et al.,* 1989). In general, there are significant genetic

differences among the three varieties in both production and quality traits. The var. *hondurensis* was found to be the most fast-growing variety during the first 10 years after planting in low-land tropical areas and has most variable stem form and branching behaviour. The var. *caribaea* has the lowest growth rates (Greaves and Kemp, 1978). Breeding strategies have also been developed for the species, particularly the var. *hondurensis* in various countries or regions (Barnes and Gibson, 1984; Gibson *et al.*, 1989).

China started to grow *P. caribaea* in the 1960s. Since then the species has showed great potential for industrial purposes in the tropical area in southern China. Top priority has been given to the species in the present national afforestation programme (Zheng *et al.*, 1994; Wang *et al.*, 1995). However, most present plantations of the species in China are derived from early introductions which were based on limited resources. An expectation for this is that the genetic base of these materials might be bottlenecked, a common problem for exotic species. Particularly for the var. *caribaea*, field trials showed little genetic variation for morphological traits while large variation was found within the other two varieties (Pan, 1991). Recent research indicated that this is probably caused by the narrow genetic base of the early introduction of the var. *caribaea* (Wang *et al.*, 1995). It was believed that to investigate the range-wide genetic variation, to acquire more widely based genetic resources and to develop a breeding strategy would be of key importance to the genetic improvement of the species in China.

Financed by the British ODA, an international joint research project involving the CAF, the University of Edinburgh and the Forestry Research Institute of Cuba was set up to carry out basic genetic studies with the var. *caribaea* and var. *bahamensis* and to develop a breeding strategy for the species in China. The project involved an extensive range-wide seed collection from the natural forests (Brodie, 1994), which provided material for this study. Another significant aspect of the project is the development of a breeding strategy for *P. caribaea* that may serve as a good example for breeding other important industrial species, both exotic and indigenous, in China.

1.4.2 Questions addressed

This PhD study consists of a series of separate investigations which employ genetic markers such as isozyme and DNA markers as well as quantitative traits. Different questions can be

addressed by the use of different genetic methods. Each of the investigations addresses some specific questions which are summarised below.

Questions that can be answered with isozyme markers:

Investigation: Genetic diversity, structure and population differentiation

How much genetic diversity and variation is there in each of the varieties?

How is this variation distributed within and among populations?

How is variation distributed geographically?

How genetically differentiated are the varieties?

What are the effects of growing the species as an exotic in terms of genetic diversity and structure?

Investigation: Breeding system

What is the breeding system of the species and does this differ among populations?

Questions addressed with DNA markers:

Investigation: Application of DNA techniques in identification of unknown source

Can the DNA markers be used to identify unknown sources of *P. caribaea*?

Questions addressed with quantitative traits:

Investigation: Quantitative genetic analysis of a combined provenance/progeny trial

What is the extent of the adaptive genetic variation within and among populations (provenances)?

What are the genetic parameters estimated from a field trial for the quantitative traits?

And a general question addressed:

Investigation: Breeding strategy

What is the implication of these results for tree improvement programmes in China?

1.4.3 Objectives

As a major part of the whole tree improvement project, the purpose of the thesis is to answer these questions and hence to provide useful genetic information needed by the breeding programme, with the use of genetic markers such as isozyme and DNA markers, as well as quantitative traits.

2. Genetic structure and patterns of variation of natural populations of two varieties of *P. caribaea* using isozyme markers

2.1 Introduction

2.1.1 Diversity of genetic variation

Genetic variation is somewhat ambiguous if the methods used to detect the genetic variation are not specified. Genetic variation of forest trees can be generally categorised into 3 types: quantitative variation, allozyme variation and molecular (DNA) variation. Quantitative variation is revealed by quantitative traits which are usually assumed to be controlled by many genes and possibly at many loci and influenced by environments. Allozyme variation and DNA variation are associated with allozyme variants and DNA variants, respectively, and both of these variants are normally known as genetic markers, which allow inference of the genotype of individuals (seeds or complete plants) at one or many Mendelian gene loci from their variable phenotypes, i.e., electropherograms of protein, particularly enzymes, or of DNA fragments.

The detection of these types of genetic variation employs distinct methods. The results from these different methods do not necessarily correspond to each other. The allozyme and DNA variation generally represent the selectively neutral traits and the quantitative variation is selectively adaptive. To obtain information on genetic variation for the breeding programme requires the use of genetic marker and quantitative traits as tools to investigate genetic variation. Nonetheless, the need to be able to monitor genetic variation throughout the breeding programme, to ensure that there is not dramatic loss of variation during the phases of introduction, domestication and improvement, also requires the use of genetic markers. Furthermore, the conservation and gene management of the genetic resources of the species need to use the genetic markers to assess, monitor and conserve the genetic diversity.

2.1.2 Use of isozymes as genetic markers

Genetic markers are important tools for forest tree improvement. Genetic markers are usually selectively neutral and therefore the variation revealed by them is selectively neutral variation. They help us to know about the population history and hence to determine populations which

probably passed through genetic bottlenecks, to estimate the degree of isolation between populations in their natural range and to devise appropriate sampling strategies for breeding programmes.

Traditionally, population geneticists have employed unusual morphological phenotypes, proteins, or secondary compounds as genetic markers. Due to the difficulties in inferring genotypes from phenotypes, morphological characters and secondary compounds such as flavonoids and terpenes fail to be good candidates for genetic markers. Environmental influences are other complicating factors in the use of such characters as genetic markers (Crawford, 1983). Feasible genetic markers employed in population genetic studies are isozyme markers and DNA markers because they allow precise inference from phenotype to genotype. Numerous methods for assessing genetic diversity and population structure of tree species using biochemical and molecular markers have been developed during the past decades (Fineschi *et al.*, 1991; Khasa, *et al.*, 1994). Of these methods isozyme analysis is still the least expensive for the study of population genetics.

Isozyme markers are detected by the technique called protein electrophoresis. This method takes advantage of the fact that nondenatured proteins with different net charges migrate at different rates through starch or acrylamide gels (or other supporting media such as cellulose acetate strips) to which an electric current is applied. The charge characteristics stem primarily from the three amino acids with positive side chains (lysine, arginine, and histidine) and the two with negative side chains (aspartic acid and glumatic acid). The net charge of a protein, which varies with the pH of the running conditions, determines the protein's movement toward the anode (positive pole) or cathode (negative pole) in the gel. Protein size and shape also can interact with pore size in the electrophoretic matrix to influence migrational properties. Starch gel electrophoresis is most widely adopted because of low cost, safety and ease of use.

The procedure begins with the extraction of water-soluble proteins from a particular tissue (leaves, buds, roots etc.). The extract from each individual is soaked onto a paper wick, and a number of such wicks are placed side by side along a slit (the origin) cut into the gel. The gel is placed in a buffer tray connected to an electric power supply, and electrophoresis proceeds over several hours. The gel is then removed and horizontally sliced. The gel slices are

incubated with histochemical stains specific for the enzymes under assay. Each stain contains a commercially available substrate for the enzyme, necessary cofactors, and an oxidized salt (usually nitro-blue tetrazolium, NBT). At the position(s) in the gel to which an enzyme has migrated, a reaction is catalysed, the substrate is oxidized and the salt is reduced to a blue precipitate producing visible discrete bands. The band profile is the "zymogram" pattern for the enzyme, and usually can be interpreted in simple genetic terms (Avise, 1994).

The use of isozyme markers is reported here and the following Chapter, and the use of DNA markers and quantitative traits are reported in Chapters 4 and 5, respectively.

2.1.3 Advantages and drawbacks of isozyme markers

Studies of isozyme variation in forest trees have several advantages over other measures (Hamrick, 1989; Hamrick *et al.*, 1991; Schaal *et al.*, 1991). The method is relatively inexpensive and can be applied to most species; most isozyme loci have discrete Mendelian inheritance and are codominant. Therefore allele frequencies can be calculated directly and hence genetic interpretation of natural populations can be easily obtained without concern for environmental effects compared to quantitatively inherited traits; isozyme loci can be analysed in every population or across related species, allowing direct comparison of degree and pattern of the genetic variation between different populations or species; a large number of enzymatic loci can be assayed using small quantities of material; many loci express at all stages of the life cycle; isozymes can be resolved for most forest tree species regardless of habitat, size or longevity; the methods are relatively simple, thus a large amount of samples can be analysed simultaneously.

The major drawbacks of isozyme markers are that they are not random samples of the whole genome because they represent primarily structural genes. Isozyme markers only detect genetic variation encoding proteins, therefore the variation detected by isozyme markers is not precisely the full genetic variation. Sometimes the interpretation of the banding pattern is somewhat ambiguous; and the technique is quite variable from lab to lab. Another problem is that isozyme markers are usually restricted to different numbers and sets of loci for different organisms.

2.1.4 Applications of isozyme markers in forest genetics and tree improvement

Isozyme markers are useful for looking at the levels of variation, the distribution of variation within and among species and populations, evolutionary factors affecting genetic structure such as mating systems, gene flow and selection (Hamrick, 1989). Assessment of allozyme variability has been helpful in inferring historical events that have shaped the genetic structure of populations (Hamrick *et al.*, 1981; Hiebert and Hamrick, 1983; Knowles and Grant, 1985; Kinlock *et al.*, 1986; Ennos, 1992; Wang, 1992), in studies of mating systems in natural and breeding populations and in the assessment of genetic efficiency of seed orchards (Muona *et al.*, 1990). Studies of allozyme markers have also contributed to a better understanding of natural hybridisation and phylogenetic relationships in plants (Wang *et al.*, 1991).

Applications of isozyme markers have contributed significantly to tree breeding programmes (Adams, 1981, 1983; Adams *et al.*, 1988; Cheliak *et al.*, 1987; Friedman and Neal, 1992; Miller *et al.*, 1989; Wheeler and Jech, 1992). Genetic analyses have been conducted with isozyme markers for many tree species, particularly conifer species. Detailed reviews on this were given by Mitton (1983), Yeh (1989) and Hamrick (1989).

Reviews on isozyme analysis revealed a general conclusion that plant species generally maintain relatively high amounts of isozyme variation within populations (Hamrick, 1989). Most tree species have shown moderate to high isozyme variation as expected for an organism which has a long life cycle, large range of distribution, high fecundity, outcrossing mode of reproduction and wind pollination. Hamrick (1989) reviewed over 100 taxa of plant species including forest tree species and found that woody plants, in general, maintain more variation within their populations than herbaceous species. In his review, conifer species have average values of 67.7 percent of polymorphic loci, 2.29 number of alleles per locus and 0.207 gene diversity compared with 36.8 percent, 1.69 and 0.141, respectively, averaged for all plant species.

The distribution of isozyme variation within and among populations is commonly measured with two values, either Wright's (1951) F_{st} or Nei's (1973) G_{st}. F_{st} measures variance in allele frequencies among populations relative to the standardised variance based on mean allele

frequencies ($F_{st} = \sigma^2 / \overline{pq}$). It is calculated on each allele at a locus. G_{st} is the genic diversity due to variation among populations (D_{st}) divided by the total diversity (H_t), i.e., ($G_{st} = D_{st}/ H_t$). These two values are equivalent when they are calculated over loci. The distribution of isozyme variation among populations is largely determined by several evolutionary factors such as selection, effective population size, seed and pollen dispersal and the interactions among these factors (Hamrick, 1989; Yeh, 1989). The life history and ecological characteristics, such as long life cycle, predominant outcrossing, high fecundity and wind pollination, that mean a large proportion of the genetic variation detected by isozyme markers resides within populations of forest trees and only a small proportion of genetic variation is accounted for among populations.

2.1.5 Genetic studies of *P. caribaea*

As described in Chapter 1, *P. caribaea* is one of the most important and widely planted softwoods in the tropics and subtropics. It plays an important role in the world's plantation forestry. This is particularly true in China since recent rapid economic development stimulates the commercial plantation forestry. The var. *caribaea* is one of the most important economic timber species in Cuba, accounting for near a quarter of present total forest area (Brodie, 1994). The var. *bahamensis* is also an important commercial tree species. Major timber and pulp products were obtained from this variety in the Bahamas (Lamb, 1973). The great potential in plantation forestry resulted in extensive commercial exploitation and genetic research on the species over the world.

However, the majority of previous research has focused on quantitative genetic analyses which have been conducted by establishing a series of field trials such as provenance trials, progeny trials, etc. and only the var. *hondurensis* has been intensively studied (Matheson *et al.*, 1989; Zheng *et al.*, 1994; Nikles *et al.*, 1978). These genetic tests, in general, evaluate the adaptive variation but indicating nothing about the population history nor the genetic structure of the population. Studies on population structure using genetic markers have rarely been conducted on *P. caribaea,* particularly var. *caribaea* and var. *bahamensis.* Until recently, very few such studies had been conducted. Matheson *et al.* (1989) studied breeding systems and genetic structure of some Central American Pine populations by use of isozyme markers, but only two populations of var. *bahamensis* and three geographically close populations of var.

hondurensis were assayed in their study. This small number of populations is obviously not sufficient to represent the true genetic variation and distribution within the variety particularly with a highly discontinuous (for var. *bahamensis*) and wide ranging (for var. *hondurensis*) distribution. In addition, var. *caribaea* was not included in their study. Nelson *et al.* (1994) reported chloroplast DNA variation among and within taxonomic varieties of *P. caribaea* and *P. elliottii*, but their study focused on the genetic relationships among taxa and relatively small sample sizes were used (from 2 to 20 individuals for each variety). No details were presented on population differentiation within a variety. In China, although a few local genetic tests have been conducted, the useful and applicable genetic information for a breeding programme is rarely available.

To secure long-term commercial benefit from the species, a breeding strategy is essential for the genetic improvement programme. Central to the breeding programme is knowledge of the pattern and degree of genetic variability within and between populations of each variety. Populations which are fragmented or differentiated will require a different sampling strategy than if they were highly uniform in the introduction. Information about gene flow and breeding system is of critical importance in designing a suitable breeding strategy for a tree improvement programme. Data on population genetic structure provides basic genetic information on *ex situ* use and *in situ* conservation of the genetic resources of the species.

2.1.6 Aims of study

An understanding of population genetics and reproduction biology is essential to the design of sampling, genetic conservation, management, and breeding programmes in tree improvement (Bawa, 1976; Yeh, 1989; Khasa *et al.*, 1994). Given little or no knowledge about the variation and genetic structure of var. *bahamensis* and var. *caribaea,* and the feature of different patterns of natural distribution for the two varieties, how the sampling strategy should be designed, how the natural variation could be utilised, what has been the effect on the genetic structure of introduced or domesticated populations and the need to monitor the genetic changes during the progress of the improvement programme are important issues for the breeding programme. Isozyme markers are the most appropriate tools which can be used to obtain genetic information to answer these questions. Using selectively neutral isozyme markers to investigate the genetic structure and patterns of variation facilitates the

understanding of population genetics of the species and hence is useful for the breeding programme.

In this study, isozyme markers are used to look at genetic diversity and population structure of two varieties (var. *bahamensis* and var. *caribaea*) of *P. caribaea.* The amount of isozyme variation and its distribution between and among populations, and the differences between varieties are investigated. Meanwhile the study is also intended to look at the genetic influence of germplasm transfer (tree introduction) and domestication on genetic structure of the species, and potentially, to look at the possibility of application of isozyme markers in identification of unknown (or suspicious) sources.

2.2 Methods and materials

2.2.1 Seed collection, handling and storage

Two varieties, var. *caribaea* and var. *bahamensis,* were analysed in this study. Seeds of var. *bahamensis* were supplied by the OFI. These seeds were collected in the 1970s and handled and stored by the OFI (Baylis and Barnes, 1989) and used for international trials conducted in many tropical countries including China. Although the materials used for marker study and for quantitative study are not exactly the same, some of the same populations of var. *bahamensis* were used in both studies. Some of these populations were also used in the early field trials in China.

As one of the main activities of the ODA funded UK-China forestry project "Tree improvement and early establishment of *P. caribaea* in China", an extensive seed collection of var. *caribaea* was made from the natural forests in Cuba by the Edinburgh Centre of Tropical Forests (ECTF), the Institute of Ecology and Resource Management (IERM) in Summer, 1994. The seed collection was well planned and fully documented (Brodie, 1994). The overall purpose of this seed collection is to form the base populations for use in China and priority was placed on the collection from a large number of trees from natural forest with a sampling strategy designed to capture as much genetic variation as possible (Gibson, 1993; Brodie, 1994). Seeds of var. *caribaea* were collected from 12 natural populations (Provenances) and one seed orchard population in Cuba (Table 2.1).

Cone sacks collected from the field went through a pre-curing stage for approximately a week since cones from each tree varied in their maturity. Cone sacks were placed in a dark, well ventilated room until the cones in the sacks had dried and green cones were no longer present. After the pre-curing, sacks were placed out in the sun for further drying, to stimulate cone opening and seed release. Seeds then were extracted from opened cones. Seeds were then cleaned and packaged and transported to Britain and stored. A small amount of these seeds were used for electrophoresis. Identification of families/clones was well kept throughout the whole process (Brodie, 1994).

The Chinese sample was provided by the Forestry Institute of CAF (Chinese Academy of Forestry). The only information known is that the seeds were collected from Zhanjiang in the Leizhou peninsula which is one of the major areas producing seeds of *P. caribaea* in China. No details of collection are available.

Table 2.1 Summary of seed collection data of *P. caribaea* var. *caribaea* and *P. caribaea* var. *bahamensis*

Site name	Site code	Lat. (N)	Long. (W)	Alt. (m)	No. of trees
var. *caribaea*					
Cajalbana	CAJ 0101-0120	22°47'	83°27'	300	20
El Buren	BUR 0201-0215	22°45'	83°27'	70	15
Los Palacios*	PAL 0301-0315	22°36'	83°14'	20-30	15
Galalon	GAL 0401-0413	22°41'	83°26'	150-200	13
La Guira	GUI 0501-0510	22°41'	83°21'	80-100	10
La Jagua*	JAG 0601-0618	22°43'	83°38'	200-280	18
Juan Manuel*	MAN 0701-0714	22°38'	83°33'	70-150	14
Quinones	QUN 0801-0804	22°51'	83°10'	150-200	4
Cayajabos	CAY 0901-0906	22°51'	83°55'	150	5
Malas Aguas[(1)]*	SOR 1001-1117	22°41'	83°53'	50	43
Isla de la Juventud*	IDJ 2001-2016	21°25' 21°46' 21°43'	83°00' 83°02' 83°55'	50-100	16
Marbajita[(2)]*	MBJ 2101-2117	22°49'	83°28'	50-70	17
La Guabina	GUB 2201-2205	22°48'	83°46'	150	5
Zhanjiang	China	unknown	unknown	unknown	unknown
var. *bahamensis*					
San Andros	69/88/1-10	24°54'	78°01'	5	10
Adelaide	77/88/1-10	25°00'	72°26'	10	10
East New Providence	78/88/1-10	25°01'	77°24'	5	10
Little Abaco	39/77/1-10	unknown	unknown	unknown	unknown
High Rock	07/78/1-10	unknown	unknown	unknown	unknown
Byfield	Australia	unknown	unknown	unknown	unknown

[(1)]: Seed orchard; [(2)]: managed as seed stands *: Cuban populations of var. *caribaea* analysed

2.2.2 Sample size

Five natural populations and one plantation population (from Australia) of the var. *bahamensis* were analysed (Table 2.1).

For the var. *caribaea*, because the aim of this isozyme study, which is mainly to explore genetic variation at the marker level, differs from the overall purpose of the entire project stated above and the available funding and time for this study are limited, use of all of the populations sampled in the collection is therefore unnecessary. Only those geographically distinct populations were selected for the isozyme analysis. Five of the 12 natural populations, the seed orchard population and the plantation population in China were sampled and electrophoretically assayed.

Each population was represented by 35-52 (except one by only 17) seeds for the var. *bahamensis* and at least 50 seeds for the var. *caribaea* from bulked seedlots. The smaller sample size for var. *bahamensis* is due to the relatively low germination rates of seed.

2.2.3 Electrophoresis

Prior to electrophoresis, seeds were germinated at room temperature until the radicle emerged at least 5 mm long. Sufficient numbers of seeds were prepared for germination to provide enough seeds for electrophoresis. About 100 seeds for each population were usually prepared for germination in this study. Populations with low germination rate were re-prepared with larger numbers of seeds.

Diploid tissue was used for enzyme extraction and starch horizontal gel (11%) was used to separate allozymes. Germinated seeds were excised by removing the seed coat and endosperm and then individually homogenised in Extraction Buffer (Appendix 2.1). Enzyme extracts were absorbed onto filter paper wicks which were loaded onto the starch gel. Electrophoresis was run for 5 hours at 75 mA. Two gel buffer systems were used for different enzyme systems and their compositiona are given in Appendix 2.1.

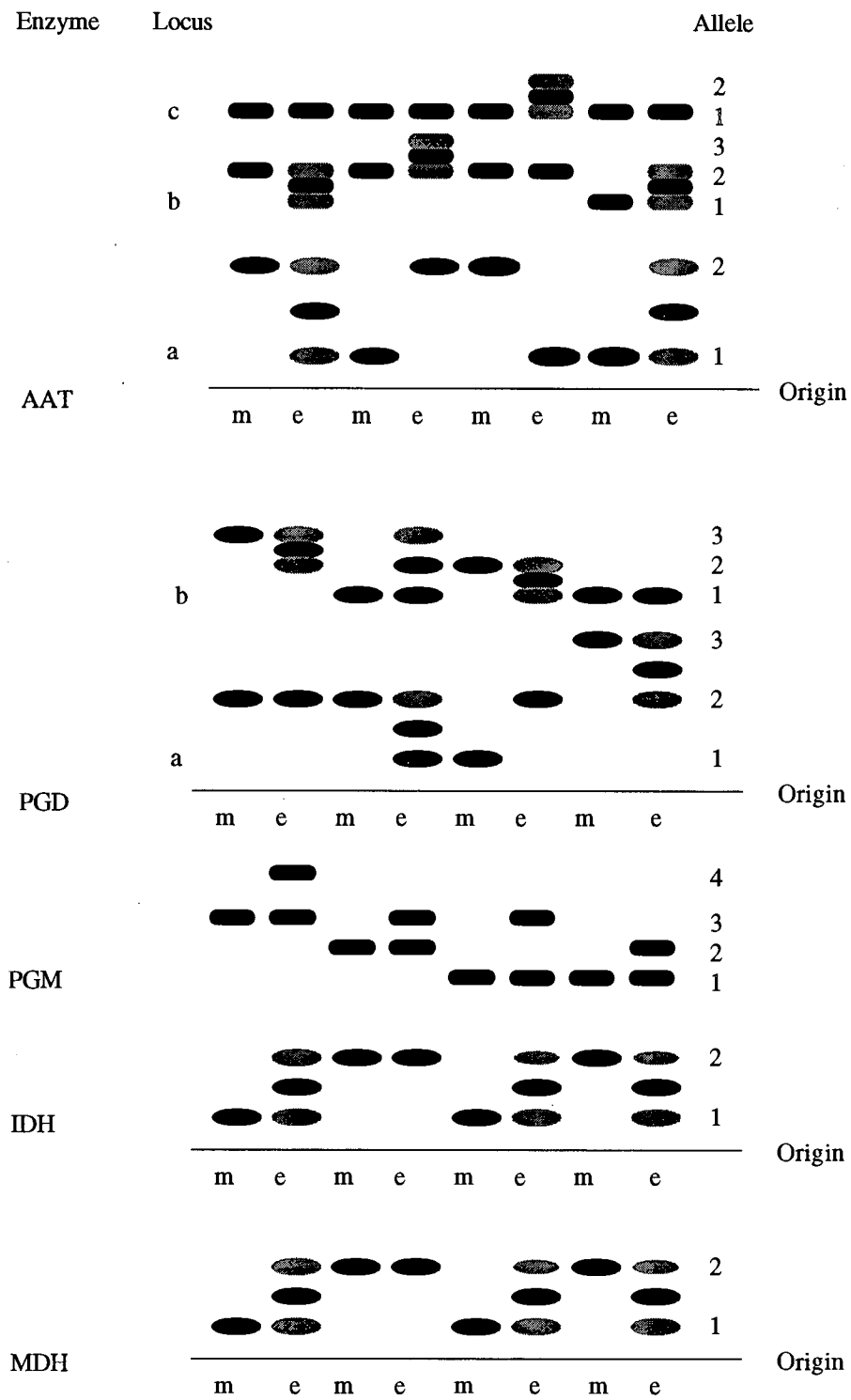

Figure 2.1 Diagrams of AAT, 6PGD, PGM, IDH and MDH (m: megagametophyte; e: embryo). The loci of the same enzyme are designated with a, b, and c from the slowest migrating locus to the fastest one (from origin to anode). Similarly, alleles of the same locus are designated with 1,2,3 and 4. For instance, three loci of AAT were recorded as AAT-a, AAT-b and AAT-c in turn from the origin to anode. PGM and IDH were stained on the same gel. Data on embryos were used in the study of breeding system (See Chapter 3).

Staining follows the method described by Cheliak and Pitel (1984) which was considered to be the most inexpensive way. All the enzyme systems were assayed simultaneously with the same batch of samples. Enzyme loci were chosen if they were constantly clearly resolved. Eight loci of 5 enzymes, AAT (Aspartate aminotransferase, E.C.2.6.1.1), IDH (Isocitrate dehydrogenase, E.C.1.1.1.42), MDH (Malate dehydrogenase, E.C.1.1.1.37), 6PGD (6-phosphogluconate dehydrogenase, E.C.1.1.1.44) and PGM (Phosphoglucomutase, E.C.5.4.2.2), which were clearly resolved and were consistent throughout the assay, were stained and scored for both varieties. The designation of loci and alleles in a gel is given in Figure 2.1.

2.2.4 Data analysis

Bands were recorded for each enzyme locus and the genotypes were designated with numbers which were put into the EXCEL worksheet. Data analysis was carried out with the BIOSYS (Swofford and Selander, 1989), FSTAT (Goudet, 1995) and GENEPOP (Raymond and Rousset, 1995a) packages.

The data analysis was conducted separately for each variety firstly to look at the genetic diversity and population differentiation within each variety and then followed by the analysis combining the two varieties together to look at the genetic relationship between the two varieties and their genetic divergence. All the analyses were based on the 8 loci recorded for both varieties.

2.2.4.1 Tests of hypothesis

- Hardy-Weinberg exact tests

A deviation from Hardy-Weinberg proportions indicates either selection, population mixing or nonrandom mating, and its detection is one of the first steps in the study of population structure. Population structuring and selfing can result in heterozygote deficiency. A Chi-square test is usually used to test such deviations. However, the Chi-square test is suspect in cases where expected frequencies of some genotypic classes are low (Sokal and Rohlf, 1969) and in cases where the sample size is small (Leven, 1949). An alternative is the exact tests which are appropriate even when many rare alleles are present (Guo and Thompson, 1992;

Chakraborty and Zhong, 1994; Rousset and Raymond, 1995; Raymond and Rousset, 1995b). The Hardy-Weinberg test is performed by use of the GENEPOP package (Raymond and Rousset, 1995a).

- Genotypic linkage disequilibrium

This test was also carried out by use of the GENEPOP programs. The programs create all contingency tables of genotypic counts for each possible pair of loci within each population and then perform a probability test (Fisher exact test) for each table using Markov chain methods (Rousset and Raymond, 1995). The null hypothesis (H_o) in the test is 'Genotypes at one locus are independent from genotypes at the other locus'.

2.2.4.2 Measures of genetic variation within populations

The following measures are estimated to describe population genetic diversity and variation. Most of these measures were calculated by the BIOSYS programme.

- Allele frequency (single locus)

It is the basic parameter of population genetics. Assuming H-W equilibrium,

$$p_i = \frac{N_{ii} + 1/2 \sum_{j=1}^{n} N_{ij}}{N}$$

p_i is the frequency of the ith allele (A_i), N_{ii} and N_{ij} are the numbers of A_iA_i and A_iA_j genotypes observed, respectively, N is the total number of individuals analysed. The allele frequencies were estimated with the maximum-likelihood method for all loci and all populations (Hedrick, 1985).

- Percentage of polymorphic loci

This statistic is usually reported in isozyme surveys. It is at best a rough guide to the level of genetic variation in a sample (usually when the most frequent allele has frequency less than 0.95 or 0.99 a locus is considered as 'polymorphic').

$$p = \frac{x}{m}$$

where x is the number of polymorphic loci in a sample of m loci.

- Average number of alleles per locus (n_a)

It has the obvious merit of emphasising one component of diversity, allelic richness.

- Effective number of alleles (Hedrick, 1985)

It is an inverse function of the theoretical homozygosity.

$$n_e = \frac{1}{\sum_{i=1}^{n} p_i^2}$$

where n_e is the effective number of alleles at a locus.

- Genetic diversity (Heterozygosity)

The most widespread measures of genetic variation in a population.

$$H = 1 - \sum_{i=1}^{n} p_i^2$$

$$= \sum_{i=1}^{n} p_i(1-p_i)$$ [polymorphic index (Hamrick and Allard, 1972); or gene diversity (Nei, 1975)].

n is the number of alleles. In contrast with the actual number of alleles per locus which reflects the allelic richness, the estimates of effective number of alleles and the genetic diversity are measures of allelic evenness because the allelic frequencies and the number of alleles are both taken into account.

2.2.4.3 Measures of inter-population genetic variation

- The genetic distance of Nei

$$D_N = -ln(I_N) = -lnJ_{xy} + \tfrac{1}{2} lnJ_x + \tfrac{1}{2} lnJ_y$$

where $I_N = \frac{J_{xy}}{\sqrt{J_x Jy}}$; $J_{xy} = \sum_{i=1}^{n} p_{i.x} p_{i.y}$; $J_x = \sum_{i=1}^{n} p_{i.x}^2$; $J_y = \sum_{i=1}^{n} p_{i.y}^2$ and $p_{i.x}$ and $p_{i.y}$ are the frequencies of the ith allele in populations x and y. For multiple loci, average values are used. UPGMA (Unweighted pair group method) was used to construct the phylogeny trees for each variety and for the two varieties pooled together.

- *F* statistics (Wright, 1965, 1978; Nei, 1977)

F-statistics were devised to measure population subdivision. A subdivided population has three distinct levels of structure: individual organisms (I), subpopulations (S), and the total population (T). The basic formula is:

$$(1 - F_{is})(1 - F_{st}) = (1 - F_{it})$$

Where F_{is} and F_{it} are fixation indices which measure the deviation from H-W proportions within subpopulations and in the total population, respectively. F_{st} measures the genetic differentiation among subpopulations. These can be estimated as:

$$F_{is} = 1 - H_I / \overline{H}_S$$
$$F_{it} = 1 - H_I / H_T$$
$$F_{st} = 1 - \overline{H}_S / H_T$$

where H_I is the observed heterozygosity of an individual averaged over subpopulations; H_s is the expected heterozygosity of an individual in an equivalent random mating subpopulation and is averaged over subpopulations to get $\overline{H}_S$; H_T is the expected heterozygosity of an individual in an equivalent random mating total population which is obtained from the pooled allele frequency in the total population.

For multiple loci, these measures were estimated as weighted averages across all loci. The F_{st} estimated from this formula was also used for the indirect estimation of gene flow among populations.

In addition, a measure of Hamilton's (1971) relatedness was also calculated using Queller and Goodnight's (1988) estimator of relatedness: $R = 2 F_{st} / (1+F_{it})$. This measure estimates the average relatedness of individuals within samples when compared to the whole (Goudet, 1995). These estimates were performed by the FSTAT programs.

- Genic differentiation among populations

This test was performed by using GENEPOP (Raymond and Rousset, 1995a). It tests the allelic distribution of alleles in the various populations. The null hypothesis (H_o) tested is "the allelic distribution is independent across populations" which corresponds to an absence of population differentiation. For each locus an unbiased estimate of the P-value of the

probability test (or Fisher exact test) is performed (Raymond and Rousset, 1995b) for all pairs of populations for all loci.

2.2.4.4 Measures of inter-population gene flow

An indirect method using Wright's formula (1951) was employed to estimate gene flow among populations, which assumes an island model of population structure. It was demonstrated, however, that the method provides reasonable estimates of *Nm* under a wide variety of population structures (Slatkin and Barton, 1989). The formula is:

$Nm=(1-F_{st})/(4F_{st})$,

where N is the effective population size and m is the rate of gene flow. *Nm* estimates the number of migrants per generation.

2.2.4.5 Measures of inbreeding

Levels of inbreeding were estimated by the fixation index $F=1-(H_o / H_e)$ for each allele of each polymorphic locus over all populations, and summarised as weighted means across all loci by F_{is} (Weir and Cockerham, 1984). These values were obtained from the calculations of *F*-statistics. The F_{is} values were also estimated for each population over all loci. The fixation index can be equated with an inbreeding coefficient if it is primarily determined by mating behaviour (Wright, 1969).

2.3 Results

2.3.1 *P. caribaea* var. *caribaea*

2.3.1.1 Banding patterns

The four enzyme systems assayed each showed a wide variety of banding phenotypes (Figure 2.1). These were distinguished by both position and relative intensity of bands. Figure 2.1 shows the putative genotypes associated with each banding pattern of AAT (Dimeric enzyme with 3 loci, 2 alleles for loci a and c, 3 alleles for locus b), PGD (Dimeric enzyme with 2 loci, 3 allele for each locus), PGM (Monomeric with 4 alleles), IDH (Dimeric with 2 alleles) and MDH (Dimeric with 2 alleles, 2 loci). It was proved that a slow-migrating zone with PGM is identical to IDH by comparing the banding pattern of these two enzyme systems, then both IDH and PGM were stained on a same gel to increase the experiment efficiency (Figure 2.1).

The PGD-a locus is normally fixed for the second allele although very occasionally the first and the third alleles were observed. Four alleles of PGM including a rare allele (the 4th allele) were observed. Two loci were observed for MDH, but the slow locus which is not constantly resolved for diploid tissue was unusable for mating system estimation (Figure 2.1).

2.3.1.2 Hardy-Weinberg equilibrium

For Cuban populations, most loci are in H-W equilibrium, but PGD-b was in heterozygote deficit in most populations except MAN and JAG (Appendix 2.2). AAT-a in PAL and IDJ and AAT-b in MAN and JAG, and PGM in MAN are also in disequilibrium. The Chinese population is in disequilibrium for all loci but AAT-a, PGD-a and PGM.

2.3.1.3 Linkage disequilibrium

Within most populations, the majority of loci were independent of each other. For the Cuban populations, only three pairs of loci showed significant linkage disequilibrium in three different populations, i.e., PGD-a and PGD-b in MBJ, AAT-a and MDH in IDJ, and AAT-a and PGM in SOR. However, in the Chinese population, four pairs of loci showed significant linkage disequilibrium, i.e., AAT-a and IDH, IDH and MDH, AAT-a and PGD-a, and IDH and PGD-a. The tables of probabilities are presented in Appendix 2.3. In the global test, three pairs of loci showed significant ($0.01 < P < 0.05$) linkage disequilibrium. They are IDH and PGD-a, AAT-a and PGM, PGD-a and PGD-b.

2.3.1.4 Genetic diversity and variability

2.3.1.4.1 Gene frequencies

All loci showed polymorphism in at least one population (Table 2.2). However, AAT-c and PGD-a were completely monomorphic (fixed with one allele) in most Cuban populations. AAT-c fixed with its allele 1 and PGD-a fixed with its allele 2. AAT-b was highly monomorphic but not as completely as AAT-c. Most loci segregated for two or three alleles, but PGM segregated for four alleles. The least frequent and the fastest-migrating allele of the PGM locus, which was present in all Cuban populations, did not appear in the Chinese population.

Table 2.2 Allele frequencies in populations 1 through 7 of var. *caribaea*

Locus	Allele	PAL (1)	MAN (2)	MBJ (3)	JAG (4)	IDJ (5)	SOR (6)	CHN (7)
AAT-a	*(N)*	*80*	*57*	*102*	*108*	*174*	*63*	*88*
	1	0.3	0.456	0.412	0.301	0.385	0.413	0.409
	2	0.7	0.544	0.588	0.699	0.615	0.587	0.591
AAT-b	*(N)*	*80*	*57*	*121*	*131*	*177*	*172*	*88*
	1	0	0.018	0	0	0.003	0	0.722
	2	0.988	0.895	1	0.992	0.992	0.994	0.273
	3	0.013	0.088	0	0.008	0.006	0.006	0.006
AAT-c	*(N)*	*80*	*57*	*121*	*131*	*177*	*172*	*88*
	1	0.994	1	1	1	1	1	0.972
	2	0.006	0	0	0	0	0	0.028
IDH-1	*(N)*	*80*	*57*	*102*	*110*	*148*	*269*	*62*
	1	0.169	0.132	0.201	0.145	0.108	0.188	0.355
	2	0.831	0.868	0.799	0.855	0.892	0.812	0.645
MDH-b	*(N)*	*80*	*57*	*127*	*110*	*148*	*249*	*68*
	1	0.1	0.105	0.15	0.036	0.03	0.11	0.125
	2	0.9	0.895	0.85	0.964	0.97	0.89	0.875
PGD-a	*(N)*	*80*	*57*	*126*	*21*	*148*	*160*	*74*
	1	0	0	0.012	0	0	0	0.088
	2	1	1	0.984	1	1	0.994	0.912
	3	0	0	0.004	0	0	0.006	0
PGD-b	*(N)*	*80*	*57*	*126*	*131*	*176*	*340*	*92*
	1	0.313	0.342	0.444	0.431	0.33	0.529	0.554
	2	0.256	0.57	0.23	0.271	0.389	0.404	0.272
	3	0.431	0.088	0.325	0.298	0.281	0.066	0.174
PGM	*(N)*	*80*	*57*	*102*	*110*	*148*	*277*	*68*
	1	0.081	0.088	0.142	0.109	0.108	0.13	0.007
	2	0.387	0.447	0.353	0.541	0.426	0.444	0.279
	3	0.481	0.421	0.471	0.336	0.436	0.406	0.713
	4	0.05	0.044	0.034	0.014	0.03	0.02	0

2.3.1.4.2 Genetic variability within population

The mean number of alleles per locus (n_a) varied very slightly among populations ranging from 2.1 to 2.4 and the overall mean was 2.29. The values of n_a are almost the same (2.3) for Cuban populations except that population JAG has a slightly smaller value of 2.1. The Chinese population has a slightly greater value (2.4). Similarly, the effective numbers of alleles per locus (n_e) for different populations were very close to each other. The largest value of n_e was found in the Chinese population (1.64) and the smallest was found in Cuban population JAG (1.55) (Table 2.3).

Table 2.3 Genetic variability in populations of var. *caribaea* based on 8 loci (standard errors in parentheses)

Population	Mean sample size per locus	Mean no. of alleles per locus	Mean effective no. of alleles / locus	Percentage of loci polymorphic[1]	Mean Heterozygosity	
		n_a	n_e	*P (%)*	*Observed* H_o	*Expected*[2] H_e
1. PAL	80.0 (.0)	2.3 (.3)	1.60	62.5	.236 (.081)	.274 (.094)
2. MAN	57.0 (.0)	2.3 (.4)	1.57	75.0	.237 (.074)	.286 (.086)
3. MBJ	115.9 (4.1)	2.3 (.4)	1.67	62.5	.252 (.082)	.297 (.097)
4. JAG	106.5 (12.8)	2.1 (.4)	1.55	50.0	.223 (.088)	.250 (.096)
5. IDJ	162.0 (5.3)	2.3 (.4)	1.60	50.0	0.212 (.083)	.253 (.101)
6. SOR	212.8 (30.9)	2.3 (.3)	1.57	62.5	.254 (.084)	.274 (.091)
mean		***2.26***	***1.59***	***60.0***	***0.202***	***0.272***
7. China	78.5 (4.2)	2.4 (.2)	1.64	87.5	.273 (.074)	.350 (.065)
Grand Mean		**2.29**	**1.60**	**64.29**	**0.241**	**0.283**

[1] frequency of the most common allele does not exceed 95%.
[2] unbiased estimate (Nei, 1978)

The values for *P* ranged from 50% to 87.5 %. The average percentage of polymorphic loci (at 95% criterion) over Cuban populations was 60%, indicating that 5 out of 8 loci were polymorphic on average. The Chinese population had the highest values of *P* and n_a. The natural population JAG had the smallest values of *P* (50%) and n_a. Obviously, the Chinese population showed greater polymorphisms than the natural populations.

The mean expected gene diversity (H_e) varied from 0.250 in population JAG to 0.350 in the Chinese population. The average for the 6 Cuban populations was 0.272 but 0.283 when including the Chinese population. The average proportion of observed heterozygotes was always lower than that expected under H-W equilibrium in all populations. This tells us that heterozygote deficit exists in the populations as demonstrated in the section of H-W equilibrium tests.

It is obvious that the Chinese population has greater polymorphism and gene diversity than the the Cuban populations. This is due to the higher evenness of allelic frequencies across all loci even fewer alleles were observed at the PGM locus (Table 2.3).

2.3.1.5 Genetic structure and population differentiation

2.3.1.5.1 Genic differentiation among populations

The results showed that the allelic distribution is not always independent in various populations and for different loci, indicating that there are significant differences of allele frequencies among populations (Appendix 2.4). There are no significant differences of allelic frequencies at the AAT-c locus among most populations because of its high fixation. PGD-b has no independent allelic distribution in all populations showing a high population differentiation at the locus. Other loci varied from pair to pair of populations. For all Cuban populations, when paired with the Chinese population, the distributions of nearly all loci (except AAT-c and MDH) were not independent. This implies that the Chinese population significantly differed from the others.

2.3.1.5.2 F statistics and gene flow

The average value of F_{is} over all loci for the natural populations was 0.148, i.e. there was 14.8 % heterozygote deficit relative to H-W proportions within the sub-populations (Table 2.4). The weighted mean of F_{it} over all loci was 0.165 showing 16.5 % heterozygote deficit in the variety as a whole. All loci but the PGD-a locus have positive values for F_{it} and F_{is} indicating heterozygote deficits. The negative value for the PGD-a locus indicates a homozygote deficit at the locus.

The population substructuring (differentiation) was examined by the F_{st} value. An overall F_{st} value of 0.02 indicated about 98 % of the genetic variability resided within populations; in other words, about 2 % of genetic variation was among the populations. Although the overall F_{st} value was not large, the significance test showed it was significantly larger than zero. The overall heterogeneity among populations was also found in the allelic distribution.

The F_{st} value was slightly larger when including the seed orchard (0.024) and became more than three times larger when including the Chinese population (0.076) than the F_{st} value for only the 5 natural populations (0.02) (Table 2.4). This indicates that the Chinese population has differentiated far more than the Cuban populations. The pairwise populations F_{st} estimates also showed the large difference between the Chinese population and the Cuban populations (Table 2.5). Other values of pairwise F_{st} between natural populations were rather small and varied slightly.

Table 2.4 *F*-statistics (excluding Chinese population and seed orchard population, standard error based on 500 bootstraps) for var. *caribaea*

		F_{it}	F_{st}	F_{is}	R
*AAT*_a	Mean	0.179 **	0.008 *	0.172 **	0.014
	s.e.	0.042	0.008	0.041	0.013
*AAT*_b	Mean	0.632 **	0.07 **	0.591 **	0.097
	s.e.	0.241	0.051	0.223	0.066
IDH	Mean	0.075	0.008	0.068	0.015
	s.e.	0.053	0.008	0.047	0.015
MDH	Mean	0.125 *	0.042 **	0.086	0.077
	s.e.	0.06	0.021	0.049	0.035
*PGD*_a	Mean	-0.001	0.012	-0.013	0.024
	s.e.	0.002	0.006	0.006	0.011
*PGD*_b	Mean	0.218 **	0.027 **	0.196 **	0.046
	s.e.	0.047	0.023	0.045	0.036
PGM	Mean	0.118 **	0.01 **	0.11 **	0.018
	s.e.	0.027	0.012	0.029	0.021
Over all loci	***Mean***	***0.165 ** ***	***0.02 ** ***	***0.148 ** ***	***0.035***
	s.e.	***0.032***	***0.008***	***0.027***	***0.013***
Incl. Seed orchard	***Mean***	***0.15 ** ***	***0.024 ** ***	***0.128 ** ***	***0.043***
Incl. all populations	***Mean***	***0.202 ** ***	***0.076 ** ***	***0.136 ** ***	***0.127***

*: 5% significant level; **: 1% significance level

Table 2.5 Pairwise populations F_{st} estimations and gene flow (upper diagram gene flow, lower diagram F_{st}) for populations of var. *caribaea*

	PAL	**MAN**	**MBJ**	**JAG**	**IDJ**	**SOR**
PAL		4.428	29.6	16.386	19.199	4.957
MAN	0.05		5.939	5.18	12.872	17.316
MBJ	0.01	0.04		12.018	14.267	10.973
JAG	0.02	0.05	0.02		20.955	8.957
IDJ	0.01	0.02	0.02	0.01		10.541
SOR	0.05	0.01	0.02	0.03	0.02	
China	0.21	0.19	0.23	0.23	0.21	

The inter-population gene flow for the natural populations of var. *caribaea* was 12.25 (based on F_{st} = 0.02). Pairwise population estimates were also obtained (Table 2.5). These values varied from 4.43 to 29.6 Populations with larger values of F_{st} have smaller estimates of gene flow.

2.3.1.5.3 Genetic distance

Estimates of genetic distance are presented in Table 2.6. Values of *D* varied from 0.004 to 0.132 with a overall mean of 0.042. Populations PAL and MBJ, JAG and IDJ have the smallest genetic distance (0.004), whereas the Chinese population has much greater distances from all other populations. These results were better visualised by UPGMA cluster tree (Figure 2.2).

Table 2.6 Genetic distances (Nei, 1978) among populations of var. *caribaea*

Population	1 PAL	2 MAN	3 MBJ	4 JAG	5 IDJ	6 SOR	
2 MAN, Cuba	0.023						
3 MBJ, Cuba	0.004	0.018					
4 JAG, Cuba	0.005	0.017	0.008				
5 IDJ, Cuba	0.005	0.007	0.006	0.004			
6 SOR, Cuba	0.019	0.006	0.009	0.01	0.009		***Mean***
7 China	0.125	0.119	0.117	0.132	0.127	0.12	***0.042***

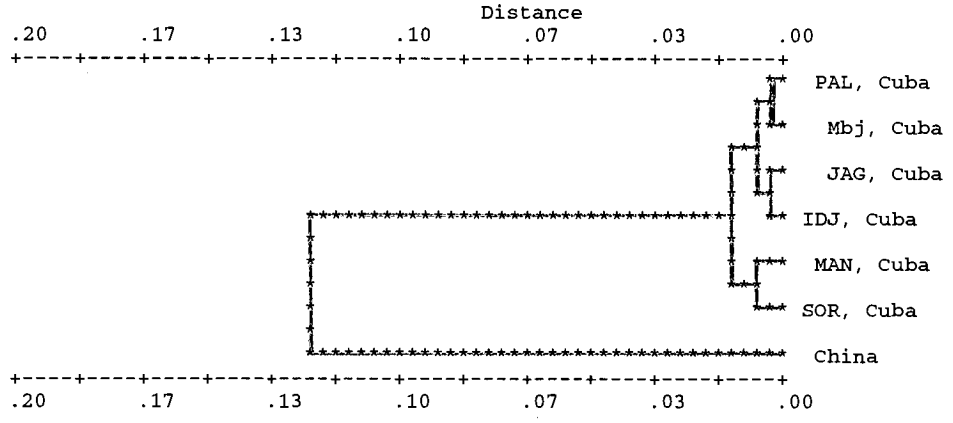

Figure 2.2 UPGMA cluster tree for populations of var. *caribaea*, coefficient used is unbiased genetic distance (Nei, 1978)

The natural populations PAL and MBJ, JAG and IDJ grouped together in the first cycle, then at the second cycle, the above two groups joined together, and MAN and SOR grouped. Then

these two groups joined and finally joined the Chinese population. The great difference between the Chinese population and Cuban populations was clearly shown in the cluster tree.

2.3.1.6 Inbreeding

As shown in Table 2.4, F_{is} values varied across loci with a weighted average of 0.136, showing 13.6 % heterozygote deficit within the sub-populations on average. The mean values of F_{is} were also variable among the populations with a range from 0.09 in the seed orchard population in Cuba to 0.221 in the Chinese population (Table 2.7). The mean value over all Cuban populations was 0.139 with the highest in population MAN and the smallest in SOR (seed orchard). The Chinese population had a much higher F_{is} (0.221) than all Cuban populations did. This means that the variety may have a higher heterozygote deficit in an exotic environment than in its native conditions.

Table 2.7 Fixation index estimated for each population of var. *caribaea* over all loci

Population	PAL (1)	MAN (2)	MBJ (3)	JAG (4)	IDJ (5)	SOR (6)	***mean***	CHN (7)	**Grand mean**
F_{is}	0.139	0.173	0.156	0.107	0.167	0.090	***0.139***	0.221	**0.150**

2.3.2 *P. caribaea* var. *bahamensis*

2.3.2.1 Banding pattern

Isozyme banding patterns of var. *bahamensis* were similar to those of var. *caribaea* as shown in the Figure 2.1.

2.3.2.2 Hardy-Weinberg equilibrium

Most loci showed H-W equilibrium. PGM and PGD-b showed significant heterozygote deficit in the Australian population (Byfield) and in the San Andros population (Appendix 2.5).

2.3.2.3 Linkage disequilibrium

No significant linkage disequilibrium was found among loci for the var. *bahamensis* (Appendix 2.6).

2.3.2.4 Genetic diversity and variability

2.3.2.4.1 Gene frequencies

Allelic frequencies varied among loci and among populations (Table 2.8). PGD-a was completely fixed with the second allele (a middle-migrating allele) in 4 of the 6 populations and the third allele appeared in the other two population as a rare allele (frequencies <0.01). MDH also showed high monomorphism but not as complete as the PGD-a locus. Similar evenness of allele frequencies across populations was observed at most loci except PGD-b and PGM. These two loci have 3 or 4 alleles and the most common or the main allele changed among populations at these loci.

Table 2.8 Allele frequencies in populations 1 through 6 of var. *bahamensis*

		Population					
Locus		**Little Abaco**	**High Rock**	**Byfield, Australia**	**San Andros**	**Adelaide**	**East New Providence**
AAT-a	***(N)***	***25***	***62***	***51***	***52***	***45***	***50***
	1	0.2	0.113	0.206	0.048	0.2	0.06
	2	0.8	0.887	0.794	0.952	0.8	0.94
AAT-b	***(N)***	***25***	***62***	***51***	***52***	***45***	***50***
	1	0	0.024	0	0	0	0.01
	2	0.98	0.968	0.971	0.99	1	0.97
	3	0.02	0.008	0.029	0.01	0	0.02
AAT-c	***(N)***	***25***	***62***	***51***	***52***	***45***	***50***
	1	0.88	0.871	0.922	0.904	0.922	0.98
	2	0.12	0.129	0.078	0.096	0.078	0.02
IDH	***(N)***	***18***	***58***	***51***	***40***	***46***	***49***
	1	0.028	0.172	0.069	0.313	0.261	0.245
	2	0.972	0.828	0.931	0.688	0.739	0.755
MDH	***(N)***	***18***	***29***	***46***	***40***	***34***	***38***
	1	0.056	0.017	0.011	0	0	0.013
	2	0.944	0.983	0.989	1	1	0.987
PGD-a	***(N)***	***18***	***58***	***51***	***40***	***46***	***49***
	1	0	0	0	0	0	0
	2	1	0.991	0.99	1	1	1
	3	0	0.009	0.01	0	0	0
PGD-b	***(N)***	***18***	***58***	***51***	***40***	***46***	***49***
	1	0.25	0.095	0.245	0.112	0.261	0.143
	2	0.222	0.371	0.206	0.463	0.413	0.398
	3	0.528	0.534	0.549	0.425	0.326	0.459
PGM	***(N)***	***18***	***58***	***46***	***40***	***34***	***41***
	1	0.028	0.052	0.217	0.125	0.191	0.378
	2	0.361	0.371	0.283	0.162	0.603	0.524
	3	0.556	0.526	0.446	0.7	0.206	0.061
	4	0.056	0.052	0.054	0.013	0	0.037

2.3.2.4.2 Genetic variability within populations

The mean number of alleles per locus (n_a) ranged from 1.9 in Adelaide to 2.5 in High Rock with an average of 2.27. The Australian population Byfield has a value of 2.4 which is above the average (Table 2.9).

The effective number of alleles per locus (n_e) varied among loci with an overall mean of 1.49, ranging from 1.43 in San Andros to 1.55 in both Adelaide and Byfield. The Australian population Byfield had a value of 1.55 which was larger than the overall mean (Table 2.9).

Table 2.9 Genetic variability in populations (Standard errors in parentheses) of var. *bahamensis* based on 8 loci

Population	Mean sample size per locus	Mean no. of alleles per locus (n_a)	Mean effective no. of alleles per locus (n_e)	Percentage of loci polymorphic* P (%)	Mean heterozygosity	
					Observed H_o	*expected*** H_e
1. Little Abaco	20.6 (1.3)	2.3 (.3)	1.47	62.5	0.238 (.078)	0.243 (.086)
2. High Rock	55.9 (3.9)	2.5 (.3)	1.47	62.5	0.214 (.066)	0.249 (.080)
3. Byfield, Au.	49.8 (.8)	2.4 (.3)	1.55	62.5	0.167 (.054)	0.248 (.093)
4. San Andros	44.5 (2.2)	2.1 (.4)	1.43	50.0	0.173 (.060)	0.225 (.086)
5. Adelaide	42.6 (1.9)	1.9 (.3)	1.55	62.5	0.210 (.071)	0.261 (.094)
6. East New	47.0 (1.7)	2.4 (.3)	1.47	50.0	0.204 (.080)	0.227 (.091)
Mean		**2.27**	**1.49**	**58.33**	**0.201**	**0.242**

** A locus is considered polymorphic if the frequency of the most common allele does not exceed .95*

*** Unbiased estimate (Nei, 1978)*

The values of *P* varied from 50.0 % to 62.5 % with an overall mean of 58.33 %. The *P* value of the Australian population Byfield was 62.5 %, equal to the values of the natural populations except the East New Providence and the San Andros both of which had a value of

50.0 % (Table 2.9). The Australian population did not show a big difference from the natural populations and was above the average level.

Gene diversity at H-W equilibrium for the 6 populations varied from 0.225 to 0.261 with a mean value of 0.242 (Table 2.9). The population San Andros, which had smaller values of P, n_a, and n_e than other populations, had the smallest gene diversity. Similarly, the population Adelaide, which has larger P and n_e, had the largest H_e; however, the n_a value was the smallest. The gene diversity of the Australian population presented an intermediate value but was slightly larger than the overall mean.

2.3.2.5 Genetic structure and population differentiation

2.3.2.5.1 Genic differentiation among populations

Heterogeneity of allele frequencies across populations was significant ($P<0.05$) at many loci except the loci AAT-b, MDH and PGD-a which were highly monomorphic (Appendix 2.7). At the PGM and PGD-b loci, almost all pairs of populations had different allelic distributions except the populations Little Abaco and High Rock. The Australian population Byfield was significantly different compared to natural populations of San Adros, East New Providence and Adelaide, but smaller differences were found when compared to the natural populations of Little Abaco and High Rock. Difference of allelic distribution was only found at the IDH locus for the pair of Little Abaco and High Rock and at PGD-b and PGM for the pair of Little Abaco and Byfield. Adelaide and High Rock only differ significantly at the PGD-b and PGM loci. Other pairs of populations were found to be significantly different at three or more.

2.3.2.5.2 F-statistics and gene flow

The F_{is} values for the natural populations were variable among loci with an average of 0.163 over all loci, indicating 16.3 % heterozygote deficit relative to H-W proportion within a population (Table 2.10). The negative values of F_{is} at AAT-b and MDH indicate homozygote deficit at these loci although the values were very small.

Considering the pooled population as a whole, values of the fixation index F_{it} were generally larger than F_{is} at the various loci surveyed. Like the F_{is}, the negative values of F_{it} at AAT-b, MDH and PGD-a indicate homozygote deficit at these loci although they were small and were

probably not different from zero. The overall mean of F_{it} across loci for the natural populations was 0.227 and implies 22.7 % heterozygote deficit in the variety as a whole population.

Table 2.10 *F*-statistics for each locus over all populations of var. *bahamensis* (Exclude the Australian population. Standard error based on 500 bootstraps)

		F_{it}	F_{st}	F_{is}	R
*AAT*_a	Mean	0.061	0.034 **	0.028	0.065
	s.e.	0.086	0.021	0.084	0.038
*AAT*_b	Mean	-0.014	-0.001	-0.014	-0.001
	s.e.	0.006	0.005	0.004	0.011
*AAT*_c	Mean	0.035	0.014 *	0.021	0.029
	s.e.	0.094	0.02	0.09	0.038
IDH	Mean	0.04	0.021 **	0.019	0.042
	s.e.	0.065	0.028	0.056	0.052
MDH	Mean	-0.012	0.014	-0.026	0.029
	s.e.	0.008	0.022	0.02	0.044
*PGD*_a	Mean	-0.001	-0.008	0.007	-0.015
	s.e.	0.001	0.004	0.003	0.008
*PGD*_b	Mean	0.3 **	0.013 *	0.291	0.02 **
	s.e.	0.03	0.014	0.038	0.022
PGM	Mean	0.317 **	0.183 **	0.166	0.281 **
	s.e.	0.092	0.08	0.096	0.111
Overall	***Mean***	***0.227 *****	***0.078 *****	***0.163*****	***0.13 *****
	s.e.	***0.072***	***0.054***	***0.071***	***0.09***
Incl. all population	***Mean***	***0.236*****	***0.065*****	***0.183*****	***0.106***

*: 5% significance level; **: 1% significance level

Similar variation was found in the values of relatedness. An overall mean value of 0.13 of relatedness indicated that there was about 13 % relatedness between individuals in a natural population.

An overall mean F_{st} value of 0.078 indicated 7.8 % of the total variation was found among the natural populations, in other words, 92.2 % of genetic variation resided within populations. The significance test showed that this value was significantly larger than zero, indicating a strong population substructuring within var. *bahamensis*.

Estimates of F_{st} for all pairs of populations are given in Table 2.11. The largest F_{st} value (0.122) was found between the populations East New Providence and the Little Abaco. Results showed 12.2 % inter-population genetic variation for this pair of populations. Values

for other pairs of populations were variable with the lowest value found between the Australian population and Little Abaco. Although the value was negative and close to zero, what it implies is no inter-population differentiation regarding this pair of populations.

The inter-population gene flow for the natural populations of var. *bahamensis* was 2.96 (based on F_{st} = 0.078) which was much lower than that detected among the natural populations of var. *caribaea* (12.25). Gene flow was also estimated for each pair of populations (Table 2.11). The values varied from pair to pair of populations ranging from 42.39 for the pair of Little Abaco and High Rock to 1.8 for the pair of East New Providence and Little Abaco. Gene flow between the Australian population and the natural populations is meaningless because of the huge geographic distance. It is obvious that a larger amount of gene flow occurred between populations that have smaller F_{st} values.

Table 2.11 Pairwise populations F_{st} estimations and gene flow (below: F_{st} ; above: gene flow) for populations of var. *bahamensis*

	Little Abaco	**High Rock**	**Byfield**	**San Andros**	**Adelaide**	**East New Providence**
Little Abaco		42.393	-	2.790	2.354	1.798
High Rock	0.0059		-	6.677	3.655	2.732
Byfield	-0.0081	0.0139			-	-
San Andros	0.0822	0.0361	0.0799		3.765	2.730
Adelaide	0.0960	0.0640	0.0843	0.0623		17.130
East New Providence	0.1221	0.0838	0.1014	0.0839	0.0144	

2.3.2.5.3 Genetic distance

Table 2.12 shows the estimates of genetic distances calculated over all loci for all possible combinations of pairs of populations. The values of genetic distance ranged from as small as zero for the Little Abaco population and the Australian population Byfield to 0.054 for populations Little Abaco and East New Providence with an overall mean of 0.027. The Australian population Byfield was almost identical to the Little Abaco.

These results are best visualised in Figure 2.3. The UPGMA cluster tree clearly showed the genetic relationships among populations. The most identical populations Little Abaco and Byfield joined together in the first cycle and were then joined by the High Rock, then the San Andros. Adelaide and East New Providence were grouped together and then joined with the

other group. The cluster tree seemed to match the geographic distribution of the populations. The Adelaide and East New Providence are in the same island and they are grouped together in the cluster tree. These results also suggest that the Australian population is probably originated from the Little Abaco.

Table 2.12 Estimates of genetic distance (Nei (1978) unbiased genetic distance) for populations of var. *bahamensis*

Population	1	2	3	4	5	6
1 Little Abaco						
2 High Rock	0.003					
3 Byfield, Australia	0.000	0.006				
4 San Andros	0.035	0.014	0.033			
5 Adelaide	0.044	0.027	0.037	0.030		**mean**
6 East New Providence	0.054	0.036	0.043	0.039	0.006	**0.027**

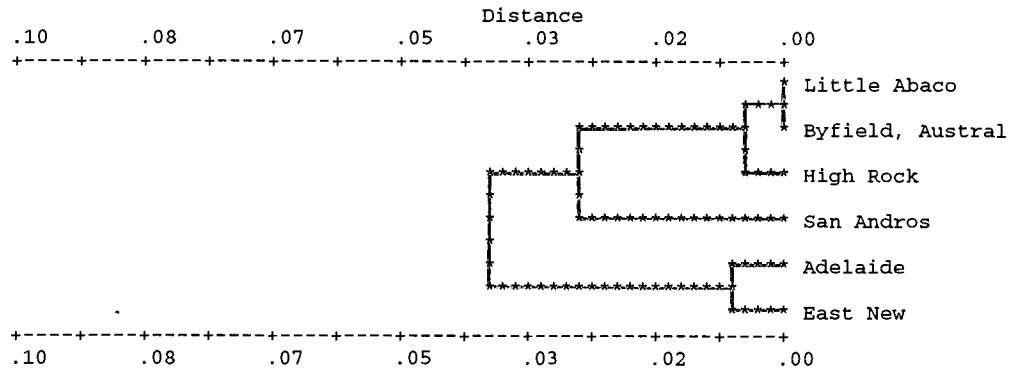

Figure 2.3 Cluster tree using unweighted pair group method (Coefficient used: Nei (1978) unbiased genetic distance) for populations of var. *bahamensis*

2.3.2.6 Inbreeding

Table 2.13 Estimates of fixation index over all loci for each population of var. *bahamensis*

	Little Abaco	High Rock	San Andros	Adelaide	East New Providence	***Mean***	Byfield, Australia	***Grand Mean***
F_{is}	0.002	0.142	0.209	0.196	0.11	***0.164***	0.327	***0.197***

As shown in Table 2.10, the F_{is} values varied across loci with a mean value of 0.183, indicating 18.3 % heterozygote deficit within sub-populations on average. Estimates of F_{is} for each population also showed variable values (Table 2.13). The Australian population Byfield had the largest F_{is} (0.327) and the little Abaco had the smallest (0.002). This was also

observed with var. *caribaea* for which the population grown in exotic conditions presented a much higher fixation index than that grown in the native environments.

2.3.3 Genetic comparison and relationship of the two varieties

2.3.3.1 Genetic diversity and variability within populations

The genetic diversity and variability within populations were generally similar for both varieties. The values of n_a and P were nearly the same for both varieties but were more variable within var. *bahamensis*. The mean value of H_e was larger for var. *caribaea* than for var. *bahamensis*, indicating that the genetic diversity within populations for the var. *caribaea* is somewhat larger than that for var. *bahamensis* (Table 2.14).

Table 2.14 Genetic variability and inbreeding for each population of var. *caribaea* and var. *bahamensis* based on 8 loci

Variety	*Population*	n_a	P	H_e	F_{is}
caribaea	PAL, Cuba	2.3	62.5	0.274	0.139
	MAN, Cuba	2.3	75	0.286	0.173
	Mbj, Cuba	2.3	62.5	0.297	0.156
	JAG, Cuba	2.1	50	0.25	0.107
(island)	IDJ, Cuba	2.3	50	0.253	0.167
	Mean[a]	**2.26**	**60**	**0.272**	**0.148**
(Seed orchard)	SOR, Cuba	2.3	62.5	0.274	0.09
	Mean[b]	***2.26***	***60***	***0.272***	***0.139***
	Chinese	2.4	87.5	0.35	0.221
	Mean[c]	**2.29**	**64.29**	**0.283**	**0.150**
bahamensis	Little Abaco	2.3	62.5	0.243	0.002
	High Rock	2.5	62.5	0.249	0.142
	San Andros	2.1	50	0.225	0.209
	Adelaide	1.9	62.5	0.261	0.196
	East New	2.4	50	0.227	0.11
	Mean[d]	***2.24***	***57.5***	***0.241***	***0.164***
	Australian	2.4	62.5	0.248	0.327
	Mean[e]	**2.27**	**58.33**	**0.242**	**0.197**

a: mean including only natural populations; b: mean excluding the Chinese population; c: mean including all populations; d: mean excluding the Australian population and the Little Abaco population (estimates may be biased by use of small sample size); e: mean including all populations but the Little Abaco population.

2.3.3.2 Inbreeding

The var. *bahamensis* had higher mean values of F_{is}, F_{it}, F_{st} and R than the var. *caribaea*, indicating that the var. *bahamensis* experiences higher inbreeding (Figure 2.4). This is

probably due to its isolated natural distribution, which is mainly in 4 separated islands. The mean value of F_{is} was larger and more variable among populations in var. *bahamensis* than in var. *caribaea* (Table 2.14). The values of F_{is} in the natural populations of var. *caribaea* ranged from 0.107 to 0.173 with an overall mean of 0.148. In contrast, the range of F_{is} values in var. *bahamensis* was from 0.11 (the extremely small value of 0.002 for the little Abaco populations which might be due to the small sample size was excluded) to 0.209 with an overall mean of 0.164 which was larger than that of var. *caribaea.*

2.3.3.3 Population differentiation

2.3.3.3.1 F-statistics

The values of F_{st} were 0.078 and 0.02 in var. *bahamensis* and var. *caribaea,* respectively (Figure 2.4). Both values were significantly ($P < 0.01$) larger than zero indicating significant population differentiation within both varieties. The values of F_{st} within var. *bahamensis* was nearly as large as 4 times of that within var. *caribaea,* indicating that stronger population differentiation occurred in *bahamensis* than in *caribaea.* This may also be attributed to the isolated distribution of var. *bahamensis.*

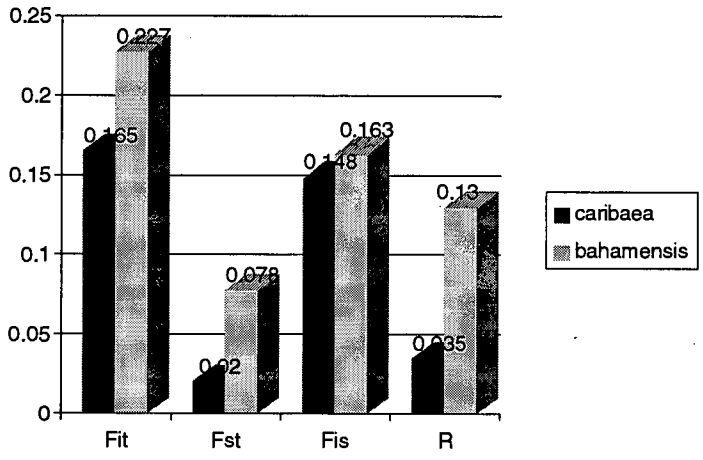

Figure 2.4 Comparisons of *F*-statistics for the two varieties based on 8 loci

2.3.3.3.2 Genetic distance

Genetic distances were calculated by pooling the natural populations together to look at the genetic relationships among the natural and *ex situ* populations and between the two varieties

(Table 2.15). It is clear that the natural and *ex situ* populations of var. *bahamensis* were most similar to each other (smallest distance 0.009), and that the Chinese population was very different from the natural populations. The Chinese population was also very different from the var. *bahamensis*. Visualisation of these difference is presented in the Figure 2.5. The difference between the two varieties is clear although the genetic distance between them is not too far. The UPGMA cluster tree of unbiased genetic distance (Nei, 1978) well matches the taxonomic classification. Chinese material is however an exception.

Table 2.15 Genetic distances (Nei, 1978) among *ex situ* populations and natural populations (as a whole population for each variety) of var. *caribaea* and var. *bahamensis*

Population	1 Cuban var. *caribaea*	2 Chinese var. *caribaea*	3 Natural var. *bahamensis*
1 Cuban var. *caribaea*			
2 Chinese var. *caribaea*	0.122		
3 Natural var. *bahamensis*	0.020	0.155	
4 Australian var. *bahamensis*	0.017	0.153	0.009

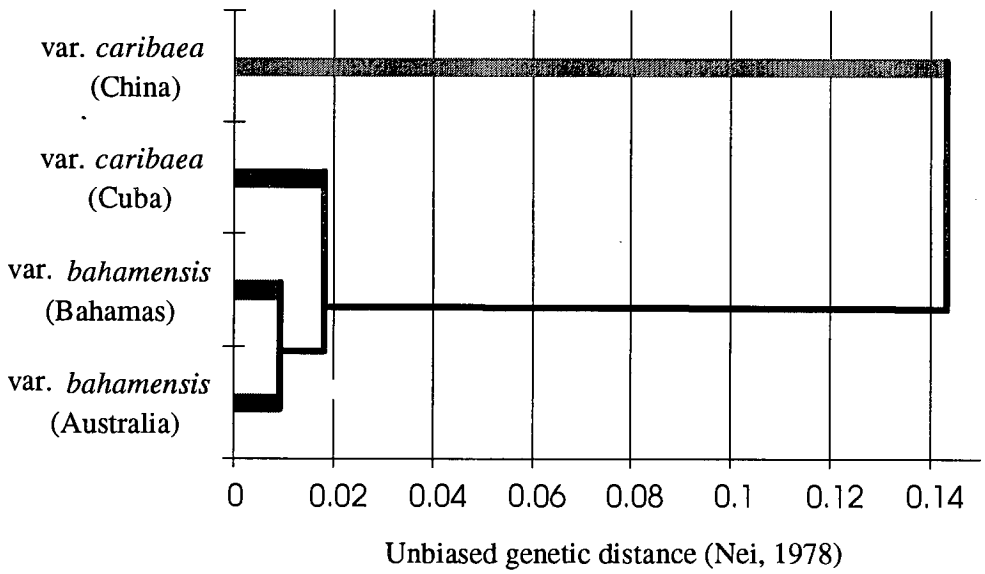

Figure 2.5 UPGMA cluster tree for the natural populations (each variety as a single population) and their *ex situ* populations

UPGMA cluster analysis was also conducted for all populations of the two varieties and the cluster tree is shown in Figure 2.6. It is clear that the genetic differentiation exists between the

two varieties and between the populations within each variety. The populations are more variable within the var. *bahamensis* than within the var. *caribaea*.

The cluster tree showed very good correspondence to the taxonomic classification. All populations of the same variety were clustered into the same group except the Chinese population.

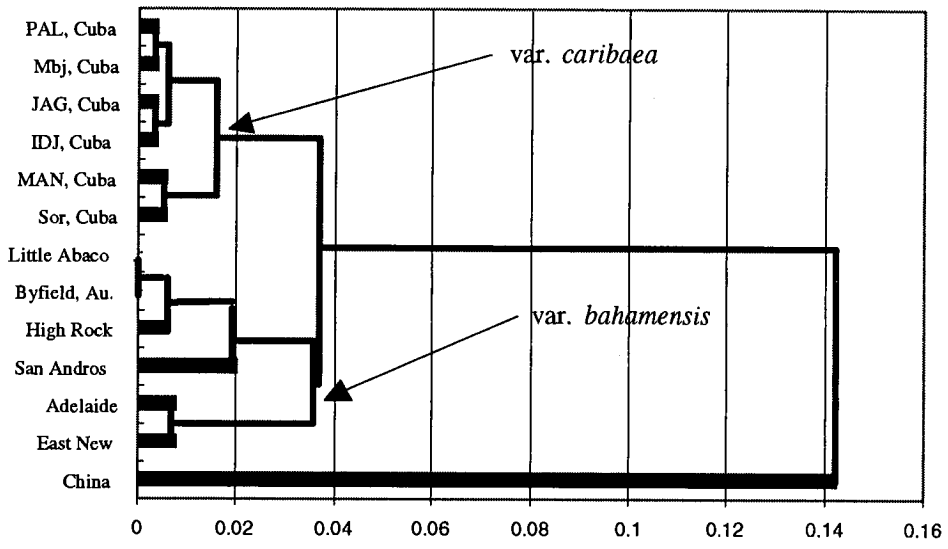

Figure 2.6 UPGMA cluster tree for all populations of the var. *caribaea* and var. *bahamensis*

2.3.3.4 Genetic changes of populations grown *ex situ*

As shown in Table 2.15 and Figure 2.5, the *ex situ* population of the var. *bahamensis* is similar to the natural populations in terms of genetic variability, but experiences a higher heterozygote deficit from the H-W proportion. Similarly, the Chinese material also experienced a higher deviation from H-W proportion with heterozygote deficit compared with other natural populations of the var. *caribaea*. Moreover, the Chinese material differed much from the natural populations in other measures investigated. The genetic distances between the Chinese population and the natural populations have the largest values.

Within the var. *caribaea*, the largest F_{is} (0.221) was found in the Chinese population. Similarly, within the var. *bahamensis*, the F_{is} value of the Australian population (0.327) is much larger than those of the natural populations. These revealed that populations grown in exotic environments had significantly higher F_{is} values.

2.3.3.5 Comparisons of Chinese material with Cuban and Bahaman materials

The genetic distances between the Chinese population and the Cuban natural populations were much greater than those between natural populations. The Chinese population also differed greatly from the populations of var. *bahamensis*. The fourth allele of the PGM locus which appeared in the Cuban populations as a rare allele was not present in the Chinese population. In addition, seeds of the Chinese material look different more or less from those of natural materials of both varieties. They have more variable sizes than the seeds from natural populations. This evidence raises suspicions about the true identity of the Chinese material.

Application of DNA techniques may help to obtain the actual identity of the Chinese material (See Chapter 4).

2.4 Discussion

2.4.1 Hardy-Weinberg equilibrium

In var. *caribaea*, significant deviations from H-W expectations were found for one polymorphic locus PGD-b in most populations. PGD-b had significantly fewer heterozygotes than expected H-W proportions in populations (PAL, MBJ, IDJ, SOR and CHN) in which it was polymorphic ($P < 0.01$). Deviations were also found within the Chinese population at most polymorphic loci ($P < 0.02$). The causes for such observed deficits may be one of several possibilities that are not mutually exclusive. First, a Wahlund effect or population substructuring (the inadvertent sampling across heterogeneous subpopulations), can lead to heterozygote deficiencies (Wahlund, 1928). This explanation is possible given the large spatial scale of the natural populations in the sampling scheme. However, this effect might be influenced by the ability of pollen and seed dispersal of the species. These factors could contribute to a Wahlund effect within the population. Second, experimental error cannot be discounted. Failure to identify heterozygotes will inevitably lead to skewed data. Third, heterozygote deficiencies are common if inbreeding (selfing or consanguineous mating) occurs

in the populations. In such a case, however, deficiencies should affect all loci equally. No heterozygote deficiencies were detected at other loci for most of the populations. Further evidence of no significant inbreeding can be obtained from the investigation on breeding systems (See Chapter 3). However, this explanation seems likely to be true for the Chinese population because deviations were found across loci. A fourth cause is that heterozygotes may be selected against. This alternative is a possibility because the deficiency was observed consistently across populations for the PGD-b locus.

In var. *bahamensis*, no consistent deviation was found across loci and populations, therefore, the second cause, the random experimental errors, seems to be the cause for the deviation from H-W equilibrium.

2.4.2 Linkage disequilibrium

Significant genotypic linkage disequilibrium was found in the Chinese material. One possible cause is mixture of populations with different gametic frequencies. A second possibility is the occurrence of a sufficient intensity of natural selection in favour of certain heterozygous genotypes to overcome the natural tendency for linkage equilibrium. A final cause is random drift, in which some chromosomes may drift to high frequency purely by chance. As the true identity of the Chinese material has been questioned as mentioned earlier, the first cause of linkage disequilibrium seems to be true if the Chinese material is of hybrid origin.

2.4.3 Genetic diversity and variability

Allele frequencies were found to be variable among populations for the two varieties. For some loci, such as PGD-b and PGM, the main allele changed from population to population. Exact tests of genetic differentiation also showed heterogeneity of allelic distributions among populations. Such heterogeneity was mainly due to the more common alleles of some loci, such as AAT-b, AAT-c, PGD-a, PGM for var. *caribaea* and AAT-b, MDH, PGD-a and PGM for var. *bahamensis*, being fixed in one or more of the populations.

The two varieties both contained a moderate to high genetic variability compared to other conifer species. Other pine species showed variable values of genetic variability measures (Ledig, 1986). The values for H_e ranged from 0.0 for *P. torreyana* to 0.362 for *P. taeda* and

the values for *P* varied from 0 % for *P. torreyana* to 100 % for *P. sylvestris, P. nigra, P. palustris, P. rigida* and *P. taeda* (a locus was determined to be polymorphic if any allelic variant was observed). The value of H_e for conifers as a whole is 0.21 (Hamrick *et al.*, 1981). Obviously, comparison of the genetic variability with other close taxa is more informative. Matheson *et al.* (1989) studied genetic variation of some Central American pine populations including *P. oocarpa, P. maximinoi,* and var. *bahamensis* and var. *hondurensis* of *P. caribaea.* Their results, based on 11 loci, showed variable values of n_a, *P* and H_e, ranging from 1.7 (in *P. maximinoi*) to 2.4 (in *P. caribaea* var. *bahamensis*) for n_a, and from 62 % (*P. maximinoi* and *P. oocarpa*) to 77 % (*P. oocarpa*, var. *bahamensis* and var. *hondurensis*) for *P* (5% criteria for percentage of polymorphic loci), and from 0.17 (*P. maximinoi* and *P. oocarpa*) to 0.29 (var. *hondurensis*) for H_e. The values obtained from this study are quite comparable to their results, but the values of expected gene diversity are larger than the average of theirs.

One caution that should be borne in mind is that the polymorphism revealed by the three measures, especially the *P* values (percentage of polymorphic loci) between different studies may not be always comparable, because very frequently, only the loci which showed variation were of interest to the investigator and those with no variation (monomorphic) were therefore considered non-informative, and consequently were not scored during analysis. This strategy is also of some economic interest, since staining for many loci which have no variation is unlikely to be cost-effective and therefore is likely to be inefficient.

2.4.4 Population differentiation and gene flow

Population differentiation was found to be significant within both varieties. The mean values of F_{st} indicated that 2 % of the genetic variation was distributed among the 5 Cuban natural populations of var. *caribaea*, and 7.8 % of the total genetic variation in var. *bahamensis* was among the 5 natural populations. That is to say, the majority of the genetic variation resides within populations. These results agree with data on many other pine species which generally indicated relatively little interpopulational differentiation with values ranging from 2 % for lodgepole pine and jack pine (Dancik and Yeh, 1983) to 12 % for ponderosa pine (O'Malley *et al.*, 1979). The values obtained from this study are low to moderate compared to the above values for other species. They showed no particular evidence to infer any prominent

differentiation process like random genetic drift especially for the var. *caribaea*. The continuous distribution of var. *caribaea* in fairly large stands (although sometimes mixed with other species) must have maintained a large enough effective population size to avoid the effect of genetic drift and some gene exchanges between contiguous populations must have occurred. However, these factors might have less effects on the island population (IDJ) of var. *caribaea* due to the isolation. This may similarly apply to the var. *bahamensis* which has an isolated discontinuous distribution.

The above hypothesis can be also supported by the mean value of gene flow (Nm) (Slatkin, 1987). The mean value of gene flow (based on 8 loci) among Cuban natural populations was 12.25 and was 2.96 among natural populations of var. *bahamensis*. The value for var. *bahamensis* was similar to those of other conifer tree species (Govindaraju, 1988). However, the value for var. *caribaea* was higher than values for other coniferous species. These estimates indicated that gene flow among populations was sufficiently high to prevent major effects of genetic drift on population structure, particularly within var. *caribaea*.

The var. *bahamensis* exhibited higher genetic differentiation than var. *caribaea*. Since the populations of var. *bahamensis* sampled in this study represented different islands of the natural range, the population differentiation detected is probably due to this isolated distribution. The reason for this may be that the genetic exchange between populations in different islands is small compared to a continuous distribution. These results agree with that from our previous research in which significant genetic variation among populations (provenances) was found in biometric traits of the var. *bahamensis* (Zheng *et al.*, 1994), although this concordance may not suggest a necessary association between the isozyme variation (selectively neutral) and the quantitative variation (selectively adaptive). It was also found that a positive correlation exists between genetic and geographic distances. Adjacent populations from the same island were more similar and tended to group into the same group in the UPGMA cluster tree, e.g. the Adelaide population and the East New Providence population from the New Providence island (Figure 2.5), because the gene flow within the island is probably higher than between islands.

For var. *caribaea*, relatively little research has been conducted due to the limitations of seed supply. The Oxford Forestry Institute co-ordinated an international provenance trial in the

1970s and few results are available so far. Pan (1991) reported non-significant provenance variation in growth traits of var. *caribaea*, and he attributed this to the narrow natural distribution of the variety. Wang *et al.* (1995) recently reviewed the provenance trials carried out in China and found that the conclusion of non-significant genetic variation was not well grounded because only two provenances from two adjoining areas were included in the trials. The variation detected might not represent the range-wide genetic variation. In contrast, this study used six natural populations including the population from the Island of Pines, Cuba. It was found that the population differentiation was statistically significant ($P < 0.01$) although it is not as large as that within var. *bahamensis*. The island population was more different from the mainland populations which may be also attributed to the isolation. This is similar to var. *bahamensis*.

2.4.5 Genetic relationships between varieties and phylogenetics

Nikles (1967) studied the genetic variation among the three varieties of *P. caribaea* by analysis of growth, morphology and xylem monoterpenes and found that var. *bahamensis* is very similar to var. *caribaea* while var. *hondurensis* is more distinct. Recently, Nelson *et al.* (1994) studied the cpDNA variation among the varieties and found that var. *caribaea* was most distinct and that the other two varieties were similar to each other. This result is in contrast with the above conclusion by Nikles. Matheson *et al.* (1989) investigated the isozyme variation of some Central American pine populations and found that var. *bahamensis* was quite different from populations of var. *hondurensis* although var. *caribaea* was not included in their study. Although the var. *hondurensis* was not included in our study, the results showed that var. *caribaea* and var. *bahamensis* are genetically similar, tending to agree with the phylogenetic relationships reported by Nikles. Clearly, a complete phylogenetic relationship of the varieties may be made by employing more powerful genetic markers (e.g. DNA markers) and by sampling more populations from each variety.

2.4.6 Inbreeding

The seed orchard population of the var. *caribaea* showed much less deviation from the H-W proportions compared to other natural populations. This may be attributed to the orchard design which aimed at minimising the inbreeding in the reproduction. The materials used in the clonal orchard consist of different origins and the ramets of the same clone have been well

designed in the clonal deployment to reduce inbreeding (Brodie, 1994). Another study on the mating system of the seed orchard, which is included in this thesis, has revealed the predominant outcrossing in the mating system which agrees very well with the results from the present study. The island population of var. *caribaea* showed a higher inbreeding than all natural populations except the MAN population. The population Little Abaco of var. *bahamensis* had an exceptionally low F_{is} value compared to other populations. This might be a biased estimate since the sample size of the population was much smaller (17) than others (around 50).

The *ex situ* populations of both var. *caribaea* and var. *bahamensis* have shown much higher inbreeding than their natural populations. Although a question arose about the actual identity of the Chinese material, the Australian population of var. *bahamensis* showed similar genetic variability to the natural populations. The reason for *ex situ* populations experiencing a higher level of inbreeding needs to be further explored.

2.4.7 Influence of tree introduction and domestication

P. caribaea is widely planted over the world as an exotic plantation species. However, the genetic influence of the introduction and domestication of the species is not known. Wang *et al.* (1995) reported that the seed production of the species is lower in China than other tropical pines and suggested that this be probably related to specific genetic and physio-ecological conditions and that further research into the reproductive biology of the species is required. Burley (1976) pointed out that the low seed production is a common phenomenon in tropical pines, especially when introduced to an exotic environment.

Evidence from our research showed that significant inbreeding occurs in the introduced populations for both varieties. The exotic population showed greater inbreeding than natural populations. When the species was introduced outside its natural range, the trees were exposed to a new environmental condition which might be quite different from their natural habitats. Under such new conditions, some biological and physiological processes of the trees may be affected. Pollen maturing and the process of pollination may be different from the natural condition.

2.4.8 Identification of unknown source

The results from this study showed the great difference between the Chinese material with both var. *caribaea* and var. *bahamensis*. Therefore, it is not surprising that a question would be put forward, i.e. what is the identity of the Chinese material?

There are two possible answers, i.e. mislabelling or hybridisation. The only information known is that the seeds were provided by a local forest agency, no detailed collection information is available about the material. There are two other related variety and species, var. *hondurensis* and *P. elliottii*. They could be mislabelled as var. *caribaea* if the local collector was not well trained in this field.

The other possible reason may be that the mother trees have been hybridised by either the var. *hondurensis* or *P. elliottii* by natural pollination. Hybridisation between them is fairly easy as indicated by some experiments carried out in southern China (Wang *et al.*, 1995). Since the seeds were unlikely to be collected from a single tree, the first answer appears to be more probable.

Certainly, a firm conclusion can not be made only based on the isozyme analysis. Analysis of other traits such as morphological traits and DNA markers is needed. Ideally, to distinguish the *P. elliottii* and *P. caribaea* and their varieties, the best way may be to find unique DNA markers for different species and varieties. However, to our knowledge, no such study has been undertaken. Nelson *et al.* (1994) studied the cpDNA variation of these two species including their varieties. They were not capable of finding a unique marker for *P. elliottii* or for var. *caribaea* of *P. caribaea*, but by comparing the genotypic frequencies of the cpDNA, *P. elliottii* might be identified because it has more than one genotype, while only one genotype was found in var. *caribaea* (Nelson *et al.*, 1994). However, this is not a complete proof and the method needs a larger sample size. Some other markers are needed.

This work is one of the investigations of this thesis and is presented in Chapter 4.

3. Breeding systems

3.1 Introduction

3.1.1 Breeding systems

Knowledge of the breeding system facilitates our understanding of the genetic variation and population structure of species. Mating systems play a crucial role in shaping the genetic composition of populations, because they not only determine genotypic frequencies in subsequent generations but also affect population parameters such as neighbourhood size, gene flow and selection. Many plant populations deviate from assumptions of random mating which are commonly made in theoretical models of genetic transmission (Clegg, 1980). Plants exhibit a variety of mating systems which include self-fertilisation, outcrossing or a mixture of both. The consequences of these different mating systems will markedly affect the structure and evolution of populations (Allard, 1975; Clegg, 1980; Hedrick, 1990). Loveless and Hamrick (1984) found that the distribution of allozyme variation within and among populations was strongly controlled by the mating systems of the plants.

In forestry practices, knowledge of the mating system is of importance for tree breeders. Significant inbreeding depression in the form of decreased survival and growth can be displayed by selfed progeny of most coniferous forest species (Franklin, 1970; Sorensen and Miles, 1982). In addition, the presence of inbreeding violates some of the assumptions made in analysis of wind-pollinated progeny tests, leading to overestimation of additive genetic variance and genetic gain from selection (Namkoong, 1966; Squillace, 1974). Estimation of mating system is a primary concern in breeding and gene conservation programmes(Muona, 1989), and plays a central part of any study of selection in plant populations (Ennos, 1981). To obtain quantitative measures of the forces (selection) that control changes in gene and genotype frequencies in populations, precise estimates of outcrossing rate (t) vs self-fertilisation rate (s=1-t) are required.

3.1.2 Quantifying mating systems

Traditionally, mating systems of plant populations have been studied using observations on floral morphological traits and the behaviour of pollinators, and by controlled crosses.

However, information derived from such studies is insufficient and cannot be used for quantitative estimates of mating system parameters.

Isozyme markers provide a valuable tool for studies of mating systems of forest trees because: 1) populations are often polymorphic for many enzyme loci, thus providing the investigator with a large set of markers in virtually any species; 2) allozymes are often codominant so that all genotypic classes can be identified directly from their allozyme phenotypes without progeny testing (Shaw *et al.*, 1981). Allelic isozymes (allozymes) have much facilitated the determination of how gametes carrying particular gene markers are brought together in both natural and experimental plant populations (Allard, *et al.*, 1975). The use of electrophoretically detectable genetic markers and the development of appropriate genetic models have greatly facilitated detailed quantitative studies of plant mating systems (Brown *et al.*, 1985).

During the last decade, considerable attention has been paid to the development of analytical procedures to evaluate better existing electrophoretical data (Wheeler and Jech, 1992). Such procedures include: 1) rare marker: tracking of pollen distribution from a single clone possessing a rare allele; 2) paternity analysis; 3) mating system models. Of these, the mathematical models are most commonly used.

In conifers, allozymes are particularly useful because of their high polymorphism and the presence of two types of seed tissues: the haploid megagametophyte tissue (maternally derived) and the diploid embryo tissue. Census of the megagametophytes, which are identical to the maternal gametic contributions to the embryos, allows easy inference of the maternal genotype of each progeny array (Tigerstedt, 1973; Ritland and El-Kassaby, 1985). Therefore, the paternal contribution at a given locus can be inferred by comparing the isozyme phenotype of the megagametophyte to that of the corresponding embryo. Progeny testing is then unnecessary.

3.1.3 Mating systems in seed orchards

Seed orchards have become a predominant source for the production of conifer seed for the increasing demands for genetically improved seed. The breeding systems of conifers in seed

orchards have drawn increasing scientific and managerial attention (Weir and Zobel, 1975; Adams and Joly, 1980). Electrophoresis has been used extensively to investigate issues such as clonal identification, pollen contamination levels, mating systems, supplemental mass pollination (SMP) verification, patterns of gene flow and levels of pollen competition in seed orchards. In particular, mating system analysis has been of the greatest interest to population geneticists. Studies of outcross fertilisation patterns within seed orchards have been the most recent focus of mating system studies.

Allozyme segregation patterns in progeny arrays have been used to estimate the mating system in both natural stands and seed orchards of coniferous trees (Morgante *et al.*, 1991; Matheson *et al.*, 1989; Rudin and Lindgren, 1977; Adams and Joly, 1980; Moran *et al.*, 1980; Shen *et al.*, 1981; Muller-Stark, 1982; Shaw and Allard, 1982; Ritland and El-Kassaby, 1985; Paule *et al.*, 1993). The low rates of selfing found in seed orchards and the lack of difference between single-locus and multilocus estimates of selfing (Shaw and Allard, 1982) have indicated that orchard design has been successful in preventing consanguineous matings (Ritland and El-Kassaby, 1985).

3.1.4 Mating systems in *P. caribaea*

An earlier study on natural populations of the two varieties var. *hondruensis* and var. *bahamensis* of *P. caribaea* obtained slightly low values of outcrossing rates compared to other conifer species, which varied considerably among populations and species and averaged around 90% (Matheson *et al.*, 1989). The explanations for this may be as follows. First, there may be true selfing in a greater amount of the natural populations sampled in their study although the breeding systems and floral morphology of the species appear to be similar to other pine species which exhibit high outcrossing rates (Wheeler and Jech, 1992). Secondly, there could be less inbreeding depression in these taxa, thus producing a higher proportion of selfed progeny in seed collections. Finally the populations may be genetically substructured; for example, biparental inbreeding generates effective selfing in the progeny.

3.1.5 Aims of study

In Chapter 2, the isozyme variation and its distribution within and among populations was analysed. Such distribution of genetic variation is influenced by several factors including

mating system, pollen and seed dispersal, and the adaptive response of species to selective forces (Hamrick, 1989). To understand how evolutionary forces combine to produce such a population genetic structure of *P. caribaea*, quantitative estimates of these factors are required. Isozyme loci provide a useful tool as genetic markers to investigate such influences. In this Chapter, a study of the mating system of var. *caribaea* is reported. The purpose of this study is to estimate the outcrossing rate in both natural and artificial populations and to explain their evolutionary influence on the population structure. The implication of mating systems for seed production is also examined. Questions that can be answered in this study are as follows. Are there differences of mating systems between natural populations? Are there differences between natural and seed orchard populations? Is there inbreeding and what are the practical implications in terms of seed production?

3.2 Materials and methods

3.2.1 Seed material

In this study, three natural populations and one seed orchard population were sampled (Table 3.1). Two of the three natural populations, JAG and MBJ, were located on the Cuban mainland and the other, IDJ was located on an island (Isle of Youth). The seed orchard population, SOR was located on the mainland. The details of the seed collection for natural populations have been described in the previous Chapter and Brodie (1994).

Table 3.1 Summary of seed collection data of the populations of *P. caribaea* var. *caribaea*

Site name	Site code	Lat. (N)	Long. (W)	Alt. (m)	No. of trees
Isla de la Juventud[1]	IDJ 2001-2016	21°25' 21°46' 21°43'	83°00' 83°02' 83°55'	50-100	16
La jagua	JAG 0601-0618	22°43'	83°38'	200-280	18
Marbajita[2]	MBJ 2101-2117	22°49'	83°28'	50-70	17
Malas Aguas[3]	SOR 1001-1117	22°41'	83°53'	50	43

[1]: Island population; [2]: managed as seed stands; [3]: clonal seed orchard

The JAG population is a mixed pine forest. *P. caribaea* var. *caribaea* dominated on the better soils, typically the ridge tops and valley bottoms. Associated vegetation included *Byrsonina crassifolia, Sorghastrum stipoides* and *Paspalum* spp. The MBJ poulation was managed as seed stand. *P. caribaea* var. *caribaea* is almost pure in the upper canopy. Samples of the IDJ

population were collected from various sites including scattered stands on the island. The forest type on the island is mixed pine and broad-leaved forest, dominated by *P. caribaea* and *P. tropicalis*. Sites had been heavily degraded by clear felling and selective logging. The absence of experienced tree climbers also produced a bias on poor tree selection during cone collection (Brodie, 1994).

For the seed orchard population, seeds were collected from a clonal seed orchard in Malas Aguas, Cuba. The site of the seed orchard was isolated from natural stands and plantations of *P. caribaea*. The nearest plantation of the same species is 5 km to the south of the seed orchard. The predominant wind direction is north west and the probability of pollen contamination was assumed to be low (Brodie, 1994). The orchard consists of 109 clones of plus trees selected from four provenances in Pinar del Rio Province and Topes de Collantes, with a majority of plus trees (68 out of 109) from the seed stands MBJ.

Forty-three clones out of 109 in the orchard were sampled and most of them were represented by 3 or 4 ramets. Eleven clones are represented by only one ramet due to insufficient seed production of those clones (Appendix 3.1). Seeds of individual clones were kept separate and their identity was retained. The planting design of the seed orchard is random placement of single tree plots. There are approximately 200 blocks, covering an area of 185 ha. The ramets in each block were placed randomly with the restriction that ramets of the same clone were placed at an interval at least 30 m apart (Brodie, 1994). There exists a one-row boundary of *Eucalyptus citriodora* as a buffer zone for the seed orchard. The strategy of sampling clones in the seed orchard is that clones which had performed above the progeny average in volume production were selected for seed collection based on the data from progeny tests.

Methods used for seed handling and storage are detailed in Chapter 2.

3.2.2 Sample size

For the seed orchard population, 43 individual clones were sampled (Appendix 3.1). For the natural populations, 16 families for IDJ, 18 families for JAG and 17 families for MBJ were sampled, respectively. For each of the individual trees sampled, at least 6 seeds were analysed

electrophoretically. The exact number of seeds assayed varied slightly across loci due to some unrecordable banding for some samples (Table 3.2 & Table 3.6).

3.2.3 Electrophoresis

The same method of electrophoresis as that described in Chapter 2 was followed in this study. Unlike the study reported in Chapter 2 which used only diploid embryos of seeds, the haploid material was also used in order to infer maternal genotype. The haploid endosperm and diploid embryo from at least 6 seeds of each family were individually homogenised in the extraction buffer and loaded onto adjacent positions on the gel (i.e., haploid, diploid, haploid.... and so on) for easy comparison.

For the seed orchard population, four enzymes PGM, IDH, MDH and PGD-b, which were clearly resolved and consistent throughout the assay, were chosen for analysing the mating system. In addition to the above four loci, AAT-a was recorded for the three natural populations. Sufficient numbers of seeds were prepared for germination to provide enough seeds for electrophoresis. About 20 seeds for each family were usually prepared for germination in this study. Very few families having extreme low germination rate were re-prepared with large numbers of seeds.

3.2.4 Data analysis

Isozyme phenotypes were recorded for both the haploid endosperm and the diploid embryo of each seed assayed. Mendelian inheritance and lack of linkage were assumed as for many other conifer species (Adams and Joly, 1980). Maternal genotype was inferred at each of the four or five loci from the haploid megagametophytes of the progeny array which are genetically identical to the maternal parent. The probability of correct identification of maternal genotype from n is $1-(0.5)^{n-1}$ (Tigerstedt, 1973; Cheliak *et al.*, 1983). In this study, at least 6 megagametophytes were scored per family giving a probability of 0.97 of correct identification of maternal heterozygotes which is sufficiently high.

Data on progeny-genotype arrays from the individual families were used to estimate pollen allele frequencies and outcrossing parameters for each population by a joint maximum likelihood method (Ritland, 1986, 1990). Because of the restriction in the computer

programme, the 4th allele, the least frequent allele, of PGM was combined with the allele having the nearest mobility.

Estimates of single-locus outcrossing rate (t_s) and multilocus outcrossing rate (t_m) in each population and the pollen allele frequencies (p) were computed with the computer programs written by Ritland (1990, 1995). The programs use the Maximum Likelihood method for the multilocus estimation. The program also gave estimates of multilocus outcrossing rate and outcrossing-pollen gene frequencies for individual trees (t_{mi}) in the sample. At an individual level, the estimated outcrossing rate is the female one and is not necessarily equal to the male outcrossing rate (Ross and Gregorius, 1983; Ritland and El-Kassaby, 1985). The (minimum variance) average single-locus inbreeding coefficient of maternal parents (F) was also estimated by the program. The method makes the normal assumptions of the mixed mating model.

Variances of estimates were obtained by conducting 100 to 500 bootstraps in which resamplings were performed between families. Family estimates were conducted by two methods; 1) fixing the probability of pollen gene frequency (p), 2) allowing pollen gene frequency to vary among families.

Regression of outcrossing-pollen allele frequency on the additive value of the ovule genotype was conducted to obtain a direct measure of effective selfing caused by mating with relatives (Ritland, 1985). Additive values (0, 0.5, 1.0) were assigned to the maternal genotypes (aa, Aa, AA), respectively, where A is the most common allele and a is the class of other alleles at the same locus. The assigned additive value was the independent variable and the inferred outcrossing-pollen allele frequency for each individual tree was the dependent variable. Regressions were obtained for each locus and averaged over loci.

3.3 Results

3.3.1 Banding pattern

Banding patterns of the 5 loci chosen for this study were detailed in Chapter 2 (Figure 2.1). Generally, the bands of diploid embryos are relatively weaker than those of the endosperms. It

was found that the intensity of bands of diploid tissue are slightly stronger when the embryos grew longer, but it did not give much improvement for some loci.

3.3.2 Estimation of outcrossing rates in a clonal seed orchard

3.3.2.1 Gene frequencies in pollen and ovule pools

The multilocus estimates of allele frequencies in the pollen pool (for outcrossing pollen parents) are compared with the corresponding maternal parent frequencies at each locus in Table 3.2. Difference between the allele frequencies of pollen and ovule pools was tested by *t*-test at all loci, and no significant difference was revealed. Differences ranged from 0.004 for the 3rd allele of PGD-b to 0.066 for alleles of IDH. Sufficient polymorphism at these loci for estimation of inbreeding parameters is apparent.

3.3.2.2 Population estimates of outcrossing rates

Single-locus outcrossing estimates (t_s) (Table 3.2) were examined by comparing the bounds of confidence intervals to see the departure from the complete outcrossing (t=1). All single-locus estimates showed non-significant departure from complete outcrossing although the outcrossing rate at the PGM locus was apparently lower than at other loci.

Heterogeneity tests of single-locus estimates of outcrossing rates showed no significant differences among loci. The estimates ranged from 0.90 at PGM to 1.007 at MDH. The minimum variance mean of single-locus outcrossing rates (t_s = 0.941) and the arithmetic mean (t_s = 0.962) differ from the multilocus estimate of outcrossing rate (t_m = 0.985) by 4.4% and 2.3%, respectively. The single-locus estimates suggest 4% to 6% selfing while the high value of the multilocus outcrossing rate suggests a very low level of selfing (1.5%) in the orchard.

The spatial homogeneity in pollen pool allele frequencies over maternal trees was examined by means of a Chi-square test (Table 3.2). All loci showed no evidence to reject the null hypothesis assumed in the model, i.e., the pollen was equally distributed among individual maternal trees. There was no significant ($P > 0.05$) difference between multilocus and mean single-locus outcrossing rates.

Table 3.2 Gene frequency estimates in ovule and pollen pools, single-locus and multilocus estimates of outcrossing rate [Multilocus population estimates were conducted on 4 loci, 43 clones, 326 individuals. Standard errors (in parentheses) based upon 500 bootstraps. t^b is arithmetic mean; t^c is minimum variance mean; t_m is multilocus outcrossing rate].

Locus/allele		Size	Ovule pool	Pollen pool	Diff.[A]	χ^{2} [C]	t_s
IDH	1	259	0.163 (.037)	0.229 (.026)	-0.066 ns	3.02	0.977 (.079)
	2		0.837 (.037)	0.771 (.026)	-0.066 ns		
PGD-b	1	247	0.500 (.036)	0.488 (.045)	0.012 ns	16.14	0.964 (.106)
	2		0.279 (.049)	0.295 (.031)	-0.016 ns		
	3		0.221 (.049)	0.217 (.032)	0.004 ns		
PGM	1	275	0.140 (.034)	0.119 (.030)	0.021 ns	10.41	0.900 (.088)
	2		0.465 (.049)	0.446 (.030)	0.019 ns		
	3		0.395 (.050)	0.435 (.040)	-0.040 ns		
MDH	1	241	0.070 (.026)	0.131 (.035)	-0.061 ns	1.41	1.007 (.254)
	2		0.930 (.026)	0.869 (.035)	0.061 ns		
t^b							0.962 (.132)
t^c							0.941 (.053)
t_m		326					0.985 (.050)
$t_m - t^c$							0.044 (.028)
χ^2 of t_s [B]							0.000 (.000)

ns: Non-significant

[A] *Values of t-test for the difference between ovule and pollen pool allele frequencies were calculated as:*

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_{\bar{x}_1 - \bar{x}_2}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{2S_p^2}{n}}}$$

where n is the number of bootstraps used for the estimation of standard error, S_p^2 is the pooled variance for the differences and calculated as (when df for both estimates are equal):

$$S_p^2 = \frac{df_1\sigma_1^2 + df_2\sigma_2^2}{df_1 + df_2} = \tfrac{1}{2}(\sigma_1^2 + \sigma_2^2)$$

where σ_1^2 and σ_2^2 are the variances of the allele frequencies in ovule and pollen pool respectively.

[B]*: The heterogeneity test of single-locus outcrossing rates is given by:*

$$\chi^2_{k-1} = \sum_{i=1}^{k} I_i(t_i - t^c)^2$$

with k-1 degree of freedom; the information measure (I_i) is the inverse of the variance (V_i) for the ith estimate of t (Kahler et al., 1984).

[C]*: The df of the χ^2 is 1 and 9 for a diallelic and a triallelic locus respectively plus an approximation of 1/n (n is the number of loci used in multilocus estimation) (Ritland, 1991).*

3.3.2.3 Estimates of outcrossing rates for individual clones

Estimates of outcrossing rate for the 43 clones assayed are given in Table 3.3. These values were calculated in two different ways: 1) estimation of t_m only, keeping p (outcrossing pollen gene frequency) constant at population estimates, i.e. assuming that allele frequencies in the pollen pool are the same for each maternal parent; 2) joint estimation of t_m and p, which can only be performed with multilocus data (Ritland and El-Kassaby, 1985). The mean estimate of outcrossing rate using the first method was smaller than that given by the second method (Table 3.3). The values of female outcrossing rates obtained showed a wide fluctuation among clones for both methods, ranging from 0.41 to as large as 2. The mean estimates of both methods were greater than the population multilocus estimates. The standard errors of the multilocus estimates for individual clones are much higher than those of population estimates.

Table 3.3 Multilocus estimates of outcrossing rate for individual clones, 6 seeds of each clone were used (Standard errors based on 500 bootstraps).

	Fix p		Joint			Fix p		Joint			Fix p		Joint	
Clone	t_m	s.e.	t_m	s.e.	Clone	t_m	s.e.	t_m	s.e.	Clone	t_m	s.e.	t_m	s.e.
1001	0.9	0.39	2	0.07	1025	0.86	0.37	1.86	0.3	1067	0.82	0.33	0.82	0.5
1002	1.24	0.37	1.47	0.31	1027	1.06	0.53	0.55	0.64	1068	0.87	0.5	0.71	0.66
1005	0.98	0.53	1.54	0.28	1029	2	0	2	0	1070	0.85	0.4	0.67	0.45
1007	1.25	0.45	1.14	0.37	1030	0.21	0.17	0.97	0.59	1071	2	0	2	0
1008	0.97	0.58	0.7	0.65	1032	0.6	0.59	2	0	1072	1.32	0.41	1.37	0.3
1011	2	0	2	0	1038	0.56	0.3	2	0	1078	0.82	0.38	0.79	0.61
1012	1.03	0.41	1.96	0.36	1040	1.37	0.42	2	0	1083	0.93	0.71	1.57	0.27
1014	0.83	0.5	1.66	0.47	1044	2	0.24	1.47	0.26	1091	0.41	0.36	0.57	0.72
1015	1.13	0.21	1.41	0.26	1048	0.6	0.33	2	0.03	1100	1.31	0.5	2	0
1016	1.39	0.4	2	0	1050	0.86	0.4	0.68	0.48	1105	0.88	0.42	0.96	0.33
1018	1.18	0.49	1.17	0.39	1051	0.98	0.29	1.53	0.28	1106	0.76	0.48	2	0
1021	0.9	0.47	0.51	0.47	1052	0.82	0.31	1.78	0.29	1110	1.15	0.43	1.18	0.39
1022	0.93	0.41	2	0.19	1055	0.8	0.29	1.44	0.35	1117	2	0	2	0
1023	1.01	0.51	1.57	0.27	1056	0.66	0.4	2	0					
1024	2	0	2	0	1060	2	0	2	0	**Mean**	**1.09**		**1.49**	

The distributions of individual outcrossing rates obtained from the two methods (fix p and joint t, p) are compared in Table 3.4. It seems that keeping the pollen allele frequency p constant among parent trees was more realistic than allowing it to vary, since most individual outcrossing rates were distributed around 1 whereas many joint estimates departed a long way from 1.

Table 3.4 Distribution of estimates for female outcrossing rate of individual clones, for intervals with upper limit

keep p			*Joint t, p*		
Group limit	*Frequency*	Percentage	*Group limit*	*Frequency*	Percentage
0.25	1	2.33%	0.25	0	0.00%
0.50	1	2.33%	0.50	0	0.00%
0.75	4	9.30%	0.75	7	16.28%
1.00	18	41.86%	1.00	4	9.30%
1.25	8	18.60%	1.25	3	6.98%
1.50	4	9.30%	1.50	4	9.30%
>1.50	7	16.28%	>1.50	25	58.14%
Mean t_m	**1.09**		**Mean t_m**	**1.49**	

The regression coefficients of outcrossing-pollen allele frequency on ovule genotype varied widely among loci (Table 3.5), ranging from -0.199 at PGD-b in the SOR population to 0.212 at IDH in MBJ. The regression was found to be significant only at PGD-b in the SOR population.

Table 3.5 Regression of pollen allele frequency on maternal genotype for each locus (showing regression coefficients and probabilities).

	IDJ		JAG		MBJ		SOR	
Locus	*b*	*P*	*b*	*P*	*b*	*P*	*b*	*P*
AAT	0.006	0.466	0.012	0.912	0.076	0.584		
IDH	-0.006	0.339	-0.009	0.961	0.212	0.058	-0.089	0.294
MDH	0.002	0.738	0.041	0.850	0.036	0.707	-0.183	0.184
PGD-b	0.003	0.560	0.145	0.332	-0.116	0.330	-0.199	0.045*
PGM	0.002	0.752	0.022	0.868	0.053	0.455	-0.025	0.712
Mean	***0.001***		***0.042***		***0.052***		***-0.124***	

b: regression coefficient; P: significance probability of regression
**: Significant at 5 % level.*

3.3.3 Breeding system in natural populations

3.3.3.1 Gene frequencies in pollen and ovule pools

The multilocus estimates of allele frequencies in the outcrossing pollen pool are compared with the corresponding frequencies in the ovule pool for three natural populations in Table 3.6. The difference of allele frequencies between ovule and pollen pools was checked by *t*-test. For most of the loci no significant differences were detected. Significant difference was only found at the PGM locus in populations IDJ and MBJ.

Table 3.6 Multilocus estimates of allele frequencies in pollen and ovule pools. (Standard errors based upon 100 bootstraps)

Population	Locus	Allele	n	Pollen		Ovule		Difference/
				freq.	s.e.	freq.	s.e.	Significance
	AAT-a	1	149	0.323	0.053	0.500	0.085	0.177
		2		0.677	0.053	0.500	0.085	-0.177
	IDH	1	149	0.128	0.033	0.094	0.049	-0.034
		2		0.872	0.033	0.906	0.049	0.034
IDJ	MDH	1	149	0.039	0.022	0.063	0.038	0.024
		2		0.961	0.022	0.938	0.038	-0.023
	PGD-b	1	149	0.315	0.044	0.313	0.053	-0.002
		2		0.429	0.055	0.469	0.080	0.040
		3		0.256	0.058	0.219	0.081	-0.037
	PGM	1	149	0.208	0.047	0.031	0.025	-0.177*
		2		0.447	0.069	0.500	0.077	0.053
		3		0.345	0.103	0.469	0.071	0.124
	AAT-a	1	98	0.343	0.055	0.278	0.086	-0.065
		2		0.657	0.055	0.722	0.086	0.065
	IDH	1	110	0.191	0.030	0.111	0.049	-0.080
		2		0.809	0.030	0.889	0.049	0.080
JAG	MDH	1	110	0.029	0.014	0.028	0.021	-0.001
		2		0.971	0.014	0.972	0.021	0.001
	PGD-b	1	110	0.461	0.054	0.444	0.073	-0.017
		2		0.272	0.050	0.222	0.060	-0.050
		3		0.267	0.051	0.333	0.065	0.066
	PGM	1	110	0.100	0.024	0.139	0.051	0.039
		2		0.515	0.043	0.583	0.083	0.068
		3		0.385	0.045	0.278	0.081	-0.107
	AAT-a	1	84	0.453	0.079	0.441	0.087	-0.012
		2		0.547	0.079	0.559	0.087	0.012
	IDH	1	110	0.197	0.041	0.206	0.058	0.009
		2		0.803	0.041	0.794	0.058	-0.009
MBJ	MDH	1	110	0.116	0.034	0.206	0.077	0.090
		2		0.884	0.034	0.794	0.077	-0.090
	PGD-b	1	110	0.481	0.053	0.471	0.084	-0.010
		2		0.166	0.054	0.206	0.052	0.040
		3		0.353	0.042	0.324	0.074	-0.029
	PGM	1	110	0.032	0.014	0.206	0.087	0.174*
		2		0.431	0.066	0.265	0.073	-0.166
		3		0.538	0.070	0.529	0.089	-0.009

*: Significant at 5 % level.

3.3.3.2 Population estimates of outcrossing rates

Population estimates of single and multilocus outcrossing rates are presented in Table 3.7. In the IDJ population, the single-locus estimates of outcrossing rate ranged from 65.4 % at PGD-b to 84.9 % at IDH, from 86 % at AAT-a to complete outcross at PGM in JAG, and from 67.8 % at AAT-a to complete outcross at PGM in MBJ. The heterogeneity tests between the single-locus estimates showed no significant difference in any of the three populations. The single-locus values, the means and the multilocus value were lower in the island population (IDJ) than in the two mainland populations.

Table 3.7 Single-locus and multilocus estimates of outcrossing rate [Multilocus population estimates based on 5 loci, standard errors based on 500 bootstraps. t^b is arithmetic mean; t^c is minimum variance mean; t_m is multilocus outcrossing rate].

Locus	IDJ	s.e.	χ^2	JAG	s.e.	χ^2	MBJ	s.e.	χ^2
AAT-a	0.798	0.124	9.36*	0.86	0.181	6.84*	0.678*	0.133	7.18*
IDH	0.849	0.173	0.79	0.942	0.277	0.52	0.827	0.169	3.55
PGD-b	0.654*	0.116	13.44	0.888	0.111	1.33	0.919	0.097	8.70
PGM	0.762*	0.095	26.21*	1.049	0.097	19.56*	1.004	0.223	4.55
MDH	-			-			0.959	0.138	10.13
t^b	0.766*	0.127		0.935	0.167		0.877	0.152	
t^c	0.798*	0.052		0.954	0.066		0.907	0.065	
t_m	0.894*	0.038		0.984	0.046		1.009	0.086	
$t_m - t^c$	0.096	0.02		0.029	0.058		0.102	0.049	
χ^2 of t_s	0.0709			0.0792			0.1406		

- maximum likelihood estimation did not converge due to the monomorphism at this locus.
** Significant departure from complete outcrossing (t=1.0) at 5 % level.*

In the island population, single-locus values at PGD-b and PGM and the multilocus estimate were significantly different from complete outcrossing (t=1.0). Significant departure was found at AAT-a in the MBJ population and no significant departure was found in the JAG population. The significantly lower outcrossing rates indicate that the island population practices more inbreeding than the other two natural populations.

The spatial homogeneity in pollen pool allele frequencies over maternal trees was examined by means of Chi-square tests (Table 3.7). The null hypothesis of homogeneity of pollen allele frequencies was rejected at the PGM locus in the IDJ and JAG populations and at AAT-a in all three populations. All other loci showed no evidence to reject the null hypothesis. The

difference between the single and multi-locus estimates was examined by *t* test and no significant difference was found in any of the three populations.

3.3.3.3 Estimates for individual families

Multilocus estimates of *t* (outcrossing rate) and *p* (pollen allele frequency) were calculated for each individual tree (Table 3.8). The estimates of *t* are female outcrossing rates. The mean estimates of *t* were greater than the population estimates of *t* as given in Table 3.7. Mean values of *t* in all three populations were greater than the unity. The mean estimate of *t* using fixed *p* was smaller than that given by allowing *p* to vary among trees. In the same way as for population estimates, the mean value of *t* for the island population was smaller than that for the other two continental populations.

Table 3.8 Multilocus family estimates: 1)Pollen gene frequencies constrained to population values; 2) *t* and *p* jointly estimated. (Standard errors based upon 500 bootstraps).

IDJ Population						JAG Population					
		Fix p		*Joint p & t*				*Fix p*		*Joint p & t*	
Family	n	*t*	(SE)	*t*	(SE)	Family	n	*t*	(SE)	*t*	(SE)
2001	6	0.87	0.40	0.83	0.59	601	8	1.48	0.32	1.77	0.31
2002	6	2.00	0.42	2.00	0.00	602	6	0.39	0.39	0.19	0.18
2003	9	1.49	0.48	2.00	0.19	603	6	0.96	0.35	0.98	0.55
2004	9	0.94	0.36	0.88	0.51	604	6	2.00	0.11	1.94	0.24
2005	10	0.92	0.30	0.94	0.57	605	6	1.10	0.49	0.87	0.59
2006	10	0.81	0.26	1.22	0.64	606	6	0.87	0.57	2.00	0.03
2007	10	0.85	0.22	0.96	0.62	607	6	0.77	0.39	0.72	0.46
2008	10	0.54	0.27	0.76	0.50	608	6	0.93	0.55	0.86	0.59
2009	10	1.02	0.33	2.00	0.00	609	6	1.12	0.43	1.34	0.37
2010	9	0.89	0.40	0.84	0.42	610	6	2.00	0.00	2.00	0.00
2011	10	0.81	0.42	0.80	0.59	611	6	0.85	0.40	2.00	0.20
2012	10	0.85	0.20	0.76	0.53	612	6	1.05	0.51	1.16	0.50
2013	10	1.14	0.35	1.66	0.54	613	6	2.00	0.00	2.00	0.00
2014	10	1.09	0.48	1.02	0.50	614	6	0.72	0.32	0.69	0.52
2015	10	1.02	0.35	0.90	0.54	615	6	0.79	0.43	0.84	0.56
2016	10	1.37	0.34	0.92	0.55	616	6	1.11	0.52	2.00	0.24
						617	6	2.00	0.00	2.00	0.00
						618	6	2.00	0.33	2.00	0.22
Mean		**1.04**	**0.35**	**1.16**	**0.46**			**1.23**	**0.34**	**1.41**	**0.31**

(Continued)

MBJ population					
		Fix p		*Joint p & t*	
Family	n	*t*	(SE)	*t*	(SE)
2101	6	2.00	0.00	2.00	0.00
2102	6	0.61	0.27	0.55	0.40
2103	6	0.75	0.31	2.00	0.38
2104	6	1.57	0.37	2.00	0.00
2105	6	1.01	0.52	1.02	0.55
2106	6	2.00	0.00	2.00	0.00
2107	6	0.89	0.39	2.00	0.38
2108	6	0.46	0.28	2.00	0.53
2109	6	1.20	0.47	1.41	0.34
2110	6	2.00	0.00	2.00	0.00
2111	6	2.00	0.00	2.00	0.00
2112	6	0.90	0.40	0.75	0.34
2113	6	2.00	0.00	2.00	0.00
2114	6	2.00	0.00	2.00	0.00
2115	6	2.00	0.06	1.90	0.30
2116	6	0.88	0.44	0.72	0.39
2117	6	0.64	0.27	0.56	0.29
Mean		**1.35**	**0.22**	**1.58**	**0.23**

There was wide among-tree fluctuation of estimates of female outcrossing rate, ranging from 0.39 to 2. However, the large values of estimates of t obtained by allowing p to vary among trees indicate that the case may not be the true situation in the JAG and MBJ populations.

The regressions of outcrossing-pollen allele frequency on ovule genotype varied widely among loci (Table 3.5). Significant regression was found in no case for the three natural populations over all loci.

Table 3.9 Mean observed fixation indices of parental population (F) and equilibrium fixation index (F_e). [Standard errors based on 500 bootstraps]

	Population								
	IDJ		**JAG**		**MBJ**		**SOR**		***Mean***
	Fixation index	s.e.	Fixation index	s.e.	Fixation index	s.e.	Fixation index	s.e.	
F	-0.18	0.129	-0.161	0.14	-0.089	0.109	-0.41	0.133	***-0.21***
F_e	0.056		0.0081		-0.0045		0.006		***0.0164***

Wright's (1965) fixation index (F) for the parental generation and the expected fixation index at inbreeding equilibrium ($F_e=(1-t_m)/(1+t_m)$), i.e. when all H-W equilibrium assumptions apart from random mating are met, were estimated in Table 3.9. Fixation indices for the maternal parents ranged from -0.089 in MBJ to -0.41 for the seed orchard with an overall

mean of -0.21. All of these values were negative and were not significantly different from zero. The equilibrium fixation indices were greater, with a mean of 0.016. All maternal F values were not significantly different from the corresponding F_e values.

3.4 Discussion

3.4.1 Population estimates of outcrossing rate

The multilocus estimates of allele frequencies in the pollen pool were statistically similar to that in the ovule pool at most of the loci assayed. Only at the PGM locus in IDJ and MBJ populations were the differences significant. This could be caused by differential pollen production among adult plants (Müller-Stark and Ziehe, 1984; Schoen and Stewart, 1986; Muona, 1989), by selection at the gametophytic level (Apist *et al.*, 1989; Morgante, *et al.*, 1991), or by non-random pollination which may have resulted from restricted pollen dispersal in the population (Ennos and Clegg, 1982; Brown *et al.*, 1985). Gene flow into the population may also contribute to this (Godt and Hamrick, 1995). However, the heterogeneity Chi-square test did not detect any significant difference between single-locus estimates in any of the four populations. This rejects the possibility of restricted pollen dispersal among trees within the populations. The non-significant heterogeneity in single-locus outcrossing rate of the populations indicates non-significant departures from the assumptions of the mixed-mating model.

In all populations, the estimated single-locus outcrossing rates were always lower than the multilocus estimates, but not significantly so. Single-locus estimates of outcrossing rate are known to be lowered by any form of inbreeding in addition to selfing, such as consanguineous mating due to family substructuring of the populations (Shaw and Allard, 1982; Ennos and Clegg, 1982; Ellstrand and Foster, 1983). It is very common in many species that the single-locus estimates of outcrossing rate vary substantially over loci (Shaw *et al.*, 1981). However, alleles at all loci are transmitted in the same gametes and outcrossing rates must therefore be the same for all loci. The observed variability may reflect the large variances that are common with single-locus estimation. Random effects may contribute to the heterogeneity commonly observed in experimental studies (Shaw *et al.*, 1981). Another possible explanation for the

disparate estimates is the invalidity for some loci of one or more of the assumptions made in formulating the mixed-mating model.

Multilocus estimates, on the contrary, are much more robust to any violation of the mixed-mating model assumptions (Morgante *et al.*, 1991). Information on the amount of inbreeding other than selfing therefore can be obtained by comparison of single- and multi- locus measures (Ritland and Jain, 1981). The close correspondence of these two estimates should indicate true self-fertilisation rather than consanguineous matings (Shaw and Allard, 1982). The difference between the weighted (minimum variance) average of single-locus and the multilocus estimate ranged from 0.044 (Table 3.2) in the seed orchard population to 0.102 (Table 3.7) in the MBJ population. The differences between single and multilocus estimates of outcrossing rate were not statistically significant for any population, which indicates that estimated selfing caused by biparental inbreeding (mating between related individuals) was negligible. The smallest difference was found in the seed orchard and indicates that the detected 1.5 % selfing is probably caused by true selfing rather than biparental inbreeding or consanguineous mating. The evidence of biparental inbreeding was somehow stronger in the IDJ and MBJ populations. The explanation for this may be that in the seed orchard, the possible factors which are causing inbreeding depression, such as substructuring or family clustering which often exist in natural populations may be reduced by the deployment of clones in the orchard which has been carefully designed to minimise the inbreeding depression. The wider spacing and random placement of ramets in the orchard should allow freer air movements and mixing of pollen and thus promote outbreeding.

However, it should be noted that multilocus estimates of outcrossing rate can also be biased downwards when a small number of loci is used (Ritland and Jain, 1981). Since this is the case in this study, two other methods were used to determine whether and to what extent the estimates of outcrossing have been biased downwards by the violations of the assumptions of the mixed-mating model.

Firstly, the hypothesis of homogeneity of the pollen pool allele frequencies over the maternal trees was examined by the Chi-square test, which compares the number of genotypes of the progeny array for each maternal genotype. The local heterogeneity in the pollen pool may result from spatial substructuring of the population and restricted pollen dispersal, and

therefore can also be related to consanguineous matings, or tree to tree variation in the outcrossing rate (Brown *et al.*, 1975). In this study, significant heterogeneity of pollen allele frequency was only found at the PGM locus in the IDJ and JAG populations and at the AAT-a in all three natural populations. The cause for this is unlikely to be the tree to tree variation since all loci should be affected in the same way if it is the case (Knowles *et al.*, 1987). If family substructuring is the cause, there should be a correlation between the presence of the heterogeneity and the departure from the random mating at each locus. However, the outcrossing rates at the PGM locus were found to be the highest in population JAG although a smaller value was observed in population IDJ. It therefore can be concluded that spatial heterogeneity of the pollen pool allele frequencies does not contribute significantly to the lowering of the outcrossing rate, at least in the JAG population. The cause for this heterogeneity is therefore probably the random effects of the sampling.

A second method for directly estimating the effective selfing caused by consanguineous matings, is the regression of outcrossing-pollen allele frequency (obtained from the joint estimation of t_m and p for individual trees) on the additive value of the ovule genotype (maternal genotype inferred from its megagametophyte segregation pattern). According to Ritland (1985) the amount of selfing caused by consanguineous matings is best estimated by the expected slope of this regression line. The regression showed non-significant regression coefficients for all loci for the three natural populations and was only significant at the PGD-b locus with a negative value for the seed orchard population. Thus the regression of pollen allele frequencies on maternal genotypes also reveals that, with the exception of the above mentioned locus, mating with relatives is negligible. This therefore implies that the selfing detected is probably true selfing. A significant amount of true selfing exists in the island population.

In general, results from this study suggested that *P. caribaea* var. *caribaea*, as a taxonomic unit, has a high outcrossing rate which does not statistically significantly differ from the complete outcrossing (t=1.0). The exception is the island population (IDJ). The outcrossing estimates for the island population (IDJ) were significantly lower than the two mainland natural populations and the seed orchard population in which outcrossing estimates did not significantly differ from the unity (t=1.0). This indicates that the island population experiences higher inbreeding than other populations, which supports the similar results reported in the

previous Chapter. The estimated selfing rates in the IDJ population were 10.6 % and 20.2 % for multilocus and single-locus estimates respectively. In contrast, the MBJ population, which was heavily managed as a seed stand, and the seed orchard have the highest values of outcrossing rates. The natural population JAG has estimates of outcrossing rates similar to that of the two managed populations (seed orchard and JAG).

Many studies of mating systems with tropical tree species have revealed that many tropical trees are adapted for outcrossing (Loveless, 1992). Outcrossing rates in conifer seed orchards appear to be greater, on average, than 0.9, ranging from 0.84 to 0.99 (Adams and Birkes, 1989). Studies also have revealed that outcrossing rates are higher in seed orchards than in natural populations of the same species (Shaw and Allard, 1982; Rudin *et al.*, 1986; Muona and Harju, 1989) or even the same clones (Szmidt, 1987). Findings from the present study agree with the above generalisations. Although no study on mating system of the var. *caribaea* has been reported yet, multilocus estimates of outcrossing rates obtained from an earlier study on the other two varieties of the same species ranged from 15 % to 7 % selfing for var. *bahamensis* and from 11 % to 8 % for var. *hondurensis*, based on data assayed for 11 loci (Matheson *et al.*, 1989). These estimates fall between the values obtained from the present study, slightly higher than the results for the seed orchard population and the mainland natural populations but slightly lower than the estimates for the island population.

Given the evidence that, as discussed above, there is no significant violation of the assumptions of the mixed-mating model on which all these estimates are based, and the sampling (seed collection) scheme (Brodie, 1994), the most probable explanation for the above findings from this study is that the management history of the populations has important effects on the breeding system in the population. The very low true selfing rate of the seed orchard and the heavily managed population MBJ may be attributed to the genetic management such as genetic thinning and silvicultural measures which may effectively promote pollen movement and avoid family substructuring in the population. The higher selfing rate in the island population may be due to the low density of the stands, e.g. scattered individuals resulted from selective logging (Brodie, 1994) which increased the probability of self-pollination. Such effects of density on selfing rate have also been found in other tree species (Knowles *et al.*, 1987)

3.4.2 Outcrossing rate for individual trees

As far as the multilocus estimates of t and p obtained for individual trees are concerned, it must be remembered that the t values are the female outcrossing estimates and are not necessarily equal to the male outcrossing rates (Horovitz and Harding, 1972; Gregorius *et al.*, 1987). The estimates of female t varied greatly among individual trees in all populations. Such a large variation of outcrossing estimates among individual clones in seed orchards has often been observed and may range from 0.5 to 1.0 (Shaw and Allard, 1982; Omi and Adams, 1986; Moran *et al.*, 1980; Erickson, 1987).

Large variation of outcrossing among individual families in the population indicates either significant departure from mixed-mating model assumptions or sampling errors due to the small sampling size. Possible reasons for the violation of assumptions are restricted and non-random outcross pollen dispersal, and true differences in outcrossing rate among trees (Brown, *et al.*, 1985). The individual estimates of female outcrossing which obviously exceed 1 may not be biologically realistic, unless negative assortative mating is involved (Cheliak *et al.*, 1983). However, they should provide reasonable estimate of average outcrossing rate over all trees (Ritland and El-Kassaby, 1985). As the violation of assumptions was ruled out by the population estimates as discussed above, the variation is not due to true biological variation between trees in individual outcrossing rates but to the sampling error attributable to the small progeny array. In our study a small progeny array size of 6 was used and the family estimates of outcrossing rates based on this small number are probably biased, although the number for inferring maternal genotype is statistically sufficient. The much greater standard errors of the female outcrossing estimates of individual clones (Tables 3.3 & 3.8) than that of population estimates also reveal statistical inefficiency of the small size of progeny arrays. This would be improved by simply increasing the size of progeny arrays.

Comparison of the distribution of female outcrossing rates (Tables 3.4 & Table 3.8) obtained from the two methods suggests that the first method is more realistic because its overall mean estimate is close to 1 and there are fewer clones with outcrossing rate larger than 1. This further supports random mating and rejects restricted pollen dispersal in the seed orchard.

Information on tree to tree variation in outcrossing rate can facilitate practical seed orchard management. The orchard manager can tailor harvest regimes or supplement mass pollination

(SMP) activities based on estimates of self-fertility and the mating system parameters. Studies have also been used to evaluate orchard management techniques such as top pruning (Omi and Adams, 1986), bloom delay (El-Kassaby *et al.*, 1986), flower stimulation (Wheeler and Jech, 1992) and effects of biological variables such as crown position (Shen *et al.*, 1981; El-Kassaby *et al.*, 1986; Adams and Birkes, 1989) and reproductive phenology (El-Kassaby *et al.*, 1988) on mating systems. However, since the large tree to tree variation observed in the seed orchard is probably due to the statistical inefficiency caused by small sample size rather than to the biological factors, there is therefore no actual need to carry out any specific management measures such as mentioned above in the seed orchard.

3.4.3 Fixation index

The expected fixation index at inbreeding equilibrium (Allard *et al.*, 1968) was computed on the basis of the multilocus estimates of outcrossing rate in all the populations as $F_e=(1-t_m)/(1+t_m)$. The equilibrium inbreeding coefficients were low, as expected for an outcrossing species. If the mating system is the sole determinant of genotypic proportions, the F_e values should be similar to the F values of the adult plants. This is probably true for the populations in this study. Although no significant difference was found between the F and F_e values, the parental generation always shows a much lower fixation index than the filial one (F values for the progeny population were estimated in another study which was reported in the previous Chapter). This reflects a significant reduction in homozygosity between the two phases of the life cycle, which could be due to the selection against inbreds. This phenomenon has been often reported in conifers, which usually show a noticeable inbreeding and particularly selfing depression (Sorensen and Miles, 1982; Sorensen, 1982; Griffin and Lindgren, 1985) attributed to the action of recessive lethal and deleterious alleles.

3.4.4 Practical implication for seed orchard management

In conclusion, any selfing detected in the clonal seed orchard and managed stands is most probably caused by true selfing. Self-fertilisation can have severe effects on seed production as many selfed zygotes may either not germinate or die as seedlings in some species. The inbreeding depression caused by consanguineous mating is not as severe as self-fertilisation, but still has important effects on tree breeding programs. The inbred seedlings produced by consanguineous matings may survive past the seedling stage, but would still suffer from loss

of heterozygosity, which leads to inferior growth performance and productivity as mature trees. Obviously, the determination of consanguineous matings is of significant importance for seed orchards and thus measures can be taken to control this inbreeding depression. Findings from this study indicate that an effective experimental design and clonal deployment strategy in seed orchards can largely protect seed production from inbreeding depression caused by consanguineous matings. However in natural populations, particularly those subjected to intensive logging and lowering of population density, selfing rates may be raised leading to the production of a significant proportion of progeny that may suffer from inbreeding depression.

4. Application of DNA techniques in indentification of an unknown seed source of *Pinus caribaea* in China

4.1 Introduction

4.1.1 Identity of the Chinese material

In Chapter 2, the Chinese material labelled as var. *caribaea* was distinguished as genetically distinct from the Cuban sources. The UPGMA cluster analysis based on isozyme data assayed for 8 loci showed that the material taken from China has the largest genetic distance from materials taken from Cuba. Also the Chinese material was detected as significantly different from var. *bahamensis* analysed in that study. Suspicions about the true identity of the Chinese material arose. It could be a different species which was wrongly labelled or a hybrid between related species. The possibility of being *P. elliottii* or its hybrid with the var. *caribaea* was considered in Chapter 2. In order to confirm or refute these suspicions, further genetic evidence is required. In particular, a species/variety-specific marker is needed. The use of rapidly advancing DNA technology could definitely help to solve this problem.

4.1.2 Use of DNA markers

4.1.2.1 Advantages

DNA markers are powerful tools for investigating population genetics because they provide information on the genotype directly although the method is much more expensive than the isozyme method. Isozyme variation may not provide a complete measure of nucleotide variation in the genome because the loci are limited in number and restricted to the coding regions of the genome. In addition, silent mutations and some amino acid substitutions are not detectable (Boscherini *et al.*, 1994). In contrast, DNA methods allow direct analysis of both coding and non-coding regions of all three genomes of plants and are capable of providing a vast range of genetic markers. DNA markers have demonstrated obvious advantages on occasions when traditional genetic markers were deficient (Leigh Brown, 1989; Hartl and Clark, 1989; Strauss *et al.*, 1992; Szmidt and Wang, 1991; Wagner, 1992):

- A large number of polymorphisms (some of which are selectively neutral and some of which are under selection) can be identified in any taxon

- Widely differing levels of polymorphism can be studied.
- DNA analyses allow investigation of not only coding, but also non-coding variation.
- Both Mendelian and non-Mendelian markers can be identified, because DNA resides in chloroplasts, mitochondria and nuclei.
- It is usually possible to determine the mutational differences among DNA variants (e.g., point mutations, insertions/deletions, or rearrangements), which strengthens population analyses.

In addition, DNA markers do not vary among tissue types or developmental stages of the plant because the assays are based on the DNA itself and not the products of genes. DNA can be isolated from almost any plant part, thus enabling detection from many tissue types and at most developmental stages. Another obvious advantage is that DNA markers are not affected by environmental variation.

It should be borne in mind that DNA polymorphisms are not a panacea: other genetic markers are sometimes preferable because of their lower costs in personnel, reagents, laboratory facilities, and time. These are the disadvantages of DNA markers.

4.1.2.2 Chloroplast DNA and RFLP

Among the three genomes, cpDNA is by now the best known molecular genetic unit in forest trees. It is a circular molecule of 120-210 kb in size, tightly packed with structural genes (Downie and Palmer, 1992; Sugiuna, 1992).

The relative simplicity and compact size of cpDNA makes it amenable even for detailed molecular analysis. Complete cpDNA sequences have been studied for several plants (Ohyana *et al.*, 1986; Shinozaki *et al.*, 1986; Hiratsuka *et al.*, 1989). Gene and restriction cpDNA maps have been published for many plants including some conifers (Strauss *et al.*, 1988; Lidholm and Gustafsson, 1991). The knowledge of these cpDNA sequences and maps facilitates investigation of a variety of questions regarding changes in gene content, genome structure and rates of gene evolution (Clegg *et al.*, 1986; Clegg and Zurawski, 1992).

The cpDNA in most conifers has been found to be paternally inherited (Neal *et al.*, 1986; Szmidt *et al.*, 1987; Szmidt *et al.*, 1988a; Neal *et al.*, 1989; Neal and Sedoff, 1989; Wagner

et al., 1989; Stine and Keathley, 1990). The paternal cpDNA transmission implies that it can be used as a source of genetic markers for analysis of gene flow via pollen. Its uniparental inheritance, which does not undergo recombination (Chiu and Sears, 1985) makes it possible to infer a great deal of historical information preserved in cpDNA sequences (Whittemore and Schaal, 1991). The ancestry of an organellar haplotype will remain recognisable even after many generations of sexual reproduction (Whittemore and Schaal, 1991). The non-recombinant nature of cpDNA markers makes them a powerful tool in tracing the long-term effects of natural hybridisation (Avise and Saunder, 1984). A hybrid is expected to possess an identical, or nearly identical paternally inherited cpDNA pattern of one parent. This advantage has been fully realised by plant geneticists and cpDNA markers have been frequently applied in genetic documentation of purported hybrid taxa (Wagner *et al.*, 1987; EL-Kassaby *et al.*, 1988; Szmidt *et al.*, 1988a; Govindaraju *et al.*, 1989; Ernst *et al.*, 1990; Wagner *et al.*, 1991; Filppula *et al.*, 1992; Sigurgeirsson, 1992).

Chloroplast DNA in plants is conservative in most respects (Zurawski and Clegg, 1987). The evolution rate of chloroplast genes is estimated to be several times slower than nuclear genes (Wolfe *et al.*, 1987). Therefore, cpDNA sequence change is appropriate to resolve plant phylogenetic relationships at higher taxonomic levels but probably insensitive for detection of differentiation at an intra-specific level (Ritland and Clegg, 1987; Palmer *et al.*, 1988; Szmidt, 1991; Clegg and Zuraski, 1992). Chloroplast DNA analysis has been mostly used to assess phylogenetic relationships in conifer genera such as *Pinus*, *Picea* and *Pseudotsuga* (Szmidt *et al.*, 1988b; Strauss and Doerksen, 1990; Strauss *et al.*, 1990; Wang *et al.*, 1991a; Sigurgeirsson and Szmidt, 1992).

One of the straightforward and most frequently used DNA methods for detecting DNA variation is that of RFLP (restriction fragment length polymorphisms) analyses. RFLPs are simple Mendelian genetic markers which result from various types of mutations and rearrangements of the DNA. The concept of RFLP mapping of complex genomes was first described by Botsein *et al.* (1980). It has been described by many authors (Beckmann and Soller, 1983, 1986a, 1986b; Landry and Michelmore, 1987; Soller and Beckmann, 1983; Tanksley *et al.*, 1989). In RFLP detection, the first step is to isolate DNA from the organism of study and cleave the DNA with one or more restriction endonucleases. These are bacterial enzymes cleaving double-stranded DNA at unique recognition sequences, usually 4-8

nucleotides in length. Insertions and deletions of small segments of DNA or the gain or loss of a restriction site are two types of RFLPs which are easily detected by the techniques of Southern blotting and probe hybridisation. Following digestion with restriction enzymes, DNAs are fractionated electrophoretically on agarose gels. The DNA is then denatured and single-stranded DNA molecules are transferred and covalentaly linked to nylon membranes (blots). Small DNA fragments, called probes, are then radiolabelled and allowed to hybridise to their complementary DNA sequences bound to the blots. The radioactive DNA hybrids on the blot are visualised by autoradiography.

4.1.2.3 Applications in forest trees particularly *P. caribaea*

The DNA markers, such as RFLPs or other DNA variants allow for monitoring genetic polymorphisms in cases where appropriate genetic analyses have identified allelic variants. Chloroplast DNA has been utilised to study phylogenetic relationships among species (Szmidt, 1991) and for solving specific problems, like the classification of seedlots which originate from introgression zones (Szmidt *et al.*, 1988a).

The application of DNA polymorphisms in forest tree populations has received only limited attention in comparison with such application in other organisms. One of the useful applications of DNA markers is germplasm identification which is a common problem in tree improvement. A simple application would be the clonal identification in which two ramets are to be identified as members of the same or different clones. The more complex problem is to identify seed source of a bulked seed collection because it involves estimation of gene frequency. Identification of the seed source of the Chinese populations is the major problem encountered in this study.

The idea of using DNA techniques to identify the Chinese material is to find species/variety-specific markers for each of the varieties, by looking at the cpDNA variation within each of the three varieties and the Chinese material, and possibly other related species such as *P. elliottii* if material is available. If such a species/variety marker could be found, the identity of the Chinese material can be confirmed. The use of cpDNA variation of *P. caribaea* has already been applied to distinguish species and varieties (Nelson *et al.*, 1994). In their study, obvious cpDNA variation was found among the varieties of *P. caribaea* and other closely related pine species. The var. *caribaea* was distinguished from the other two varieties of the

same species by comparing their cpDNA genotypes. *P. elliottii*, a species closely related to *P. caribaea*, was also distinguished from *P. caribaea* and its related species *P. taeda* because of its high polymorphism of cpDNA. Similar techniques could be applied to identify the Chinese material since it was considered as *P. elliottii* or its hybrid with the var. *caribaea*.

4.1.2.4 Aims and strategy to be adopted

This study is, by the application of DNA techniques, intended to answer the question raised from the isozyme analysis: what is the true identity of the Chinese material? As the time and funding resource were very limited, this study was very restricted in terms of sampling and completeness. The RFLPs were used as markers to distinguish the various taxa in question.

4.2 Material and method

4.2.1 Preparation of probe

Restriction fragment pPCB28 cloned from the chloroplast genome of *Pinus contorta* Dougl. (Lindholm and Gustafsson, 1991) was prepared with the standard gene cloning method (Sambrook *et al.*, 1989). *E. coli* bacteria of which the plasmid DNA contained the above fragment were cultured in the LB (Luria-Bertani Medium) medium. The antibiotic for the bacteria culture is Ampicillin used at 60 μg/ml. Plasmid DNA was prepared from the bacteria (Procedures in Appendix 4.1). The digestion of plasmid DNA was carried out with *Bam*HI restriction enzyme at 37 °C overnight. Then the fragment pPCB28 of the plasmid DNA was separated on a mini argarose gel (1%). The correct bands were extracted from the gel and stored for use as probe.

4.2.2 Seed material and DNA extraction

Seeds were germinated on filter paper until the radicle emerged. The seed coat was removed and the whole seed (both the endosperm and embryo) was used for DNA extraction. Total cellular DNA was extracted from bulked seeds of the var. *caribaea* from Cuba and China, and of the var. *hondurensis* in order to compare the patterns of cpDNA of each variety. In order to look at the patterns of the cpDNA variation within the Chinese material and the Cuban material, DNA was also extracted from individual seeds with CTAB buffer (Appendix

4.2). About 7 seeds were prepared for each sample. However, only 4 of each sample were finally used in the Southern transfer.

4.2.3 DNA digestion and separation

The extracted total cellular DNA was digested with *Bam*HI restriction enzyme at the amount of 0.3 µl per µl DNA. The digested DNA fragments were then electrophoretically separated on an agarose gel (1%) for blotting.

4.2.4 Southern blotting

The fractionated DNA fragments were transferred to nylon membrane by the method of Southern blotting (Appendix 4.3) and the blot was stored for hybridisation.

4.2.5 Hybridisation

Only one cpDNA marker was assayed by probing southern pine *Bam*HI restriction fragments. The *Bam*HI fragment (pPCB28) was cloned from the chloroplast genome of *P. contorta* Dougl. The blot was baked at 80 °C for 2 hours before hybridisation to bind the fragments irreversibly to the membrane. The probe obtained from bacteria plasmid DNA was radioactively labelled as described by Nelson *et al.* (1994). Labelled DNA probe was hybridised to the membrane-bound *Bam*HI fragments at 65 °C overnight. Prehybridisation and hybridisation solutions were as used by Nelson *et al.* (1994). Following the hybridisation, the membrane were washed in 2 x SSC and 0.1% SDS (Appendix 4.4). The luminescent membrane was radio-photographed for visualisation.

4.3 Results

4.3.1 Banding patterns

The *Bam*HI-pPCB28 genotypes are shown in Figure 4.1. The bands of the bulked samples have much stronger intensities than those of individual seeds due to the different amounts of DNAs extracted from the material. The banding patterns of the bulked seed samples showed all possible sizes of restriction fragments shown by individual seeds.

The differences between the Chinese material and the Cuban material were obvious. The Chinese material showed three different genotypes while the Cuban material showed only one genotype. The three genotypes shown by the Chinese material are 2.5-7.9-10.5, 10.5 and 2.5-4.0 (Genotypes are denoted by the sizes in kilobase pairs of their restriction fragments. Hyphens are used to separate the restriction size classes.). The one genotype shown by the Cuban material is 10.5. The one genotype shown by the Cuban material was also shared by the Chinese material (Table 4.1).

Table 4.1 *Bam*HI-pPCB28 genotypic counts for *P. caribaea* including three varieties, *P. elliottii* and *P. taeda* [the left part of the table based on Nelson *et al.* (1994)].

Genotype	*P. caribaea*			*P. elliottii*	*P. taeda*	Chinese	Cuban
	bahamensis	*caribaea*	*hondurensis*				
(2.6)-10.5[a]		9		32		1	2
2.5-(2.6)-7.9	10		20	8			
2.5-7.9-10.5				2		1	
2.5-4.0-10.5				1			
2.5-4.0					77	1	
Total	***10***	***9***	***20***	***43***	***77***	***3***	***2***

[a] *(2.6) is a band at low intensity reported by Nelson et al. (1994) but not found in this study.*

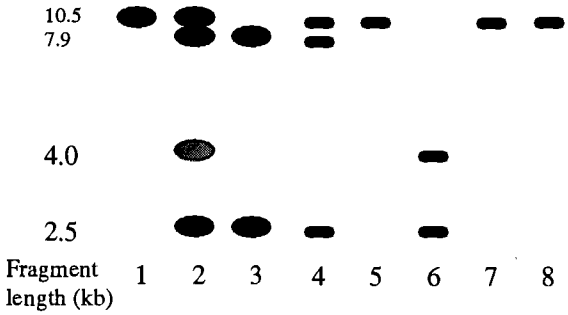

Figure 4.1 Diagram showing results obtained from the autoradiogram. The figure shows the BamHI-pPCB28 genotypes of bulked seeds and individual seeds. Lanes 1 to 3 are bulked seed samples and are var. *caribaea* from Cuba, Chinese material and var. *hondurensis* respectively. Lanes 4 to 8 are individual seed samples. 4 to 6 are Chinese material and 7 to 8 are var. *caribaea* from Cuba.

4.3.2 Difference between the Chinese and the Cuban samples

As Figure 4.1 shows, the cpDNA genotypes are obviously different between the Chinese and the Cuban samples. The Chinese material was so variable that three seeds assayed in the study

showed 3 different genotypes, while the Cuban samples appeared to have the same one genotype for the two seeds. Differences were also observed for the bulked seed samples. The Cuban sample showed only one band and the Honduran sample had two bands while the Chinese sample had four bands. This evidence supports the results obtained from the isozyme analysis that the Chinese material genetically differs from the Cuban material.

4.3.3 Identity of the Chinese sample

P. elliottii was not analysed in the present study due to unavailability of seeds; direct comparison of the Chinese material with *P. elliottii* is therefore impossible. However, if we compare these results with that from Nelson *et al.* (1994), it becomes clear that the Chinese material is most probably *P. elliottii* because of the highly variable cpDNA genotypes in the Chinese material. The other two genotypes shown in *P. elliottii* by Nelson *et al.* (1994) did not appear in the Chinese material, which could be simply due to the small number of seeds used in the present study. The Chinese material also showed a cpDNA genotype which is specific to *P. taeda* (Table 4.1).

To identify the Chinese seed sample, a larger sample size would need to be used. This requirement was not met in this study due to the limitations on time and funds. Also more taxa, including all varieties of *P. caribaea* and the related species like *P. elliottii* and *P. taeda* need to be assayed with cpDNA.

The results obtained from this study further suggested that the Chinese sample was not the true var. *caribaea.*

4.4 Discussion

Although only two samples and only 2 or 3 seeds of each were assayed in this study, the comparison (Table 4.1) of these results with those obtained by Nelson *et al.* (1994) provided enough evidence that the Chinese sample is neither the true var. *caribaea*, nor the var. *bahamensis* or the var. *hondurensis*. The var. *bahamensis* and var. *hondurensis* share no genotype with the Chinese material. The true var. *caribaea* has only one cpDNA genotype. While this is found in the Chinese material, 2 other genotypes are also found in addition.

The Chinese material shares two genotypes with *P. elliotii* and one genotype with each of *P. taeda* and *P. caribaea* var. *caribaea*. Because the Chinese material showed highly variable cpDNA pattern which is similar to the *P. elliottii*, the most probable identity of the Chinese material is *P. elliottii*. If this is not true, the possibility of being a hybrid between *P. elliottii* and *P. caribaea* var. *caribaea*, or between *P. elliottii* and *P. taeda*, or a mixture of them is then to be considered. Obviously, a firm conclusion on the identity of the Chinese material cannot be made so far based on the present evidence, but the conclusion that the Chinese material is neither the true var. *caribaea* nor other two varieties of the *P. caribaea* remains.

Further studies need to be conducted in order to identify accurately the Chinese sample. A large sample size should be used to allow effective estimation of gene frequencies. Also all varieties of the *P. caribaea* and related species need to be compared in the study.

Since mtDNA was found to be more variable at the intra-species level due to its large rearrangements (Palmer, 1990), it may be useful for distinguishing low level taxa within species. If variety-specific mtDNA makers can be found for each of the varieties, the Chinese material could be easily distinguished. Furthermore, the maternal inheritance for mtDNA and paternal inheritance for cpDNA (Neale and Sederoff, 1989) in conifers make them potential molecular markers for identifying a hybrid between two species. The possibility of being a hybrid of those varieties or related species can be evaluated by comparing their mtDNA and cpDNA patterns. If the Chinese material is of hybrid origin, then its mtDNA should share genotypes with its maternal parent, and its cpDNA should share genotypes with its paternal parent. If it is the true var. *caribaea* or mislabelled *P. elliottii*, the mtDNA of the Chinese material should contain the same genotype as that of trees from the natural stands, or of trees of *P. elliottii*.

5. Quantitative genetic studies using morphological traits

5.1 Introduction

5.1.1 Quantitative genetic variation

Genetic variation affects such diverse traits as DNA, allozymes and visible genetic abnormalities. Such traits have a relatively simple pattern of inheritance because they are influenced by relatively few genes. Morphological traits are more complex traits that are influenced by alleles at several or many loci. These types of traits are defined as polygenic traits, and they are more or less strongly influenced by the environment. The first category of traits is often selectively neutral and is useful for inference of population history and evolution while the second is selectively adaptive and includes economically important traits. Genetic variation manifested by these different types of traits may have different implications for applied tree breeding programmes. The information derived from these traits may also be exploited for different tasks in a breeding programme.

It is often assumed that marker genes can be used to predict the amount of quantitatively inherited variation. However, the amount of variation may well differ between quantitative traits and marker loci (Milligan *et al.*, 1994). The level of genetic variation in quantitative traits depends on a balance between mutation and selection, or between different selective pressures (Barton and Turelli, 1989). Variation at isozyme and other marker loci may mostly be governed by mutation and drift (Kimura, 1983). The level of genetic differentiation between populations at neutral loci depends on a balance between migration and genetic drift (Hartl and Clark, 1989).

Analysis of quantitative traits is another somewhat different morphological approach to measuring genetic variation. The genetic basis of this variation can be determined from common garden experiments. An alternative morphological approach is quantitative genetic analysis in which the transmission of morphological traits from parents to offspring of selected crosses is scored. Such studies involve a series of crosses (either controlled or open-pollinated), with the resulting progeny often grown in different environments to determine the

degree of environmental versus genetic control of polygenic morphological traits. Quantitative traits are analysed by either a least-squares or a maximum-likelihood procedure.

The results from a quantitative genetic analysis not only indicate the degree to which a phenotypic trait is under genetic or environmental control, but also determine such parameters as narrow-sense heritability and genetic covariance between characters, which are important in estimating the potential for evolutionary change within populations. They are obviously also important in breeding programmes for a forest tree species. These studies also estimate the opportunity for response to selection (Namkoong, 1979; Falconer, 1989). A quantitative genetic analysis can thus provide information on a critical and central aspect of forest tree breeding and conservation genetics.

The disadvantage of quantitative genetic analyses is that they are very time consuming, particularly for the long life-cycled forest trees; they require careful experimental design and data analysis; they cannot be conducted in species that will not readily cross or that have extremely long life spans.

Variation for quantitative traits which appear to be of evolutionary importance (e.g. size, fitness components, rate of growth, etc.) does not follow single-locus, Mendelian patterns of inheritance. In other words, there is no direct correspondence between particular phenotypes and genotypes as is apparent for the other genetic variants (e.g. allozyme and DNA variants). The importance of genetic factors in quantitative (polygenic) traits can be proved by the following evidence. First, when artificial selection is practiced over generations, the mean value of virtually any quantitative trait can be altered. Such a response to selection indicates an underlying genetic determination of the trait. Second, related individuals are generally more similar to each other in phenotype than unrelated individuals. This fact can be used to estimate the amount of genetic variation for the trait under consideration. The third indication that genetic factors are involved in quantitative traits is the actual location on specific chromosomes of the genes that affect a particular trait. Location of genes affecting quantitative traits has generally only been possible in organisms with special chromosomal stocks, such as *Drosophila* and wheat (Hedrick, 1985).

Forest tree improvement is applied when control of parentage is combined with other forest management activities, such as site preparation or fertilization, to improve the overall yields and quality of products from forests land (Zobel, 1984). Tree improvement is effective only when it consists of the combination of all silvicultural and tree breeding skills of the forester to grow the most valuable forest products as quickly as possible and as inexpensively as possible. It is obvious that most decisions that are made in implementing a tree improvement programme depend on quantitative genetic analysis. Such decisions include Programme initiation/continuation decisions; selection decisions; breeding decisions; testing decisions; production decisions; and allocation /deployment decisions (Talbert, 1992).

There are two aspects to any successful tree improvement programme. The first is the short-term genetic gain. It relates to obtaining an immediate gain of desired products as rapidly and as effectively as possible. This is achieved by intensively applying genetic principles to operational forestry programmes to produce better-quality, better-adapted and higher-yielding tree crops. Maximum gains are achieved by the use of a few super genotypes as parents to supply planting stocks for operational programs. The second aspect is the long-term genetic variability. It is concerned with the long-term need to provide the broad genetic base that is essential for continued progress over many generations. These two aspects are both of great importance for a successful improvement programme. However, the two aspects are always in negative relationship to each other and optimisation between long-term genetic variability and short-term genetic gains is therefore of great importance to achieve these two goals.

The breeding strategy is another core component for an improvement programme. It governs all the breeding activities. To develop a good breeding strategy for a specific tree species for a specific region, it is essential to obtain basic genetic information by employing other genetic methods discussed above such as the use of genetic markers. Only when a better understanding of the genetic constitution of the species is obtained, can a better breeding strategy be developed.

5.1.2 Using integrated multiple purpose tests in tree breeding

Traditionally, forest genetic tests have been conducted sequentially with successive stages; first the species trial, then the provenance trial which usually includes two steps, range-wide

test initially and then limited-range test based on the information resulting from the range-wide test. The provenance trial is then followed by progeny trials and clonal tests, etc. However, as the limitations of poorly adapted or narrowly based tree introductions became apparent (Brewbaker, 1990; Burley and Adlard, 1992; Hughes, 1989), and the economical demands for rapid value return by foresters, coupled with the special feature of long life cycle of almost all forest tree species, usually several decades, the disadvantages of the lengthy traditional procedure have been highlighted. In practice, there is often a strong economic pressure to reduce the testing interval between these stages in a modern tree improvement programme. Various alternatives have been proposed as a solution in which a common feature is to combine two or more stages in the testing sequence, such as species/provenance tests (Wang, *et al.*, 1992; Wang and Zheng, 1993; Zheng and Wang, 1993a, b), provenance/progeny tests (Kanowski and Nikles, 1989b; Crockford, *et al.*, 1989), or the Breeding Seedling Orchard (BSO) combining progeny tests with seed production (Barnes and Mullin, 1989).

In a tree breeding programme, the ultimate extension of this is an experiment that combines provenance testing, progeny testing and seed production in a single trial. This requires that design of the trial, timing of assessments, and timing of thinning in relation to seed production must all be compatible, and consistent with the breeding strategy. The multiple purpose of the trial, i.e. testing and seed production, allows genetically improved seeds to be made available within a short time, while achieving many of the aims of a long-term breeding programme. The thinning of the trial can be flexible depending on the breeding strategy to be adopted. Most of the literature reports estimates of genetic parameters at family and individual levels with no recognition of provenance information. The analysis of the trial requires estimation of a genetic variance component for provenance which is of great help for developing the breeding strategy and for subsequent thinning of the trial (*e.g.* constructing a selection index).

Despite the advantages discussed above for integration, the experimental design and the statistical analysis of the integrated genetic tests are complicated. Firstly, the multiple purpose of the tests, particularly being the breeding population, requires that the population size be sufficiently large, i.e. a large number of trees need to be included in a single test, and hence a large experimental area is needed. Trials over large areas introduce problems of environmental heterogeneity. This introduces problems of environmental heterogeneity within replications and leads to the difficulty of eliminating the effect of such heterogeneity within the

replication, and consequently biases the estimation of environmental variance in the analysis. A second drawback is the undesigned non-orthogonality (i.e. unbalanced representation) of testing entries across sites or replications (Matheson, 1989b). It is almost impossible in practice to keep all the testing units such as provenances or families equally represented if a large number of entries is used, particularly for species as exotics. This leads to another difficulty in the analysis of the experiment.

Fortunately, advances in our understanding of the quantitative genetics of trees (Namkoong and Kang, 1990; Kanowski, 1993), and in statistical methodology and computing power, have greatly enhanced our capacity to derive more precise answers from genetic tests. Much effort has been directed towards coping with the environmental heterogeneity and the non-orthogonality: various field designs of varying composition have been addressed in both theoretical and empirical studies (Lambeth, 1986; Loo-Dinkins *et al.*, 1990; McCutchan *et al.*, 1989); statistically efficient designs, such as incomplete blocks, are widely used (Loo-Dinkins, 1992; Matheson, 1989b); early selection from short-term testing in the greenhouse or nursery (deSouza *et al.*, 1992; Lowe and van Buijtenen, 1989), or at a close spacing on relatively uniform, high quality sites (Li, *et al.*, 1992). The development of computer software directly relevant to the estimation of genetic parameters from forest genetic data sets (GENSTAT REML or SAS VARCOMP) has greatly facilitated the efficient and appropriate statistical analyses. Even though the analysis has not been a problem for such experimental designs mentioned above, large data sets still take some time to analyse. It is anticipated that the rapid hardware development (eg calculation speed) will lead to great reduction in computing time.

Of those integrated genetic tests, the combined provenance/progeny test is one of the most widely used in tree improvement (Kanowski and Nikles, 1989). This experiment combines provenance testing, progeny testing and seed production together and has the functions of testing, breeding and production. The combined provenance/progeny test has been used in several breeding programmes for important industrial species, particularly for exotic fast growing species such as *Eucalyptus* and *Pinus* species (Zheng *et al.*, 1993, Zheng *et al.*, 1994, Wang *et al*, 1992). This method is particularly useful for exotic fast growing species because it allows genetically improved seed supply to be available in a relative short time, and along with it a long-term breeding programme is also able to be conducted.

Although the complexity of the analysis and genetic interpretation of the results as mentioned above, the advantages outweighed the disadvantages. Overcoming difficulties largely relies on the availability of knowledge on experimental design and statistical analyses, and access to computer statistical packages.

5.1.3 Tree improvement of the variety in China

As introduced earlier in this thesis, the species *P. caribaea* has been very successfully grown in China as an exotic species and has shown great potential for industrial purposes in the tropical area in southern China. In recent years, *P. caribaea* has been increasingly planted in Southern China on a large scale. However, it appears that there are obvious differences among the three varieties in growth rates, stem form and pest and disease resistance. Species trials have shown that in addition to fast growth rates, *P. caribaea var. bahamensis* has excellent stem form and more resistance to the shoot tip moth (Pan, 1991).

However, the genetic base of the material used in the above provenance trials was relatively narrow, particularly *P. caribaea* var. *bahamensis* which was introduced only since the 1980s, although significant provenance variation has been shown in the trials. Furthermore, founder effect and genetic bottleneck may exist for the plantations of the variety in China. A range-wide provenance trial combined with open-pollinated progeny tests of *P. caribaea* var. *bahamensis* was carried out as the initial step of a genetic improvement programme for the variety in China. Although no relevant genetic information on such as genetic structure and population differentiation was available, the existing literature indicates that the genetic variation for quantitative traits may be large. Nevertheless, the environmental condition in Southern China where the variety is grown varies considerably from place to place. Therefore, based on the genetic information from this experiment, a breeding strategy that accommodates these problems would be essential to the success of the genetic improvement for the species in China.

5.1.4 Aims of the study

This study is a field experiment of a range-wide provenance trial combined with open-pollinated progeny tests of *P. caribaea* var. *bahamensis*. It was established as the initial step

of the genetic improvement programme for the species in China, and is intended to serve as the breeding population of the variety and to form the basis for a seed orchard when appropriately thinned. The general purposes of the study are to investigate the amount and the pattern of the genetic variation existing within var. *bahamensis* at various levels, to estimate genetic parameters at provenance, family, within-family and individual levels, and on the basis of these estimates to determine the best selection method and develop an optimal breeding strategy for the species. In other words, more specifically, the objectives of this field experiment are:

a. To investigate the possible regional variation;
b. To study between-provenance variation;
c. To study within-provenance (family) variation;
d. To create a breeding population in the form of breeding seedling orchards using identified half-sib families;
e. To create a seedling seed orchard to produce genetically improved seeds for commercial plantations;
f. To estimate genetic parameters of *P. caribaea* var. *bahamensis* in China for important economical traits;
g. To establish gene conservation stands.

5.2 MATERIALS AND METHODS

5.2.1 Establishment of the trials

5.2.1.1 Seed

In June 1986, FAO studied the forest genetics and seed production of *P. caribaea* var. *bahamensis* in the Bahamas, in which a new international provenance trial of the species was recommended. In June 1988, the Overseas Development Administration of the United Kingdom funded a six-month project for the OFI in cooperation with the Forest Officer in the Department of Lands and Surveys to collect the seeds recommended by Barnes in 1986 (Baylis and Barnes, 1989).

Seeds were collected from the four islands (Table 5.1) and accordingly grouped into four regions. According to the specific populations from which seeds were collected, seeds were

divided into 14 provenances. The seeds were kept identified at individual tree level (family identity).

Seeds used in this combined provenance/progeny trial are all from the above collection. The total number of half-sib families is 121 and number of families within a provenance varied from different provenances. Details of provenances and number of families within each provenance are given in Table 5.1. Seed collections are mapped in Figure 5.1.

5.2.1.2 Experimental site

The experiment was established at the Forestry Institute of Hepu county, Guangxi Province, located at 21°41' N, 109°11'E in the coastal low land, where frost does not usually occur. It experiences a northern tropical climate. The lateritic red-yellowish sand soil is very deep. *Eucalyptus*, *P. caribaea* and *P. elliottii* and *Acacia* are major plantation species in this area. Winter is usually dry from December to February and most rainfall is in summer. July and February are the hottest and coldest months respectively. Mean annual temperature is about 22.3°C, and mean annual rainfall is about 1651 mm.

5.2.1.3 Experimental design

To make full use of the limited materials, the trials are composed of two parts. The main part is the combined provenance/progeny trial which is also to be managed as a BSO (Breeding Seedling Orchard) and to be ultimately converted into a SSO (Seedling Seed Orchard). The other, as a supplementary part, is the gene conservation stand which makes use of all the extra seedlings after establishing the main part. Different designs were used for each of the two parts.

Although the environmental condition at the experiment site is fairly uniform, the original purpose of the design for the provenance/progeny trial was to use a Lattice Square (LATTSQ) (Cochran and Cox, 1957). The advantage of this design is that it reduces the environmental effects in both row and column directions. It also guarantees that the experiment can be analysed as a RCB design when the requirements for the balance of treatments cannot be met (Matheson, 1989a, b).

Table 5.1 Details of provenances and number of families within each provenance

Provenance (Code)	Seed ID No.	Lat.(°N)	Long.(°W)	Alt.(M)	Families
Abaco Island					
Cedar Harbour (1)	658/1-10	26°53'	77°39'	10	10
Norman Castle (2)	668/1-10	26°45'	77°26'	5	10
Central Abaco (3)	678/1-10	26°26'	77°05'	5	5
Sandy Point (4)	688/1-10	26°02'	77°12'	10	10
Andros Island					
San Andros (5)	698/1-10	24°57'	78°01'	5	10
Staniard Creek (6)	708/1-10	24°50'	77°55'	5	10
Roker Cay (7)	718/1-10	24°07'	77°44'	2	8
Kemps Bay (8)	728/1-10	24°06'	77°36'	5	9
Grand Bahama Island					
Freeport (9)	738/1-10	26°32'	78°45'	5	3
South Riding (10)	748/1-10	26°40'	78°13'	10	9
Maclean's Town Cay(11)	758/1-10	26°34'	77°55'	2	7
Little Harbour Cay (12)	768/1-10	26°33'	77°53'	2	10
New Providence					
Adelaide (13)	778/1-10	25°00'	77°26'	10	10
East New Prov. (14)	788/1-10	25°01'	77°24'	5	10

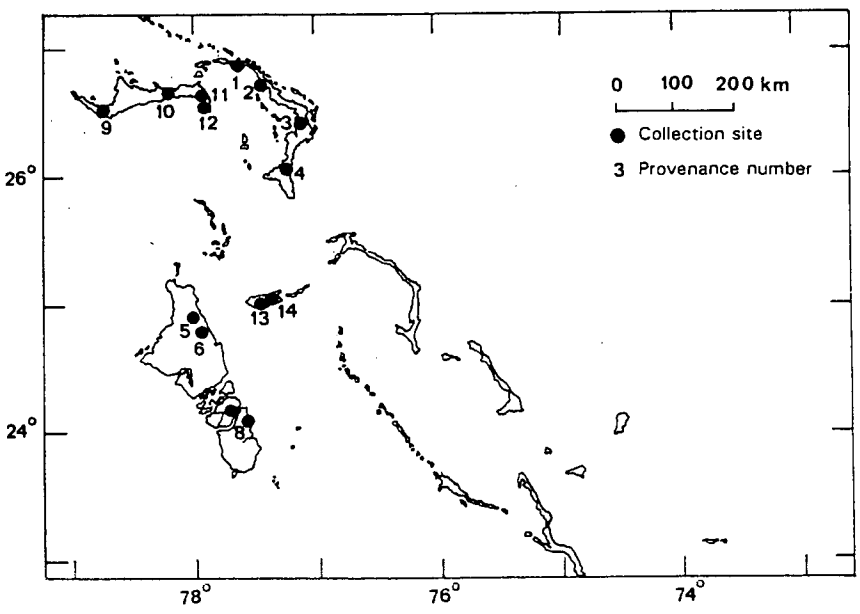

Figure 5.1 Natural distribution of *Pinus caribaea var. bahamensis* and the provenance locations

An 11 (rows) × 11 (colums) LATTSQ design was used with 6 replicates and 121 treatments and a 4-tree size of plot. Each of the 121 treatments is a different family. Initial spacing is 3 m by 3 m. A single-row plot was used for the gene conservation stand, each plot of which represents a single family.

5.2.1.4 Nursery and planting

Seeds were sown in August 1990 and the standard methods for pines in South China were used to raise containerised seedlings (Pan and You, 1991). Seedlings were planted out in May 1991.

The site was wholly and intensively prepared. Planting holes were dug at a size of about 60 x 50 x 40 cm. Normal managemental measures were taken as usually applied for pine plantations.

5.2.2 Measurements and data storage

The trial was planted in May, 1991 and is measured annually at the end of each growing season. The latest data available for this study was from the measurements done in December, 1993 at the age of 2.5 years. Height (m), DBH (cm) (at ground level for the first year after planting), Crown Width (m) (measured with two cross directions of the crown), damage by pests and disease are measured. However, because the pest and disease damage are slight and have occurred on only a few trees, analysis of these two traits has not so far been completed. All data were stored in dBASE files.

5.2.3 Statistical methodology

5.2.3.1 Computer software

Genstat 5 (Genstat 5 Committee, 1990) was used to carry out the statistical analysis, i.e. ANOVA, estimation of genetic parameters and some calculations. In addition, QBasic language was used for programming to pre-handle the data. Excel® is used for general spreadsheet handling and graphic presentation. Missing values were estimated and the analysis was adjusted by the GENSTAT software.

5.2.3.2 Statistical model and ANOVA

5.2.3.2.1 Standard model

An attempt to analyse the trial as a LATTSQ (Lattice square) was unsuccessful due to the imbalance of treatments which is probably caused by the misplacement of the treatments in the replicates. The Lattice Square design requires that each pair of treatments must only appear together in the same row or column once (Cochran and Cox, 1957). However, the advantage of this design is that it can be analysed as a RCB design (Kempthorne, 1952; Matheson, 1989a, b), in which each replication of the LATTSQ design is considered as a complete block. The standard model for the combined provenance/progeny trial is a form of mixed model (Land, *et al.*, 1987; Kanowski and Nikles, 1989).

$$X_{ijkl}=\mu+R_i+P_j+F_{j:k}+(PR)_{ij}+(FR)_{j:ki}+\varepsilon_{ijkl} \qquad (5.1)$$

Where

X_{ijkl} is the phenotypic value of the *l*th individual of the *k*th family from the *j*th provenance in the *i*th replication;

μ is the fixed term overall mean;

R_i is the fixed effect of the *i*th replicate;

P_j is the effect of the *j*th provenance (assumed random);

$F_{j:k}$ is the effect of the *k*th family in the *j*th provenance (assumed random);

$(RP)_{ij}$ is the effect of the interaction between the *i*th replicate and the *j*th provenance;

$(FR)_{j:ki}$ is the effect of the interaction between the *i*th replication and the *k*th family of the *j*th provenance;

ε_{ijkl} is the effect of the *l*th tree within the *k*th family of the *j*th provenance in the *i*th replication. It includes the effect of errors.

i=1, ...R (R is the number of replications);

j=1,...P (P is the number of provenances);

k=1,...f (f is the number of families within each provenance);

l=1,...n (n is the number of trees per plot)

The ANOVA corresponding to the model (5.1) is given in Table 5.2.

Table 5.2 The ANOVA and expected mean squares for a balanced combined provenance/progeny trials on individual tree.

Source of variation	D.f.	Expected mean squares
Replicate	*R-1*	$\sigma_w^2 + n\sigma_{fR}^2 + nf\sigma_{PR}^2 + nfP\sigma_R^2$
Provenance	*P-1*	$\sigma_w^2 + n\sigma_{fR}^2 + nR\sigma_f^2 + nf\sigma_{PR}^2 + nfR\sigma_P^2$
Provenance x Replicate	*(P-1)(R-1)*	$\sigma_w^2 + n\sigma_{fR}^2 + nf\sigma_{PR}^2$
Family/Prov.	*P(f-1)*	$\sigma_w^2 + n\sigma_{fR}^2 + nR\sigma_f^2$
Family/Prov. x Replicate	*P(f-1)(R-1)*	$\sigma_w^2 + n\sigma_{fR}^2$
Trees within plot	*PRf(n-1)*	σ_w^2
Total	***PRfn-1***	

σ_R^2 : Variance component of replicate

σ_P^2 : Variance component of provenance

σ_{PR}^2 :Variance component of the interaction of provenance by replicate

σ_f^2 : Variance component of family within provenance

σ_w^2 : Variance component of trees within plot

5.2.3.2.2 Adjustments for the model

It may not be appropriate to test the significance of provenance differences if there are significant provenance by replicate effects. Kanowski and Nikles (1989) proposed that an exact test for provenance effects will only be possible if provenance by replicate interactions are negligible, or ignored. Westfall (1992) suggested that in a single-site experiment when errors are probably under control, the provenance by replicate interaction can be pooled with the plot error. A preliminary analysis showed that the interaction is very small and not significant (5% level), i.e. the interaction can be ignored. In addition, the computational time and the demands for numerical space (Genstat 5 Committee, 1990) are much reduced when pooling the provenance by replicate interaction. This model is used to estimate variance components.

$$X_{ijkl}=\mu+R_i+P_j+F_{j:k}+(FR)_{j:ki}+\varepsilon_{ijkl} \qquad (5.2)$$

The ANOVA of this model is given in Table 5.3.

Table 5.3 The ANOVA and expected mean squares for a balanced combined provenance/progeny trials on individual tree.

Source of variation	D.f.	Expected mean squares
Replicate	*R-1*	$EMS_r=\sigma_w^2+n\sigma_{fR}^2+nfP\sigma_R^2$
Provenance	*P-1*	$EMS_p=\sigma_w^2+n\sigma_{fR}^2+nR\sigma_f^2+nfR\sigma_P^2$
Family/Prov.	*P(f-1)*	$EMS_f=\sigma_w^2+n\sigma_{fR}^2+nR\sigma_f^2$
Family/Prov. x Replicate	*P(f-1)(R-1)*	$EMS_{fr}=\sigma_w^2+n\sigma_{fR}^2$
Trees within plot	*PRf(n-1)*	$EMS_w=\sigma_w^2$
Total	***PRfn-1***	

Since *P. caribaea* var. *bahamensis* is mainly distributed in four isolated islands, comparisons of seed sources from different islands (regions) were conducted. To test the significance among islands, an extra effect of island was included in the statistical model. To simplify the analysis, the two interaction terms are ignored in the model and the analysis is based on plot means.

$$X_{ijkl}=\mu+I_h+R_i+P_j+F_{j:k}+\varepsilon_{ijkl} \qquad (5.3)$$

This model is used to test significance only. Variance components are not estimated from this model. The corresponding ANOVA is given in Table 5.4.

Table 5.4 The ANOVA for significance test of regions (differences between islands) for a balanced combined provenance/progeny trials on plot means.

Source of variation	D.f.	Expected mean squares
Replicate	*R-1*	$EMS_r=\sigma_e^2+Pf\sigma_R^2$
Region	*I-1*	$EMS_i=\sigma_e^2+R\sigma_f^2+Rf\sigma_P^2+RP\sigma_I^2$
Provenance/Region	*I(P-1)*	$EMS_p=\sigma_e^2+R\sigma_f^2+Rf\sigma_P^2$
Family/Provenance	*I(P-1)(f-1)*	$EMS_f=\sigma_e^2+R\sigma_f^2$
Residuals	*I(P-1)(f-1)(R-1)*	$EMS_e=\sigma_e^2$
Total	***RIPf-1***	

5.2.3.3 Estimation of variance components

One of the most important aspects of the analysis of data from random or mixed models is the estimation of variance components. An obvious technique for estimation of variance components is to take the mean squares in the ANOVA as estimates of their respective

expectations and to estimate the individual components using the obvious linear combinations of the expected mean squares. The computations for this method of estimation, generally called ANOVA, are quite simple and are most frequently used traditionally.

The special feature of this experiment is that it is unbalanced. The size of the *i*th provenance (f_i), varies from provenance to provenance. This case in which samples are of unequal size occurs commonly in family studies in human and animal genetics and in the social sciences (Snedecor and Cochran, 1967). In forestry experiments it normally occurs only when established that way or when very poorly adapted provenances are included in the test (Land *et al.*, 1987). In this experiment the imbalance is mainly caused by using unequal number of families within each provenance. However, making full use of the limited test material is particularly important for the genetic improvement programme of this exotic species. All materials have been included although some of the provenances have very few families.

The total family number is $F=\Sigma f_i$. With equal f_i (=f), the mean squares between classes (Replicate, Provenance, Family/Provenance) were found to be unbiased estimates of the respective expected mean squares. With unequal f_i, the corresponding coefficient of f is f_0, where

$$f_0 = \frac{1}{(P-1)}\left(F - \frac{\sum f_i^2}{F}\right) \quad \text{(Snedecor and Cochran, 1967).} \tag{5.4}$$

$$= \bar{f} - \frac{\sum (f_i - \bar{f})^2}{(P-1)F} \tag{5.5}$$

F: Total number of families; P: Number of provenances;

f_0: Adjusted number of families within provenance

The equation (5.5) shows that f_0 is always less than the arithmetic mean $\bar{f}$ of the f_i although usually only slightly less.

The above equation (5.4) was given for the case in which only one stratum is involved in sampling. In the case of a combined provenance/family trial in which two error strata are involved, more complicated modification is needed. Adjustments are made not only for the family number but also replication number and within-plot tree number (due to missing trees or plots).

There are certain drawbacks to the ANOVA method for estimating variance components, such as the fact that the variance components may take on negative values. This was proved by a pre-analysis of the trial. In recent years several other techniques for estimation of variance components have been growing in favour, especially for use with unbalanced data (Kennedy and Gentle, 1980). One of the most widely used methods is maximum likelihood.

Likelihood-based methods for estimating variance components are rapidly gaining favour among animal breeders and practitioners. Of these methods restricted maximum likelihood (REML) estimation, in which the likelihood function is taken to be that associated with a set of error contrasts, is becoming the method of choice (Harville and Callanan, 1990). However, the computations required to implement REML estimations are very extensive, prohibitively so in many cases. In general, closed-form expressions for the REML estimation do not exist, and the estimates must be computed by using an iterative numerical method to locate the maximum of the likelihood (Harville and Callanan, 1990).

The REML procedure of GENSTAT 5 is used to fit a variance component model by residual (or restricted) maximum likelihood, together with the command VCOMPONENT which is used to define the model for REML (GENSTAT 5 Committee, 1990).

5.2.3.4 Relationship between growth performance and provenance origin

Simple correlations are conducted between all growth traits and latitude and longitude of the provenances.

5.2.3.5 Estimate of genetic parameters

Variance components and adjusted coefficients are calculated based on the model 5.2. Definitions and calculations of genetic and phenotypic variances, heritabilities, genetic correlations and responses to selection are given in the following sections (Namkoong, 1979; Falconer, 1989; Land *et al.*, 1987).

5.2.3.5.1 Variance

Calculations of variances based on the definitions described in Table 5.2 are given below.

Additive genetic variance of individual trees $=\sigma_A^2$

Total phenotypic variance of individual trees $=\sigma_T^2=\sigma_w^2+\sigma_{fR}^2+\sigma_f^2+\sigma_p^2$

Additive genetic variance of provenances $=\sigma_p^2$

Phenotypic variance of provenances $=\sigma_{Tp}^2=EMS_p/nRf=\sigma_w^2/nRf+\sigma_f^2/f+\sigma_{PR}^2/R+\sigma_P^2$

Additive genetic variance of families $=\sigma_f^2$

Phenotypic variance of families $=\sigma_{Tf}^2=EMS_f/nR=\sigma_w^2/nR+\sigma_{fR}^2/R+\sigma_f^2$

Additive genetic variance within families $=\sigma_A^2-\sigma_f^2=(1-r)\sigma_A^2$, since the genetic variance between families is $\sigma_f^2=r\sigma_A^2$ (r is 1/2 for full-sibs and 1/4 for half-sibs).

Phenotypic variance within families$=\sigma_{Tw}^2=\sigma_w^2+\sigma_{fR}^2$

5.2.3.5.2 Heritability

Four types of narrow-sense heritabilities are estimated.

(i) Individual heritability (Mass selection)

$$h_i^2=\frac{\sigma_A^2}{\sigma_T^2}=\frac{\sigma_f^2/r}{\sigma_w^2+\sigma_{fR}^2+\sigma_f^2+\sigma_p^2} \qquad (5.6)$$

When open-pollinated families are considered to be true half-sib families, and assuming no epistatic effects, $F_{j:k}$ is the general combining ability of the *k*th parent, so $F_{j:k}$ is equal to 1/2 of the breeding value of the *k*th parent (Hodge and White, 1992; Hodge and Purnell, 1993), thus $Var(F_{j:k})=Var(BV/2)=\sigma_A^2/4$ (i.e. 1/4 of the additive genetic variance. $\sigma_f^2=\sigma_A^2/4=r\sigma_A^2$).

However, only when the progenies are true half-sibs and there is no inbreeding, *r* is equal to 1/4; in the wild population, there is often some self-pollination and therefore *r* is increased. It was found that *r* is between 0.27 to 0.39 (Namkoong, Snyder and Stonecypher, 1966) and 1/3 is usually used (Bridgwater 1992). This paper assumes the value of *r* is 1/3.

(ii) Family (mean) heritability

$$h_f^2 = \frac{\sigma_f^2}{\sigma_{Tf}^2} = \frac{\sigma_f^2}{EMS_f / nR} = \frac{\sigma_f^2}{\sigma_w^2 / nR + \sigma_{fR}^2 / R + \sigma_f^2} \quad (5.7)$$

(iii) Within-family heritability

$$h_w^2 = \frac{(1-r)\sigma_A^2}{\sigma_{Tw}^2} = \frac{(1-r)\sigma_f^2 / r}{\sigma_w^2 + \sigma_{fR}^2} = \frac{(1/r - 1)\sigma_f^2}{\sigma_w^2 + \sigma_{fR}^2} \quad (5.8)$$

(iv) Provenance (mean) heritability

$$h_P^2 = \frac{\sigma_P^2}{\sigma_{TP}^2} = \frac{\sigma_P^2}{EMS_P / nRf} = \frac{\sigma_P^2}{\sigma_w^2 / nRf + \sigma_f^2 / f + \sigma_{PR}^2 / R + \sigma_P^2} \quad (5.9)$$

5.2.3.5.3 Phenotypic and genetic correlation

Phenotypic and genetic correlations between traits are calculated using the following formulae respectively.

Phenotypic correlations are estimated on individual tree values.

$$r_{p_{xy}} = \frac{\sigma_{p_x p_y}}{\sqrt{\sigma_{p_x}^2 \sigma_{p_y}^2}} \quad (5.10)$$

where r_p is the phenotypic correlation coefficient, σ_{PxPy} is the cross-product of the phenotypic values of traits x and y; σ_{Px}^2 and σ_{Py}^2 are phenotypic variances of traits x and y respectively.

Genetic correlations are estimated with the following equation:

$$r_{A_{x,y}} = \frac{\sigma_{A_x A_y}}{\sqrt{\sigma_{A_x}^2 \sigma_{A_y}^2}} = \frac{\sigma_{f_x f_y} / r}{(1/r)\sqrt{\sigma_{f_x}^2 \sigma_{f_y}^2}} = \frac{\sigma_{f_x f_y}}{\sqrt{\sigma_{f_x}^2 \sigma_{f_y}^2}} \quad (5.11)$$

where $r_{A_{x,y}}$ is the genetic correlation coefficient, $\sigma_{A_x A_y}$ is the cross product of the additive genetic values of traits x and y; $\sigma_{A_x}^2$ and $\sigma_{A_y}^2$ are additive genetic variances of traits x and y, respectively. The genetic cross-products are estimated from the following ANCOVA.

Table 5.5 ANCOVA and expected mean squares of traits x and y

Source of variation	D.f.	Expected mean products (EMP) of trait (x+y)
Replicate	R-1	$EMP_r=\sigma_{w_xw_y}+n\sigma_{fR_xfR_y}+nPf\,\sigma_{R_xR_y}$
Provenance	P-1	$EMP_p=\sigma_{w_xw_y}+n\sigma_{fR_xfR_y}+nR\sigma_{f_xf_y}+nRf\,\sigma_{P_xP_y}$
Family/Provenance	P(f-1)	$EMP_{f(p)}=\sigma_{w_xw_y}+n\sigma_{fR_xfR_y}+nR\sigma_{f_xf_y}$
Family/Prov. x Replicate	P(f-1)(R-1)	$EMP_{f(p).R}=\sigma_{w_xw_y}+n\sigma_{fR_xfR_y}$
Within plot error	PRf(n-1)	$EMP_w=\sigma_{w_xw_y}$
Total	**PRfn-1**	

$\sigma_{R_xR_y}$: Covariance component of replicate

$\sigma_{P_xP_y}$: Covariance component of provenance

$\sigma_{fR_xfR_y}$:Covariance component of the interaction of family (within provenance) by replicate

$\sigma_{f_xf_y}$: Covariance component of family within provenance

$\sigma_{w_xw_y}$: Covariance component of trees within plot

where $\sigma_{f_xf_y}=[EMP_{f(p)}-EMP_{f(p).R}]/nR$

The EMP for each stratum in the ANCOVA table can be deduced as follows:

Since $Var(x+y)=Var(x)+Var(y)+2Cov(xy)$,

then $EMP=Cov(xy)=[Var(x+y)-Var(x)-Var(y)]/2=[EMS(x+y)-EMS(x)-EMS(y)]/2$

5.2.3.6 Selection response

Selection responses to different selection methods are estimated by using the formula:

$$R=i\sigma h^2 \tag{5.12}$$

where i is the selection intensity, σ is the standard deviation of phenotypic variance, and h^2 is the heritability of the selection method.

The selection responses for relevant selection methods are therefore as the follows:

$$R_p=i\sigma_p h_p^2=i\sqrt{V_p}h_p^2 \tag{5.13}$$

$$R_f=i\sigma_f h_f^2=i\sqrt{V_f}h_f^2 \tag{5.14}$$

$$R_i=i\sigma_i h_i^2=i\sqrt{V_i}h_i^2 \tag{5.15}$$

$$R_{wf}=i\sigma_{wf} h_{wf}^2=i\sqrt{V_{wf}}h_{wf}^2 \tag{5.16}$$

5.2.3.7 Age to age correlation

The different measurements of a trait at different ages were considered as separate traits. Genetic correlation for a trait between ages is calculated with the following equation.

$$r_{A_{1,2}} = \frac{\sigma_{A_1 A_2}}{\sqrt{\sigma^2_{A_1}\sigma^2_{A_2}}} = \frac{\sigma_{f_1 f_2}}{\sqrt{\sigma^2_{f_1}\sigma^2_{f_2}}} \qquad (5.17)$$

where traits 1 and 2 are the measurements of the same trait at age 1 and age 2.

5.3 Results

5.3.1 Genetic variation

5.3.1.1 Differences between regions (islands)

The mean of a region is calculated on all provenances from the same island. There are significant differences among regions in all three traits (Table 5.6). Abaco island ranked top followed by New Providence, Grand Bahama and Andros island respectively. This ranking order applies to height, diameter and crown width. (Table 5.7).

Table 5.6 ANOVAs of H (height), DBH (Diameter at breast height) and CW (Crown width) measured in December, 1993 at age of 2.5 years. Number of missing values indicated in the parenthesis (Significance test for one-tailed hypotheses).

Trait	Source of variation	d.f.	s.s.	m.s.	v.r.
H	Replicate	5	17.91516	3.58303	45.86
	Region	3	17.23699	5.74566	7.63**
	Prov./Reg.	10	7.52342	0.75234	3.45**
	Fam./Prov./Reg.	107	23.35064	0.21823	2.79**
	Residual	600	46.87955	0.07813	
	Total	**725**	**112.90575**		
DBH	Replicate	5	70.2703	14.0541	43.03
	Region	3	35.1218	11.7073	7.73**
	Prov./Reg.	10	98.5658	1.5149	1.64
	Fam./Prov./Reg.	107	194.3211	0.9212	2.82**
	Residual	595(5)	412.0512	0.3266	
	Total	**720(5)**	**810.3302**		
CW	Replicate	5	18.00579	3.60116	80.45
	Region	3	2.94108	0.98036	4.28*
	Prov./Reg.	10	2.28842	0.22884	2.40*
	Fam./Prov./Reg.	107	10.21642	0.09548	2.13**
	Residual	588(12)	26.32136	0.04476	
	Total	**713(12)**	**59.56277**		

$F_{0.05(1),3,10}=3.71$; $F_{0.05(1),10,107}=1.93$; $F_{0.05(1),107,600}=1.28$
$F_{0.01(1),3,10}=6.55$; $F_{0.01(1),10,107}=2.50$; $F_{0.01(1),107,600}=1.41$

Table 5.7 Average values of the 3 traits measured for the four regions at age 2.5 years (± Standard error)

Region	Height (m)	DBH (cm)	CW(m)
Abaco Island	3.52 ± 0.36	5.51 ± 0.69	2.08 ± 0.28
New Providence	3.48 ± 0.35	5.46 ± 0.70	2.00 ± 0.27
Grand Bahama	3.36 ± 0.33	5.34 ± 0.65	1.99 ± 0.29
Andros Island	3.14 ± 0.40	4.98 ± 0.82	1.91 ± 0.27
Mean	**3.36 ± 0.40**	**5.30 ± 0.76**	**1.99 ± 0.29**

Results show that the eastern islands, i.e. Abaco and New Providence islands, have higher growth rates than the western islands Grand Bahama and Andros islands. Abaco island has highest mean growth rates. This is probably due to the higher rainfall in the island than in other southern islands. On average, the mean height of Abaco island is 12% larger than that of Andros island, and so are the mean DBH and CW, with 10.6% and 8.9% larger respectively.

The comparisons for each trait at different ages are illustrated in the following Figures (Figures 5.2 to 5.4). Diameter of 1991 is the ground level diameter. It is clear that Abaco island has the highest growth rate at all ages whereas Andros island has the lowest.

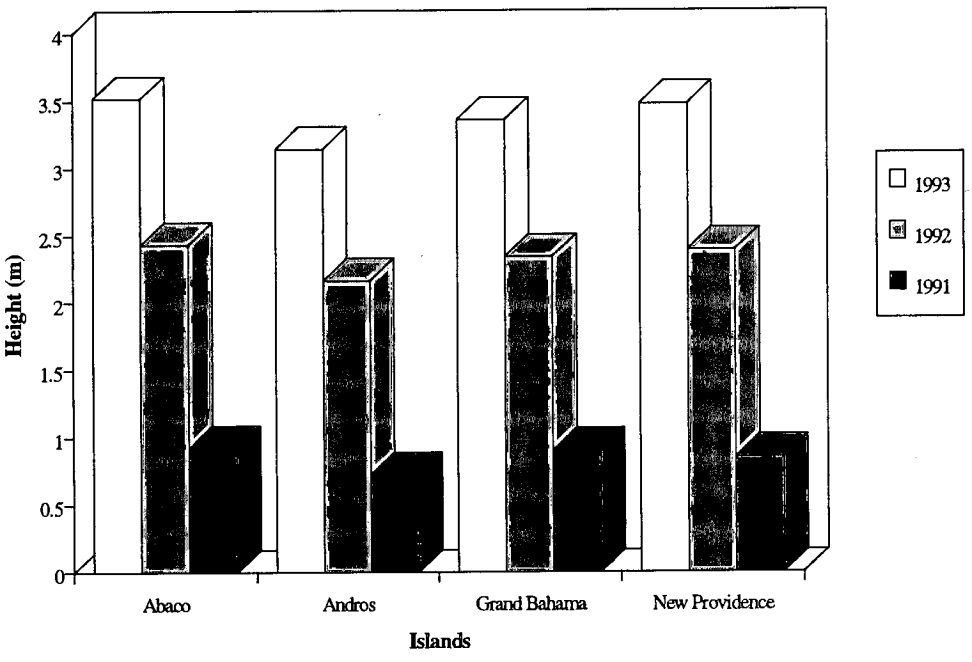

Figure 5.2 Regional comparisons of height at different age

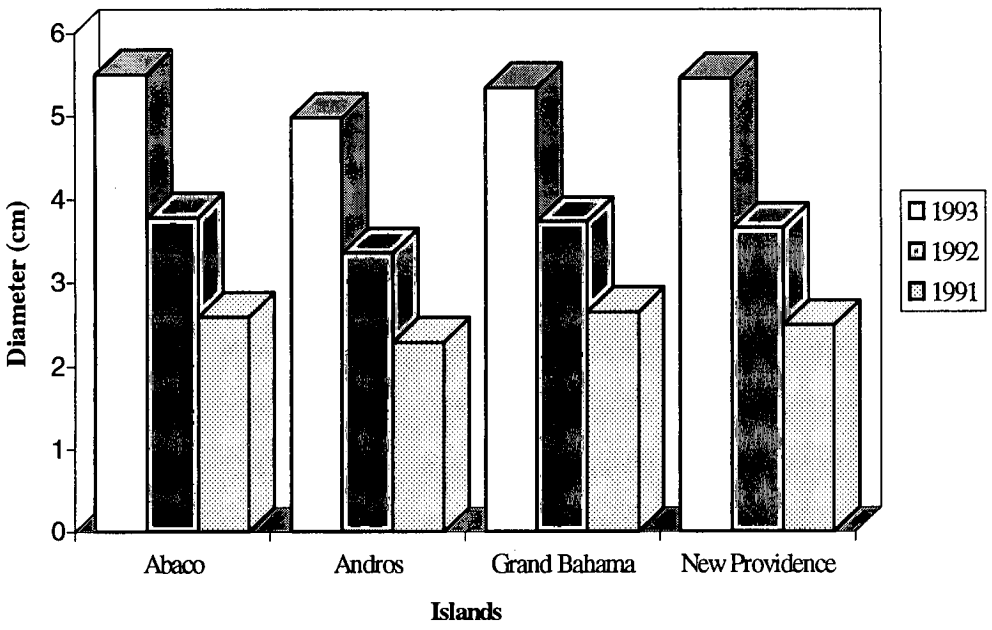

Figure 5.3 Regional comparisons of diameter at different age

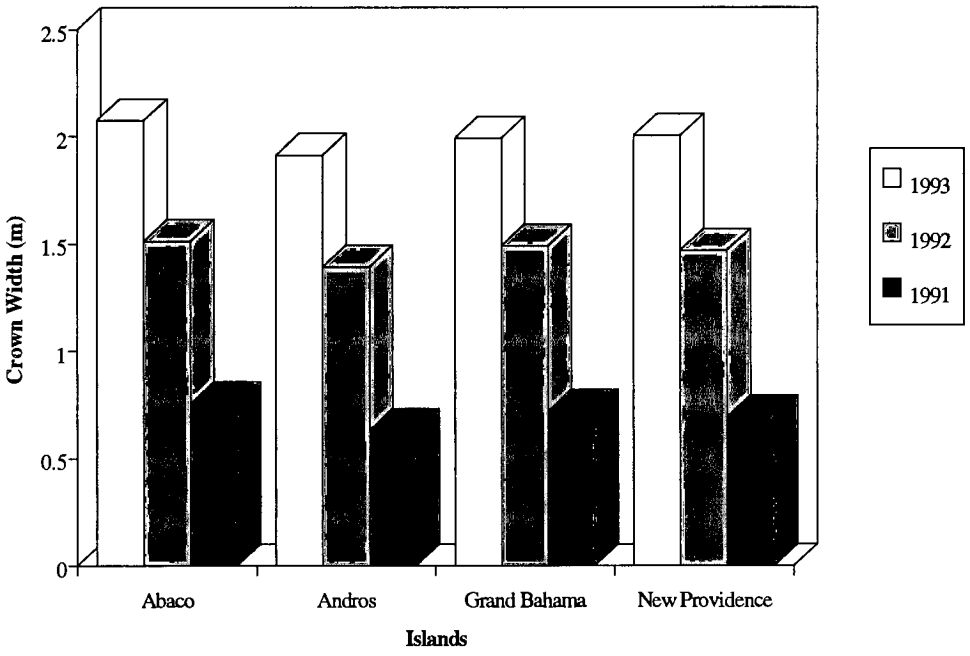

Figure 5. 4 Regional comparisons of crown width at different age

5.3.1.2 Differences between provenances

ANOVAs show that significant provenance variation within region exists in height and crown width (Table 5.6). Provenances from Abaco island and from New Providence have higher growth rates than those from the other two islands. Provenances from Andros have the lowest growth and most of them are ranked at the bottom (Table 5.8; Figure 5.8).

At age 2.5 years, provenance Adelaide has the highest height (3.59 m) and is followed by the provenances Sandy Point (3.57 m) and Norman Castle (3.55 m). The lowest height is of provenance Kemps Bay (2.98 m). Provenances Roker Cay and Staniard Creek have lower heights (3.01 and 3.20 m, respectively). The mean height of provenance Adelaide is 20.5% higher than that of the lowest provenance Kemps Bay (Table 5.8; Figure 5.5).

The ranking of DBH is generally the same as of height. Provenances from Abaco and New Providence have bigger diameter. The biggest diameter is of provenance Norman Castle (5.57 cm) and followed by Sandy Point (5.56 cm) and Adelaide (5.53 cm). The smallest diameter is of the same provenance as of height, i.e. provenance Kemps Bay (4.75 cm). Like in height, the provenances Roker Cay and Staniard Creek have smaller diameter (4.83 and 5.00 cm, respectively). The diameter of provenance Norman Castle is 17.3% larger than that of the Kemps Bay (Table 5.8; Figure 5.6).

The range of crown width is smaller than that of height. All provenances from Abaco are ranked at the top while the ones from New Providence have narrower width than some of the other two islands. This factor would be important when the tree form is taken into account in the selection (Table 5.8; Figure 5.7).

The range of height between provenances within each island range from 0.16 to 0.32 which are of Abaco and Andros island, respectively.

The growth rates are probably related to adaptation to the rainfall within the natural distribution, which decreases from north to south. Provenances from Abaco island are adapted to higher rainfall and those from Andros island to lower rainfall. The provenance Adelaide from New providence has the highest height growth although it has a lower latitude than those of the provenances from Grand Bahama island. This is probably due to its lower longitude.

Table 5.8 Growth data of each provenance at 2.5 years after planting (numbers in parentheses are ranking order; ± Standard error)

Region	Provenance	Height (m)	DBH (cm)	CW (m)
Abaco Island	Cedar Harbour	3.48±0.35 (4)	5.47±0.67 (4)	2.25±0.31 (1)
	Norman Castle	3.55±0.32 (3)	5.57±0.63 (1)	2.14±0.29 (2)
	Central Abaco	3.41±0.41 (6)	5.35±0.77 (6)	2.11±0.25 (3)
	Sandy Point	3.57±0.38 (2)	5.56±0.77 (2)	2.05±0.26 (4)
Andros Island	San Andros	3.34±0.32 (9)	5.29±0.68 (8)	2.00±0.27 (8)
	Staniard Creek	3.20±0.38 (12)	5.00±0.73 (12)	1.91±0.27 (11)
	Roker Cay	3.01±0.41(13)	4.83±0.92 (13)	1.85±0.27 (13)
	Kemps Bay	2.98±0.38 (14)	4.75±0.88 (14)	1.88±0.27 (12)
Grand Bahama	Freeport	3.26±0.38 (11)	5.28±0.66 (9)	1.83±0.30 (14)
	South Riding	3.30±0.34 (10)	5.18±0.68 (11)	1.93±0.32 (10)
	Maclean's Town Cay	3.36±0.33 (8)	5.32±0.63 (7)	2.05±0.29 (5)
	Little Harbour Cay	3.43±0.29 (5)	5.26±0.58 (10)	2.05±0.26 (6)
New Provid.	Adelaide	3.59±0.32 (1)	5.53±0.68 (3)	1.99±0.30 (9)
	East New Providence	3.37±0.35 (7)	5.39±0.72 (5)	2.02±0.25 (7)

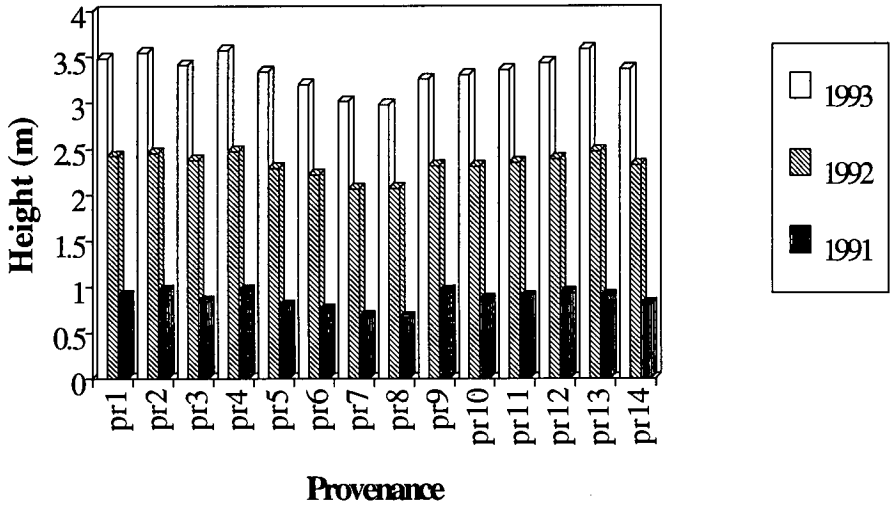

Figure 5.5 Comparisons of height at different ages among provenances

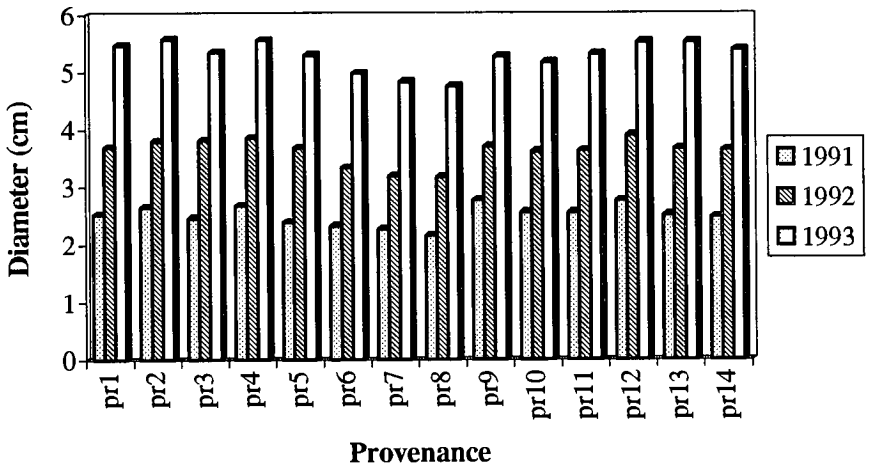

Figure 5.6 Comparisons of diameter at different ages among provenances

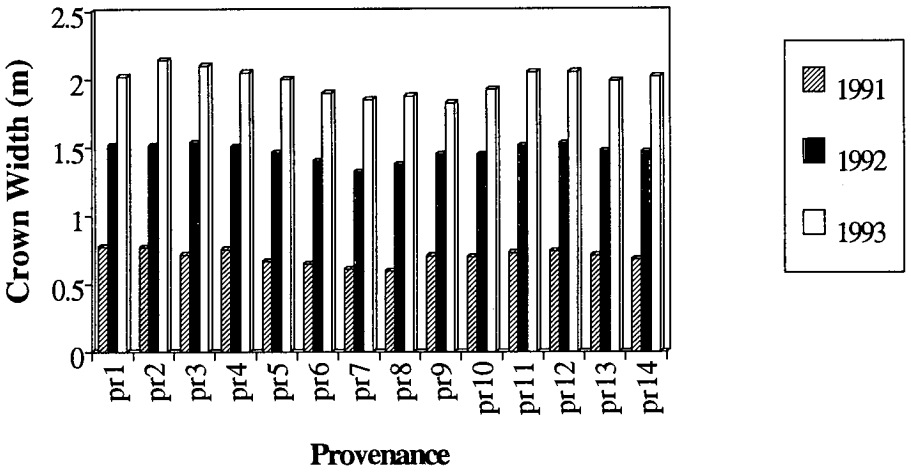

Figure 5.7 Comparisons of crown width at different ages among provenances

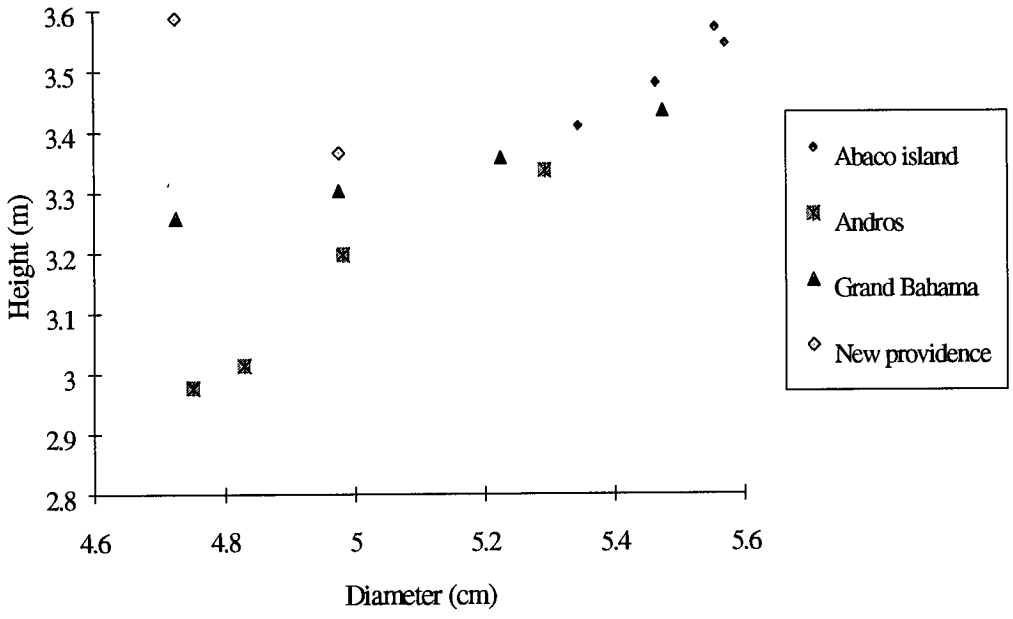

Figure 5.8 Plotted diagrams of provenances with height versus diameter at age 2.5 years.

5.3.1.3 Differences between families within provenances

Results show that there are significant differences between families within provenances for all three traits. Fast-growing families are mostly from Abaco island and New Providence. Family 6582, from provenance Cedar Harbour of Abaco island, has the highest height growth (3.88 m), followed by families 6882 (3.87 m), 6881 (3.86 m) from provenance Sandy Point of Abaco, 7788 (3.80 m) and 7784 (3.80 m) from provenance Adelaide of New Providence.

The top five families for DBH are 6882 of provenance Sandy Point (6.19 cm), 7683 of Little Harbour Cay (6.11 cm), 7784 of Adelaide (5.97 cm), 6680 of Norman Castle (5.96 cm) and 7880 of East New Providence (5.96 cm). The top five families for crown width are mostly of provenance Norman Castle and one of Cedar Harbour. This suggests that provenance Norman Castle may be selected in breeding for narrow crown characteristics (Table 5.9).

Table 5.9 Top 10 families and bottom 10 families for H (height), DBH and CW (crown width)

Rank	Provenance	Family	H	S.E.	Provenance	Family	DBH	S.E.	Provenance	Family	CW	S.E.
1	Cedar Harbour	6582	3.88	0.44	Sandy Point	6882	6.19	1.24	Norman Castle	6685	2.34	0.34
2	Sandy Point	6882	3.87	0.72	Little Harbour Cay	7683	6.11	0.74	Norman Castle	6681	2.28	0.41
3	Sandy Point	6881	3.86	0.45	Adelaide	7784	5.97	0.68	Norman Castle	6683	2.27	0.36
4	Adelaide	7788	3.80	0.40	Norman Castle	6680	5.96	0.98	Norman Castle	6686	2.26	0.42
5	Adelaide	7784	3.80	0.43	East New Providence	7880	5.96	1.11	Cedar Harbour	6585	2.25	0.40
6	Adelaide	7789	3.77	0.46	Adelaide	7782	5.93	0.94	Little Harbour Cay	7683	2.25	0.40
7	Little Harbour Cay	7683	3.73	0.34	Norman Castle	6681	5.90	0.89	Norman Castle	6689	2.21	0.33
8	Norman Castle	6686	3.73	0.35	Sandy Point	6889	5.86	1.09	Maclean's Town	7581	2.20	0.36
9	Adelaide	7783	3.71	0.34	Sandy Point	6887	5.85	0.75	East New Providence	7880	2.20	0.38
10	Norman Castle	6680	3.70	0.40	Maclean's Town	7583	5.83	0.74	San Andros	6986	2.19	0.39
112	Roker Cay	7184	2.98	0.58	Staniard Creek	7083	4.58	0.70	South Riding	7482	1.79	0.31
113	Staniard Creek	7085	2.95	0.50	Staniard Creek	7085	4.47	1.06	Freeport	7386	1.77	0.38
114	Staniard Creek	7083	2.90	0.46	Kemps Bay	7288	4.46	0.81	Cedar Harbour	6587	1.76	0.32
115	Roker Cay	7189	2.90	0.43	Staniard Creek	7083	4.43	0.86	Kemps Bay	7285	1.75	0.35
116	Kemps Bay	7281	2.87	0.53	Kemps Bay	7287	4.40	1.25	East New Providence	7881	1.74	0.30
117	Kemps Bay	7287	2.83	0.60	Roker Cay	7187	4.30	1.28	Kemps Bay	7282	1.73	0.41
118	Roker Cay	7187	2.78	0.57	South Riding	7483	4.26	1.11	Staniard Creek	7083	1.72	0.25
119	Kemps Bay	7288	2.75	0.51	Roker Cay	7181	4.26	0.89	Roker Cay	7189	1.72	0.23
120	Roker Cay	7181	2.66	0.33	Roker Cay	7189	4.18	0.89	Roker Cay	7182	1.68	0.36
121	Kemps Bay	7285	2.58	0.54	Kemps Bay	7285	3.57	1.28	Staniard Creek	7084	1.68	0.24
	Mean		**3.36**	**0.74**	**Mean**		**5.31**	**1.06**	**Mean**		**2.00**	**0.37**

Slow growing families are all from Andros island. Family 7285 from provenance Kemps Bay of Andros has the lowest height growth (2.58 m). Others having slow growth are families 7181 (2.66 m) from provenance Roker Cay and 7288 (2.75 m) from provenance Kemps Bay. Families with smallest DBH are also from provenances of Andros island.

5.3.2 Relationship between growth performance and provenances' origins

Calculations of the correlation between growth rates and provenances' origins show that all the three traits are significantly correlated to the latitudes of the provenances (Table 5.10). The positive correlations mean that the northern provenances grow faster than southern ones. This coincides with the rainfall distribution which decreases from north to south.

The negative values of the correlations with longitude imply that the eastern provenances grow faster than the western ones, although only the correlation of CW is significant.

Table 5.10 Single factor correlation between growth rates and latitude and longitude of provenances at age 2.5 years

Trait	Latitude (°N)	Longitude (°W)
Height(m)	0.610424886*	-0.36592
Diameter(cm)	0.584494243*	-0.32071
Crown-width(m)	0.584266071*	-0.53243*

*: at 5% level.

5.3.3 Genetic parameters

5.3.3.1 Variance components and adjusted coefficients of components

All VCOMPs and associated standard errors are estimated by the REML procedure. Results are summarised in Tables 5.11 and 5.12. Detailed output from estimation is given in Appendix 5.1. Since each trait has a different number of missing values, the REML estimates the coefficients of variance components differently (Table 5.11).

Table 5.11 Matrix of adjusted coefficients of variance components estimated with REML (Data measured in December, 1993)

	Variance Stratum	σ^2_p	$\sigma^2_{f(p)}$	σ^2_{fR}	σ^2_w
Height	Provenance	186.69	23.24	3.88	1
	Prov.Fam	0	23.08	3.86	1
	Prov.Fam.Rep	0	0	3.89	1
	Within plot	0	0	0	1
DBH	Provenance	191.06	23.05	3.88	1
	Prov.Fam	0	22.89	3.85	1
	Prov.Fam.Rep	0	0	3.89	1
	Within plot	0	0	0	1
CW	Provenance	186.86	22.61	3.86	1
	Prov.Fam	0	22.59	3.84	1
	Prov.Fam.Rep	0	0	3.87	1
	Within plot	0	0	0	1

Prov.Fam: Family within provenance; Prov.Fam.Rep: Family within provenance by replicate interaction

Table 5.12 Estimated Components of Variance and standard errors for each trait, data measured in December, 1993

Trait	Height		DBH		CW	
Source of variation	Stratum Variance	Variance Component	Stratum Variance	Variance Component	Stratum Variance	Variance Component
Prov.	6.683380	0.03133±0.01	14.2271	0.0565±0.03	1.51696	0.00602±0.003
Prov.Fam	0.830478	0.02296±0.01	3.39432	0.0948±0.02	0.340516	0.00862±0.002
Prov.Fam. Rep	0.301558	0.02787±0.01	1.22694	0.1211±0.02	0.146831	0.01611±0.002
Within plot	0.193030	0.1930±0.01	0.755731	0.7557±0.02	0.084524	0.08452±0.002

5.3.3.2 Heritability

Variance components are estimated with the REML procedure. Heritabilities (Narrow sense) are estimated for each trait at provenance, family, within-family and individual (Mass selection) levels (Table 5.13).

Table 5.13 Estimates of heritabilities from data measured in different years.

	Trait										
	Height				DBH				CW		
Age	4/91	11/91	12/92	12/93	4/91	11/91	12/92	12/93	11/91	12/92	12/93
h_p^2	0.88	0.91	0.79	0.88	0.79	0.87	0.75	0.76	0.90	0.79	0.74
h_f^2	0.79	0.72	0.68	0.64	0.50	0.63	0.53	0.64	0.40	0.43	0.57
h_i^2	0.63	0.37	0.33	0.25	0.17	0.29	0.25	0.28	0.10	0.13	0.22
h_{wf}^2	0.76	0.38	0.27	0.21	0.13	0.25	0.18	0.22	0.08	0.10	0.17

It is clear that for each of the traits, the provenance heritabilities have the highest values and are followed by the family mean heritabilities, individual heritabilities and within-family heritabilities. This suggests that if the selections are to be made under the same selection intensity at each of the four levels separately, considerable improvements for the growth traits can be made by first the provenance selection, followed by the family selection and individual selection and that the within-family selection would make the least improvement.

The heritability of crown width is the smallest at all four levels. At provenance and family levels, the heritability of height is higher than or equal to that of diameter, while at individual

and within-family levels the heritability is always bigger than that of height (Figure 5.9). It means that at different selection levels, the different traits have different genetic control.

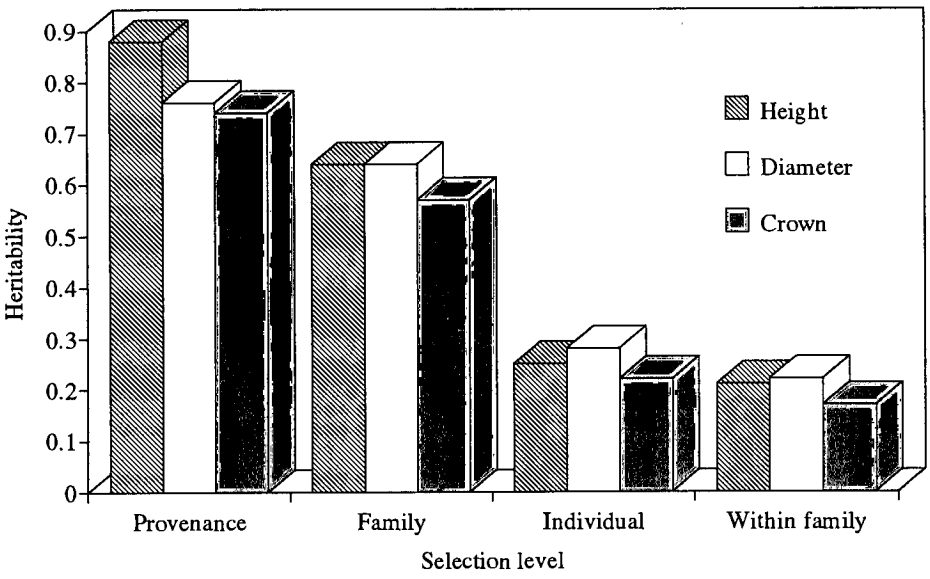

Figure 5.9 Heritabilities of height at provenance, family, individual and within-family levels in 1993.

However, for any trait, selection would also be made with the use of information of relatives, i.e. the combined selection using provenance, family and individual information. The individual values are adjusted with their provenance means and family means. An index can be constructed for each individual by combining individual value, family mean and provenance mean for one or more traits. Therefore, higher genetic gains can be achieved.

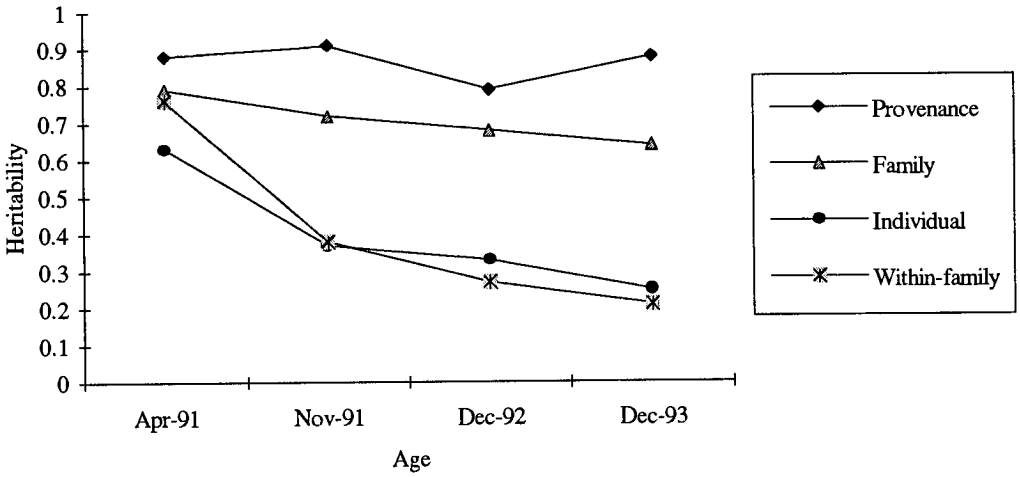

Figure 5.10 Heritabilities of height at provenance, family, individual and within-family levels at different ages

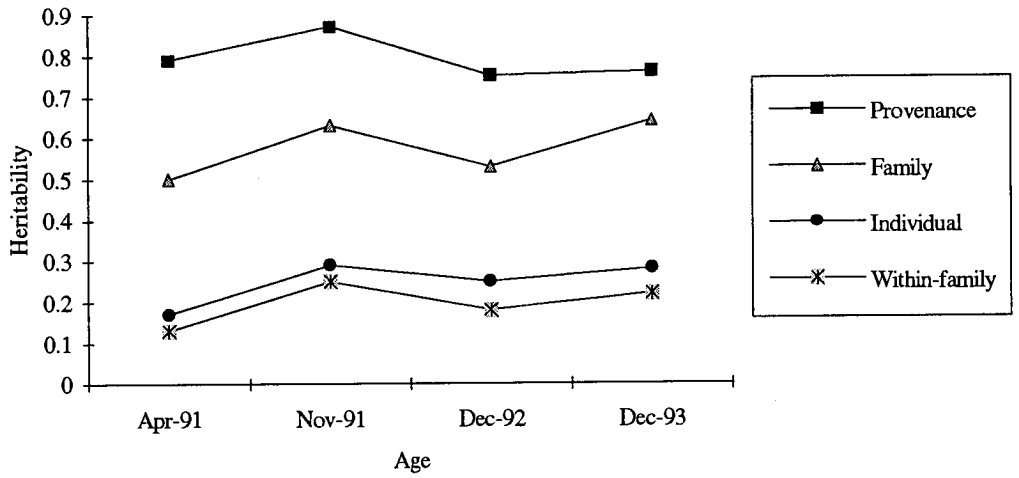

Figure 5.11 Heritabilities of diameter at provenance, family, individual and within-family levels at different ages.

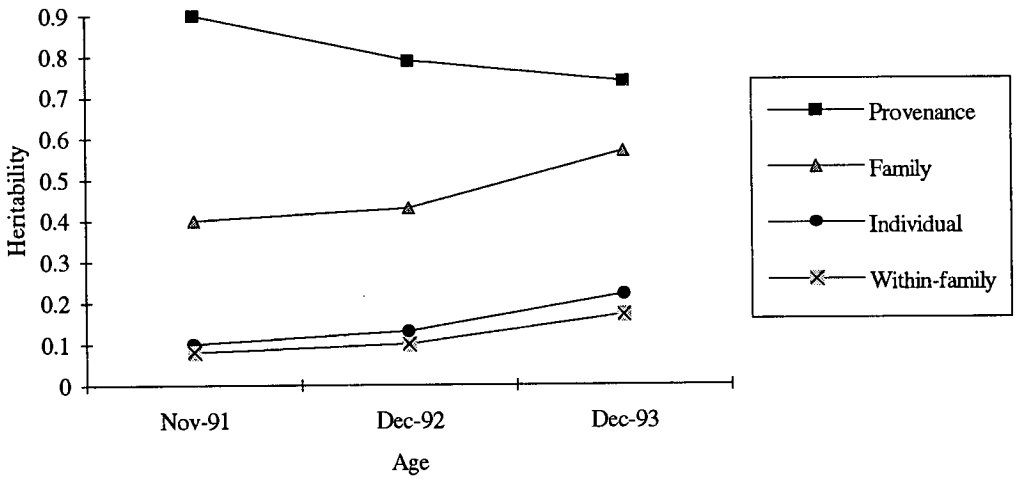

Figure 5.12 Heritabilities of crown width at provenance, family, individual and within-family levels at different ages.

The heritabilities of height and diameter fluctuate from age to age while those of the crown width change steadily over the ages. The heritabilities of height and diameter have a generally descending trend with the age. The provenance heritability of crown width decreases with the age and the other three heritabilities of crown width increase with the age (Figures 5.10 to 5.12). This implies that the genetic control of the detected variation is still not stable at these early ages. The variation in heritabilities is probably due to the variable growth performance over the years thus suggesting that selection at these early years may be less reliable. There may be still some maternal effects at these early ages and the effects are decreasing with the age.

5.3.3.3 Phenotypic correlations at age of 2.5 years

The phenotypic correlation between height and diameter is statistically significant and has the highest correlation coefficient (Table 5.14). The phenotypic correlation coefficients between height and CW and diameter and CW are smaller. The coefficient between DBH and CW is the lowest.

5.3.3.4 Genetic correlations at age of 2.5 years

All genetic correlation coefficients between each pair of traits are accordingly bigger than those phenotypic ones and have the same correlation trends as in phenotypic values. All genetic correlation coefficients are statistically significant at the 0.01 level. (Table 5.14). Height is most strongly correlated with DBH and then CW. The DBH and CW have the weakest correlation although it is significant.

Table 5.14 Phenotypic (above diagonal) and genetic (below diagonal) correlation coefficients matrix among traits.

	Height	DBH	CW
Height		0.8175**	0.5479**
DBH	0.8580**		0.4808**
CW	0.6281**	0.6083**	

**: 1% level

5.3.4 Selection response

To compare the relative selection responses of different selection methods, selection intensity is assumed to be the same for all methods. Thus, the selection response is determined by the magnitude of the heritability and the phenotypic variance at each of the four selection levels. The relative responses can be obtained by substituting the corresponding stratum variance and heritability into equations (5.13) to (5.16). The results are presented in Table 5.15.

Table 5.15 Relative selection responses to different selection methods.

	Height	DBH	CW
$R_p=i\sigma_p h_p^2=i\sqrt{V_p}h_p^2$	$2.275i$	$2.866631i$	$0.91142i$
$R_f=i\sigma_f h_f^2=i\sqrt{V_f}h_f^2$	$0.5832i$	$1.179116i$	$0.33262i$
$R_i=i\sigma_i h_i^2=i\sqrt{V_i}h_i^2$	$0.1373i$	$0.310149i$	$0.0843i$
$R_{wf}=i\sigma_{wf} h_{wf}^2=i\sqrt{V_{wf}}h_{wf}^2$	$0.0923i$	$0.191252i$	$0.04942i$

Suppose 1/3 of the total population of the trial is to be selected for the first thinning, i.e. standardised selection intensity i=1.091 (Cotterill and Dean, 1990), the relative selection responses to different selection methods are shown in Figure 5.13.

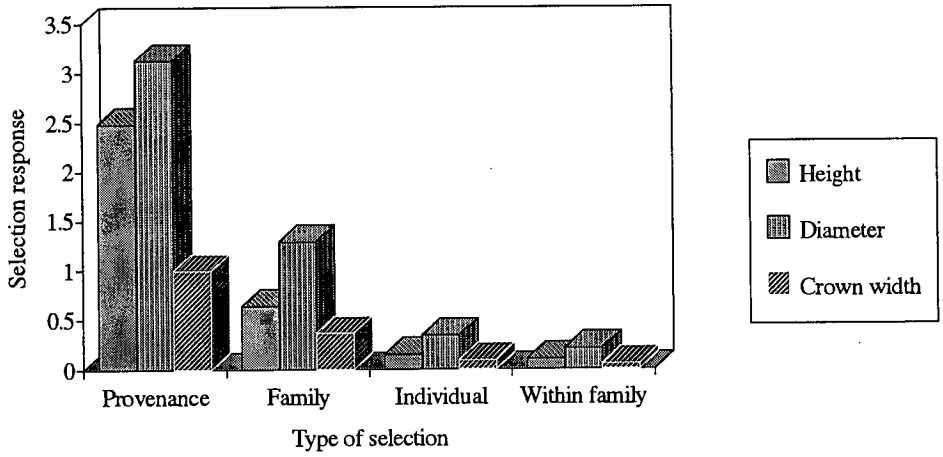

Figure 5.13 Comparisons of selection responses of different selection methods supposing selection intensity i=1.091 (i.e. 1/3 proportion of the population is selected)

It can be seen from Figure 5.13 that selection responses of diameter are obviously bigger than those of height and crown width whereas the selection responses of crown width are the smallest ones. This indicates that it may be more appropriate to use the diameter as a representative trait for growth traits in constructing an index for multi-trait selection.

Clearly, the provenance selection will obtain the highest genetic gain and followed by the family selection, the individual selection (mass selection) within-provenance and within-family selection in order (Table 5.15). The selection responses to different selection methods are in the following manner: $R_p > R_f > R_i > R_{wf}$.

With different selection intensities, corresponding genetic gains are shown in the Figures from 5.14 to 5.16. As Land (1987) proposed, a sophisticated combined selection index can be constructed for either single trait or multitraits. With this index selection method, the highest genetic gain is likely to be obtained.

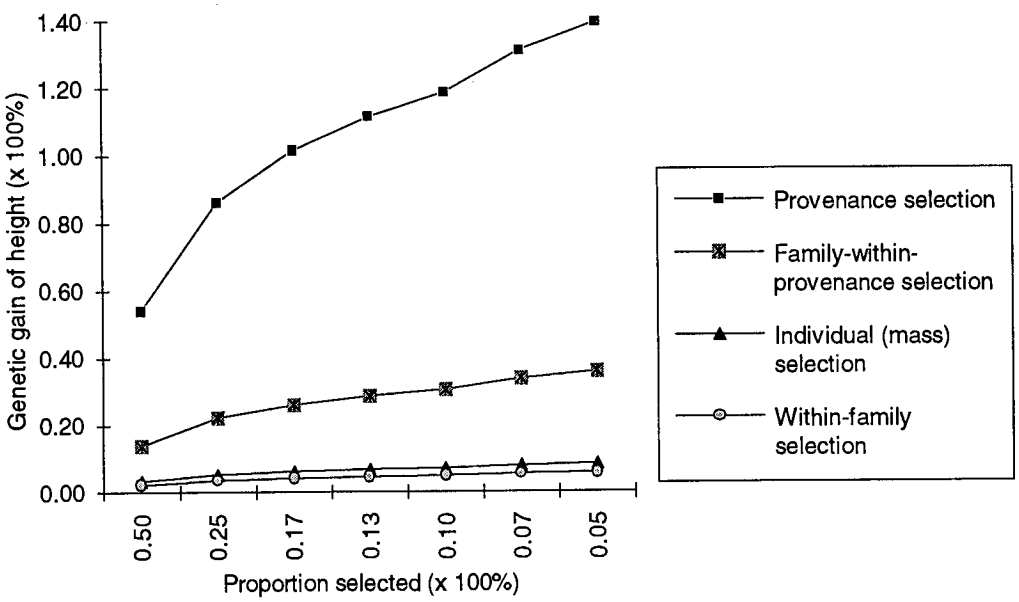

Figure 5.14 Genetic gains of height in 1993 with different selection intensity for the four selection types

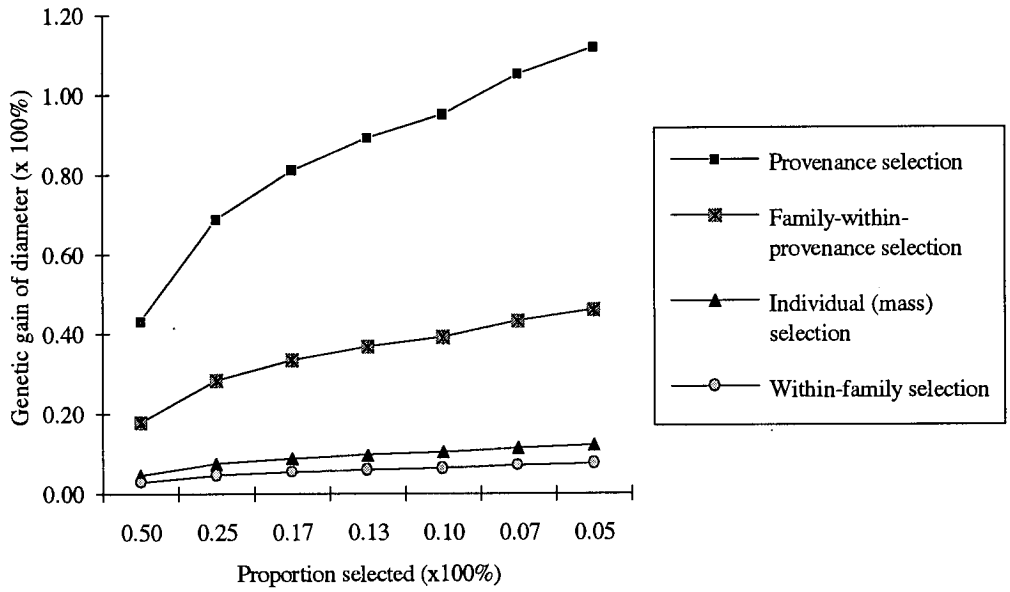

Figure 5.15 Genetic gains of diameter in 1993 with different selection intensity for the four selection types

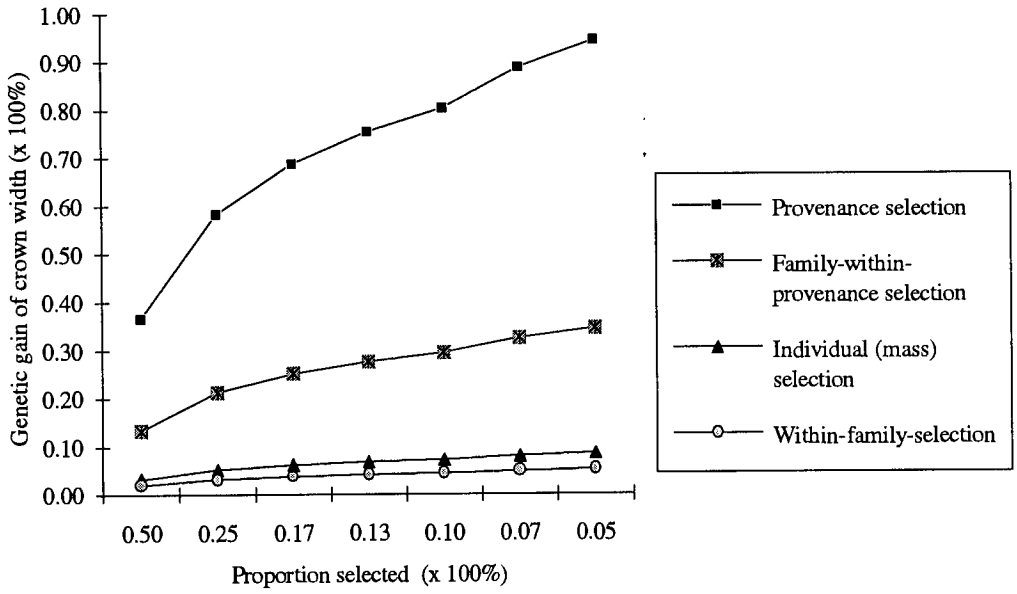

Figure 5.16 Genetic gains of crown width in 1993 with different selection intensity for the four selection types

5.3.5 Year to year stability

Genetic correlations between each pair of ages are given in Table 5.16. Although most of the correlations are significant, the correlations of height are higher than those of diameter and crown width while the values for crown width are the smallest. This implies that use of height and diameter as selection traits is more reliable than use of crown.

Correlations between the planting age and the other ages are always the smallest for all traits. An exception is that the correlation between the planting age and 6 months after planting are obviously significant for height. This is probably due to the nursery effect. The seedlings do not adapt to the new growing condition at once and their performance is still related to the nursery performance.

The values of genetic correlation coefficients generally tend to be larger as the tree grows larger.

It must be remembered that the performance of the seedlings in the nursery does not necessarily reflect the later performance after planting. To investigate optimum selection age for early selection observations for several more years are needed.

Table 5.16 Age to age genetic correlations of each traits

	Dec-93	Dec-92	Nov-91	Apr-91	Dec-92	Nov-91
	DBH				CW	
Dec-93	1	0.4358**	0.4649**	0.0302**	0.4287**	0.4715**
Dec-92	0.5225**	1	0.5086**	0.0094		0.3718**
Nov-91	0.5529**	0.3868**	1	0.3282**		
Apr-91	0.2171*	0.1593	0.7459	1		
	Height					

5.4 DISCUSSION

5.4.1 Genetic variation and implication for provenance selection

Success in the establishment and productivity of a forest tree plantation is determined largely by the species used and the source of seed within the species. The first step in all breeding programmes has traditionally been the selection of provenance to utilise the available geographic variations (Zobel and Talbert, 1984). For any tree improvement programme of any species, selection of the proper provenance or seed source can obtain the largest, cheapest and fastest gains no matter how complicated and comprehensive the breeding techniques.

Results showed that significant genetic variation exists either among the regions (islands) or among the provenances (for height and crown width) within a region. For the planting programme of *P. caribaea* var. *bahamensis* in China, the selection of islands is the initial step for obtaining genetic gain. Clearly the Abaco island and the New Providence island are the desirable regions for provenance selection which should be the second stage.

One objective of selecting trees from the best provenance is to collect the best alleles into the base population and to increase their frequency by breeding without having to return to an unimproved population (Namkoong, 1979). The most practical way of breeding of *P.*

caribaea var. *bahamensis* is to start selection at the highest possible level, i.e., selection of provenances from desirable regions.

The immediate application of these results is to guide the seed importing for the national afforestation programme in which *P. caribaea* has been taken as a major fast-growing species for industrial purposes. Furthermore, the future breeding is to be conducted on the basis of provenances selected from Abaco and New Providence islands.

Early species trials have shown that the proper area for planting var. *bahamensis* is mainly in the tropics and south subtropics covering from Hainan Island to southeastern Fujian province (Pan, 1991; Pan and You, 1991). Clearly, there is large environmental variation existing in the areas. Therefore more sites representing various types of environmental conditions need to be selected to conduct multisites testing to investigate G × E interactions. Moreover, genetic parameters can be estimated for different environmental conditions and thereby realistic genetic gains from the breeding programme can be predicted.

5.4.2 Genetic parameters and implication for breeding strategy

5.4.2.1 Breeding method

The overall purpose of a breeding programme is to obtain the highest genetic gain in a given time. Although the provenance heritability is the highest among the four types of heritabilities, the family heritability is obviously higher than those of individual (mass selection) and within-family selection. This means that considerable genetic improvement can also be achieved through family selection in addition to the provenance selection. In terms of short-term breeding, the breeding purpose is to produce genetically improved seeds for operational use as quickly as possible. Rapid improvements would be obtained by the use of the best families selected from the best provenances within the islands of Abaco and New Providence. A clonal seed orchard can be established by grafting those selected materials, thus to accelerate the production of genetically improved seeds. However, as an exotic species, the emphasis has traditionally been placed on the establishment of SSO by forward selection directly made in the trial.

It is of key importance to the success of the genetic improvement of of *P. caribaea* var. *bahamensis* that the genetic diversity and variation must be maintained. Any breeding work based only on a narrow genetic base would lead to inbreeding and genetic depression in the sense of long-term breeding. Intensive selection from the trial would lead to a narrowing of the genetic base although a higher gain is obtained in the first generation. The MPBS seems to best meet this long-term purpose (Barnes 1984; Barnes and Mullin, 1989). In the MPBS, this trial is taken as a sub-population of the multiple populations and managed as a BSO. To maintain the genetic diversity and genetic variation, selections are to be conducted only within families before mass production of seeds from the trial. Seeds which have a wider genetic base will be collected for the continued breeding programme. After this seed collection, with provenance and family within provenance selections the trial can be finally converted from a BSO into a SSO. Improved seeds from this SSO will be delivered for operational use.

It is strongly recommended that a short-term breeding programme, which is aimed at higher genetic gain and quick improvements, is carried out in parallel with a long-term one, which is aimed at maintaining genetic diversity and long-term gains. The short-term programme can be started with a few intensively selected and highly performing families whereas the long-term one should keep or create genetic variation as much as possible. Even those worst performing provenances and families should be retained in the long-term programme.

5.4.2.2 Selection strategy

The optimal breeding system is determined by the mode of reproduction, operational costs, time costs and types of gene action (Namkoong, 1979). Heritability (h^2) and selection differential (S) are mostly used to estimate genetic gain for any system. The simplest form of breeding is rapid and cheap mass selection. However, clearly it may not be as efficient and profitable as those complicated methods.

The method of combined index selection is likely to be the most effective way to achieve the highest genetic gains (Zobel and Talbert, 1984; Land *et al.*, 1987; Cotterill and Dean, 1990; Namkoong, 1979). This method is probably practically applicable although it requires more computational knowledge.

It is often desired in a tree improvement programme to improve all traits of interest thus selections are made for several traits simultaneously. Methods of improving one trait at a time in a tandem sequence or of independent culling to arrive at the same level of genetic gain have been shown to be less efficient than integrated index selection (Namkoong, 1979; Land *et al.*, 1987; Cotterill and Dean, 1990). Theoretical development and practical application of this method have been demonstrated in detail by Land *et al.* (1987) and Cotterill and Dean (1990).

An index integrating all possible information of relatives, provenances and families, and adjustment on block mean is most relevant to this combined progeny trial. Table 5.15 gives different selection responses to different selection methods. However, the expected genetic gain from combined index selection is likely to be the highest.

If there is a large demand for quick improvements of the species, this comprehensive method, combined index selection, is therefore well justified. For this trial selection indices will be constructed when the selective thinning is to be done.

Results showed that the genetic correlations between each pair of traits are significant and positive. Thus it is possible to conduct indirect selection for one trait based on the other traits while genetic gain can be obtained for both traits. In a breeding programme, it is important to improve uncorrelated traits simultaneously. In that case, a comprehensive trait may be used for a group of related traits, for example, volume can be considered a comprehensive growth trait which combines height and diameter. However, for the practical simplification, the sectional area is usually used as growth trait in index selection (Cotterill and Dean, 1990). Other traits used for index selection are most frequently the wood density and stem form. These traits are important economical traits for timber production which is the main purpose of plantation forestry of *P. caribaea* in China.

5.4.3 Age to age stability and implication for early selection

Results of age to age correlation showed that the genetic correlations of growth rates between ages are generally increasing with increasing age, implying that the trees' performance becomes more stable as they grow up. It is not possible to determine the optimal age for selection at present since only 2.5 year's observations are probably not sufficient to estimate juvenile and mature correlations. Therefore it would be unwise to investigate the possibility of early selection based only on these short-term observations. Continued observation is important to determine the optimal selection of age. It is commonly believed that the selection age for growth traits for fast growing species should be no less than 1/2 of the rotation age (Zobel and Talbert, 1984). For *P. caribaea* the selection age may be more than 10 years. However, if there is a strong juvenile to mature correlation, it is then possible to make selections at earlier ages.

5.4.4 Analysis of the combined provenance/progeny test

5.4.4.1 Balanced design

For a standard balanced design of a progeny trial, clearly defined estimates of both genetic and environmental variance components can be easily obtained. However, most tree experiments require extensive areas, which are expensive to establish and maintain, and by their blocking requirements, either incorporate large environmental heterogenities or prohibit testing on specific sites (Namkoong, Kang and Brouard, 1988). The design and analysis of unbalanced experiments for both estimations of VCOMP and testing purposes for forestry are of great advantage to improve the selection. It is believed that with the rapid development of both computer hardware and software, dealing with this unbalanced design would be greatly facilitated.

It is possible to estimate very different kinds of heritabilities from the same experimental data but for application to different breeding situations (Namkoong *et al.*, 1988). In the case of this study, the heritabilities are estimated at provenace, family, within-family and individual (mass selection) levels.

5.4.4.2 Unbalanced design and alternative methods for analysis

Imbalance is very often a source of analytical problems in the experimental design. Many analytical softwares can accommodate a reasonable number of missing trees (e.g. GENSTAT, SAS etc.) A more serious problem is unequal numbers of families representing each provenace. Land *et al.* (1987) suggested such imbalance is best avoided, but it often happens practically.

Although the estimate of variance components can be made without bias by modifying the coefficient f, the method may not be always practically suitable. It is not fully efficient unless the variance component is small (Snedecor and Cochran, 1967).

Apart from the method used in this study, there are two general approaches that have been traditionally used. One simple way is to conduct an overall analysis using provenance means and then to consider families within each provenace separately (Kanowski and Nikles, 1989; Crockford *et al.*, 1989). The number of families is that in the most poorly represented provenance. In this case, it is likely that the provenance containing only two families should be excluded from the analysis. However, this method is likely to overestimate provenance VCOMP and thereby overestimates provenance heritabilities due to the unequal representation of provenances and the inclusion of family VCOMP.

The other alternative is to conduct analyses of variance repeatedly on a large number of randomly-drawn balanced subsets of the data and the resultant estimates are pooled (Land *et al.*, 1987). Nevertheless, this approach needs much computational work and is not convenient. The error associated with VCOMP estimates will be increased because the d.f. for families within provenances and for family plots are reduced.

It was argued that such combined designs limit their values for provenance comparison and parameter estimation (Crockford *et al.*, 1989; Simons, 1992; Kanowski, 1993) and that the purposes of provenance ranking and genetic parameter estimation may be better served by separate tests. However, the principle of limiting the purpose of particular genetic tests to precise objectives has wider applicability (Foster, 1992), and may be facilitated by the judicious choice of genetic material (Foster and Shaw, 1988). Alternative analyses discussed above may also be used to overcome this limitation.

5.4.4.3 Influence of plot size

Since 4-tree plots are used in the experimental design, the within-plot variation might be somewhat increased by including environmental variability within a plot. Thus the total genetic variance of individual trees (σ_A^2) is increased, therefore the individual heritability and genetic gain might be overestimated. An alternative way is to conduct the analysis repeatedly by selecting one tree from each plot at random and the results are averaged. However, there is uncertainty over how many times would be acceptable. A sampling error would be also associated with the random selections.

Theoretically, the way to overcome the problem of microsite differences within plots is to reduce the plot size. The ultimate is, of course, the single-tree plot. However single-tree plots would raise other problems, such as missing plots, difficulty in laying out the experiment and the competition with unlike adjacent plots (Libby and Cockerham, 1980; Matheson, 1989a, b). To overcome these problems, Libby and Cockerham (1980) proposed the use of non-contiguous plots instead of single-tree plots, i.e. there could be several trees of each family located at random within a replication rather than being located together as a row or square plot.

Lambeth, Gladstone and Stonecypher (1983) suggested that within-family selection is more difficult if family members are not adjacent to one another, and the thinning by family while retaining reasonably uniform spacing is difficult. They argue for non-contiguous plots because of their advantage over multi-tree plots because the plot members are exposed to the microsite variation within replicates (Blocks) in a random manner.

Matheson (1989b) proposed that non-contiguous plots have no advantage over single-tree or contiguous multi-tree plots. Although non-contiguous plots reduce the effect of within-plot environmental variability, they do not resolve the problems in laying out the trial and in thinning. He suggested that the use of an incomplete block design, such as a lattice design, would have a high degree of control over error.

In the case of the genetic improvement programme of *P. caribaea* var. *bahamensis* in China, the trial is also to be managed as a BSO in which genetic thinnings will be carried out. The

single-tree plot is unlikely to be preferred. The non-contiguous plot does not overcome the problems raised by the single-tree plot. It seems to have no advantage over the multi-tree contiguous plot.

5.4.4.4 Simplification of the analysis by conducting it on plot-means

It is very often the case in field trials that some trees in a plot are missing and that the data files are very large. Therefore, the ANOVA based on individual values is not convenient in terms of computer memory size and computational time. In many cases the ANOVAs and ANCOVAs can not be conducted due to the huge size of the data. However, to conduct the ANOVAs on plot means rather than individual values can overcome these problems. The within plot error was thereby pooled into the experimental error which includes family (within provenance) by replication interaction. The equation (5.1) is then changed to:

$$X_{ijkl}=\mu+R_i+P_j+F_{j:k}+(PR)_{ij}+\varepsilon_{ijkl} \qquad (5.18)$$

With this model, analyses were done using plot-means, and estimates of within-plot variance for each trait are calculated separately. The ANOVA and expected mean squares are given in Table 5.17.

Table 5.17 The ANOVA and expected variance components for a balanced combined provenance/progeny trials on plot-means.

Source of variation	D.f.	Expected mean squares
Replicate	*R-1*	$EMS_r=\sigma_e^2 + f\sigma_{PR}^2 + Pf\sigma_R^2$
Provenance	*P-1*	$EMS_p=\sigma_e^2 + R\sigma_f^2 + f\sigma_{PR}^2 + Rf\sigma_P^2$
Replicate x Provenance	*(R-1)(P-1)*	$EMS_{rp}=\sigma_e^2 + f\sigma_{PR}^2$
Family/Provenance	*P(f-1)*	$EMS_{f(p)}=\sigma_e^2 + R\sigma_f^2$
Residuals	*P(f-1)(R-1)*	$EMS_e=\sigma_e^2$
Total	***RPf-1***	

σ_e^2 : *Residual variances including within plot variance and family by replicate interaction.*

where $\sigma_e^2=\sigma_w^2/n+\sigma_{fR}^2$ (Westfall 1992). The estimation of σ_w^2 requires the ANOVA to be conducted on individual values which is often limited by the available CPU time. Assuming that family (within provenance) by replicates interaction is negligible, i.e. σ_{fR}^2 is ignored, then $\sigma_w^2=n\sigma_e^2$. This would overestimate the σ_w^2 if family by replicate interaction is not negligible.

5.4.4.5 Genetic correlations and linkage disequilibrium

The type and magnitude of the genetic parameters that can be estimated from a provenance/progeny test are affected by the parental sampling design used (Land *et al.*, 1987). Relatedness among parental trees, environmental differences between the origins of parents, and linkage disequilibrium among provenances are all determined by the parental sample.

The coefficient of relationship r has been adjusted from 1/2 to 1/3 to overcome the overestimation of the heritability. However, the environmental differences between provenances of the parent trees in the test will affect the size of the VCOMP of provenances (σ_p^2) and the possibility of linkage disequilibrium. Range-wide provenance/progeny tests may therefore not be as suitable as limited-range tests for selection and breeding, particularly for species with wide distributions. Obviously, the estimated genetic gains can be increased by using a wide-range test, because the total genetic variance (σ_A^2) will be increased and the test mean will be decreased by including some poorly adapted sources. Furthermore, effects of linkage disequilibrium will be confounded with provenance effects (Harvey and Townsend, 1985; Land *et al.*, 1987), thus the actual genetic variance and genetic gains will be less than estimated. For an operational breeding programme, the simplest way to analyse the trials would be to use only a subset of the data containing information for provenances within an acceptably-adapted breeding zone.

5.5 Conclusion

Through 2.5 years' observation, it was found that significant genetic variations are existing in growth traits at region (island), provenance and family levels. The growth rates are strongly correlated to the latitudes of provenances. Northern provenances which are adapted to higher

rainfall grow faster than southern ones. Abaco island and New Providence should be considered the regions for seed importing for the current national afforestation programme because of their fast growth. Seed importing from Andros island should be avoided. If a narrow crown is the breeding purpose, the New Providence island may be selected because it has fast growth rate and a relatively narrow crown.

The highest genetic gain can be obtained by the simplest and cheapest way of provenance selection from the trial. Then family selection can be carried out within the selected provenance. Genetic gain from family selection is the second highest. Genetic gains from Mass selection and within- family selection are the third and the least, respectively. Continuous observation of the trial is needed to investigate the traits of wood quality and stem form, which are likely to be less correlated with each other. Then an integrated index can be constructed for multitrait selection. The genetic gain from such index selection integrates all relatives' information and incorporates all traits of interest and will be the highest one. A comprehensive trait representing the strongly correlated traits, for instance the diameter (or sectional area) representing growth traits, can be used in constructing the multitrait selection index. Optimal selection age for growth traits cannot be determined yet because we are unable to estimate the juvenile and mature correlation at these early ages.

A long-term breeding programme, which is aimed at maintaining high genetic diversity and a wide genetic base, needs to be carried out in parallel with a short-term programme, which is aimed at high genetic gains and quick improvements. The long-term programme must retain as many provenances and families as possible to make full use of existing genetic variations among them whereas the short-term programme should make use of the best families within the best provenances by intensive selection.

To increase the precision of the estimation of genetic parameters and genetic gains, limited range provenances should be used instead of range wide provenances and a range wide provenance test should be conducted previously as a first stage test, because the heritabilities and expected genetic gains are probably overestimated due to the great provenance differences. The imbalance of the design should also be avoided by trying to use equal number of families within each provenance as much as possible. To investigate the genotype by

environment interactions, multi-site tests are needed to evaluate the cross-sites performance of the testing materials.

6. Breeding strategy

6.1 Introduction

Forest tree breeding is aimed at increasing the quantity and quality of forest products. In the past, traditional tree breeding efforts have been directed more towards immediate operational success than at the conceptual representation and strategic planning of breeding activities (Kanowski, 1993). It was not until the 1980s that conceptual models of tree improvement programmes and strategic plans for breeding activities were widely reported (Cotterill, 1984; Kang and Nienstaedt, 1987; Mullin and Park, 1992; Namkoong *et al.*, 1980; 1988; White *et al.*, 1992).

Classical tree breeding programmes are usually based on initial selection of a number of plus trees selected from natural stands and plantations, followed by full-sib progeny tests and recurrent selection from the tests. Subsequently clonal seed orchards may be set up for commercial seed production. However, inflexibility of traditional breeding methods has often been apparent. There are difficulties with incorporating new material into the breeding programme, balancing short-term genetic gain with the maintenance of long-term genetic variability and in exploitation of the genotype by environment interaction (GEI) (Barnes, 1984; Barnes and Mullin, 1989; Barnes *et al.*, 1995). These shortcomings are more serious for advanced generations, and particularly for exotic species where the introduction of new genetic material is of obvious importance.

To overcome the problems encountered in traditional breeding programmes, a few breeding strategies employing a structured breeding population have been proposed so that the genetic variability can be maintained and is available for breeding whilst short-term genetic gain is obtained. Typical examples are: hierarchy of populations, such as Nucleus Breeding and Elite Breeding, which promotes an emphasis on superior material (Jackson and Turner, 1972; James, 1977; Cameron *et al.*, 1988; Cotterill and Cameron, 1989; Cotterill, 1989; Cotterill *et al.*, 1989; White, 1992); multiple population breeding system (MPBS) which emphasizes interpopulational diversity for diverse traits and environmental adaptability (Namkoong, 1976; Burdon and Namkoong, 1983; Namkoong *et al.*, 1988; Barnes, 1984; Barnes and Mullin, 1989; Barnes *et al.*, 1995; White, 1992; Eriksson *et al.*, 1993); and breeding groups usually described as Sublines Breeding, which focuses on controlling relatedness among members of

the production population and hence controlling the inbreeding in the offspring to be planted operationally (van Buijtenen, 1976; Burdon *et al.*, 1977; van Buijtenen ad Lowe, 1979; McKeand and Bieneke, 1980; Burdon and Namkoong, 1983; White, 1992). In practice a combination of these breeding structures is often necessary (Burdon, 1986) in some ongoing breeding programmes (Kanowski, 1993). Obviously, structured breeding strategies are best suited for advanced generation breeding.

So far, *P. caribaea* including all three varieties has proved to be a very successful plantation species in China. Genetic structure and patterns of variation have been investigated for two of the three varieties. How to make use of the genetic variation for breeding purposes, how to use effectively the limited genetic resources and in what way to breed *P. caribaea* to optimise genetic gains in the short- and long- term are questions to be answered before implementing any breeding activity for the species. A breeding strategy making use of the results reported in this thesis is necessary to guide the breeding programme in China. This chapter presents a breeding strategy which covers all the three varieties with particular emphasis on var. *bahamensis* because a quantitative genetic analysis for the variety is also reported in the previous Chapter. Emphasis is placed on the first generation breeding.

6.2 Multiple population breeding system and its relevance to the Chinese situation

6.2.1 Current breeding and its features

Pinus caribaea has been an important industrial tree species in Southern China. The introduction, domestication and breeding of the species have been described in Chapter 1. The early emphasis of breeding of *P. caribaea* was on evaluation of genetic variation in growth, adaptability and form of varieties and provenances. Based on that information, varieties and sources are being recommended for commercial planting in Southern China. Several seed orchards or seed production stands were established based on phenotypic selection from field trials and plantations. Seed orchards were set up following the traditional method of plus-tree selection and establishment of grafted clonal seed orchards. Seed production stands were converted from plantations through selective thinning (Wang *et al.*, 1995). However, the genetic gains obtained were small due to the inappropriate selection criteria used. Moreover, the genetic base of these resources is limited because most of the resources were derived from the early introductions. This limitation may lead to loss of genetic diversity and to inbreeding

depression, which will eventually cause the loss of genetic gain from the long-term point of view.

It was not until the early 1990s that a comprehensive cooperative tree programme on the species was started, involving international and national organizations, and provincial and local forestry organizations, with coordination by the Chinese Academy of Forestry (CAF). The lack of a well defined breeding strategy is an urgent issue to be solved in the present improvement programme. The main focus of the programme is now placed on genetic improvement on a nation-wide scale and on breeding, testing and selection for long-term improvement.

The main objective of the breeding strategy is to obtain maximum genetic gain within a given time while maintaining genetic diversity and variability for long-term continuous productivity. In other words, any effort for breeding *P. caribaea* in China should be made to develop an efficient, cost effective breeding strategy to ensure both short- and long-term benefits for the tree breeders involved.

As an exotic species, the breeding of *P. caribaea* in China has several unique features. Firstly, the programme involves multiple varieties. All three varieties of *P. caribaea* are to be genetically improved. A second feature is that there are limited genetic resources available for the breeding programme. Most of the present resources are derived from the early introductions which were found to be of limited genetic base. Thirdly, the programme deals with diverse environments and multiple traits. Environmental conditions change substantially from place to place in the area for planting. Because *P. caribaea* is grown mainly for timber and pulp production, multiple traits such as wood density, fibre length, growth rates, stem form and pest resistance are to be improved. Finally and importantly, financial constraints affect the breeding strategy to be used.

As a result, a primary criterion for the breeding strategy is that it be flexible for current and future breeders and be flexible in accommodating the diverse objectives for different end products and for different environmental conditions. Another concern is that whenever new information becomes available, the breeding plan should be flexible enough to incorporate the information. Finally, a less expensive but effective breeding method is preferred. Among the

three commonly used breeding strategies mentioned earlier in this chapter, MPBS appears to be the most appropriate one for breeding *P. caribaea* in China at least for the first generation breeding. For advanced generation breeding, the Nucleus breeding can be incorporated with the MPBS.

Information obtained from the genetic structure and breeding system studies reported in the previous chapters is useful for the present breeding programme in sampling natural variation, designing mating systems, managing seed orchards and structuring breeding populations. On the other hand, results from the quantitative genetic analysis facilitate decisions on breeding methods, experimental design and selection strategy.

6.2.2 Outline of the Multiple Population Breeding System

6.2.2.1 Population structure

In the multiple population breeding system, separate populations from a single species are grown in different environments or to produce trees with different combinations of traits. This system is designed to handle multiple objective breeding, including the objective of adapting to multiple environments and the objective of improving multiple traits (Namkoong *et al.*, 1988; Barnes 1984). The main idea of MPBS is that breeding activities concentrate on separated sub-populations for different purposes, (for instance, different traits and environments), so that the genetic variability within the species can be maintained and even created for long-term breeding while rapid improvement is obtained by the breeding within each sub-population called BSO (Breeding seedling seed orchard). The objective is to maintain or create genetic diversity between these sub-populations while applying heavy selection pressure within them. The theory is that a different gene complex is maintained or created in each sub-population. The multiple population preserves genetic variation by the conservation of genes in packages in usable genotypes rather than in trees with characteristics which were found in the natural forests and which have no immediate silvicultural use (Barnes, 1984).

6.2.2.2 Breeding seedling orchard

The BSO is central and essential in the proposed multiple population breeding strategy and combines the functions of breeding, testing, selection and seed production (in the form of archives, progeny tests, plantations and seed orchards) which are being found to be

prohibitively expensive and time consuming in conventional breeding programmes. It can vary from being simple and small to being complex and large depending on the importance of the species, the potential value of the material for breeding and its value for commercial seed production. It has great flexibility in function, size and number. Ideally, separate sub-populations (BSOs) can be created by selection of families and individuals which perform best in each environment. Sub-populations of plus trees can be kept separate by environment with the aim of exploiting and developing GEI even where it has not been measured. Similarly, sub-populations can be created by selection of different traits which are to be improved.

Each BSO is treated as a sub-population in the MBPS for the first generation. The BSO has to be designed to fulfill all four functions as much as possible since there are always some compromises between each function. The BSO will be thinned progressively and selectively, using available genetic information until just the seed producers remain and it is from this population that the founder members of the next generation are selected. The thinnings could be optional depending on the design (spacing) used for the BSO, i.e. the degree of competition among trees in the BSO. The intensity of the breeding method used for the sub-population will depend upon the importance of the species, the importance of the sub-population itself, the construction of the sub-population and the population size.

6.2.2.3 Advantages and disadvantages

The advantages of the MPBS are the flexibilities in setting up sub-populations, incorporating new materials, breeding for different environments and for multiple traits and in maintaining and creating genetic variability for long-term breeding. A major disadvantage of the MPBS is that the maintenance of multiple populations may be expensive. This disadvantage can be avoided by establishing a composite breeding seedling orchard (CBSO) (Barnes *et al.*, 1995) for advanced generations or by a shift to other breeding methods like Nucleus Breeding in which emphasis can be placed on a few superior genotypes (Cotterill and Dean, 1990).

6.3 Proposed breeding programme

6.3.1 Resource population (Base population)

Resources of *P. caribaea* available for the breeding programme in China comprise experimental populations and production plantations such as commercial plantings. The

genetic base of these resources is of limited scope as often occurs with an exotic species. Although the species is well known and many breeding programmes exist around the world, inclusion of selections from natural populations is crucial to the breeding programme in China. Extra materials need to be selected from its natural populations to broaden the genetic base of the resource population.

6.3.1.1 P. caribaea var. *bahamensis*

Resources for var. *bahamensis* include an extensive seeds collection provided by OFI. It is a range-wide seed collection and should contain most of the natural genetic variation. This material is the major genetic resource for breeding the variety. More selections will be made from plantations in China and possibly imported from other breeding programmes in non-native countries.

6.3.1.2 P. caribaea var. *caribaea*

Similar to var. *bahamensis*, an extensive seed collection of the var. *caribaea* was recently available for the breeding programme as the major genetic resource. It consists of 220 open-pollinated families representing the natural stands and a clonal seed orchard in Cuba and land races developed in Brazil and China. The plantation area of this variety is the largest among the three varieties.

6.3.1.3 *P. caribaea* var. *hondurensis*

The major resources available are those provenance trials and plantations of early introductions. Plus tree selection from experimental stands and commercial plantings can be carried out to form the breeding population. Most breeding programmes of *P. caribaea* around the world focus on this variety due to its fast growth and large genetic variation. Therefore incorporation of materials from other countries may be useful for rapid progress in improvement.

6.3.2 Breeding populations

6.3.2.1 Number of sub-populations

The number of sub-populations which should be created will depend upon the resources available, the diversity of the species and the environmental and utilization spaces which need to be filled. The more sub-populations there are, the quicker can be the response for a strain to produce what is required in a new or changed environment. In this programme, sub-populations are set up for the objective of breeding for different environments.

Based on the genetic information available so far for the species, the potential area for planting *P. caribaea* covers different environments which are mainly in Hainan, Guangdong, Guangxi and Fujian provinces. For var. *caribaea*, studies (reported in previous chapters) of selectively neutral isozyme makers revealed obvious variation among populations, although little genetic variation was shown in selective quantitative traits in early provenance trials using a small number of provenances. In contrast, the other two varieties showed substantial genetic variation for quantitative and adaptive traits. Considering the suitable planting zone and its variable site conditions, and the amount of genetic variation detected so far in the field trials, for the first generation, it is desired that at least three sub-populations (BSOs) are needed for the var. *bahamensis*, two populations for var. *caribaea* and two for var. *hondurensis*. Sites for these BSOs will be chosen in Leizhou Peninsula, Southeast Guangdong or southwest Fujian, and Hainan to represent various environmental conditions.

However, the MPBS is expensive due to the multiple populations involved. Restricted by the availability of both resource material and finances, the breeding programme has, so far, only established one BSO for each of var. *bahamensis* and var. *caribaea* in Guangxi and Guangdong respectively. Therefore the breeding work proposed will be based on these single breeding populations. More sub-populations will be established when finance and resources (particularly exotic material) become available. These will become essential when the programme moves to the advanced generations.

It is one of the advantages of MPBS that new material can be incorporated at any time. This means that new breeding populations can be established when resources and finance are available.

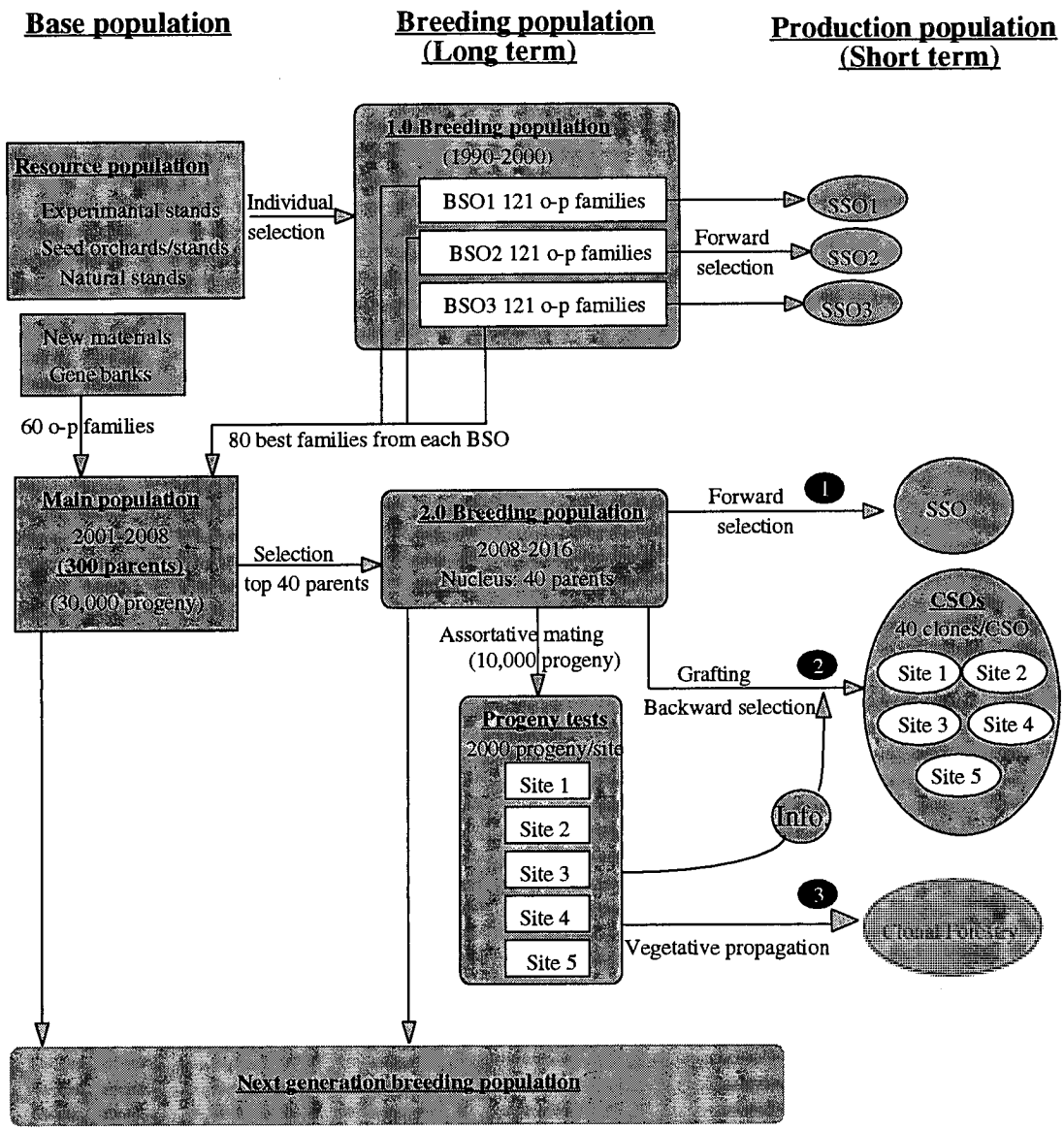

Figure 6.1 Schematic diagram of 1st to 2nd generation breeding strategy for var. *bahamensis* in China. First generation SSO (converted from BSO) will be established in 2000. Second generation breeding population will be set up from 2001 (main population) to 2008 (Nucleus). Three options for production populations depending on finance and vegetative propagation techniques are shown. Breeding strategies for the other two varieties will follow similar procedures. O-p refers to open pollination, CSO refers to clonal seed orchard, BSO refers to breeding seedling orchard and SSO refers to seedling seed orchard.

6.3.2.2 Population size

Population size of the breeding population has an important effect on either short- or long-term genetic gain of a breeding programme. A large starting population is desirable for fixing some favourable alleles, particularly those with low gene frequencies, and to avoid their loss by

chance (genetic drift) from generation to generation (Robertson, 1960). Larger population sizes are needed to ensure fixation of rare alleles, whereas smaller populations are required to achieve fixation of common alleles or even those of intermediate frequencies (Kang, 1979; White, 1992). An effective population size of about 50 is enough to ensure fixation of all but rare desirable alleles through many generations (Robertson, 1960; Bulmer, 1985). Neutral alleles and uncommon genes could also be fixed with this size of population (Kang, 1979). However, assumptions made in the theoretical studies may not be practically realistic in real breeding. As a consequence, larger initial population sizes (200 to 500) have commonly been used in many breeding programmes (White, 1992; Citations therein).

Ideally, the number of individual founder trees for breeding *P. caribaea* in China should be as large as possible especially for the first generation of breeding, given the large genetic variation existing within the varieties and the difficulties in obtaining genetic material, an obvious problem met with in many exotic species. Since a BSO is the main component of the MPBS in the first generation, the initial number of families in each BSO largely depends on the availability of resources. 120 to 250 open-pollinated half-sib families will be used in each of the BSOs for the three varieties. The BSO for the var. *bahamensis* which is located in Hepu, Guangxi province used 121 families and the other two BSO of similar size to be set up would increase the total breeding population size to about 360. One BSO for the var. *caribaea* has just been established in Guangdong in 1995 and used 220 open-pollinated families, all the resources available so far. Two more BSOs need to be set up in future and a small number of families could be used since genetic variation detected within this variety is smaller than that in var. *bahamensis* and a large number, 220, is already used in one BSO. For the var. *hondurensis*, two BSOs of size of 150 parents will be established.

6.3.2.3 Establishment and management of sub-populations

Limited genetic resources is a major limitation of the breeding programme. Although selection from the present plantations will ensure a certain amount of genetic gain, a disadvantage pertaining to this is that the narrow genetic base of the genetic resources would lead to a risk of inbreeding and hence loss of genetic gain in terms of long-term breeding. It is desirable that the establishment of a BSO includes a range-wide collection from the natural populations designed to sample as much natural genetic variation as possible. Provenance and family identity need to be retained, because these materials are also to be used for genetic testing.

Sites will be selected where the surrounding plantation species are not the same species. At the Hepu site where one BSO of var. *bahamensis* was established, different species, mainly *Eucalyptus*, were planted around the BSO. The site of the first BSO of var. *caribaea* was in a lake-like condition in which 6 small islands isolated by water were used for planting the BSO.

Each BSO will be designed as a combined provenance/progeny test and planted across different sites. The advantages of this integrated test have been discussed in Chapter 6. One such BSO has already been set up for var. *bahamensis* and var. *caribaea*. The one for var. *bahamensis* was planted on a single site while the one for var. *caribaea* was subdivided into six groups according to their geographic origin and planted on six water-isolated islands in a lake-like area. Both BSOs were isolated from plantations of the same species. The idea for this subdivision is that isolation is good for the purpose of testing where there is a need to evaluate family performance within provenances and good for seed production. These subgroups are not suitable to serve as multiple populations in the breeding strategy because they are in the same area and do not represent different site conditions for future planting, therefore they cannot be used for evaluation of GEI. For these two varieties, new sub-populations are to be set up using the same resources but inclusion of new material is desired, particularly for var. *bahamensis* because no selection from plantations was included in the first BSO. Obviously there will be no material left from the first BSO for new sub-populations. The first BSO will be used as a major resource for the new sub-populations. Seeds will be collected around year 2,000 when the BSO becomes mature.

For the var. *hondurensis*, there are two options of setting up the BSOs. One way is that used for the other two varieties, that is to conduct a range-wide seed collection from the natural populations. The other is to carry out a traditional plus tree selection from the present plantations in China and possibly import material from non-native countries. The latter seems to be more practicable according to the present situation in China.

This capability of establishing sub-populations at various times is another advantage of the MPBS.

Since a large number of families is used in the BSOs, the Incomplete Blocks experimental design is recommended. As discussed in Chapter 5, this type of design minimizes the area

required and secures the future statistical analysis. Multi-tree plot is used to avoid imbalance caused by missing plots. For example, the BSO of var. *bahamensis* established in Hepu consists of 121 open-pollinated families. A Lattice Square design was used with 3 by 3 trees within each plot and initial spacing of 2 by 1.5 m and 5 replicates were used.

The BSOs will be selectively thinned to avoid tree to tree competition and to remove inferior genotypes before they become a seedling seed orchard. This can be done through 2 to 3 optional thinnings but at least one is needed.

Figure 6.2 presents the management scheme for the var. *bahamensis*. For the var. *caribaea*, a similar design was used but with a larger number of families included in the BSO (220 families). The BSO of var. *caribaea* was established in 1995, five years later than var. *bahamensis*. Each of the activities for the BSO of var. *caribaea* will be conducted 5 years later corresponding to var. *bahamensis*. Similar designs will also be used for the BSOs of var. *hondurensis*.

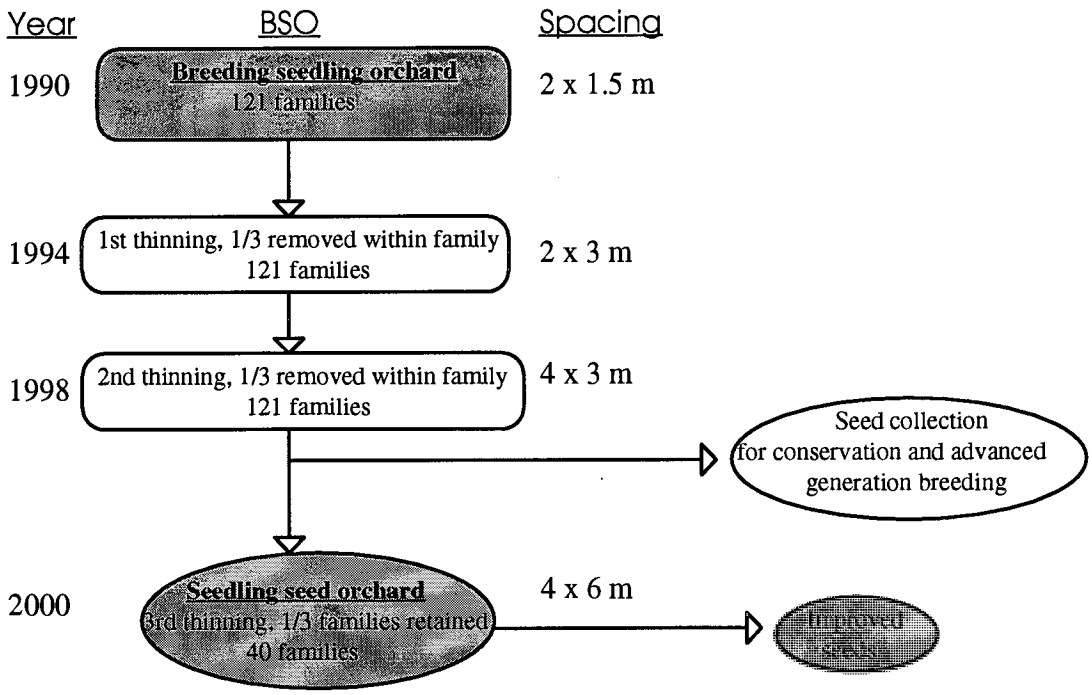

Figure 6.2 Procedures for converting the BSO of var. *bahamensis* established in Hepu county to first generation SSO.

6.3.3 Mating design

Chapter 3 has reported the study of the breeding system of var. *caribaea* and has shown that the species is predominantly outcrossing and that in a seed orchard inbreeding depression can be kept at a reasonably low level by appropriate orchard design. A high outcrossing rate was also found with the other two varieties (Matheson *et al.*, 1989). These mean that no supplemental pollination is needed for seed production in the seed orchards. In the first generation, a simple mating design, with an open-pollinated scheme will be used in the BSOs. It is cheap, easy and effective. This mating design makes use of general combining ability (GCA) rather than the specific combining ability (SCA) used by controlled pollination.

In the second and advanced generations, controlled pollination e.g. assortative mating (best with best) (Carson, 1986; Lindgren, 1986; Cotterill *et al.*, 1987; 1989) will be used in the nucleus population where the focus is on fewer superior genotypes. This can maximize genetic gain per cross and ensure an increase in the frequency of desirable alleles.

6.3.4 Selection strategy

6.3.4.1 Traits to be assessed

The most important economic traits of *P. caribaea* include wood density, fiber length, growth rates, and form quality as the major use of the *P. caribaea* in China is for timber and pulp production. The adaptability refers to the pest resistance (tip moth is the major pest). In addition to improvement in such traits as wood quality and growth rates, pest resistance and stem form (foxtail) especially for var. *hondurensis* need also to be taken into account as they have strong direct effects on those economically important traits. The large genetic variation found in these traits ensures genetic gains by selection.

6.3.4.2 Selection method

The purpose of the breeding programme is to improve the important economic traits mentioned above. Thus an effective way which can maximize the improvement for all the traits as quickly as possible is needed. As found in many other tree species, there are strong genetic correlations among growth rates of *P. caribaea* (Table 5.13). The wood density and stem form are usually negatively correlated with growth rates for most tree species. These

mean that genetic improvement cannot be made for a trait by selection for another trait, i.e., indirect selection. There are many selection methods available for improving single traits such as tandem selection, independent culling, etc. (Cotterill and Dean, 1990). However, these need long selection intervals and are not the optimal way regarding resources required and the gains returned. There is a need for the breeding strategy to balance alternative selection methodologies (Cotterill, 1986a, 1986b; Kanowski, 1993). The development of index selection specifically for tree improvement (Cotterill and Dean, 1990; Magnussen, 1990, 1991) has greatly facilitated the selection methods.

Combined index selection is probably the most effective way to achieve maximum gains for all traits of interest simultaneously (Cotterill and Dean, 1990). Thus the index selection will be carried out in each of the BSOs. All trees in the BSO will be ranked according to their index values. Construction of the index will be based on the quantitative genetic analysis of the combined provenance/progeny test. It makes use of not only the information of individuals, but also of provenances and families. The environmental effects on individual performance will be adjusted with their block means (Cotterill and Dean, 1990). Traits involved in the index construction can be weighted with their economic importance.

As the breeding programme moves to advanced generations, new sub-populations can be set up for improving single traits like wood density or stem form, so that a simple selection method can be employed within that sub-population. This is another advantage of the flexibility shown by the MPBS.

6.3.4.3 Selection intensity and genetic gains

Selection intensity may vary among different sub-populations depending on the variability within the sub-populations and the initial spacing used. Take the BSO of var. *bahamensis* as an example, 1/3 of the total number of families (i.e., 40 families) will be finally retained in each of the BSOs to produce seeds as a SSO (Seedling Seed Orchard). This intensity ensures a certain amount of genetic gain while it does not cause loss of too much genetic diversity from the selection. The genetic gains from various selection intensities for var. *bahamensis* were predicted in Chapter 5 (Figures 5.14 to 5.16). For example, if 1/3 of the total population is selected, the genetic gain for tree height will be more than 20% for family selection and about 80% for provenance selection. Even higher genetic gains can be obtained when

combined index selection is used because it combines information on provenance, family and individual. However, one should be aware that there is a compromise between genetic gain and diversity. A higher selection intensity leads to greater genetic gains but lower genetic diversity remaining. It is advisable to obtain greater genetic gains by setting up a clonal seed orchard using the few best individuals of the best families from the best provenance. However, a decision needs to be made to justify the balance between the investment and return (gains) of the extra CSO.

6.3.4.4 Early selection

Due to the long rotation intervals and juvenile-mature transition for many important traits of industrial trees, indirect selection is widely used. Early selection has received most attention (Barralho *et al.*, 1992; Burdon, 1989; Kang, 1991; Kremer, 1992).The rotation interval of the plantation of *P. caribaea* in China is 15 years. For growth traits, the selection at a half rotation age namely 7 to 8 years can provide a reliable prediction for the rotation age. However, for wood density, it is likely that evaluation after 10 years of age would be reliable although there is no existing information available on this. At least half rotation age is recommended for many fast growing plantation species (Zobel and Talbert, 1984). Continuous measurements for assessment of the age:age correlation need to be done for *P. caribaea.* Investigations of the physiology and early development of the trees would help the application of early selection.

6.3.5 Production population

6.3.5.1 Seedling seed orchard

Seedling seed orchards (SSO) are the major bodies in the first generation to produce improved material for operational planting. The SSO will be directly converted from the BSO by several selective thinnings. To ensure a reasonable genetic gain and fairly good control of inbreeding (or genetic diversity), 40 families and 3 trees per family will be retained in the SSO. The provenances performing better may have larger representation in the SSO. This method is cost-effective and especially suits the situation in China for the first generation breeding of exotic species because of its simplicity, ease and inexpensiveness (Figures 6.1 & 6.2).

6.3.5.2 Clonal seed orchard

Despite the emphasis that the MPBS places on BSOs in the first generation, clonal seed orchards may still have an important place in advanced generations. Clonal seed orchards are probably the most effective way of producing improved seeds for operational use. Whenever a selection is made in the breeding population, a clonal seed orchard using the few best individuals can be set up for rapid production of genetically improved seeds (see Figure 6.1).

6.3.6 Insurance of genetic diversity

For an exotic species, it is important to maintain a wide genetic base for future breeding. Although the MPBS is designed to maintain genetic diversity for long-term breeding, the genetic thinnings conducted in the BSOs (conversion of the BSO to SSO) will inevitably lead to some loss of genetic diversity. Because most of these resources are exotic and have not yet been extended to operational plantings, it is often practically desirable and easily applicable to establish gene banks (or gene archives). They are usually established along with the field trials using the extra planting stock from the experiments. In some sense, this type of gene bank is somewhat like an insurance for the genetic diversity for any accident that might happen to the experiments in the future. One such gene bank has been set up for each of the var. *bahamensis* in Hepu County and var. *caribaea* in Guangdong province. Additional ones will be set up along with each new BSOs. Gene banks would be of great value to the long-term genetic improvement and should be designed to collect as many genotypes as possible. Selections for advanced generations breeding will introduce new materials from it. Single line plots with no replication will be used in the gene bank to make it easy to manage.

6.3.7 Regeneration of breeding cycle

It is still early to decide the breeding method for the second generation since genetic information on GEI and genetic parameters is not available. However, the Nucleus Breeding is recommended for the next generation because superior genotypes are established in the first generation and the focus is on these superior genotypes in order to obtain maximum genetic gains. Two options are advisable for establishment of the base population in Nucleus Breeding. Since maintaining multiple populations is expensive for more generations, a large single population for each variety will be established for the second generation as the base population of the Nucleus Breeding. This base population is equivalent to the Composite

Breeding Seedling Orchard (CBSO) (Barnes *et al.*, 1995) but the breeding will focus on a small part (the nucleus) out of it. Materials will be selected from the BSOs and new materials will be introduced from any other resources available. This population will be located in the mainland. A second such base population (or CBSO) for the next generation may also be established on Hainan island if financial constraints disappear. Then the Nucleus breeding will be carried out within each of these base populations. The other option is that each BSO will be taken as the base population and Nucleus Breeding will be conducted within each of them. This option still needs to maintain the multiple populations. There are three alternatives for setting up production populations for the second generation. These are SSO converted from the Nucleus, CSO and mass vegetative propagation (Figure 6.1).

In the first generation, half-sib family identities are retained and selection is done by a combination of between-family genotypic and within family phenotypic selection. In the second and advanced generations, there is complete pedigree control with genotypic selection between full-sib families and phenotypic (or even genotypic if it is clonal) selection within families using heritabilities, genetic correlations and general and specific combining abilities. Pedigree control is used not only for family selection, but also for controlling the genetic contribution to the next generation, ensured through controlled pollination and full-sib families selection and therefore there is no chance of genetic leakage through stray pollen contamination.

6.3.8 Application of biotechnology and hybridisation

New biotechnology is rapidly moving forward and its wide application in forestry research and practical forestry has been demonstrated in many countries for many species such as *in vitro* propagation. The technique of *in vitro* mass propagation has been implemented with Eucalypt species in China. A mass production plant has been established for *Eucalyptus* breeding and has already shown great economic advantage. When the breeding programme moves to the advanced generations, applications of biotechnology used for other species may be considered for *P. caribaea*. The potential benefit from application of biotechnology is quick multiplication of desired genotypes and cheap investment although the development and establishment of such technology itself is expensive.

Hybridisation between *P. caribaea* and *P. elliotii* has been very successful. In Australia, these two species are being replaced by their hybrids which showed superior growth, wood quality and adaptabilities (Nikles and Robinson, 1989; Dieters, 1996). In China, the hybrid of these two species is under field testing and the potential has not been evaluated yet.

6.4 Discussion

This chapter proposes an effective, simple and relatively easy breeding strategy for the first generation and a relatively intensive breeding strategy for the next generation. The theoretical background for adopting this breeding strategy is the large genetic variation and the very low proportion of self-pollination of the species. The MPBS provides the flexibility that accommodates the major problem caused by limited genetic resources for breeding. It has various flexiblilities in responding to new materials and new demands and is suitable for breeding exotic species.

Three BSOs for var. *bahamensis* and two for each of the other two varieties are proposed. However, so far only two of such BSOs have been established. Five more BSOs need to be set up. This is important for the breeding programme. Setting up multiple BSOs under various site conditions is also intended to study the genotype by environment interaction which is important for the breeding programme. Therefore, establishment of more BSOs is the major task for the breeding programme in the near future. As the breeding programme moves to advanced generations, the potential of applications of biotechnology and hybridisation need to be evaluated and studies on such areas should be conducted in parallel with the breeding programme.

As given in the example (Figures 6.1 & 6.2), it will take at least 10 years, from 1990 to 2000 for var. *bahamensis* and from 1995 to 2005 for var. *caribaea*, to establish the first generation seedling seed orchard in which improved seeds will be produced for operational use. Three selective thinnings (forward selections) will be conducted within the BSOs to convert them into SSOs in which 40 families will be retained. Although the SSO was recommended in the breeding strategy, grafted clonal seed orchards using the best few individuals selected from the BSOs may be set up for rapid production of improved seeds. But the extra investment for these CSOs need to be balanced with the returns (genetic gains).

Combined selection index is to be used for multiple-trait selection to maximize the genetic gains for all traits simultaneously and to make use of the genetic information from provenances, families and individuals which are obtained from the integrated multiple function experiments. Although this method is theoretically advantageous, in practice, there are still difficulties in implementing the method. A major problem caused by index selection in the breeding population is the non-uniform spacing which complicates the continuous assessment of the populations and affects the final seed production. Some restrictions may be needed to apply the index selection.

For second generation breeding, three options are proposed to set up second generation production populations depending on planting requirement and the financial situation. Both forward (convert the Nucleus to SSO) and backward (set up CSO through progeny test) selections can be adopted to select materials for the establishment of seed orchards. If vegetative propagation methods can be developed for the species, clonal forestry by mass vegetative propagation is desirable.

7. General discussion

7.1 Results and implications

7.1.1 Difficulties of breeding Caribbean pines in China

P. caribaea has been grown in China for a few decades and quite a few genetic studies have been conducted on the species, but there is very little information available from the literature in China. Wang *et al.* (1995) have comprehensively summarised the introduction of *P. caribaea* in China, but details of genetic tests and breeding remain incomplete. The review presented in this thesis detailed the genetic tests and breeding in China and summarised two major problems as difficulties in tracking down the origin of the seed sources used and the limited genetic base of current resources for breeding. These two problems affect the progress of genetic improvement. Another crucial issue is that the lack of a well defined breeding strategy and the basic genetic information on which the breeding strategy can be based were recognised.

7.1.2 Genetic structure and patterns of variation

There is a tremendous number of genetic studies on genetic structure and patterns of variation of forest tree species. However, the majority of these are on temperate species. Such a study of tropical tree species has only been a recent fashion as the importance of tropical forests has been of increasing global concern. Hamrick *et al.* (1992) summarised from a review of isozyme studies of more than 300 woody taxa that geographic range is the best predictor of levels of isozyme variation in long-lived woody species. They found that breeding systems play a smaller role in predicting the level of genetic diversity among populations of woody plant species. They concluded that while life history and ecological characteristics of woody plants play a significant role in shaping the present-day genetic structure of species and their populations, the evolutionary history of the species probably plays an equally significant role. Their generalised conclusion seems to apply to the results obtained from present studies very well.

Results obtained from the isozyme studies showed that there are great differences of genetic diversity and population differentiation between the two varieties studied. Genetic variation among populations was found to be greater in var. *bahamensis* (F_{st} = 0.078) than in var.

caribaea (F_{st} = 0.02) while within-population variation was found to be smaller in var. *bahamensis* (H_e = 0.241) than in var. *caribaea* (H_e =0.272) (Table 2.14). This is probably caused by the wide and disconnected natural distribution of var. *bahamensis*. In contrast, var. *caribaea* has a smaller geographic range in its natural distribution, and therefore it contains less genetic variation. However, as far as the island population, which is disconnected from the mainland distribution, is concerned, a large difference was also found between this isolated population and the mainland populations. Compared to other natural populations, the island population of var. *caribaea* has a smaller genetic diversity (Table 2.3). The mean genetic diversity of the four mainland natural populations was 0.277 with the greatest value of 0.297 whereas the diversity of the island populations was 0.253. This may be attributable to the effects of a genetic bottleneck because the samples were taken from heavily degraded and selectively logged forests which may have decreased the population size drastically. As found within many other tree species, a large proportion of isozyme variation detected within *P. caribaea* resided within populations (92.2 % for var. *bahamensis* and 98 % for var. *caribaea*). Only a small amount of the genetic variation revealed by isozyme markers was distributed among populations.

F_{st} or G_{st} was used to measure the levels of allozyme differentiation among populations. A small value of F_{st} suggests a minor degree of differentiation among populations. However, there may exist substantial differentiation of allelic frequencies among populations (Yeh, 1989). Therefore, results from F_{st} should be interpreted with some caution because they do not indicate the lack of significant variation in allelic frequencies among populations; instead, they indicate a low genetic variation among populations relative to variation within populations. It is recommended that F_{st} is used to measure population differentiation together with the heterogeneity analysis of allelic frequencies to overcome the above shortcomings. There is another concern with comparing levels of genetic differentiation within a species. In some cases, sampled populations are separated by several hundred kilometres, in other cases by only a few kilometres. Therefore the results might be quite different if the sampling strategy is different. In this study, the populations refer to provenances which are sampled from different geographic and climatic regions. The geographic distances among populations were greater for the var. *bahamensis* than for the var. *caribaea*. The var. *caribaea* has a relatively limited natural range compared with that of var. *bahamensis*.

7.1.3 Genetic change of domesticated populations

The maintenance of genetic diversity is important for an exotic species for long-term productivity. Genetic diversity is usually linked to population size. Theoretically, loss of genetic diversity occurs when population size is small for several generations (genetic drift), when a population is initiated from a small number of colonists (founder effect) and when a population suffers rapid declines in size (genetic bottlenecks). To examine the genetic change in *ex situ* populations, one Australian population of var. *bahamensis* and one Chinese population of var. *caribaea* were analysed and compared with their natural populations. Unfortunately, the Chinese material was identified as a distinct species rather than var. *caribaea*, thus the comparison becomes meaningless in this regard. For the var. *bahamensis*, the Australian population did not differ much from the natural populations in terms of genetic diversity but demonstrated a higher inbreeding (Table 2.14). For example, the genetic diversity under Hardy-Weinberg equilibrium (H_e) for the Australian population was 0.248 compared with a value of 0.241 for the natural populations. Other measures of within-population genetic variability for the Australian population were also at a comparable level with the natural populations. Surprisingly, the fixation index for the Australian population (F_{is} = 0.327) was almost as great as two times of the average of the natural populations (F_{is} = 0.164, excluding the Little Abaco population which has a very small value due to a small sample size). The reason for this finding was not clear since the breeding system was not analysed for this population, and no details on population history were available. Further study is needed to find out the reasons.

It is common in practice that an exotic species was introduced initially with a few individuals and once the potential of the species was found the plantation area of the species expanded rapidly with most individuals being reproduced from the initial introduction. Thus founder effects may occur. Although the genetic consequences of decreasing population size are predicted by theory, empirical data documenting these effects are rare (Godt *et al.*, 1996). Monitoring of genetic change for a domestication programme would help effectively to prevent loss of genetic diversity caused by founder effects and to maintain the genetic variability for long-term breeding.

7.1.4 Mating systems

Breeding system was analysed only for var. *caribaea* in the present study. The outcrossing rate detected (ranging from 0.894 to 0.985) was close to complete outcrossing (t = 1.0). Such a high outcrossing rate has often been found for many other wind-pollinated conifers and pines (Loveless, 1992; Adams and Birkes, 1989; Rudin *et al.*, 1986; Muona and Harju, 1989; Ritland, 1985; Shaw and Allard, 1982). The small proportion of self-fertilisation detected in this study was identified as true selfing by comparing the single-locus outcrossing estimates and the multilocus estimates. Matheson *et al.* (1989) obtained similar results for several tropical pines including var. *hondurensis* and var. *bahamensis*, but they failed to distinguish whether the self-fertilisation was true selfing or caused by biparental inbreeding. Results from this study suggest that an effective design of clonal deployment in the seed orchard can reduce the amount of inbreeding. Similar findings have also been reported for other coniferous species (Shaw and Allard, 1982; Rudin *et al.*, 1986; Muona and Harju, 1989; Szmidt, 1987).

We found no evident association of the outcrossing rate with genetic diversity measures of populations. No significant difference in outcrossing rate was found between the natural populations and the seed orchard populations, but the breeding systems have been influenced significantly by the management history of the populations. For example, the IDJ population of the var. *caribaea* in an island has a significantly lower outcrossing rate than other natural populations. The management of this population such as clear felling and heavy selective logging has probably affected the pollen production and distribution, hence pollination. Similar results have been reported for other species (Forest Genetic Resources Work Group, 1991; Knowles *et al.*, 1987). A practical implication for natural forests of this result is that spatial isolation of trees due to forest decline or degradation may increase the level of inbreeding and its concomitant effects.

The level of inbreeding is determined not only by the nature of the reproductive system, but also by family structure, which itself is influenced by pollen and seed dispersal characteristics. Inbreeding in natural populations was found to be higher for var. *bahamensis* (F_{is} = 0.163; R = 0.13) than for var. *caribaea* (F_{is} = 0.148; R = 0.035) (Figure 2.4). This is probably affected by the pollen and seed dispersal of the var. *bahamensis* which may be attributed to the island model of natural distribution.

7.1.5 Identity of the Chinese material

The most striking results obtained are that sufficient evidence, both by isozyme analysis and cpDNA analysis, was found to reject the identity of Chinese material as var. *caribaea.* The importance of this discovery cannot be overestimated given the important role the species plays in forest industries in China. The most possible explanation for this is mislabelling, the material taken from China having been mislabelled from another species, most probably *P. elliottii* as shown by the cpDNA evidence. This is possible because *P. elliotii* and *P. caribaea* are very close relatives and many characteristics are similar for both species. Seed orchards or seed production stands of both species are usually located in the same area within short distances. This is particularly true in Zhanjiang district where the Chinese samples were taken. Unfortunately, it has not been possible to track down to the details of the seed collection.

This mislabelling might have caused confusion regarding the correct material for planting programmes. Because no details on seed collection are available, there are two assumptions for the confusion caused by the mislabelling. First, only the samples used in this study were mislabelled. This is a conservative assumption. In this case, there is not much to worry about. It is necessary to examine the identities of seeds provided by the same person or institution. Another assumption is that all seeds from the same supplier were mislabelled and quite possibly the mislabelled seeds were distributed around the country and might be already planted. This is serious and measures should be taken immediately to track down the seed distribution and to correct the identity for any possible users of these mislabelled seeds, because these two species have distinct planting areas in which environment varies considerably although they both grow well in some areas such as Zhanjiang district.

7.1.6 Quantitative genetic variation

Large provenance variation of var. *bahamensis* was found for three growth rates investigated as expected for a species having a discontinuous natural distribution. Substantial genetic gains can be obtained from provenance selection from desirable regions (or islands), i.e. Abaco Island and New Providence. If this is coupled with family selection, further progress can be obtained. The genetic variation detected for growth rates provides essential information for developing a breeding strategy. However, there are a few points that should be considered when applying these results. Firstly, heritabilities estimated for the growth traits may be

overestimated due to the use of range-wide tests. Hence, the predicted genetic gains should be interpreted with some caution. Secondly, the heritabilities fluctuated from year to year indicating that the levels of genetic control of the traits are still not stable. This may be due to the young age of the trees. Further observations for more years on the experiment may help to choose the optimal selection age. Thirdly, only growth rates such as height and diameter were investigated in the study. Wood density and stem form, which are also important economic traits for the species, were not studied due to the young age, therefore information on these two traits should be derived as the trees grow bigger. Pest resistance needs also to be included in future assessment. Heritability and genetic correlations need to be studied for these traits, which will provide information important to the decision of selection strategy in the breeding programme. A final and important point is that these results were based on a single site trial. Therefore it was impossible to look at the effects of GEI. Because the site conditions in the overall planting area vary considerably, it is necessary to set up multiple sites experiments so that the GEI can be investigated.

7.1.7 Breeding strategy

The breeding strategy proposed in this thesis takes advantage of the flexibility of MPBS in which new sub-populations can be established at any time and new material can be incorporated in the programme. As financial shortage and limitation of source material are the problems, a series of sub-populations cannot be established at the same period, therefore such flexibility is important for the breeding programme. Generally, as the above constraints exist and outcrossing rate is sufficiently high as detected in these studies, simple and cheap breeding methods, such measures as simple mating design, open pollination and forward selection, direct conversion of BSO to SSO are advisable for the first generation breeding. Under this strategy, minimum investment is required while genetic gains and diversity necessary for long-term breeding are optimised to a satisfactory level. As the breeding programme moves to the second and advanced generations, intensive breeding methods could be applied. Nucleus Breeding was proposed for the next generation. Flexibility still exists for adopting other breeding methods when needed. Assortative mating and clonal seed orchard were recommended for advanced generations to capture as much genetic gain as possible and as rapidly as possible. These methods make use of specific combining ability (SCA) rather than general combining ability (GCA) used in the first generation. One important thing to be borne

in mind is that no matter how good a breeding strategy is developed, a good administration for the whole breeding programme, research and production is of equal importance to the technical aspects. This is particularly important for the situation in China as it has long suffered from such problems.

7.2 The use of genetic markers and quantitative traits and their limitations

7.2.1 Comparability of isozyme variation among different studies

Undoubtedly, isozyme markers are very useful and provide basic tools for studies of population genetics of forest trees, particularly for investigation of evolutionary process, genetic structure and breeding systems. However, several points need to be considered when comparing results from isozyme analysis among different studies. First, the number of loci used in the studies varies among studies. Isozyme loci surveyed in a study were often not equivalent to the overall mean of a species (Yeh, 1979) indicating that many loci should be surveyed to obtain reliable estimate of genetic variation in forest trees. Lewontin (1974) estimated that as many as 100 loci may be needed to obtain reliable estimates of gene diversity. Unfortunately, most isozyme studies did not approach this level of sampling. Moreover, different numbers and sets of loci were often used in different studies, so that the overall results may vary well among studies. Generally, more loci should give more reliable results as they provide more genetic information. In the present studies, only 5 loci were used for mating system analysis and 8 loci used for genetic structure analysis. However, it is very common that many loci cannot be consistently resolved in a study, as was the case in the present study. Thus the use of such a large number of loci as 100 becomes impossible.

A second point is the techniques used in isozyme analysis. Various techniques in enzyme extraction, electrophoresis and staining are commonly used in different laboratories. Thus the isozyme banding may well be affected. Hence comparability among different studies may be affected.

A third point is that genetic interpretation of the allozyme bands may be different among isozyme studies. Individuals may differ in their interpretation of allozyme bands especially for

those enzymes often showing noise bands which increase the difficulty in banding interpretation. Therefore, the comparability among studies may also be affected.

Finally, the interpretation of results from isozyme studies must be treated with care. Although the most commonly used measures of genetic diversity are number of polymorphic loci, alleles per locus and gene diversity (heterozygosity), some other measures such as effective number of alleles per locus and diversity index etc. are sometimes used. Even with the three commonly used measures, two cautions should also be taken into account when comparing these measures among studies. One is the criteria of polymorphic loci may be different for different studies although 5% and 1% are usually used. The other is that many investigators intend to use only those enzymes showing variation (polymorphism) in their study, at least in this study, concerning the effectiveness and efficiency. If an enzyme shows no variation among individuals or populations, it gives little information. On the other hand, the overall means over loci and measures based on multilocus estimates may still be comparable and indicative among studies.

Although there are obvious advantages of using isozyme markers in population genetic studies of tree species, these disadvantages of them will also limit their use. Only when the isozyme analysis is supplemented by DNA techniques and the use of quantitative genetic analyses, can a full set of powerful tools be used to obtain insights into the population genetics of forest trees.

7.2.2 Use of DNA markers

In theory, molecular markers like RFLP and RAPD are most suitable for assessing genetic variability because they represent the complete inventory of all nucleotide sequences in the DNA of each population member (Brown, 1978). Among the three genomes, chloroplast DNA is most widely used for forest trees. Because cpDNA shows most inter-species variation but little within-species variation, it is useful in the study of phylogenetic relationships among tree species. It is also useful in the identification of introgression or hybridisation because of its uniparental inheritance (El-Kassaby *et al.*, 1988). In this study, RFLP markers of cpDNA were successfully used to reject the identity of the Chinese material as var. *caribaea* with very

few seeds analysed. This suggests that cpDNA has great potential in identifying unknown sources.

However, the methods of using molecular markers are expensive. In studies of genetic variation of tree species, DNA markers will still only provide information of unknown selective effects which are probably neutral. Therefore, the use of DNA will have limited scope. It is clear that the future of population genetic studies of forest trees will still rely, to a large extent, upon the use of electrophoretical techniques together with morphological variation (Brown, 1979).

7.2.3 Use of quantitative genetic variation

Genetic variation of quantitative traits among populations (provenances) is usually detected by provenance trials by gathering seeds from different populations and observing variation in performance (height, diameter, volume etc.) in trees grown under uniform environmental conditions within one or more planting sites. Generally, like genetic variation revealed by isozyme markers, provenance testing reveals that forest tree species are highly heterogeneous, but the genetic variation is organised within and between populations in diverse ways. Unlike the isozyme variation which mainly resides within populations, the variation detected from provenance testing is largely distributed among populations. These different patterns are most probably attributable to the differences in the strength of selection.

Although there are difficulties in the analysis of quantitative traits, the use of quantitative variation is still important. It is these traits that tree breeders want to benefit from by breeding and improving them. It is not the case that the genetic analysis of quantitative traits should be overlooked due to the expensiveness, long testing intervals and complication of statistical analysis. Instead, forest geneticists and tree breeders should try to find a better way to overcome these drawbacks pertaining to quantitative genetic analyses.

Many ways, such as development of better breeding strategy, use of marker assisted selection, early selection, efficient experimental design and statistical methods, etc. have been attempted in tree breeding programmes. Among these, the development of an effective breeding strategy

is probably the most promising way to overcome the problems associated with the huge investment of money and time that need to be devoted to breeding programmes.

7.2.4 Association between isozyme variation and quantitative variation

One of the goals of population genetic studies of forest trees based on markers is inferring the amount and distribution of variation in economically and adaptively important quantitative traits. However, studies of quantitative and marker variation have rarely been combined. The genetic variation detected by isozyme markers may not be necessarily associated with that detected from provenance trials because they reveal different information. In the present studies, although the materials used in both studies are not completely the same, both isozyme and quantitative variation were investigated for the var. *bahamensis* and only isozyme variation was investigated for the var. *caribaea.*

At the variety level, when comparing the isozyme variation between the two varieties, it appeared that populations of var. *bahamensis* are more differentiated than those of var. *caribaea.* This generally agrees with the results from the quantitative analysis although different populations were used in both studies. Significant genetic variation in quantitative traits such as growth rates was found within var. *bahamensis* from a provenance/progeny trial (Zheng, *et al.*, 1994) while provenance tests showed little variation within var. *caribaea* (Pan, 1991; Lu *et al.*, 1990). This correspondence of genetic variation at the species level was also observed for a few other tree species (Hamrick, 1983). Species found to be quite variable in morphological and physiological characters such as *P. contorta* (Critchfield, 1957) and *Picea sitchensis* (Burley, 1966) also displayed moderate to high levels of isozyme polymorphisms (Wheeler and Guries, 1982; Yeh and El-Kassaby, 1980). On the other hand, morphologically uniform species such as *P. resinosa* (Wright *et al.*, 1972) and *Thuja plicata* (Minore, 1969) were found to be monomorphic at all or most isozyme loci (Fowler and Morris, 1972; Yeh, 1988).

However, this association in general does not mean that isozyme markers can be used to predict morphological variation because morphological traits are commonly believed to be controlled by many genes and their interactions and to be strongly influenced by environments

while isozyme markers are controlled by single genes and are selectively neutral. The two types of variation of the var. *bahamensis* could have the same pattern for completely different reasons, e.g. isolation and genetic differentiation by drift for isozymes, but different environments in islands and differences due to selection for quantitative traits. The lack of concordance between isozyme variation and quantitative variation detected in provenance research has also been reported for many other species (Libby and Critchfield, 1987; Moran and Adams, 1989; Namkoong and Kang, 1990).

At the population level, although high levels of genetic variation were detected for both isozyme markers and quantitative traits, the distributions of variation among and within populations are different. A large proportion of isozyme variation for the var. *bahamensis* was detected within populations (92.2%) while only a small proportion (7.8%) was distributed among populations. Contrastingly, quantitative variation detected from the provenance/progeny trial was distributed in an opposite way. Such patterns of distribution of genetic variation have often been found within many other conifer species (Yeh, 1989). Comparison for the var. *caribaea* was impossible because quantitative analysis was not conducted.

In general, isozyme variation provides little information on the pattern of distribution of adaptive quantitative traits and thus may not be very useful for describing adaptive patterns of variation in tree species. The contrasting patterns of variation can be understood as results of a balance among mutation, migration and selection (Hedrick, 1985). The efficient pollen flow (Koski, 1970; Harju and Muona, 1989) and lack of differential selection forces are responsible for the small interpopulation variation for most of the marker genes. Strong selection forces are responsible for local adaptation of quantitative traits. However, recent advances in molecular methods facilitate the determination of QTL (quantitative trait loci). Such markers can be used to obtain information about quantitative traits.

7.2.5 Paralleling studies on marker variation with quantitative variation

It is actually unknown whether genetic differentiation inferred from isozyme studies is neutral, although it is generally assumed, or if it is indicative of a selective advantage or disadvantage.

This is important because if much of the variation seen in allozymes were to be devoid of evolutionary significance (neutral), it would be difficult to justify maintaining multiple populations of trees to preserve their potentially useful genetic diversity based solely on allozyme variations. Unfortunately, isozyme studies in forest trees generally have not been accompanied by parallel work on variation in morphological traits that are known to contribute to the fitness of trees (adaptive).

Knowledge of ecological and genetic relationships among individuals and the populations they comprise (population structure) is prerequisite for efficient design of breeding programmes. A breeding strategy unfit for the population structure of a species would reduce the expected genetic gain. Ideally, in a genetic improvement programme for a species, a comprehensive investigation to draw out the population structure might proceed with genetic tests on quantitatively inherited economically important morphological traits, such as provenance or seed source trials which are designed to detect long-term adaptation and select for economically and silviculturally desirable traits.

Decades of provenance research have shown that most forest tree species exhibit considerable population divergence in genetically controlled traits of direct adaptive value. But knowledge on genetic structure and population history is relatively scarce. This is particularly true for *P. caribaea* as relatively more studies have been conducted on morphological traits but few on isozymes. In practice, it is too often that for many economically important tree species, as is the case for *P. caribaea*, genetic improvement programmes started directly with the field genetic tests and made use of the selectively adaptive genetic variation expressed by the economical traits with little knowledge of the evolutionary process of the species and the genetic structure. The disadvantages of studies of genetic structure based on phenotypic traits are that they are obviously controlled by many genes and quantitatively inherited, and are often strongly influenced by environmental variation. There remains, however, a need to conduct parallel isozyme and morphometric studies to determine to what extent data on isozymes can be used as indicators of morphological variation, or to monitor changes in gene frequency due to environmental or ecological events. Information on isozyme variation is needed because it helps to reveal population history, i.e. the impacts of drift, genetic bottlenecks and founder effects, and breeding systems, hence to facilitate designing sampling

strategy and mating decisions for breeding programmes. Information on quantitative traits is useful for testing, selection and predicting genetic gains of the breeding programme.

It is recommended, therefore, that a genetic survey of isozyme variation and genetic diversity of the species prior to or parallel with comprehensive field genetic trials would be the best way to start a genetic improvement programme for a species. It is, obviously, best to use the same materials for both isozymes and the field test.

7.3 Concluding remarks and direction for future studies

The present studies have obtained many genetic insights into an economically important tropical pine species, from basic genetic information on population genetics to a breeding strategy for an improvement programme. Specifically, the major contribution of this thesis can be summarised as follows: 1) it has presented an extensive review of the tree introduction and genetic improvement of *P. caribaea* in China; 2)for the first time the genetic structure and patterns of variation of var. *caribaea* and var. *bahamensis* have been extensively investigated by the use of isozymes as genetic markers; 3) it has investigated the genetic impact of tree introduction and domestication on genetic structure and breeding system which is a very important issue in forest genetics and tree improvement; 4) for the first time the breeding system has been investigated in natural and artificial populations of the var. *caribaea*; 5) the Chinese material which was used as var. *caribaea* has been found to be actually a different species by employing both isozyme and DNA markers; 6) the utilisation of integrated multiple function field tests of forest trees has been explored and quantitative genetic analysis has been conducted for such a trial established in China; 7) based on these basic genetic studies, an effective and dynamic breeding strategy for *P. caribaea* in China has been proposed.

However, the work is not complete and there are a few aspects that need to be strengthened in future studies. As the environmental conditions change dramatically from place to place in the vast area suitable for planting *P. caribaea*, multi-site genetic tests (combined provenance/progeny trials) should be implemented in the near future to evaluate the GEI. Hence the breeding programme can effectively make use of such interactions to promote high genetic gains for different site conditions. These multi-site tests should be established for all three varieties in various locations representing future planting conditions.

Studies on genetic structure and breeding systems employing genetic markers should be conducted with var. *bahamensis* and var. *hondurensis* using the same sources as in the quantitative genetic tests.

Improving timber and pulp production is the major goal of the breeding programme. As the trees grow bigger, evaluation of wood properties such as density and fibre length, and resistance to pest and disease need to be included in the future assessment of the field trials. In addition, studies on age to age correlation in order to determine optimal selection age for *P. caribaea* in China have obvious practical importance.

The environment is assumed to change rapidly in the future due to global climatic change. To survive rapid environmental change, species must possess genetic variation in response mechanisms. They must be able to respond to strong selection forces. A dynamic breeding strategy for the species adopted to conserve and to create genetic variability is critical to such uncertain environmental change. In addition to the methods proposed in the breeding strategy, studies on effective genetic conservation for uncertain futures *in situ* and *ex situ* should be considered.

Monitoring the genetic change in a breeding programme is of importance to prevent effectively the loss of genetic diversity, particularly for species as exotics, because to preserve genetic diversity is more difficult for exotic species due to limitation of resources. Although *P. caribaea* is not an endangered or threatened species, conservation of genetic resources of the species appears to be an important issue. The high economical merits of the species lead to demands for conserving genetic variation for long-term breeding. Moreover, the over-exploitation in the natural forests, emphasis on few superior genotypes for short-term genetic gains in plantation forestry and domestication of the species in non-native country may lead to loss of genetic variability if conservation goals are not appreciated. A conservation strategy is recommended in the breeding strategy. The goal of the conservation in the breeding programme is to preserve maximum available genetic diversity and even, potentially, to create genetic variability by adopting appropriate breeding methods.

8. Reference

Adams, W.T. 1981: Applying isozyme analyses in tree breeding programs. pp. 60-64. In: Conkle, M.T. (Technical coordinator) Proc. Symp. Isozymes. North Am. forest trees and forest insects. USDA Forest Serv., Pacific SW Forest and Range Exp. Sta., Berkeley, California, Gen. Tech. Rep. PSW-48.

Adams, W.T. 1983: Application of isozyme in tree breeding. pp. 381-400. In: *Isozyme in plant genetics and breeding, Part A.* (Eds Tanksley, S.D. and Orton, T.J.). Elsevier, Amsterdam.

Adams, W.T. and Birkes, D.S. 1989: Mating patterns in seed orchards. Proc. 20th Southern Forest Tree Improvement Conference. Charleston, South Carolina, USA. pp. 75-86.

Adams, W. T., & Joly, R.J. 1980: Allozyme studies in loblolly pine seed orchards: Clonal variation and frequency of progeny due to self-fertilisation. *Silvae Genet.*, **29**: 1-4.

Adams, W.T., Neale, D.B. and Loopstra, C.A. 1988: Verifying controlled crosses in conifer tree-improvement programs. *Silvae Genet.*, **37**: 147-152.

Adams, W.T., Strauss, S.H., Copes, D.L. and Griffin, A.R (Eds). 1992: Population genetics of forest trees. Proceedings of the International Symposium on Population Genetics of Forest Trees. Corvallis, Oregon, USA., July 31- August 2, 1990. 420 pp. Kluwer Academic Publisher. Dordrecht, Boston, London.

Allard, R.W. 1975: The mating system and microevolution. *Genetics*, **79**:115-126.

Allard, R.W., Jain, S.K. and Workman P.L. 1968: The genetics of inbreeding populations. *Adv. Genet.*, **14**: 55-131.

Allard, R.W., Kahler, A.L. and Clegg, M.T. 1975: Isozymes in plant population genetics. pp. 261-272. In: Isozymes IV: Genetics and Evolution. (ed. Markert, C.L.). Academic Press.

Annonymous, 1985: Ten years of tree introduction (performance and evaluation of 145 exotic species in Qinzhou area, Guangxi). *Qinzhou Forestry*, No. **1**: 21-34. (in Chinese).

Apist, V.J. Nakamura, R.R. and Wheeler N.C. 1989: Differential male reproductive success in Douglus fir. *Theor. Appl. Genet.*, **77**: 681-684.

Arnold, M.L, Buckner, C.M. and Robinson, J.J. 1991: Pollen-mediated introgression and hybrid speciation in Louisiana irises. *Proc. Natl. Acad. Sci. USA.*, **88**: 1398-1402.

Avise, J.C. 1994: *Molecular markers, natural history and evolution.* Chapman & Hall. New York. London. 511 pp.

Avise, J.C. and Saunders, N.C. 1984: Hybridization and introgression among species of sunfish (Lepomis): analysis by mitochondrial DNA and allozyme markers. *Genetics*, **108**: 237-255.

Barnes, R.D. 1984: A multiple population breeding strategy for Zimbabwe. In *'Provenance and Genetic Improvement Strategies in Tropical Forest Trees'*, Mutare, Zimbabwe, April, 1984. (Eds. Barnes, R.D. and Gibson, G.L.). Commonwealth Forestry Institute, Oxford and, Forest Research centre, Harare, 619-632.

Barnes, R.D. and Gibson, G.L. 1984: *'Provenance and Genetic Improvement Strategies in Tropical Forest Trees'*, Mutare, Zimbabwe, April, 1984. Commonwealth Forestry Institute, Oxford and, Forest Research centre, Harare. 663 pp.

Barnes, R.D. and Mullin L.J. 1989: The multiple population breeding strategy in Zimbabwe-five year results. In *Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry* [Proc. IUFRO Conference, Pattaya, Tailand, Nov. 1988]. (Gibson, G.L., Griffin, A.R. and Matheson, A.C. eds). Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

Barnes, R.D., White, T., Nyoka, B.I., John, S. and Pswarayi, I.Z. 1995: The composite breeding seedling orchard. Paper for IUFRO conference, Tasmania, Australia, February, 1995.

Barrett, W.H.G. and Golfari, L. 1962: Descripcion de dos Nuevas Varieda des del "Pino de Caribe". *Carib. For.* 23 (2).

Barton, N.H. and Turelli, M. 1989: Evolutionary quantitative genetics: How little do we know? *Ann. Rev. Genet.* **23**: 337-370.

Baylis W.H.B. & R.D. Barnes., 1989: International provenance trial of *Pinus caribaea var. bahamensis*. In: Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin, A.R. & Matheson, A.C.). pp. 363-365. Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

Beckmann, J.S. and Soller, M. 1983: Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. *Theor. Appl. Genet.* **67**: 35-43.

Beckmann, J.S. and Soller, M. 1986a: Restriction fragment length polymorphisms and genetic improvement of agricultural species. *Euphytica.* **35**: 111-124.

Beckmann, J.S. and Soller, M. 1986b: Restriction fragment length polymorphisms in plant genetic improvement. Oxford Surv. *Plant Mol. Cell Biol.* **3**: 196-249.

Birks, J.S. 1994: *The Combined Provenance and Progeny Test in Forest Tree Improvement Programmes*. Submitted.

Birks J.S. & Barnes, R.D. 1990: Provenance variation in *Pinus caribaea, P. oocarpa* and *P. patula spp. tecunumanii*. Tropical Forestry Papers 21. Oxford Forestry Institute. University of Oxford. 40 pp.

Birky, C.W. 1988: Evolution and variation in plant chloroplast and mitochondrial genomes. In: Plant evolutionary biology. Gottieb, L and Jain, S. (Eds). Chapman and Hall. London. pp. 23-53.

Borralho, N.M.G., Cotterill, P.P. and Kanowski, P.J. 1992: Genetic control of growth of *Eucalyptus globulus* in Portugal. II. Efficiencies of early selection. *Silvae Genet.*, **41**: 70-77.

Boscherini, G., Morgante, M, Rossi, P. and Vendramin, G.G. 1994: Allozyme and chloroplast DNA variation in Italian and Greek population of *Pinus leucodermis*. *Heredity,* **72**: 284-290.

Botsein, D. White, R.L., Skolnick, M. and Davis, R.W. 1980: Construction of genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, *32*: 641-656.

Bawa, K.S. 1976: Breeding of tropical hardwoods: an evaluation of underlying bases, current status and future prospects. In Tropical trees: variation, breeding and conservation. (eds. Burley, J. and Styles, B.T.) Academic Press, London. pp. 43-59.

Brewbaker, J.L. 1990: Genetic improvement of multiple trees for agroforestry systems. In: Proceedings, 19th World Forestry Congress, Vol. B. Montreal, Canada, 5-11 August 1990. Vienna, Austria; IUFRO. pp. 304-315.

Bridgwater F. 1992: Mating design. Finns *et al.* (eds), *Handbook of Quantitative Forest Genetics*. Kluwer Academic Publishers. pp. 69-95.

Brodie, A.W. 1994: Final report on the seed collection of Pinus caribaea in Cuba for the ODA-China project: Tree improvement and early establishment with Caribbean pine (Pinus caribaea). IERM, University of Edinburgh.

Brown, A.H.D. 1978: Isozymes, plant population genetic structure and genetic conservation. *Theor. Appl. Genet.* **52**:145-157.

Brown, A.H.D. 1979: Enzyme polymorphism in plant populations. *Theor. Popul. Biol.*, **15**: 1-42.

Brown, A.H.D. 1990: Genetic characterization of plant mating systems. pp. 145-162. In: Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S. (Eds). Plant population genetics, breeding and genetic resources. Sinauer Associates, Inc., Sunderland, Ma.

Brown, A.H.D. & Allard, R.W. 1970: Estimation of the mating system in open-pollinated maize population using isozyme polymorphism. *Genetics*, **66**: 135-145.

Brown, A.H.D., Barrett, S.C.H. & Moran, G.F. 1985: Mating system estimation forest trees: models, methods and meanings. In: Gregorius, H.R. (ed.), *Population genetics in Forestry. Lecture Notes in Biomathematics*, **60**, Spinger-verlag, Berlin, pp. 32-49.

Brown, A.H.D., Matheson, A.C. and Eldridge K.C. 1975: Estimation of the mating system of *Eucalyptus obliqua* L'Herit. by using allozyme polymorphisms. *Austral. J. Bot.* **23**: 931-949.

Bulmer, M.G. 1985: The mathematical theory of quantitative genetics. Clarendon Press. Oxford. GB. 254 pp.

Burdon, R.D. 1986: Clonal forestry and breeding strategies - a perspective. pp 645-659. In: Proceedings, IUFRO joint meeting of working parties on breeding theory, progeny testing and seed orchards. Williamsburg, VA, USA. OCT. 13-17.

Burdon, R.D. 1989: Early selection in tree breeding: principles for applying index selection and inferring input parameters. *Can. J. For. Res.* **19**: 499-504.

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

Burdon, R.D., Shelbourne, C.J.A. and Wilcox, M.D. 1977: Advanced selection strategies. pp 1133-1147. In: Proc. Third World Conf. on Forest Tree Breeding. March 21-26, Canberra, Australia.

Burley, J. and Adlard, P.G. 1992: Plantation silvicultural research and genetic improvement. In: Proceedings, IUFRO Centennial meeting. Berlin, Germany. September 1992. Vienna, Austria, IUFRO. 13 pp.

Burley, J., 1966: Genetic variation in seedling development of Sitka spruce (*Picea sitchensis*) (Bong.). *Carr. Forestry*. **39**: 68-94.

Burley, J. and Wood, P.J. 1976: A manual on species and provenance research with particular reference to the tropics. Tropical forestry papers. No. 10. Commonwealth forestry Institute, University of Oxford. 226 pp.

Cameron J.N., Cotterill, P.P. and Whiteman P.H. 1988: Key elements of a breeding plan for temperate Eucalypts in Australia. pp. 69-80. In: Proc. 10th Aust. for. Genet. Res. Work. Group 1. Dieters, M.J. and Nikles, D.G. (Eds). Oct. 30 - Nov. 4, Gympie, Queensland, Australia.

Carson, M.J. 1986: Control-pollinated seed orchards of best general combiners. A new strategy for radiata pine improvement. Proc. Plant Breeding Symposium, Lincoln, New Zealand. pp 144-149.

Carlson, J.E., Tulsieram, L.K., Glaubitz, J.C., Luk, V.W.K., Kauffeldt, C. and Rutledge, R. 1991: Segregation of random amplified DNA markers in F_1 progeny of conifers. *Theor. Appl. Genet.* **83**: 194-200.

Chakraborty, R. and Zhong, Y. 1994: Statistical power of an exact test of Hardy-weinberg proportions of geneotypic data at a multiallelic locus. *Hum. Hered.* **44**: 1-9.

Cheliak, W.M. and Pitel, J.A., 1984: techniques for starch gel electrophoresis of enzymes from forest tree species. Information report PI-X-42, Petawawa National Forestry Institute. 49 pp.

Cheliak, W.M., Morgan, K., Strobeck, C., Yeh, F.C. and Dancik, B.P. 1983: Estimation of mating system parameters in plant populations using the EM algorithm. *Theor. Appl. Genet.* **65**: 157-161.

Cheliak, W.M., Yeh, F.C. and Pitel, J.A. 1987: Use of electrophoresis in tree improvement programs. *For. Chron.* **63**: 89-96.

Chen, M.-R., Lin, W. and Yu, D. 1989: Investigation on the introduction of *Pinus caribaea* in Shantou City. No. 6: 29-35. (in Chinese).

Chiu, W.-L. and Sears, B.B. 1985: Recombination between chloroplast DNAs does not occur in sexual crosses of *Oenothera. Mol. Gen. Genet.* **198**: 525-528.

Clegg, M.T. and Zurawski, G., 1992: Chloroplast DNA and the study of plant phylogeny: present status and future prospects. In: Molecular systematics of plants. Soltis, P.S., Soltis, D.E. and Doyle, J.J. (Eds). Chapman and Hall, New York, London. pp 1-13.

Clegg, M.T., 1980: Measuring plant mating system. *Bio-science*, **30**, 814-818.

Clegg, M.T., Kahler, A.L. and Allard, R.W. 1978: Estimation of life cycle components of selection in an experimental plant population. *Genetics*. **89**:765-792.

Clegg, M.T., Ritland, K. and Zurawski, G. 1986: Processes of chloroplast DNA evolution. Academic Press Inc.

Cochran W.G. and Cox G.M. 1957: *Experiment Design*. John Wiley & Sons Inc.

Conkle, M.T. 1981: Isozyme variation and linkage in six conifer species. In: Proc. Symp. Isozymes of North American Forest Trees and Forest Insects. USDA For. Ser. Gen. tech. Rep., PSW-48, pp 11-17.

Conkle, M.T., Hodgskiss, P.D., Nunnally, L.B. and Hunter S.C. 1983: Starch gel electrophoresis of conifer seeds: A laboratory manual. (Forest Service general technical report (Final) FSGTR-PSW-64). Pacific southwest forest and range experiment station, Berkeley, CA. 24 pp.

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

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

Cotterill, P.P. 1986b: Breeding strategy: don't underestimate simplicity. pp 8-23. In: Proceedings. IUFRO joint meeting of working parties on breeding theory, progeny testing and seed orchards. Williamsburg, VA, USA. Oct. 13-17, 1986.

Cotterill, P.P. 1989: The nucleus breeding strategy. In: Proceedings, 20th Southern Forest Tree Improvement Conference. Charleston, SC, USA, 26-30 June 1989. pp. 36-42.

Cotterill P.P. and Cameron, J.M. 1989: Radiata pine breeding plan. APM forests technical report No. 89/20. 50 pp.

Cotterill, P.P. and Dean, C.A. 1990: Successful tree breeding with index selection. Melbourne, Australia; CSIRO. 80 pp.

Cotterill, P.P., Dean, C.A. and Van Wyk, G. 1987: Additive and dominance genetic effects in *Pinus pinaster*, *P. radiata* and *P. elliottii* and some implications for breeding strategy. *Silvae Genet.* **36**: 221-232.

Cotterill, P.P., Dean, C.A., Cameron, J. and Brindbergs, M. 1989: Nucleus Breeding: a new structure for rapid improvements under clonal forestry. pp 39-51. In: Gibson, G.L., Griffin, A.R. and Matheson, A.C. (Eds). Breeding Tropical Trees: Population structure and genetic improvement strategies in clonal and seedling forestry. [Proc. IUFRO Conf., Pattaya, Tailand, Nov. 1988]. Oxford Forestry Institute, Oxford, UK and Winrock International, Arlington, Virginia, USA, 1989.

Crawford, D.J. 1983: Phylogenetic and systematic inferences from electrophoretic studies. pp. 257-287. In isozymes in plant genetics and breeding, part A. (eds Tanksley, S.D. and Orton, T.J., Elsevier Science publishers B. V., Amsterdam.

Critchfield, V.B. 1957: Geographic variation in *Pinus contorta.* Publ. Maria Moors Cabot Forest. 3. 118 pp.

Crokford, K.J. Birks, J.S. and Barnes R.D. 1989: Family within provenance trials of *Pinus caribaea*, *P. oocarpa* and *P. patula* ssp. *tecunumanii*—Early results and implications for breeding strategies. In: *Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry* [Proc. IUFRO Conference, Pattaya,Tailand, Nov. 1988]. (Gibson, G.L., Griffin, A.R. and Matheson, A.C. eds). Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

Dancik, B.P. and Yeh, F.C. 1983: Allozyme variability and evolution of lodgepole pine (*Pinus controta* var. *latifolia*) and jackpine (*P. banksiana*) in Alberta. *Can. J. Genet. Cytol.* **25**: 57-64.

deSouza S.M., Hodge, G.R. and White, T.L. 1992: Indirect prediction of breeding values for fusiform rust resistance of Slash Pine parents using green house tests. *For. Sci.* **38**: 45-60.

Devey, M.E., Jermstad, K.D., Tauer, C.G. and Neale, D.B. 1991: Inheritance of RFLP loci in a loblolly pine 3-generation pedigree. *Theor. Appl. Genet.* **83**: 238-242.

Dieters, M.J. 1996: Genetic parameters for slash pine (Pinus elliottii) grown in south-east Queensland, Australia: growth, stem straightness and crown effects. *Forest Genetics.* **3**: 27-36.

Downie, S.R. and Palmer, J.D. 1992: Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Molecular systematics of plants. Soltis, P.S., Soltis, D.E. and Doyle, J.J. (Eds). Chapman and Hall, New York, London. pp 14-35.

El-Kassaby, Y.A, Sigurgeirsson, and Szmidt A.E. 1988: The use of restriction analysis of chloroplast DNA in classifying hybrid spruce seedlots. pp 67-88. In: Molecular genetics of forest trees. Hallgren, J.-E. (ed). Proceedings of the Frans Kempe symposium in Umea, June 14-16, 1988. Swedish University of Agricultural Sciences. 222 pp.

El-Kassaby, Y.A. & Ritland, K. 1986: Low levels of pollen contamination in a Douglas-fir seed orchard as detected by allozyme markers. *Silvae Genet.* **35**: 224-229.

El-Kassaby, Y.A. and Davidson, R. 1991: Impact of pollination environment manipulation on the apparent outcrossing rate in a Douglas-fir seed orchard. *Heredity.* **66**: 55-59.

El-Kassaby, Y.A., Ritland, K. & Yazidani, R. 1989: Levels of outcrossing and contamination in two *Pinus sylvestris* L. seed orchards in Northern Sweden. *Scand. J. For. Res.* **4**: 41-49.

El-Kassaby, Y.A., Yeh, F. C. and Szijlai, O. 1981: Estimation of outcrossing rate of Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco) using allozyme polymorphism. *Silvae Genet.* **30**: 182-184.

El-Kassaby, Y.A., Yeh, F.C., and Sziklai, O. 1986: Clonal and seedling seed orchards: A comparison of outcrossing rates and heterozygosity in coastal Douglas-fir using allozyme markers. pp. 410-421. In: Hatcher, A.V. and Weir, R.J. (Eds). IUFRO conference Proc. A joint meeting of Working Parties-S2.04.02, S2.04.03, and S2.03.03. Williamsburg, Virginia

Eldridge, K., Davidson, J. Harwood, C. and WYK, G. 1993: Eucalypt domestication and breeding. Clarendon Press. Oxford. 288 pp.

Ellstrand, N.C. and Foster, K.W. 1983: Impact of population structure on the apparent outcrossing rate of grain sorghum (Sorghum bicolor). *Theor. Appl. Genet.* **66**: 323-327.

Ennos, R.A., 1981: Quantitative studies of mating system in two sympatric species of Ipomea (Convolvulaceae). *Genetica,* **57**: 93-98.

Ennos, R.A. 1992: Genetic variation in Caledonian pine populations: Origins, exploitation and conservation. In: Genetic variation in European populations of forest trees. Muller-Starck, G. and Ziehe, M. (Eds). Sauerlander's Verlag, Frankfurt am main. pp 235-249.

Ennos, R.A. and Clegg, M.T. 1982: Effect of population substructuring on estimates of outcrossing rate in plant populations. *Heredity,* **48:**283-292.

Ennos, R. A. and Swales, K.W. 1987: Estimation of mating system in fungal pathogen *Crumenulopsis sororia* (Karst.). Groves using isozyme markers. *Heredity*, **59**: 423-430.

Erickson, V.J. and Adams, W.T. 1989: Mating success in a coastal Douglas-fir seed orchard as affected by distance and floral phenology. *Can. J. For. Res.* **19**:1248-1255.

Erickson, V.J., 1987: The influence of distance and flora phenology on pollen gene flow and mating system patterns in a coastal Douglus-fir seed orchard. M.S. thesis. Oregon State Univ., Corvallis, USA.

Ericksson, G., Namkoong, G and Roberds, J.H. 1993: Dynamic gene conservation for uncertain futures. *Forest Ecology and Management.* **62**: 15-37.

Ernst, S.G., Hanover, J.W. and keathley, D.E. 1990: Assessment of natural interspecific hybridization of blue and Engelmann spruce in southwestern Colorado. *Can. J. Bot.* **68**: 1489-1496.

Evans, J. 1992: Plantation forestry in the tropics. Tree planting for industrial, social, environmental and agroforestry purposes. 2nd edition. Clarendon Press, Oxford. 403 pp.

Falconer, D.S. (1989). *Introduction to quantitative genetics.* 3rd ed. Longman Science & Technical. 438 pp.

FAO, 1961: Seminar and study tour of Latin America conifers. Inst. Nac. de Investig. For., Mexico.

Filppula, S., Szmidt, A.E. and Muona, O., 1992: Genetic comparison between *Pinus syvestris* L. and *P. mugo* Turra using allozymes and chloroplast DNA. *Nord. J. Bot.* **68**: 1489-1496.

Fineschi, J.S., Malvoti, M.E., Cannata, F. and Hattermer, H.H. (Editors). 1991: Biochemical Markers in the population genetics of forest trees. SPB Academic Publishing, The Hague.

Forest Genetic Resources Work Group, 1991: Managing global genetic resources. Forest trees. Committee on managing global genetic resources: Agricultural Imperatives. Board on agriculture. National Research Councill. National Academic Press, Washington, D.C. 228 pp.

Foster, G.S. 1992: Estimating yield: beyond breeding values. In: Fins, L., Friedman, S.T., and Brotshol, J.V. (Eds) Handbook of quantitative forest genetics. Dordrecht, Netherlands; Kluwer, Forestry Series 39. pp 229-269.

Foster, G.S. and Shaw, D.V. 1988: Using clonal replicates to explore genetic variation in a perennial plant species. *Theor. Appl. Genet.* **76**: 788-794.

Fowler D.P. and Morris, R.W. 1972: Genetic diversity in red pine: evidence for low genic heterozygosity. *Can. J. For. Res.* **7**: 343-347.

Franklin, E.C., 1970: A survey of mutant forms and inbreeding depression in species of the family Pinaceae. USDA for. Serv. Res. Pap. SE-61.

Friedman, S.T. & Adams, W.T. 1985: Estimation of gene flow into two seed orchards of Loblolly pine (*Pinus taeda* L.) *Theor. Appl. Genet.* **68**: 609-615.

Friedman, S.T. and Neale, D.B. 1992: Biochemical and molecular markers. In: Bramlett, D. (Ed). Pollen management handbook, Vol. II. SE Exp. Sta., Ashville, North Carolina.

GENSTAT 5 COMMITTEE (1990). *Genstat 5 Reference Manual.* Clarendon Press, Oxford.

Gibson, G.L. 1993: Report on a study tour to the Caribbean, Central and South America, 18 July - 4 August 1994. Forest tree improvement and early establishment with Caribbean Pine (*Pinus caribaea*). Edinburgh University. 20 pp.

Gibson, G.L., Griffin, A.R. and Matheson, A.C. (Eds). 1989: *Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry* [Proc. IUFRO Conference, Pattaya,Tailand, Nov. 1988]. Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

Godt M.J. and Hamrick, J.L. 1995: The mating system of *Liatris helleri* (Asteraceae), a threatened plant species. *Heredity.* **75**: 398-404.

Godt, M.J., Johnson, B.R. and Hamrick, J.L. 1996: Genetic diversity and populations size in four rare southern appalachian plant species. *Conservation Biol.* **10**(3): 796-805.

Goudet, J. 1995: FSTAT (Version 1.2), a computer program to calculate F-statistics. *J. Hered.* **86**: 485-486.

Govindaraju, D.R. 1988: A note on the relationship between outcrossing rate and gene flow in plants. *Heredity.* **61**: 401-404.

Govindaraju, D.R., Dancik, B.P. and Wagner, D.B. 1989: Novel chloroplast DNA polymorphism in a sympatric region of two pines. *J. Evol. Biol.* **2**: 49-59.

Gray, M.W. 1989: Origin and evolution of mitochondrial DNA. *Ann. Rev. Cell Biol.* **5**: 25-50.

Greaves A., 1978: Description of Seed Sources and Collections for Provenances of *Pinus caribaea.* CFI, Tropical Forestry Paper No. 12, 98 pp.

Greaves, A. and Kemp, R.H. 1978: International provenance trials - *Pnius caribaea* Morelet. In: Progress and problems of genetic improvement of tropical forest species. Proceedings of a Joint Workshop. Brisbane, Queensland, Australia, 4-7 April, 1977. Vol 1: 302-310.

Gregorius, H-R. Ziehe, M. and Ross M.D. 1987: Selection caused by self-fertilization. I. Four measures of self-fertilization and their effects on fitness. *Theor. Popul. Biol.*, **31**: 91-115.

Griffin, A.R. and Lindgren, D. 1985: Effect of inbreeding on production of filled seed in Pinus radiata- experimental results and a model of gene action. *Theor. Appl. Genet.* **71**: 334-343.

Guo, S.W. and Thompson, E.A. 1992: Performing the exact test of hardy-Weinberg proportions for multiple alleles. *Biometrics.* **48**: 361-372.

Guries, R.P., Friedman, S.T. and Ledig F.T. 1978: A megagametophyte analysis of genetic linkage in pitch pine (*Pinus rigida* Mill.). *Heredity* **40**: 309-314.

Haigh, H. 1966: Root development in the sandy soils of Zululand. *Bosbou in Suid-Africa.* **7**: 31-36.

Hamilton, W.D. 1971: Selection of selfish and altruistic behavior in some extreme models, pp. 57-91. In Man and Beast: Comparative social behavior. Eisenberg, J.F. and Dillon, W.S. (Eds.). Smithsonian inst. Press. Washington DC.

Hamrick, J.L, Godt, M.J.W., Murawski, D.A. and Loveless, M.D. 1991: Correlations between species traits and allozyme diversity: implications for conservation biology. pp. 75-86. In: Falk D.A and Holsinger, K.E. (Eds). Genetics and conservation of rare plants. Oxford University Press. New York, Oxford.

Hamrick, J.L., Godt, M.W. and Sherman-Broyls, S.L. 1992: Factors influencing levels of genetic diversity in woody plant species. *New Forests.* **6**: 95-124.

Hamrick, J.L. 1989: Isozymes and the analysis of genetic structure in plant populations. pp 87-105. In: Soltis, D. and Soltis, O. (Eds). Isozyme in plant biology. Dioscorides Press. Washington DC.

Hamrick, J.L. and Allard, R.W. 1972: Microgeographic variation in allozyme frequencies in *Avena barbata. Proc. Nat. Acad. Sci. USA*, **69**:2100-2104.

Hamrick, J.L. and Godt, M.J. 1989. Allozyme diversity in plant species. In: Plant population genetics, breeding and genetic resources (Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir B.S., Eds). Sunderland, Massachusetts: Sinauer; 43-63.

Hamrick, J.L. 1983: The distribution of genetic variation within and among natural plant populations. pp. 335-348. In: Schonewald-Cox, C.M., Chamber, S.M., MacBryde, B. and Thomas, L. (Eds.). Genetics and Conservation. Benjamin cummings, Menlo Park, CA.

Hamrick, J.L., Mitton, J.B. and Linhart, Y.B. 1981: Levels of genetic variation in trees: influence of life history charateristics. In: M.T. Conkle, (ed), Isozyme of North-American forest trees and forest insects, 35-41. Pacific SW For. Range Exp. Sta. tech. Report, No. 48.

Harju, A. and Muona, O. 1989: Background pollination in *Pinus sylvestris* L. seed orchards. *Scand. J. For. Res.* **4**(4): 513-520.

Hartl D.L. and Clark, A.G. 1989: Principles of population genetics. Sinauer Association Inc., Sunderland, MA. 682 pp.

Harvey, W.R. and Townsend, A.M. 1985: Selection on a combination of individual, family and stand merit in provenance test. *For. Sci.* **31**:813-821

Harville, D.A. and Callanan T.P. 1990: Computational aspects of likelihood-based inference for variance components. In *Advances in Statistical Methods for Genetic Improvement of Livestock* (D. Gianola and K. Hammond Eds). 1990. Springer-Verlag.

Haufe, H.H.R. 1969: Inventario forestal complementario de los Bosques de Pino (P. caribaea Mor.) de Poptun, Peten. Empresa De Formento y Desarrolo Economic Del Peten - FY- DEP, Guatemala, C.A.

Hedrick, P.W. 1985: Genetics of populations. University of Kansas. Jones and Bartlett publishers, Inc. Boston. 629 pp.

Hedrick, P.W., 1990: The mating system and evolutionary genetics. In: Wohrmann, K. and Jain, S. (eds) *Population Biology, Ecological and Evolutionary Viewpoints.* Springer, Berlin.

Hiebert, R.D. and Hamrick, J.L. 1983: Patterns and levels of genetic variation in Great Basin bristlecone pine, *Pinus longaeva. Evolution.* **37**(2): 302-310.

Hillel, J., Schaap, T., Haberfeld, A., Jeffreys, A.J., Plotzky, Y., Cahaner, A. and Lavi, U., 1990: DNA fringerprints applied to gene introgression in breeding programs. *Genetics*. **124**: 783-789.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K. and Sugiura, M. 1989: The complete sequence of rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217**: 185-194.

Hodge G.R. and Purnell, R.C. 1993: Genetic parameter estimates for wood density, transition age and radial growth in Slash pine. *Can. J. For. Res.* **23**:1881-1891

Hodge, G.R. and White, T.L. 1992: Concepts of selection and gain prediction. Finns *et al.* (eds), *Handbook of Quantitative Forest Genetics*. pp. 140-194

Horovitz, A. and Harding, J. 1972: The concept of male outcrossing in hermaphroditic higher plants. *Heredity*, **29**: 223-236.

Hughes, C.E. 1989: New opportunities in *Leucaena* genetic improvement. In: *Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry* [Proc. IUFRO Conference, Pattaya,Tailand, Nov. 1988]. (Gibson, G.L., Griffin, A.R. and Matheson, A.C. eds). Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989. pp. 218-226.

Jackson, N. and Tuner, H.N. 1972: Optimal structure for a co-operative nucleus breeding system. *Proc. Aust. Soc. Anim. Prod.* **9**: 55-64.

Jain, S.K. 1979: Estimation of outcrossing rates: Some alternative procedures. *Crop Sci.* **19**: 23-26.

James, J.W. 1977: Open nucleus breeding system. *Anim. Prod.* **24**: 287-305.

Jeffreys, A.J. 1987: Highly variable minisatellites and DNA fingerprints. *Biochem. Soc. Trans.* **15**: 309-317.

Jeffreys, A.J., Wilson, V. and Thein, S.L. 1985a: Hypervariable "minisatellite" regions in Human DNA. *Nature* (London). **314**: 67-73.

Jeffreys, A.J., Wilson, V. and Thein, S.L. 1985b: Individual-specific "fingerprints" of human DNA. *Nature* (London). **314**: 67-73.

Joly, R.J. and Adams, W.T. 1983: Allozyme analysis of pitch x loblolly pine hybrids produced by supplemental mass-pollination. *For. Sci.* **29**: 423-432.

Kahler A.L. Gardner, C.O. and Allard, R.W. 1984: Nonrandom mating in experimental populations of Maise. *Crop Sci.* **24**: 350-354.

Kang, H. 1979: Designing a tree breeding system. In: Proc. 17th Meeting Canadian Tree Improvement Association. August 27-30, Gander, Newfoundland. pp. 51-66.

Kang, H. 1991: Components of juvenile-mature correlations in forest trees. *Theor. Appl. Genet.* **87**: 173-184.

Kang, H. and Nienstaedt, H. 1987: Managing long-term breeding stock. *Silvae Genet.*,**36**: 30-39.

Kanowski P.J. and Nikles, D.G. (1989). Family structure in provenance trials, conservation stands and genetic base populations. In *Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry* [Proc. IUFRO Conference, Pattaya,Tailand, Nov. 1988]. (Gibson, G.L., Griffin, A.R. and Matheson, A.C. eds). Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

Kanowski, P.J., 1993: Forest genetics and tree breeding. *Plant Breeding Abstracts.* **63**(6): 717-726.

Kcarnheden, A. and Engstrom, P., 1992: Genetically stable, individual-specific differences in hypervariable DNA in Norway spruce, detected by hybridization to a phage M13 probe. *Can. J. For. Res.* **22**: 117-123.

Kempthrone O. (1952). *The Design and Analysis of Experiments.* Wiley Publications in Statistics.

Kennedy, W.J. Jr. and Gentle, J.E. (1980). *Statistical Computing.* Marcel Dekker Inc.

Khasa, P.D., Cheliak, W.M. and Bousquet, J. 1994: Genetic variation in 26 populations of *Racosperma auriculiforme* and *Racosperma mangium* using allozymes. *Can. J. For. Res.* **24**: 1123-1132.

Kimura, M. 1983: The neutral theory of molecular evolution. Cambridge University Press, Cambridge, 367 pp.

Kinloch, B.B., westfall, R.D. and Forrest, G.I. 1986: Caledonian Scots pine: origins and genetic structure. *New Phytologist.* **104**: 703-729.

Knowles, P. and Grant, M.C. 1985: Genetic variation of lodgepole pine over time and microgeographic space. *Can. J. Bot.* **68**: 942-946.

Knowles, P., Furnier, G.R., Aleksiuk, M.A. and Perry, D.J. 1987: Significant levels of self-fertilization in natural populations of tamarack. *Can. J. Bot.* **65**: 1087-1091.

Koski, V. 1970: A study of pollen dispersal as a mechanism of gene flow. *Commun. Inst. For. Fenn.* **70**: 1-78.

Kozlowski T.T. and Greathouse T.E. 1970: Shoot growth and form in pine in the tropics. *Unasylva.* **24** (4): 6-14.

Kreike, J., Burg, K., Zechmeister, M., Haider, T. and Glossl, J. 1991: DNA-fringerprinting and RFLP analysis as tools to study genetic diversity in populations of fir, spruce and oak. pp. 95-107. In: Muller-Starck, G and Ziehe, M. (Eds) Genetic variation of forest tree populations in Europe. Sauerlander-Verlag, Frankfurt/Main.

Kreitman, M. 1983: Nuceotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster. Nature.* **304**: 412-417.

Kremer, A. 1992: Predictions of age-age correlations of total height based on serial correlations between height increments in maritime pine (Pinus pinaster Ait). *Theor. Appl. Genet.* **85**: 152-158.

Lamb A.F.A., 1973: Pinus Caribaea. Vol. 1, Fast Growing Timber Trees of the Lowland Tropics (Now 'Tropical Forestry Papers'). No. 6. University of Oxford, 254 pp.

Lambeth, C.C. 1986: Effective genetic testing - practical and theoretical considerations. In: Proceedings, IUFRO joint meeting of working parties on breeding theory, progeny testing and seed orchards. Williamsburg, VA, USA. 13-17 October, 1986. pp. 24-33.

Lambeth, C.C., Gladstone, W.T. and Stonecypher, R.W. (1983). Statistical efficiency of row and noncontiguous family plots in genetic tests of loblolly pine. *Silvae Genet.* **32**: 24-9

Lamprecht, H. (1989). *Silviculture in the Tropics: Tropical forest ecosystems and their tree species, possibilities and methods for their long term utilization.* Technical Cooperation - Federal Republic of Germany. 296 pp.

Land, S.B. Jr., Bongarten, B.C. and Tolliver, (1987). Genetic parameters and selection indices from provenance/progeny tests. In *Statistical considerations in genetic testing of forest trees.* Southern Cooperative Series Bulletin No. 324. 59-74.

Landry, B.S. and Michelmore, R.W. 1987: Methods and applications of restriction fragment length polymorphism analysis to plants. In: Bruening, G, Harada, J. and Hollaender, A. (Eds). Tailoring genes for crop improvement: an agricultural perspective. Plenum Press. New York. pp. 25-44.

Ledig, F.T. 1986: Conservation strategies for forest gene resources. *For. Ecol. Manage.* **14**: 77-90.

Leigh Brown, A.J. 1989: Population genetics at the DNA level: a review of the contribution of restriction enzyme studies. Oxford Surv. *Ecol. Biol.* **6**: 207-242.

Leven, H. 1949: On a matching problem arising in genetics. *Ann. math. Stat.* **20**: 91-94.

Lewin, R. 1989: Limits to DNA fingerprinting. *Science.* **243**: 1549-1551.

Lewontin, R.C., 1974: The genetic basis of evolutionary change. Columbia University Press, New York, NY.

Li, B., Williams, C.G., Carlson, W.C., Harrington, C.A. and lambeth, C.C. 1992: Gain efficiency in short-term testing: experimental results. *Can. J. For. Res.* **22**: 290-297.

Li, Y.-K., Zhan, F. and Zhang, F. 1991: Exotic species of Pinus in Fujian province, China. *J. Fujian For. Sci. and Tech.* **18**(3): 30-37. (in Chinese)

Liang, K.-N. and Zhou, W.-L., 1991: The early performance of Pinus caribaea in the east and Northwest of Hainan Island. *Forest Research.* **4**(4): 374-379. (in Chinese).

Libby, W.J. and Cockerham, C.C. (1980). Random non-contiguous plots in interlocking field layouts. *Silvae Genet.* **29**:183-90

Libby, W.J. and Critchfield, W.B. 1987: Patterns of genetic architecture. *Ann. For.* **13**: 77-92.

Lindgren, D. 1986: How should breeders respond to breeding values? pp 361-372. In: Proceedings, IUFRO joint meeting of working parties on breeding theory, progeny testing and seed orchard. Williamsburg, VA, USA. Oct. 13-17, 1986.

Lindholm , J. and Gustafsson, P. 1991: The chloroplast genome of the gymnosperm *Pinus contorta*: a physical map and a complete collection of overlapping clones. *Curr. Genet.* **20**: 161-166.

Lin, K.Z. 1992: Afforestation programme of *Pinus caribaea* in Yulin area of Guangxi. *Guangxi Forestry.* No. 5 (In Chinese).

Ling, C.-F. and Huang, T.-Z., 1991: Preliminary report on provenance trials of Pinus caribaea. *Newsletter of Forestry Sci. and Tech.* **7**: 15-17. (in Chinese).

Little, E.L. and Dorman, K.W. 1954: Slash pine (P. elliottii), including South Florida slash pine, nomenclature and description. Sta. Pap. 5th East For. Exp. Sta. 36.

Loo-Dinkins J.A. 1992: Field test designs. In: Fins, L., Friedman, S.T., and Brotshol, J.V. (Eds) Handbook of quantitative forest genetics. Dordrecht, Netherlands; Kluwer, Forestry Series 39. pp 96-137.

Loo-Dinkins J.A., Tauer, C.G. and Lambeth, C.C. 1990: Selection system efficiencies of computer-simulated progeny test field designs in loblolly pine. *Theor. Appl. Genet.* **79**:89-96.

Loveless, M.D. and Hamrick, J.L. 1984: Ecological determinants of genetic structure in plant populations. *Ann. Rev. Ecol. Syst.* **15**: 65-95.

Loveless, M.D., 1992: Isozyme variation in tropical trees: Patterns of genetic organizations. *New Forests.* **6**: 67-94.

Lowe, W.J. and van Buijtenen, J.P. 1989: The incorporation of early testing procedures into an operational tree improvement program. *Silvae Genet.*,**38**: 243-250.

Lu, P.-X., Pan, Z.-G. and Xiang, S., 1990: Variety and provenance selection of Caribbean pine in southern China. *Forest Research.* **3**(5): 454-469. (in Chinese).

Luckoff, H.A. 1964: The natural distribution, growth and botanical variation of *Pinus caribaea* and its cultivation in South Africa. Annale Univ. van Stellenbosch, 39 serie A, 1.

Luo, F.-S., Pi, W., Wan, G., Yin, Q., and Wu, Z., 1992: the geographical provenance test of *Pinus caribaea* var. *hondurensis. Acta Botanica Yunnanica.* **14**(3): 301-306. (in Chinese).

Ma, X.X. 1986: Observation of adaptability of 7 exotic pines. *Tropical Plant Research.* **2**(6): 46-50 (in Chinese).

Magnussen, S. 1990: Selection index: economic weights for maximum simultaneous genetic gain. *Theor. Appl. Genet.* **79**: 289-293.

Magnussen, S. 1991: Index selection with non-linear profit function as a tool to achieve simultaneous genetic gain. *Theor. Appl. Genet.* **82**: 305-312.

Matheson A.C., Bell J.C. and Barnes, R.D. 1989: Breeding systems and genetic structure in some Central American pine populations. *Silvae Genet.*, **38**: 107-133.

Matheson C.A. (1989a). Experiment design for multiple purpose trees. *Tree Improvement of Multiple Purpose Trees.* Vol. 2. Winrock International, Arlington, Virginia, USA.

Matheson, C.A. (1989b). Statistical methods and problems in testing large numbers of genotypes across sites. In *Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry* [Proc. IUFRO Conference, Pattaya,Tailand, Nov. 1988]. (Gibson, G.L., Griffin, A.R. and Matheson, A.C. eds). Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

McCutchan, B.G., Namkoong, G. and Giesbrecht, F.G. 1989: Design efficiences with planned and unplanned unbalance for estimating heritability in forestry. *Forest Science.* **35**: 801-815.

Mckeand, S. and Beineke, F. 1980: Sublining for half-sib breeding populations of forest trees. *Silvae Genet.*,**29**: 14-17.

Millar, C.I. and Westfall R.D. 1992: Allozyme markers in forest genetic conservation. *New Forests*. **6**: 347-371.

Miller, R.G., Conkle, M.T. and Friedman, S.T. 1989: The forest service laboratory for genetic analyses of trees. *Tree Planters' Notes*. **40**: 25-29.

Milligan, B.G., Leebens-Mack, J. and Strand, E. 1994: Conservation genetics: beyond maintenance of marker diversity. *Mol. Ecol* **3**: 423-435.

Minore, D. 1969: Effect of high soil density on seedling root growth of seven northwest tree species. USDA For. Ser. Res. Note. PNW - 112.

Mitton, J.B. 1992: The dynamic mating systems of conifers. *New forest*. **6**: 197-216.

Mitton, J.B., 1983: Conifers. pp. 443-472. In isozymes in plant genetics and breeding, part A. (eds Tanksley, S.D. and Orton, T.J., Elsevier Science publishers B. V., Amsterdam.

Moran G.F. and Adams, W.T. 1989: Microgeographical patterns in allozyme differentiation in Douglus-fir from southwest Oregon. *For. Sci.* **35**: 3-15.

Moran G.F., Bell J.C., and Hilliker, A.J. 1983: Greater meiotic recombination in male vs female gametes in Pinus radiata. *J. Hered.* **74**: 62.

Moran, G.F., Bell, J.C. and Matheson, A.C., 1980: The genetic structure and levels of inbreeding in a *Pinus radiata* D. Don seed orchard. *Silvae Genet.* **29**: 190-193.

Morgante, M., Vendramin, G.G. and Oliveri, A.M. 1991: Mating system analysis in Pinus leucodermis Ant: detection of self-fertilisation in natural populations. *Heredity*. **67**: 197-203.

Morgante, M., Vendramin, G.G. et al. 1990: Mating system analysis in Pinus Leucodermis Ant: Detection of self-fertilisation in natural populations. *Heridity*. **67**: 197-203.

Müller-Stark G. and Ziehe, M. 1984: Reproductive systems in conifer seed orchards. 3. Female and male fitness of individual clones realized in seeds of Pinus sylvestris L. *Theor. Appl. Genet.* **69**: 173-177.

Müller-Stark, G. 1982: Reproductive systems in conifer seed orchards. I. Mating probabilities in a seed orchard of Pinus sylverstris L. *Silvae Genet.*, **31**: 188-197.

Müller-Stark, G., Baradat, Ph. and Bergmann, F. 1992: Genetic variation within European tree species. *New Forests*. **6**: 23-47.

Mullin, E.K. and Park, Y.S. 1992: Estimating genetic gains from alternative breeding strategies for clonal forestry. *Can. J. For. Res.* **22**: 14-23.

Muona, O., Paule, L., Szmidt A.E. and Karkkainen, K. 1990: Mating system analysis in a central and northern European population of *Picea abies*. *Scand. J. For. Res.* **5**: 97-102.

Muona, O. 1989: Population genetics in forest tree improvement. In: Brown A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S. (eds). Plant population genetics, breeding and genetic resources. Sinauer, Sunderland. pp 282-298.

Muona, O. and Harju, A., 1989: Effective population size, genetic variability, and mating system in natural stands and seed orchards of Pinus sylverstris. *Silvae Genet.*,**38**:221-228.

Murawski, D.A., Fleming, T.H., Ritland, K. and Hamrick, J.L. 1994: Mating system of *Pachycereus pringlei*: an autotetraploid Cactus. *Heredity*. **72**: 86-94.

Nagasaka K. and Szmidt A.E., 1985: Multilocus analysis of external pollen contamination of a Scots pine (*Pinus sylverstris* L.) seed orchard. pp. 134-138. In *Population Genetics in Forestry*. Proceedings of the meeting of the IUFRO working party "Ecological and population genetics" held in Gottingen, August 21-24, 1984. (ed H.-R. Gregorius). Springer-Verlag.

Namkoong G. and Kang, H. 1990: Quantitative genetics of forest trees. *Plant breeding reviews*. **8**:139-188.

Namkoong, G, Barnes, R.D. and Burley, J. 1980: A philosophy of breeding strategy for tropical forest trees. Oxford. UK; Commonwealth forestry Institute. Tropical Forestry Papers No. 16, 67 pp.

Namkoong, G. (1979). *Introduction to Quantitative Genetics in Forestry*. Forest Service, United States Department of Agriculture, Washington D.C.

Namkoong, G. 1966: Inbreeding effects on estimation of genetic additive variation. *For. Sci.* **66**: 8-13.

Namkoong, G., 1976: A multiple-index selection strategy. *Silvae Genet.*,**33**: 199-201.

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

Namkoong, G., Kang, H.C. and Brouard, J.S. 1988: Tree breeding: Principles and strategies. New York, USA; Springer-Verlag. 180 pp.

Neal, D.B. & Adams, W.T., 1985: The mating system in natural and sheltwood stands of Douglus fir. *Theor. Appl. Genet.* **71**: 201-207.

Neale, D.B. and Sederoff, R.R. 1989: Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theor. Appl. Genet.* **77**: 212-216.

Neale, D.B., Devey, M.E., Jermstad, K.D., Ahuja, M.R., Alosi, M.C. and Marshall, K.A. 1992: Use of DNA markers in forest tree improvement research. *New Forests*. **6**: 391-407.

Neale, D.B., Marshall, K.A. and Sederoff, R.R., 1989: Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D.Don Endl. *Proc. Natl. Acad. Sci. USA*. **86**: 9347-9349.

Neale, D.B., Wheeler, N.C. and Allard, R.W. 1986: Paternal inheritance of chloroplast DNA in Douglas-fir. *Can. J. For. Res.* **16**: 1152-1154.

Nei, M. 1977. F-statistics and analysis of gene diversity in subdivided populations. *Ann. Human Genet.* **41**: 225-233.

Nei, M., 1975: Molecular population genetics and evolution. North Holland, Amsterdam and New York.

Nei, M. 1978: Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*. **89**: 583-590.

Nelson, C.D., Nance, W.L., and Wagner, D.B. 1994: Chloroplast DNA variation among and within taxonomic varieties of Pinus caribaea and Pinus elliottii. *Can. J. For. Res.* **24**: 424-426.

Newton, K.J., 1988: Plant mitochondrial genomes: organization, expression and variation. Ann. Rev. Pl. Physiol. *Mol. Biol.* **39**: 503-532.

Nikles, D.G. 1967: Comparative variability and relationship of Caribbean pine (*Pinus caribaea* Mor.) and slash pine (*P. elliottii* Engelm.). Dissert. Abstr. 27 B (11): 3738.

Nikles, D.G. 1979: Genetic improvement of tropical lowland conifers. Miscellaneous Publication. Misc/79/15. FAO: Rome.

Nikles, D.G. 1989: Early growth and potential value of *P. caribaea* var. *hondurensis* × *P. tecunumanii* and × *P. oocarpa* F_1 hybrids in Australia. In *Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry* [Proc. IUFRO Conference, Pattaya,Tailand, Nov. 1988]. (Gibson, G.L., Griffin, A.R. and Matheson, A.C. eds). Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

Nikles, D.G., Burley, J. and Barnes, R.D. (Eds). 1978: Progress and problems of genetic improvement of tropical forest species. Proceedings of a Joint Workshop. Brisbane, Queensland, Australia, 4-7 April, 1977. Vol. 1.

Nikles, D.G. and Robinson, M.J. 1989: The development of Pinus hybrid for operational use in Queensland. pp. 272-282. In: Gibson, G., Griffin, A.R. and Matheson, A.C. (Eds). *Breeding Tropical Trees: Population structure and genetic improvement strategies in clonal and seedling forestry.* [Proc. IUFRO Conference, Pattaya,Tailand, Nov. 1988]. (Gibson, G.L., Griffin, A.R. and Matheson, A.C. eds). Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

Nybom, H., Rogsad, S.H. and Schaal, B.A. 1990: Genetic variation detected by use of the M13 'DNA fingerprint' probe in *Malus*, *Prunus*, and *Rubus* (Rosaceae). *Theor. Appl. Genet.* **79**: 153-156.

O'Malley, D.M., Allendorf F.M. and Blake, G.M. 1979: Inheritance of isozyme variation and heterozygosity in *Pinus ponderosa. Biochem. Genet.* **17**: 233-250.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H., 1986: Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature.* **322**:572-574.

Omi, S.K. and Adams, W.T. 1986: Variation in seed set and proportions of outcrossed progeny with clones, crown position, and top pruning in a Douglas-fir seed orchard. *Can. J. For. Res.* **16**: 502-507.

Palmer, J.D. 1990: Contrasting modes and tempos of genome evolution in land plant organelles. *TIG.* **6**: 115-120.

Palmer, J.D., Jansen, R.K., Michaels, H.J., Chase, M.W. and Manhart, J.R., 1988: Chloroplast DNA variation and plant phylogeny. *Ann. Missouri Bot. Gard.* **75**: 1180-1206.

Pan Z.-G and You Y.-T. (1991). Introduction and cultivation of *Pinus elliottii, P. taeda* and *P. caribaea* in China. Beijing Science-Technology press, Beijing.

Pan Z.-G., 1991. Growing Exotic Trees in China. China Forestry Press, Beijing. (In Chinese), 194 pp.

Paterson, A.H., Damon, S., hewitt, J.D., Zamir, D., Rabinowitch, H.D., Lincoln, S.E., lander, E.S. and Tanksley, S.D 1991: Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. *Genetics.* **127**: 181-197.

Paule, L., Lindgren, D. & Yazidani, R. 1993: Allozyme frequencies outcrossing rates and pollen contamination in *Picea abies* seed orchard. *Scand. J. For. Res.* **8**: 8-17.

Pederick, L.A. 1968: Chromosome inversions in *Pinus radiata. Silvae Genet.* **17**: 22-26.

Peng, M. 1988: Vegetative propagation of *Pinus caribaea* by leaf cuttings. *Yunnan Forestry.* **4**: 44-48. (in Chinese).

Plumptre, R.A. 1984: *Pinus caribaea.* Volume 2: Wood properties. Tropical forestry paper No. 17. Unit of tropical silviculture, Commonwealth forestry Institute, University of Oxford. 148 pp.

Poulik, M.D. 1957: Starch gel electrophoresis in a discontinuous system of buffers. *Nature,* **180**:1477-1478.

Pring, D.R. and Lonsdale, D.M., 1985: Molecular biology of higher plant mitochondrial DNA. *Int. Rev. Cytol.* **97**: 1-46.

Queller D.C. and Goodnight, K.F. 1988: Estimating relatedness using genetic markers. *Evolution.* **43**: 258-275.

Raymond, M. and Rousset, F. 1995a: GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *J. Hered.* **86**: 248-249.

Raymond, M. and Rousset, F. 1995b: An exact test for population differentiation. *Evolution.* **49**(6): 1280-1283.

Ritland, K. and Jain, S. 1981: A model for the estimation of outcrossing rate and gene frequencies using n independent loci. *Heredity.* **47**(1): 35-52.

Ritland, K., 1983: Estimation of mating system. pp. 289-302. In: *Isozyme in plant genetics and breeding, Part A.* (eds Tanksley, S.D. and Orton, T.J.). Elsevier, Amsterdam.

Ritland, K., 1984: The effective proportion of self-fertilisation with consanguineous matings in inbred population. *Genetics.* **106**: 139-152.

Ritland, K. 1985: The genetic mating structure of subdivided populations. 1. Open-mating model. *Theor. Popul. Biol.* **27**: 51-74.

Ritland, K. 1986: Joint maximum likelihood estimation of genetic and mating structure using open-pollinated progenies. *Biometrics.*

Ritland, K. 1991: Multilocus estimation program MLT. Computer program document. Department of Botany, University of Toronto.

Ritland, K., 1990: A series of FORTRAN programs for estimating plant mating systems. *J. Heredity.* **81**:235-237.

Ritland, K. 1995: Multilocus mating system program MLTR. (Version 0.9). Computer program document. Department of Botany, University of Toronto.

Ritland, K. and Clegg, M.T., 1987: Evolutionary analysis of plant DNA sequences. *The American Naturalist.* **130**: 74-100.

Ritland, K. and El-Kassaby Y.A., 1985: The nature of inbreeding in a seed orchard of Douglas fir as shown by an efficient multilocus model. *Theor. Appl. Genet.* **71**: 375-384.

Robertson, A. 1960: A theory of limits in artificial selection. *Proc. Royal Soc.* London. B **153**: 234-249.

Rogstad, S.H., Patton, J.C. and Schaal, B.A., 1988: M13 repeat probe detects DNA minisatellite-like sequences in gymnosperms and angiosperms. *Proc. Natl. Acad. Sci. USA.* **85**: 9176-9178.

Ross M.D and Gregorius H-R. 1983: Outcrossing and sex function in hermaphrodites: a resource-allocation model. *Am. Nat.* **121**:204-222.

Rousset, F. and Raymond, M. 1995: Testing heterozygote excess and deficiency. *Genetics.* **140**: 1413-1419.

Runquist, E.W. 1968: meiotic investigations in Pinus sylvestris L. *Hereditas.* **60**: 77-128.

Rudin, D. & Lindgren, D. 1977: Isozyme studies in seed orchards. *Studia For. Suec.* **139**: 5-23.

Rudin, D., Muona, O. and Yazdani, R. 1986: Comparison of the mating system of *Pinus sylvestris* in natural stands and seed orchards. *Hereditas.* **104**: 15-19.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989: Molecular Cloning. A laboratory manual. Second edition. Cold Spring Harbor Laboratory Press.

Sax, K. and Sax, H.J. 1933: Chromosome number and morphology in the conifers. *J. Arnold. Arbor.* **14**: 356-374.

Saylor L.C. and Smith, B.W. 1968: Meiotic irregularity in species and interspecific hybrids of Pinus. *Am. J. Bot.* **53**: 453-468.

Schaal, B.A., O'kane, S.L. and Rogstad, S.H., 1991: DNA variation in plant populations. *TREE.* **6**: 329-333.

Schaal, B.A., Leverich, W.J. and Rogstad, S.H. 1991: A comparison of methods for assessing genetic variation in plant conservation biology. pp. 123-134. In: Falk D.A and Holsinger, K.E. (Eds). Genetics and conservation of rare plants. Oxford University Press. New York, Oxford.

Schoen, D.J. and Stewart, S.C. 1986: Variation in male reproductive investment and male reproductive success in white spruce. *Evolution*, **40**: 1109-1120.

Sederoff, R.R., 1984: Structural variation in mitochondrial DNA. *Adv. Genet.* **22**: 1-108.

Shaw, C.R. & Prasad, R. 1970: Starch gel electrophoresis. A compilation of recipes. *Biochem. Genet.*, **4**: 297-320.

Shaw, D.V. and Allard, R.W. 1982: Estimation of outcrossing rate for open-pollinated progenies of Douglas-fir. *Theor. Appl. Genet.* **62**: 113-120.

Shaw, D.V., Kahler, A.L. & Allard, R.W. 1981: A multilocus estimator of mating system parameters in plant populations. *Proc. Natl. Acad. Sci. USA.* **78**: 1298-1302.

Shen X.-H., Rudin, D. and Lindgren, D. 1981: Study of pollination patterns in a Scots pine seed orchard by means of isozyme analysis. *Silvae Genet.*, **30**(1): 7-15.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chengwongse, J., Obokata, J., Yamaguchi-shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada,H. and Sugiura, M. 1986: The complete nuceotide sequence of the tabacco chloroplast genome: its gene organization and expression. *EMBO J.* **5**: 2043-2049.

Sigurgeirsson, A. and Szmidt, A.E., 1992: Phylogentic and biogeographic implications of chloroplast DNA variation in *Picea. Nord. J. Bot.*

Sigurgeirsson, A., 1992: Insights into the evolution of Picea inferred from chloroplast DNA. PhD thesis. Swedish University of Agricultural Sciences, Faculty of forestry, Department of Forest Genetics and Plant Physiology. Umea, Sweden.

Simons, A.J. 1992: Genetic improvement of non-industrial trees. *Agroforestry Systems.* **18**: 197-211.

Simons, A.J. and Dunsdon, A.J. 1992: Evaluation of the potential for genetic improvement of *Gliricidia sepium.* Final report, ODA Forestry research project R4525. 176 pp. Oxford Forestry Institute, Oxford, UK.

Slatkin, M. 1987: Gene flow and the geographic structure of natural populations. *Science.* **236**: 787-792.

Slatkin, M. and Barton, N.H., 1989: A comparison of three indirect methods for estimating average levels of gene flow. *Evolution.* **43**: 1349-1368.

Slee, M.U. and Nikles D.G. (1968). Variability of Pinus caribaea Morelet in young Queensland plantations. 9th British Commonwealth For. Conf., New Delhi.

Smith D.B. and Adams, W.T., 1983: Measuring pollen contamination in clonal seed orchards with the aid of genetic markers. pp 69-77. In Proc. 17th South Conf. on Forest Tree Improv.

Snedecor G.W. and Cochran, W.G. (1967). *Statistical Methods.* 6th edition. Iowa State University Press. Ames, Iowa, USA.

Sokal, R.R. and Rohlf, F.J. 1969: Biometry. 2nd ed. W.H. Freeman and Company, San Francisoco. 859 pp.

Soller, M. and Beckmann, J.C. 1983: Genetic polymorphism in varietal identification and genetic improvement. *Theor. Appl. Genet.* **67**: 25-33.

Sorensen F.C. and Miles, R.S. 1982: Inbreeding depression in height, height growth, and survival of Douglas-fir, ponderosa pine and noble fir to 10 years of age. *For. Sci.* **28**: 283-292.

Sorensen, F.C., 1982: The roles of polyembryony and embryo viability in the genetic system of conifers. *Evolution*, **36**: 725-733.

Sorensen, F.G. and Miles, R.S., 1974: Self-pollination effects on Douglas-fir and ponderosa pine seedlings. *Silvae Genet.* **23**: 135-138.

Squillace, A.E. 1974: Average genetic correlations among offspring from open-pollinated forest trees. *Silvae Genet.*,**23**: 149-156.

Squillace, A.C. and Goddard, R.E, 1982: Selfing in clonal seed orchards of slash pine. *For. Sci.* **28**: 71-83.

Stine, M. and Keathley, D.E., 1990: Paternal inheritance of plastids in Engelmann spruce x blue spruce hybrids. *J. Hered.* **81**: 443-446.

Stonecypher, R.W. 1992: Computational Methods. Finns *et al.* (eds), *Handbook of Quantitative Forest Genetics*. Kluwer Academic Publishers. 195-228

Strauss S.H. and Conkle, M.T., 1986: Segregation, linkage, and diversity of allozymes in knobcone pine. *Theor. Appl. Genet.* **72**: 483-493.

Strauss, S.H. and Doerksen, A.H., 1990: Restriction fragment analysis of pine phylogeny. *Evolution.* **44**: 1081-1096.

Strauss, S.H. and Tsai, C.-H. 1988: Ribosomal gene number variability in Douglas-fir. *J. Hered.* **79**: 453-458.

Strauss, S.H., Bousguet, J., Hipkins, V.D. and Hong Y.-P. 1992: Biochemical and molecular genetic markers in biosystematic studies of forest trees. *New Forests.* **6**: 125-158.

Strauss, S.H., Doerksen, A.H. and Byrne, J.R., 1990: Evolutionary relationships of Douglas-fir and its relatives (genus *pseudotsuga*) from DNA restriction fragment analysis. *Can. J. Bot.* **68**: 1502-1510.

Strauss, S.H., Palmer, J.D., Howe, G.T. and Doerksen, A.H. 1988: Cloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. *Proc. Natl. Acad. Sci. USA.* **85**: 3898-3902.

Sugiura, M., 1992: The chloroplast genome. *Pl. Mol. Biol.* **19**: 149-168.

Swofford D.L and Selander, R.B. 1981: BIOSYS-1, A fortran program for the comprehensive analysis for electrophoretical data in population genetics and systematics. *J. Hered.*, **72**: 281-283.

Swofford D.L and Selander, R.B. 1989: BIOSYS-1, A computer program for the analysis of allelic variation in population genetics and biochemical systematics. Release 1.7. 43 pp. Software Manual.

Szmidt, A.E. 1987: Genetic composition of seed orchard crops. *For. Ecol. Manag.*, **19**: 227-232.

Szmidt, A.E. and Wang, X-R., 1991: DNA markers in forest genetics. pp. 79-94. In: *Genetic variation in European populations of forest trees.* (eds Muller-Starck, G and Ziehe, M.). Sauerlander's Verlag, Frankfurt am Main.

Szmidt, A.E. 1991: Phylogenetic and applied studies on chloroplast genome in forest conifers. In: Biochemical markers in the population genetics of forest trees. Fineschi, S., Malvolti, M.E., Cannata, F. and Hattemer, H.H. (Eds). SPB Academic Publishing bv, The Hague: pp. 185-196.

Szmidt, A.E., Alden, T. and Hakkgren, J.-E. 1987: paternal inheritance of chloroplast DNA in *larix*. *Pl. Mol. Biol.* **9**: 59-64.

Szmidt, A.E., El-Kassaby, Y.A., Sigurgeirsson, A., Alden, T., Lindgren, D. and Hallgren, J.-E. 1988a: Classifying seedlots of *Picea sitchensis* and *P. glauca* in zones of introgression using restriction analysis of chloroplast DNA. *Theor. Appl. Genet.* **76**: 841-845.

Szmidt, A.E., Sigurgeirsson, A., Wang, X.-R., Hallgren, J.-E. and Lindgren, D. 1988b: Genetic relationships among *Pinus* species based on chloroplast DNA polymorphism. In: Frans Kempe Symposium on Molecular genetics of forest trees. Hallgren, J.-E. (Ed). June 14-16, Umea, Sweden. pp. 33-47.

Talbert, C.B. 1992: Quantitative genetics: Why bother? In: Finns *et al.* (eds), *Handbook of Quantitative Forest Genetics*. Kluwer Academic Publishers. pp. 1-28.

Tan, S.-M. and Xu, C.-Y. 1990: An investigation of growth of *Pinus caribaea* in Guangdong Province. *Guangdong Forestry Sci. and Tech.* **2**: 24-27. (in Chinese)

Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. 1989: RFLP mapping in plant breeding: new tools for an old science. *Biotechnology*. **7**: 257-264.

Tigerstedt, P.M.A. 1973: Studies on isozyme variation in marginal and central populations of Picea abies. *Hereditas*. **75**: 47-60.

Tulsieram, L.K., Glaubitz, J.C., Kiss, G. and Carlson, J.E., 1992: Single tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. *Biotechnology*. **10**: 686-690.

Vallejos C.E. and Tanksley, S.D. 1983: Segregation of isozyme markers and cold tolerance in an interspecific backcross of tomato. *Theor. Appl. Genet.* **66**: 241-247.

van Buijtenen, J.P. and Lowe, W.J., 1979: The use of breeding groups in advanced generation breeding. pp 59-65. In: Proc. 15th South. For. Tree Improve. Conf. June 19-21, Miss. State, MS, USA.

van Buijtenen, J.P., 1976: Mating designs. pp 11-27. In: Proc. IUFRO Joint Meeting on Advanced generation breeding. June 14-18, Bordeaux, France.

Wagner, D.B. 1992: Nuclear, chloroplast, and mitochondrial DNA polymorphisms as biochemical markers in population genetic analyses of Forest trees. *New Forests*. **6**: 373-390.

Wagner, D.B., Dong, J., Carlson, M.R. and Yanchuk, A.D. 1991: Paternal leakage of mitochondrial DNA in *Pinus*. *Theor. Appl. Genet.* **82**: 510-514.

Wagner, D.B., Furnier, G.R., Saghai-Maroof, M.A., Williams, S.M., Dancik, B.P. and Allard, R.W. 1987: Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proc. Natl. Acad. Sci. USA*. **84**: 2097-2100.

Wagner, D.B., Govindaraju, D.R., Yeatman, C.W. and Pitel, J.A. 1989: Paternal chloroplast DNA inheritance in a diallel cross of Jack pine (*Pinus banksiana* Lamb). *J. Hered.* **80**: 483-485.

Wahlund, S. 1928: Zusammersetung von Populationen und Korrelation-sercheinungen von Standpunkt der Verebungslehre aus betrachtet. *Hereditas*. **11**: 65-106.

Wang H.R. and Zheng Y.Q. 1993: *Species and Provenance Selection of Eucalypts in Subtropical China*. Final report of Sino-Australian co-operative research project.

Wang H.R, Zheng Y.Q. and Wang W.H. 1992: Growing Performance of *Casuarina junghuniana* and other 6 species of Casuarinaceae in Southeast China. *Scientia Silvae Sinicae*. **28**(4), 343-347.

Wang H.R. 1994: Tree introduction, breeding and improvement with exotic species in China. Asia-Pacific Symposium on forest genetic improvement. 18-22 Oct. 1994. Beijing.

Wang, G.-M., Ye, Q., Li, C. and Wu, H. 1989: Summary of provenance trials of Pinus caribaea. *Guangdong Forestry Sci. and Tech.* **1**: 22-27. (in Chinese).

Wang, H-R., Malcolm, D.C. and Fletcher, A. 1995: *Pinus caribaea* Morelet in China: past and future. (Submitted).

Wang, H.-Y. and Lai, J.-X. 1990: Growing more *Pinus caribaea* and *P. taeda* in Guangxi. *Guangxi Forestry.* **1**: 11-12. (in Chinese).

Wang, S., Lindgren, D., Szmidt, A.E. & Yazdani, R. 1991: Pollen migration into a seed orchard of *Pinus sylverstris* L. and the methods of its estimation using allozyme markers. *Scand. J. For.Res.* **6**: 379-385.

Wang, X-R. 1992: Genetic diversity and evolution of Eurasian Pinus species. Dissertation, The Swedish University of Agricultural Science, Department of Forest Genetics and Plant Physiology, Umea.

Wang, X.-R., Szmidt, A.E., Sigurgeirsson, A., and Karpinska, B. 1991a: A chloroplast DNA story about Asiatic pines. In: Biochemical markers in the population genetics of forest trees. Fineschi, S., Malvolti, M.E., Cannata, F. and Hattemer, H.H. (Eds). SPB Academic Publishing bv, The Hague: pp. 257-270.

Weir, B.S. and Cockerham, C.C. 1984: Estimating F-statistics for the analysis of population structure. *Evolution.* **38**: 1358-1370.

Weir, R.J. and Zobel B.J. 1975: Advanced-generation seed orchards. In: Faulkner R. (ed) Seed orchards. *For. Comm. London Bull.* **54**: 118-127.

Westfall R.D. 1992: Developing Seed Transfer Zones. In: Finns *et al.* (eds), *Handbook of Quantitative Forest Genetics.* Kluwer Academic Publishers. pp. 140-194

Wheeler, N.C. and Guries, R.P. 1982: Population structure, genic diversity and morphological variation in Pinus contorta Dougl. *Can. J. For. Res.* **12**: 595-606.

Wheeler, N.C. and Jech, K.S. 1986: Estimating supplemental mass pollination (SMP) success electrophoretically. pp. 111-120. In: Proceedings Canadian Tree Improvement Association. Quebec City, Canada.

Wheeler, N.C. and Jech, K.S. 1992: The use of electrophoretic markers in seed orchard research. *New forest.* **6**: 311-328.

White, T. 1992: Advanced-generation breeding populations: Size and structure. Paper at IUFRO conference. Breeding tropical trees: resolving tropical forest resource concerns through tree improvement, gene conservation and domestication of new species. Cartegena and Cali, Colombia, 9-18 October, 1992.

White, T.L. and Hodge, G.R. 1989: Predicting breeding values with applications in forest tree improvement. Dordrecht, Netherlands; Kluwer, Forestry Sciences 33. 367pp.

White, T.L., Matheson, A.C., Boomsma, D.B. and Rout, A.F. 1992: Logistics, costs and genetic gains of five options of nucleus breeding strategies. Mt Gambier, Australia; South tree breeding association technical report 92-01. 38 pp.

Whittemore, A.T. and Schaal, B.A. 1991: Interspecific gene flow in oaks. *Proc. Natl. Acad. Sci. USA.* **88**: 2540-2544.

Whyte, A.G.D., Adams, P.M. and McEwan, S.E. 1981: Foxtailing of *Pinus caribaea* var. *hondurensis* in Fiji: frequency distribution of occurrence and wood properties. *Forest Ecology and Management.* **3**(3): 237-243.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990: DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acids Res.* **18**: 6531-6535.

Wolfe, K.H., Li, W. and Sharp, P.M. 1987: Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. *Proc. Natl. Acad. Sci. USA.* **84**: 9054-9058.

Woods, F.W., Vincent, L.W. Moscheler, W.W. and Core, H.A. 1979: Height, diameter and specific gravity of "foxtail" trees of *Pinus caribaea* Morelet. *Forest Products Journal.* **29**(5): 43-44.

Wright, J.W., Read, R.A., Lester, D.T., Merritt, C. and Mohn, C. 1972: Geographic variation in red pine, 11-year data from the North Central States. *Silvae Genet.*,**21**: 205-209.

Wright, S. 1951: The genetic structure of populations. Ann. *Eugenics*, **15**: 323-354.

Wright, S. 1965: The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution.* **9**: 395-420.

Wright, S. 1969: Evolution and genetics of populations. Vol 2. University of Chicago Press, Chicago.

Wright, S. 1978: Evolution and genetics of populations. Vol 4. University of Chicago Press, Chicago.

Wu, Z.-H. and Ye, Q.-B. 1992: Growth analysis of *Pinus caribaea var. caribaea. Hainan Forestry.* **2**: 7-11. (in Chinese).

Wu, Z.-H. and Ye, Q.-B. 1993: An investigation of seed production of *Pinus caribaea var. caribaea. Hainan Forestry.* No. **1**: 48-53. (in Chinese).

Ye, Q.B. 1990: Performance of F_2 hybrid between Pinus elliottii and P. caribaea. *Guangdong Forestry Science and Technology.* **3**: 9-12 (in Chinese).

Yeh, F.C. & Morgan, K. 1987: Mating system and multilocus associations in a natural population of *Pseuodotsuga menziesii* (mirb.) Franco. *Theo. Appl. Gen.* **73**:799-808.

Yeh, F.C. and El-Kassaby, Y.A. 1980: Enzyme variation in natural populations of Sitka spruce (*Picea sitchensis*). Genetic variation patterns among trees from 10 IUFRO provenances. *Can. J. For. Res.* **10**: 415-422.

Yeh, F.C. 1979: Analysis of gene diversity in some species of conifer. In " Isozyme of North America forest trees and forest insects". Conkle, M.T. (ed.). USDA Forest Service, Berkeley, CA. pp. 48-52.

Yeh, F.C. 1988: isozyme variation of *Thuja plicata* (Cupressaceae) in British Columbia. *Biochem. Syst. Ecol.* **16**: 373-377.

Yeh, F.C. 1989: Isozyme analysis for revealing population structure for use in breeding strategies. In: Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin,

A.R. & Matheson, A.C.). pp. 119-132. Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

Yeh, F.C., Brune, A., Cheliak, W.M., Chipman, D.C. 1983: Mating system of *Eucalyptus citridora* in a seed orchard. *Can. J. For. Res.* **13**: 1051-1055.

Zeng, Y.-T. and Yang, M.-Q. 1992: Provenance trial and pest control of *Pinus caribaea*. *Forest Research*. **5** (1): 95-99. (in Chinese).

Zhang, Y.-M. 1994: Raising seedlings and disease control of Pinus caribaea var. hondurensis. *Hainan Forestry*. No. **1**: 10-14. (in Chinese).

Zheng Y.-Q, Ennos, R. and Wang H.-R. 1994: Provenance variation and genetic parameters in a trial of *Pinus caribaea* Morelet var. *bahamensis* Barr. and Golf. *Forest Genetics*. **1** (3): 165-174.

Zheng Y.-Q. and Wang H.-R. 1993a: Species/Provenance Trial of *Eucalyptus smithii* R. Baker and Other 23 Temperate Eucalypt Species in China. In Proceedings of International workshop on 'Research on Australian Broad-leaved Tree Species in China' (ed. A.G. Brown). CSIRO, Australia.

Zheng Y.Q. and Wang Z.H. 1993b: Species Trial of Eucalypts in Chuxiong Prefecture. *Research on Australian Broad-leaved Tree Species in China* (ed. Hong J.S. & Wang H.R.). China Forest Press.

Zhong, W.-H., Chen B.Q., Zheng, S.Y., Cheng, H.C. and Li, C.M. 1993: Study of establishing 1.5 generation seed orchard of *Pinus caribaea* var. *caribaea*. *Guangdong Forestry Sci. and Tech.* No. **4**: 17-20. (in Chinese).

Zhou, X.-Y. 1991: Study on the resistance of *Pinus caribaea* to *Hemiberlesia pitysophila*. *Guangdong Forestry Sci. and Tech.* No. **1**: 35-36. (in Chinese).

Zhu, Z.-S., Ding, R.-C. and Wang, G.-M. (Eds). 1986: Caribbean pine. Guangdong Science and technology Press. 228 pp. (in Chinese).

Zobel, B. J. and Talbert, J. 1984: *Applied Forest Tree Improvement*. John Wiley & Sons.

Zobel, B. J. Wyk, G.V. and Stahl, P. 1987: Growing exotic Forests. John Wiley & Sons. New York, Brisbane, Toronto. 508 pp.

Zurawski, G. and Clegg, M.T. 1987: Evolution of higher-plant chloroplast DNA-encoded genes: implications for structure-function and phylogenetic studies. *Ann. Rev. Plant physiol.* **38**: 391-418.

9. Appendices

Appendix 2.1 Composition of extraction, gel and electrode buffers used for starch gel electrophoresis

Extraction buffer	Sodium Phosphate 0.2 M dithiothreitol 1 mg/ml		
System	**Gel buffer**	**Electrode buffer**	**Enzyme systems**
B	For 2000 ml solution: Tris 12.1 g d H_2O 1000 ml (Adjust pH to 8.5 with 0.2M Citric Acid, use as it is)	For 2000 ml solution: Sodium hydroxide 4 g Boric Acid 37 g d H_2O 2000 ml (Adjust pH to 8.0 with NaOH)	AAT
H	For 200 ml stock: Histdine-HCl 2.095 g EDTA 0.208 g d H_2O 200 ml (Adjust pH to 7.0 with 1 M Tris; Dilute stock 4:1 when using)	For 2000 ml solution: Tris 30.28 g (Adjust pH to 7.0 with 1 M Citric Acid)	IDH, MDH, PGD, PGM,

Note: Extraction buffer used for both endosperms and embryos

Appendix 2.2 Hardy-Weinberg exact tests (probability) for var. *caribaea*

Population	PAL (1)	MAN (2)	MBJ (3)	JAG (4)	IDJ (5)	SOR (6)	CHN (7)
AAT-a	0.00**	0.75	0.09	0.01	0.03*	0.52	1.00
AAT-b	1.00	0.00**	-	0.00**	1.00	0.00**	0.00**
AAT-c	-	-	-		-	-	0.00**
IDH	0.70	1.00	0.07	0.40	0.23	0.88	0.02*
MDH	0.56	0.48	0.12	1.00	1.00	-	0.01**
PGD-a	-	-	1.00	-	-	-	0.08
PGD-b	0.00**	0.14	0.00*	0.12	0.00**	0.00**	0.00**
PGM	0.53	0.01**	0.06	0.24	0.24	1.00	0.66

** significant at 5% level; ** significant at 1% level*
(H_0 = no heterozygote deficit)

Appendix 2.3 Probability of linkage disequilibrium for var. *caribaea*

Genepop (Version 2), Genotypic disequilibrium
Markov chain parameters
Dememorization: 1000
Batches: 50
Iterations per batch: 1000

Pop	code	Locus#1	Locus#2	P-value	S.E.
CHN	9	aat_a	aat_b	0.23464	0.01131
CHN	9	aat_a	aat_c	0.44966	0.00854
CHN	9	aat_b	aat_c	0.51390	0.01411
CHN	9	aat_a	idh	0.04938*	0.00342
CHN	9	aat_b	idh	0.63220	0.00752
CHN	9	aat_c	idh	0.51760	0.00392
CHN	9	aat_a	mdh	0.11832	0.00377
CHN	9	aat_b	mdh	0.82288	0.00421
CHN	9	aat_c	mdh	1.00000	0.00000
CHN	9	idh	mdh	0.00872**	0.00118
CHN	9	aat_a	pgd_a	0.03366*	0.00274
CHN	9	aat_b	pgd_a	0.49944	0.01177
CHN	9	aat_c	pgd_a	0.21492	0.00855
CHN	9	idh	pgd_a	0.00084**	0.00028
CHN	9	mdh	pgd_a	1.00000	0.00000
CHN	9	aat_a	pgd_b	0.53610	0.01308
CHN	9	aat_b	pgd_b	0.47628	0.01728
CHN	9	aat_c	pgd_b	0.79522	0.00958
CHN	9	idh	pgd_b	0.51638	0.01315
CHN	9	mdh	pgd_b	0.53802	0.01201
CHN	9	pgd_a	pgd_b	0.10928	0.00676
CHN	9	aat_a	pgm	0.06980	0.00558
CHN	9	aat_b	pgm	0.19580	0.01254
CHN	9	aat_c	pgm	1.00000	0.00000
CHN	9	idh	pgm	0.67900	0.00861
CHN	9	mdh	pgm	0.37462	0.01054
CHN	9	pgd_a	pgm	1.00000	0.00000
CHN	9	pgd_b	pgm	0.25772	0.01622

**: 5% significance leve; **: 1% significance level*
(Output for other populations is not presented due to limit of space)

P-value for each locus pair across all populations
(Fisher's method)

Locus pair	Chi2	df	P-value
aat_a & aat_b	6.816	10	0.74274
aat_a & aat_c	1.599	4	0.80906
aat_b & aat_c	1.331	4	0.85602
aat_a & idh	13.274	14	0.50510
aat_b & idh	5.301	8	0.72502
aat_c & idh	3.630	4	0.45844
aat_a & mdh	21.553	14	0.08827
aat_b & mdh	2.872	8	0.94213
aat_c & mdh	0.000	4	1.00000
idh & mdh	15.205	14	0.36431
aat_a & pgd_a	6.783	4	0.14782

aat_b	& pgd_a	1.389	2	0.49944
aat_c	& pgd_a	3.075	2	0.21492
idh	& pgd_a	14.704	6	0.02269*
mdh	& pgd_a	2.718	6	0.84330
aat_a	& pgd_b	13.800	14	0.46470
aat_b	& pgd_b	7.723	10	0.65591
aat_c	& pgd_b	1.298	4	0.86179
idh	& pgd_b	5.359	14	0.98016
mdh	& pgd_b	10.124	14	0.75303
pgd_a	& pgd_b	15.447	6	0.01705*
aat_a	& pgm	24.147	14	0.04398*
aat_b	& pgm	6.054	8	0.64122
aat_c	& pgm	0.000	4	1.00000
idh	& pgm	17.126	14	0.24951
mdh	& pgm	16.898	14	0.26165
pgd_a	& pgm	1.046	6	0.98381
pgd_b	& pgm	12.898	14	0.53457

Appendix 2.4 Genic differentiation for all pairs of populations of var. *caribaea*

Populations	*AAT*-a	*AAT*-b	*AAT*-c	IDH	MDH	*PGD*-a	*PGD*-b	PGM
PAL 1 & MAN 2	0.01**	0.00**	1.00	0.51	1.00		0.00**	0.77
PAL 1 & MBJ 3	0.03*	0.16	0.40	0.49	0.19	0.39	0.02*	0.29
PAL 1 & JAG 4	1.00	0.64	0.38	0.57	0.02*		0.01**	0.00**
PAL 1 & IDJ 5	0.07	0.72	0.31	0.09	0.01**		0.00**	0.42
PAL 1 & SOR 6	0.06	0.60	0.32	0.64	0.77	0.56	0.00**	0.03*
PAL 1 & CHN 9	0.04*	0.00**	0.21	0.00**	0.58	0.00**	0.00**	0.00**
MAN 2 & MBJ 3	0.49	0.00**		0.12	0.32	0.69	0.00**	0.23
MAN 2 & JAG 4	0.00**	0.00**		0.87	0.02*		0.00**	0.13
MAN 2 & IDJ 5	0.19	0.00**		0.50	0.00**		0.00**	0.81
MAN 2 & SOR 6	0.52	0.00**		0.17	1.00	1.00	0.00**	0.28
MAN 2 & CHN 9	0.46	0.00**	0.16	0.00**	0.69	0.00**	0.00**	0.00**
MBJ 3 & JAG 4	0.01**	0.50		0.15	0.00**	1.00	0.56	0.00**
MBJ 3 & IDJ 5	0.59	0.71		0.00**	0.00**	0.04*	0.00**	0.37
MBJ 3 & SOR 6	1.00	0.52		0.67	0.14	0.19	0.00**	0.08
MBJ 3 & CHN 9	1.00	0.00**	0.01**	0.00**	0.55	0.00**	0.00**	0.00**
JAG 4 & IDJ 5	0.05*	1.00		0.23	0.80		0.01**	0.03*
JAG 4 & SOR 6	0.04*	1.00		0.18	0.00**	1.00	0.00**	0.14
JAG 4 & CHN 9	0.03*	0.00**	0.01**	0.00**	0.00**	0.08	0.00**	0.00**
IDJ 5 & SOR 6	0.60	1.00		0.00**	0.00**	0.50	0.00**	0.56
IDJ 5 & CHN 9	0.66	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**
SOR 6 & CHN 9	1.00	0.00**	0.00**	0.00**	0.65	0.00**	0.00**	0.00**

* p < 5 %; ** p < 1 %;

Note: Empty cell indicates no data were available, or only one allele was present, or two alleles were detected but one was represented by only one copy (H_o= the allelic distribution is independent across populations).

Appendix 2.5 Hardy-Weinberg exact test for var. *bahamensis*

Population	Little Abaco	High Rock	Byfield Australia	San Andros	Adelaide	East New Providence
	P-val	P-val	P-val	P-val	P-val	P-val
AAT_a	1	0.562	0.1112	1	0.2215	1
AAT_b	-	1	1	-	-	1
AAT_c	1	0.2495	0.2567	1	1	1
IDH	-	0.206	0.1975	0.58	0.8854	0.5923
MDH	1	-	-	-	-	-
PGD_a	-	-	-	-	-	-
PGD_b	0.348	0.1048	0.0001**	0.0004**	0.0243	0.0122
PGM	0.3268	0.0121	0.0001**	0.0026**	0.0008	0.1572

Note: H_o=no heterozygote deficit

Appendix 2.6 Genotypic linkage disequilibrium of var. *bahamensis*

Genepop (Version 2), Genotypic disequilibrium

Markov chain parameters
Dememorization: 1000
Batches: 50
Iterations per batch: 1000

P-value for each locus pair across all populations
(Fisher's method)

Locus pair	Chi2	df	P-value
aat_a & aat_b	2.379	10	0.99252
aat_a & aat_c	8.012	12	0.78417
aat_b & aat_c	6.710	10	0.75248
aat_a & idh	11.879	12	0.45545
aat_b & idh	4.721	10	0.90901
aat_c & idh	12.678	12	0.39288
aat_a & mdh	0.000	8	1.00000
aat_b & mdh	7.332	8	0.50128
aat_c & mdh	0.000	8	1.00000
idh & mdh	0.000	8	1.00000
aat_a & pgd_a	3.293	4	0.50997
aat_b & pgd_a	0.000	4	1.00000
aat_c & pgd_a	2.937	4	0.56838
idh & pgd_a	0.000	4	1.00000
mdh & pgd_a	0.000	4	1.00000
aat_a & pgd_b	14.055	12	0.29720
aat_b & pgd_b	7.164	10	0.70988
aat_c & pgd_b	9.734	12	0.63931
idh & pgd_b	6.193	12	0.90605
mdh & pgd_b	8.828	8	0.35704
pgd_a & pgd_b	1.904	4	0.75333
aat_a & pgm	19.628	12	0.07446
aat_b & pgm	3.949	10	0.94960

aat_c	& pgm	15.456	12	0.21745
idh	& pgm	13.873	12	0.30890
mdh	& pgm	1.242	8	0.99621
pgd_a	& pgm	2.254	4	0.68915
pgd_b	& pgm	16.272	12	0.17908

Note: Output of probability test for each population is not presented due to limit of space. No significant linkage disequilibrium was found in any of the populatioins.

Appendix 2.7 Genic differentiation for all pairs of populations (H_o=the allelic distribution is independent across populations)

Populations	AAT-a	AAT-b	AAT-c	IDH	MDH	PGD-a	PGD-b	PGM
	Probability							
Little_ab & High_Rock	0.153	0.420	1.000	0.027*	0.550	1.000	1.000	1.000
Little_ab & Byfield_A	1.000	1.000	0.397	0.681	0.192	1.000	0.034*	0.034*
Little_ab & San_Andro	0.005**	0.545	0.778	0.000**	0.094		0.021*	0.021*
Little_ab & Adelaide	1.000	0.361	0.542	0.002**	0.119		0.000**	0.000**
Little_ab & East_New	0.012*	1.000	0.017*	0.002**	0.241		0.000**	0.000**
High_Rock & Byfield_A	0.060	0.210	0.286	0.022*	0.210	1.000	0.004**	0.004**
High_Rock & San_Andro	0.092	0.315	0.534	0.023*	0.315	1.000	0.001**	0.001**
High_Rock & Adelaide	0.084	0.272	0.271	0.130	0.272	1.000	0.000**	0.000**
High_Rock & East_New	0.247	0.605	0.003**	0.238	0.605	1.000	0.000**	0.000**
Byfield_A & San_Andro	0.001**	0.370	0.809	0.000**	1.000	1.000	0.009**	0.009**
Byfield_A & Adelaide	1.000	0.253	1.000	0.001**	1.000	1.000	0.000**	0.000**
Byfield_A & East_New	0.003**	0.843	0.103	0.001**	1.000	1.000	0.000**	0.000**
San_Andro & Adelaide	0.001**	1.000	0.801	0.488			0.000**	0.000**
San_Andro & East_New	0.764	0.425	0.034*	0.397	0.486		0.000**	0.000**
Adelaide & East_New	0.004**	0.500	0.087	0.864	1.000		0.002**	0.002**

**: 5% significant level; **: 1% significant level*

Appendix 3.1 Selected clone numbers, plus tree origin and number of ramets per clone collected

Clone	Provenance	# ramets	Clone	Provenance	# ramets
1	Marbajita	2	40	Marbajita	2
2	Marbajita	3	44	El caimito	2
5	Marbajita	1	48	Galalon	2
7	Marbajita	1	50	Marbajita	3
8	Marbajita	3	51	Marbajita	1
11	Marbajita	5	52	Marbajita	2
12	Marbajita	6	55	Marbajita	3
14	Marbajita	3	56	Marbajita	1
15	Marbajita	3	60	Marbajita	2
16	Marbajita	1	67	Marbajita	1
17	Marbajita	4	68	Marbajita	1
18	Marbajita	2	70	Marbajita	4
21	Marbajita	3	72	Marbajita	1
22	Marbajita	1	78	T. de Collantes	3
23	Marbajita	5	83	T. de Collantes	2
24	Marbajita	3	91	T. de Collantes	2
25	Marbajita	3	100	T. de Collantes	5
27	Marbajita	1	105	T. de Collantes	5
29	Marbajita	2	106	T. de Collantes	3
30	Marbajita	3	110	T. de Collantes	3
32	Marbajita	1	117	T. de Collantes	2
38	Marbajita	2			

Appendix 4.1 Procedures for preparing probe

- Growing bacteria *E. coli* in LB media
- Preparing bacteria DNA
- Digestion of the plasmid DNA
- Separating plasmid DNA with argrose gel
- Remove the correct bands and store for use as probe

Appendix 4.2 DNA extraction with CTAB buffer from individual seeds

Day 1

1. Remove scales from seed
2. Grind each seed in 50 μl of 2 x CTAB buffer in an Eppindorff tube
3. Add 400 µl of 2 x CTAB and mix throughly
4. Incubate in heat block at 65 °C for 45 minutes
5. Remove from heat block and allow to cool to room temperature
6. Add 20 µl of Proteinase K, mix well and incubate in the heat block at 37 °C for 1 hour

7. Add an equal volume of Chloroform/IAA (24:1), approx 600 μl and mix well
8. Spin at 13,000 rpm for 5 minutes
9. Decant supernatant into a fresh tube and add an equal volume of Isopropanol
10. Leave overnight at -20 °C

Day 2

11. Spin at 13,000 rpm for 5 minutes
12. Remove supernatant and resuspend the pellet in 200 μl of TE buffer until disolved (incubate at 37 °C), then add 200 μl of Phenol:Chloroform and mix well
13. Spin at 1,3000 rpm for 5 minutes, then decant the top phase into a fresh tube
14. Add 20 μl of 3M CH_3COONa and approx 1 ml of absolute ethanol mix well and place at -20 °C for 1 hour
15. Spin at 13,000 rpm for 5 minutes, drain pellet and air dry
16. Resuspend pellet in 100 μl of RNaseA and incubate at 37 °C overnight
17. Examine 10 μl on a mini gel.

SOLUTIONS

2x CTAB buffer
2% CTAB 2 g (CTAB - Hexadecyltrimethylammonium Bromide, Sigma H-5882) 1.4M NaCl (28 ml of 5M NaCl : 8.19g)
100 mM Tris-HCl pH 8.0 (10 ml of 1 M stock)
20 mM EDTA (4 ml of 0.5 M stock)
1% Polyvinylpyrrolidone 1 g
0.2% 2-mercaptoethanol 200µl

Make to 100 ml with distilled H_2O

TE buffer (pH 8.0): 10 mM Tris 1 mM EDTA adjust to pH 8.0 with HCl

RNAase stock solution (store at -20 °C)
Stock-I: 10 mg/ml in distilled H_2O
Stock-II: 1% of stock-I, i.e. 10µl of stock-I made to 1 ml.
Proteinase K stock solution: 10 mg/ml in distilled H_2O, (store at -20 °C)

Appendix 4.3 Procedures for Southern Blotting

1. Soak a filter paper in 20 x SSC buffer, Wrap the support (2 starch gene plates).
2. Place in tray with 20 x SSC, such that the surface of support and filter is above liquid, but filter is in good contact with the liquid.
3. Place 2 or 3 layers of filter paper (soaked in 20 x SSC) on the support. Flatten with glass rod.

4. Cut membrane into size of 15 cm x 11 cm for normal gel, mark it with a straight line (in pencil) 1.5 cm from the end of gel to well front, label membrane with name, date, and so on, notch in top left, soak in 20 x SSC.
5. Place gel face sown on filter papers.
6. Place membrane face down on gel.
7. Five layers of filter papers (cut to size) on membrane and paper towels + weight on filter papers.
8. Protect the paper towel fall into the buffer with Parafilm. Leave overnight.
9. Wash membrane in 2 x SSC for 15 minutes.
10. Air dry on a filter paper.
11. Bake the membrane in a oven at 80 °C for 2 hours, or, UV crosslink face down on transilluminator for 2 minutes.
12. Store at room temperature for hybridisation.

Note: always make sure of no air bubbles between layers to make firm contact.

Appendix 4.4 Procedures for DNA hybridisation

1. Probe radio-labelled with standard methods.
2. Warm pre-hybridisation buffer.
3. In 2 x SSC buffer, separate membranes with nylon mesh.
4. Roll up smoothed membranes, drain, place in tube.
5. Pour in warmed pre-hybridisation buffer slowly, avoid bubbles.
6. Put into Hybridisation oven at 65 °C for 1 hour, at rotating speed 6 (ensure correct rotation).
7. Pour in all of labelled probe, and 60 ng (350 μl) of labelled marker.
8. Hybridise overnight.
9. Decant and keep probe.
10. Washes:
    20 minutes in 4 x SSC, 1% SDS at 65 °C, repeat once
    20 minutes in 2 x SSC, 0.5% SDS at 50 °C, repeat once
    10-20 minutes in 2 x SSC at 25 °C, repeat once
11. Seal membranes in bag with a few drops of 2 x SSC.
12. Expose, develop.

Appendix 5.1 Output of GENSTAT 5 programme: estimation of variance component for data of a provenance/progeny trial of var. *bahamensis* measured in 1993

```
  52  vcomp [fix=rep;abs=fam] prov/fam+prov.fam.rep
  53  reml [pr=c,d,m,s,m,vc] h

*** Estimated Components of Variance ***
                                          s.e.
prov                          0.03133    0.01407
prov.fam                      0.02296   0.004977
prov.fam.rep                  0.02787   0.004743
*units*                        0.1930   0.005970
```

```
*** Variance/Covariance matrix of Variance Components ***
          prov  0.00019808
     prov.fam -0.00000301  0.00002478
 prov.fam.rep  0.00000000 -0.00000336  0.00002250
      *units*  0.00000000 -0.00000001 -0.00000916  0.00003565

                      prov     prov.fam prov.fam.rep      *units*


*** Approximate stratum variances ***
                                                    Effective d.f.
prov                                     6.68338         12.97
prov.fam                                0.830478        106.93
prov.fam.rep                            0.301558        595.48
*units*                                 0.193030       2090.63


* Matrix of coefficients of components for each stratum *
prov            186.69      23.24       3.88        1.00
prov.fam         0.00       23.08       3.86        1.00
prov.fam.rep     0.00        0.00       3.89        1.00
*units*          0.00        0.00       0.00        1.00


Standard error of differences:      Average          0.03591
                                    Maximum          0.03600
                                    Minimum          0.03582



  52  vcomp [fix=rep;abs=fam]prov/fam+prov.fam.rep
  53  reml [pr=c,d,m,s,m,vc] d

*** Estimated Components of Variance ***
                                          s.e.
prov                             0.05645     0.02943
prov.fam                         0.09484     0.02051
prov.fam.rep                      0.1211     0.01932
*units*                           0.7557     0.02347

*** Variance/Covariance matrix of Variance Components ***
          prov   0.0008661
     prov.fam  -0.0000496   0.0004205
 prov.fam.rep   0.0000000  -0.0000567   0.0003734
      *units*   0.0000000  -0.0000002  -0.0001416   0.0005506

                      prov     prov.fam prov.fam.rep      *units*

*** Approximate stratum variances ***
                                                    Effective d.f.
prov                                     14.2271         12.89
prov.fam                                 3.39432        106.99
prov.fam.rep                             1.22694        590.71
*units*                                 0.755731       2074.40
```

```
* Matrix of coefficients of components for each stratum *
prov           191.06      23.05       3.88        1.00
prov.fam       0.00        22.89       3.85        1.00
prov.fam.rep   0.00        0.00        3.89        1.00
*units*        0.00        0.00        0.00        1.00

Standard error of differences:      Average          0.07276
                                    Maximum          0.07348
                                    Minimum          0.07227


  52  vcomp [fix=rep;abs=fam]prov/fam+prov.fam.rep
  53  reml [pr=c,d,m,s,m,vc] c

*** Estimated Components of Variance ***
                                                 s.e.
prov                              0.006015     0.003206
prov.fam                          0.008620     0.002096
prov.fam.rep                       0.01611     0.002324
*units*                            0.08452     0.002645


*** Variance/Covariance matrix of Variance Components ***
         prov 0.000010280
     prov.fam -0.000000514 0.000004393
 prov.fam.rep 0.000000000 -0.000000839 0.000005402
      *units* 0.000000000 -0.000000002 -0.000001808 0.000006996

                        prov     prov.fam prov.fam.rep      *units*


*** Approximate stratum variances ***
                                                    Effective d.f.
prov                                    1.51696          12.90
prov.fam                               0.340516         106.94
prov.fam.rep                           0.146831         583.86
*units*                               0.0845239        2042.30


* Matrix of coefficients of components for each stratum *
prov           186.86      22.61       3.86        1.00
prov.fam       0.00        22.59       3.84        1.00
prov.fam.rep   0.00        0.00        3.87        1.00
*units*        0.00        0.00        0.00        1.00


Standard error of differences:      Average          0.02539
                                    Maximum          0.02597
                                    Minimum          0.02504


  54  stop
```

APPENDIX I PUBLISHED PAPER

FOREST GENETICS 1(3):165–174, 1994

PROVENANCE VARIATION AND GENETIC PARAMETERS IN A TRIAL OF *PINUS CARIBAEA* MORELET VAR. *BAHAMENSIS* BARR. AND GOLF.

Yong-Qi Zheng[1*], Richard Ennos[2] & Huo-Ran Wang[1]

[1)] Institute of Forestry Research, Chinese Academy of Forestry, Beijing 100091, People's Republic of China
[2)] School of Forestry, Institute of Ecology and Resource Management, University of Edinburgh Edinburgh EH9 3JU, United Kingdom
* Present address: School of Forestry, Institute of Ecology and Resource Management, University of Edinburgh Edinburgh EH9 3JU, United Kingdom

Received December 2, 1994; accepted May 2, 1995

ABSTRACT

A combined provenance/progeny trial of *P. caribaea* var. *bahamensis* was established at a single site in Guangxi province, China in 1990. Fourteen provenances were used in the trial; the number of families within provenances varied from 3 to 10 and totalled 121. Three growth traits, height, diameter and crown width were investigated. Significant genetic variation was found in growth traits among regions, provenances within region and families within provenance, with the Abaco Island and New Providence provenances superior and those from Andros Island inferior. Family heritabilities for all traits are much higher than individual heritabilities. Results suggest that provenance selection would be the first step for the genetic improvement of the species and substantial genetic gain can be achieved by family selection within provenance. Genetic correlations among traits were found to be significant. Breeding strategy for the species is also discussed in the paper.

Keywords: *Pinus caribaea* var. *bahamensis*, provenance/progeny test, provenance variation, genetic parameters, breeding strategy

INTRODUCTION

Pinus caribaea Morelet is a well known fast growing species native to South and Central America. It includes three varieties: var. *caribaea,* var. *hondurensis* and var. *bahamensis* (LAMB 1973). As a widely planted exotic plantation species, it can be found throughout the tropics and subtropics over the world from 14 °N (Nigeria) to 35 °S (South Africa) in Africa, from 18 °N (Puerto Rico) to 33 °S (Brasil) in Latin America, from 1 °N (Malaysia) to 33 °N (India) in Asia, and from 1 °S (New Guinea) to 28 °S (Australia) in Oceania (LAMB 1973; LAMPRECHT 1989).

P. caribaea var. *bahamensis* Barr. & Golf. is naturally distributed mainly in four islands in the Bahamas. Interest in this variety has long been overshadowed by *P. caribaea* var. *hondurensis* which generally out-performs it in growth except in a narrow range of environments where conditions are too temperate for *P. caribaea* var. *hondurensis* and too tropical for *Pinus elliottii* (BAYLIS & BARNES 1989). Nevertheless its excellent stem form and branch form and good growth, and also high resistance to shoot tip moth (*Rhyaciona* and *Dioryctria spp*), which has devastated *P. caribaea* var. *hondurensis* in Southeast Asian countries, means that *P. caribaea* var. *bahamensis* may have great potential over a huge area. Together with *P. caribaea* var. *caribaea*, it is the next most productive softwood after *P. caribaea* var. *hondurensis* in the tropics.

Much genetic improvement work has focused on *P. caribaea* var. *hondurensis* (NIKLES 1989; KANOWSKI & NIKLES 1989a). However, breeding work on *P. caribaea* var. *bahamensis* started relatively late due to the difficulty of its seed supply. Most breeding work on *P. caribaea* var. *bahamensis* as an exotic species has been carried out since the 1980s after intensive seed collection and coordination of an international provenance trial by the Oxford Forestry Institute (OFI) (BAYLIS & BARNES 1989).

China first introduced *P. caribaea* var. *caribaea* in the 1960s, followed by *P. caribaea* var. *hondurensis* in 1973 and *P. caribaea* var. *bahamensis* in 1983 (PAN 1991). The species showed great potential for industrial purposes in the tropical area in southern China. In recent years, *P. caribaea* has been increasingly planted in Southern China on a large scale. However, it appears that there are obvious differences among the three varieties in growth rates, stem form and pest and disease resistance. Species trials showed that in addition to fast growth rates, *P. caribaea* var. *bahamensis* has excellent

stem form and more resistance to the shoot tip moth (PAN 1991).

Traditionally, forest genetic tests have been conducted sequentially with successive species, provenance and progeny trials. However, in practice there is strong economic pressure to reduce the testing interval between these stages in a traditional tree improvement programme. A common solution is to combine two or more stages, such as species/provenance tests (WANG, ZHENG & WANG 1992; WANG & ZHENG 1993; ZHENG & WANG 1993), provenance/progeny tests (KANOWSKI & NIKLES 1989b; CROCKFORD, BIRKS & BARNES 1989), or the Breeding Seedling Orchard (BSO) combining progeny tests with seed production (BARNES & MULLIN 1989).

The ultimate extension of this is an experiment that combines provenance testing, progeny testing and seed production in a single trial. This requires that design of the trial, timing of assessments, and timing of thinning in relation to seed production must all be compatible, and consistent with the breeding strategy. The multiple purpose of the trial, i.e. testing and seed production, allows genetically improved seeds to be made available within a short time, while achieving many of the aims of a long term breeding programme. The thinning of the trial can be flexible depending on the breeding strategy to be adopted. Most of the literature reports estimates of genetic parameters at family and individual levels with no recognition of provenance information. The analysis of the trial requires estimation of genetic variance component of provenance which is of great help for developing breeding strategy and for subsequent thinning of the trial (eg constructing a selection index).

The experiment reported here is a range-wide provenance trial combined with open-pollinated progeny tests of *P. caribaea* var. *bahamensis*. It was established as the initial step of the genetic improvement programme for the species in China, and is intended to form the basis for a seed orchard when appropriately thinned. The aims of the study are to investigate the amount and the pattern of the genetic variation existing within var. *bahamensis* at various levels, to estimate genetic parameters at provenance, family, within-family and individual levels, and on the basis of these estimates to determine the best selection method and develop a optimal breeding strategy for the species.

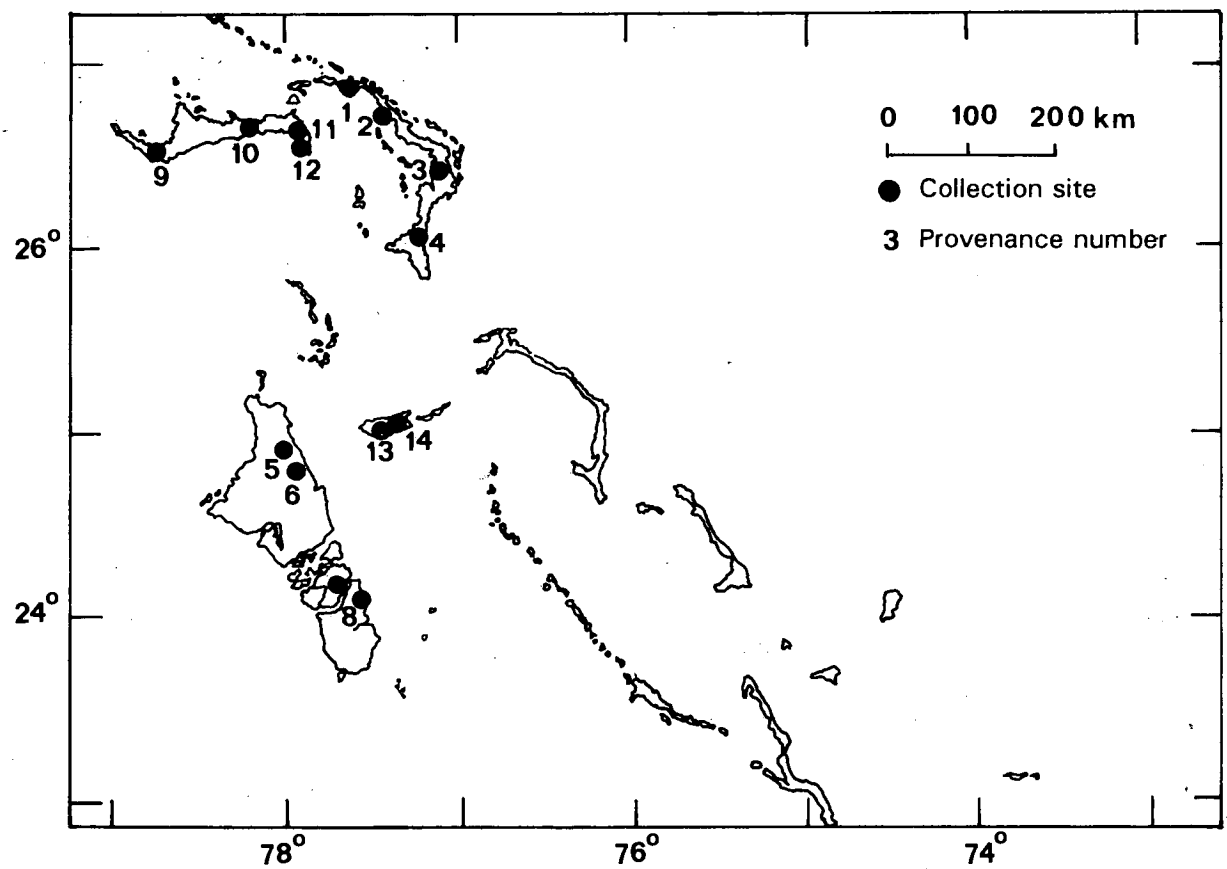

Figure 1 Location of the provenances seed collection sites of *Pinus caribaea* var. *bahamensis*

MATERIAL AND METHODS

Genetic material

Open pollinated seeds were collected from trees distributed over 14 provenances on 4 islands throughout the natural range of *P. caribaea* var. *bahamensis* (Figure 1, Table 1) (BAYLIS & BARNES 1989). Family identity was maintained. The total number of open-pollinated families grown in the trial was 121 with numbers of families per provenance ranging from 3 to 10 (Table 1).

Experimental site and design

The experiment was established at a single site at the Forestry Institute of Hepu County, Guangxi Province, located at 21°41' N, 109°11' E in the coastal lowland. The site experiences a northern tropical climate in which frost does not usually occur. Winter, from December to February, is usually dry, and most rainfall occurs in summer. July and February are the hottest and coldest months respectively. Mean annual temperature is about 22.3 °C, and mean annual rainfall is about 1,651 mm. The lateritic red-yellowish sand soil is very deep. A Randomised Complete Block (RCB) design was used with 6 replicates, 121 treatments (families) and a plot size of 4-trees. Initial spacing was 3 m by 3 m.

Establishment and data collection

Seeds were sown in August 1990, using the normal methods for raising containerised pine seedlings in South China (PAN 1991). Seedlings were planted out in May 1991. The normal management measures used in pine plantations were applied (PAN 1991). The trial was measured annually for height, stem diameter at breast height and crown width for each of the three years after establishment, at the end of each growing season. The latest measurements were carried out in December 1993, about 2.5 years after planting. The reason for measuring crown width is to facilitate selection for tree form. A narrow crown is preferred in plantation forestry for timber production.

Estimation of variance components

Since the experiment is unbalanced due to the unequal number of families within each provenance, the method of REML (Restricted Maximum Likelihood) was used to estimate variance components (KENNEDY & GENTLE 1980; HARVILLE & CALLANAN 1990). The REML procedure of GENSTAT 5 was used to fit a variance component model by restricted maximum likelihood, together with the command VCOMPONENT which is used to define the model for REML (GENSTAT 5 COMMITTEE 1990). Missing values were estimated and coefficients of variance components were adjusted for them.

A mixed model (LAND *et al.* 1987; KANOWSKI & NIKLES 1989b; WESTFALL 1992) was used for

Table 1 Details of provenances and the number of families within provenance included in provenance/progeny trial of *Pinus caribaea* var *bahamensis* in Guangxi province, China

Provenance	ID No.	Lat. (°N)	Long. (°W)	Alt. (m)	# families
Abaco Island					
Cedar Harbour (1)	658/1–10	26°53'	77°39'	10	10
Norman Castle (2)	668/1–10	26°45'	77°26'	5	10
Central Abaco (3)	678/1–10	26°26'	77°05'	5	5
Sandy Point (4)	688/1–10	26°02'	77°12'	10	10
Andros Island					
San Andros (5)	698/1–10	24°57'	78°01'	5	10
Staniard Creek (6)	708/1–10	24°50'	77°50'	5	10
Roker Cay (7)	718/1–10	24°07'	77°44'	2	8
Kemps Bay (8)	728/1–10	24°06'	77°36'	5	9
Grand Bahama Island					
Freeport (9)	738/1–10	26°32'	78°45'	5	3
South Riding (10)	748/1–10	26°40'	78°13'	10	9
Maclean's Town Cay (11)	758/1–10	26°34'	77°55'	2	7
Little Harbour cay (12)	768/1–10	26°33'	77°53'	2	10
New Providence					
Adelaide (13)	778/1–10	25°00'	77°26'	10	10
East New Providence (14)	788/1–10	25°01'	77°24'	5	10

Table 2 Expectations of mean squares for analysis of individual tree data from a balanced combined provenance/progeny trial

Source of variation	d.f	Expected mean squares
Replicate	$R-1$	$\sigma_w^2 + n\,\sigma^2_{fR} + nf\,\sigma^2_{pR} + nfP\,\sigma^2_R$
Provenance	$P-1$	$\sigma^2_w + n\,\sigma^2_{fR} + nR\,\sigma^2_f + nf\,\sigma^2_{pR} + nfR\sigma^2_P$
Provenance x replicate	$(P-1)(R-1)$	$\sigma^2_w + n\,\sigma^2_{fR} + nf\,\sigma^2_{pR}$
Family / provenance	$P(f-1)$	$\sigma^2_w + n\,\sigma^2_{fR} + nR\,\sigma^2_f$
Family/provenance x replicate	$P(f-1)(R-1)$	$\sigma^2_w + n\,\sigma^2_{fR}$
Trees within plot	$PRf(n-1)$	σ^2_w
Total	$\boldsymbol{PRfn-1}$	

σ^2_R variance component of replicate
σ^2_p variance component of provenance
σ^2_{pR} variance component of the interaction of provenance by replicate
σ^2_f variance component of family within provenance
σ^2_{fR} variance component of the interaction of family (within provenance) by replicate
σ^2_w variance component of trees within plot

For analysis of covariance, each component is replaced by the corresponding cross product

A mixed model (LAND *et al.* 1987; KANOWSKI & NIKLES 1989b; WESTFALL 1992) was used for estimation of variance components. The model used was:

$$X_{hjkl} = \mu + R_h + P_j + F_{j:k} + (PR)_{jh} + (FR)_{j:kh} + \varepsilon_{hjkl} \qquad [1]$$

where:

X_{hjkl} is the phenotypic value of the lth individual of the kth family from the jth provenance in the hth replicate;
μ is the fixed term overall mean;
R_h is the fixed effect of the hth replicate;
P_j is the effect of the jth provenance (assumed random);
$F_{j:k}$ is the effect of the kth family in the jth provenance (assumed random);
$(PR)_{jh}$ is the effect of interaction between the jth provenance and the hth replicate;
$(FR)_{j:kh}$ is the effect of the interaction between the kth family of the jth provenance and the hth replicate;
ϵ_{hjkl} is the effect of the lth tree within the kth family of the jth provenance in the hth replicate. It includes the effect of errors.
$h = 1,...,R$ (R is the number of replicates);
$j = 1,...,P$ (P is the number of provenances);
$k = 1,....,f$ (f is the number of families per provenance);
$l = 1,...,n$ (n is the number of trees per plot).

Expectations of mean squares are given in Table 2. In a carefully selected experimental site in which environmental variation is well controlled, the provenance by replicate interaction can be pooled with the plot error (WESTFALL 1992). The computational time and the demands for numerical space (GENSTAT 5 COMMITTEE 1990) are much reduced when pooling the interaction of provenance by replicate. An extra term (effect of island) was added to the model to investigate the significance of differences among regions (islands). The interaction terms were ignored to simplify the analysis which was conducted on plot means. Thus, model [1] becomes:

$$X_{hijkl} = \mu + R_h + I_i + P_{i:j} + F_{j:k} + \varepsilon_{hijkl} \qquad [2]$$

where I_i is the effect of ith region.

Estimation of genetic parameters

Definitions and calculations of genetic and phenotypic variances, heritabilities, genetic and phenotypic correlations and selection responses follow NAMKOONG (1979), LAND *et al.* (1987) and FALCONER (1989). Standard errors of heritabilities were calculated with the method given by NAMKOONG (1979).

Heritability

Family heritability was estimated as:

$$h_f^2 = \frac{\sigma_f^2}{\sigma_{Tf}^2} = \frac{\sigma_f^2}{\frac{EMS_f}{nR}} = \frac{\sigma_f^2}{\frac{\sigma_w^2}{nR} + \frac{\sigma_{fR}^2}{R} + \sigma_f^2} \qquad [3]$$

Heritability on an individual tree basis was estimated as:

$$h_f^2 = \frac{\sigma_A^2}{\sigma_T^2} = \frac{\frac{\sigma_f^2}{r}}{\sigma_w^2 + \sigma_{fR}^2 + \sigma_{pR}^2 + \sigma_f^2 + \sigma_p^2} \quad [4]$$

Where σ_{Tf}^2 and σ_T^2 are the total phenotypic variances at family and individual levels respectively, and the forms are given in Table 2.

In this paper we assumes the value of r is ⅓ (NAMKOONG *et al.* 1966; BRIDGWATER 1992), rather than the more usual ¼, to take account of localised inbreeding in the natural populations.

Phenotypic and genetic correlation

Phenotypic and genetic correlations were estimated from individual tree values, according to the following equations:

$$r_{P_{x,y}} = \frac{\sigma_{P_x P_y}}{\sqrt{\sigma_{P_x}^2 \sigma_{P_y}^2}} \quad [5]$$

$$r_{A_{x,y}} = \frac{\sigma_{A_x A_y}}{\sqrt{\sigma_{A_x}^2 \sigma_{A_y}^2}} = \frac{\sigma_{f_x f_y}}{\sqrt{\sigma_{f_x}^2 \sigma_{f_y}^2}} \quad [6]$$

where $r_{P_{x,y}}, r_{A_{x,y}}$ are phenotypic and genetic correlation coefficients, $\sigma_{P_x P_y}$, $\sigma_{A_x A_y}$ are the phenotypic and additive genetic cross products of traits x and y, $\sigma_{P_y}^2$, $\sigma_{A_x}^2$, and $\sigma_{P_x}^2$, $\sigma_{A_x}^2$ are phenotypic and additive genetic variances of traits x and y respectively. The genetic crossproducts are estimated from ANCOVA (Table 2).

Selection response

Responses to different selection methods are estimated by using the formula:

$$R = i\sigma h^2 \quad [7]$$

where i is the standardised selection intensity, σ is the standard deviation of phenotypic variance, and h^2 is the heritability of the selection method.

Age to age correlation

Measurements of a trait at different ages were considered as separate traits. Genetic correlations between corresponding traits at two ages were calculated from family means with the following equation.

$$r_{A_{1,2}} = \frac{\sigma_{A_1 A_2}}{\sqrt{\sigma_{A_1}^2 \sigma_{A_2}^2}} = \frac{\sigma_{f_1 f_2}}{\sqrt{\sigma_{f_1}^2 \sigma_{f_2}^2}} \quad [8]$$

RESULTS AND ANALYSIS

Genetic variation

Differences between regions, between provenances within regions, and between families within provenances were significant for all traits (Table 3). Provenance mean values are presented in Table 4. Family mean values are not presented due to lack of space. The growth performance of northern provenances is generally better than of the southern ones. Provenances of the fastest-growing region (Abaco Island) are around 11% larger in both height and diameter than are those of the slowest-growing (Andros Island); the difference in average crown width is around 9%. The range of variation between provenances is greater than between regions (20%, 17% and 22% for height, diameter and crown width, respectively). Differences between the best and worst families are
50%, 73% and 39% for height, diameter and crown width respectively.

Genetic parameters

Variance components and heritabilities

Variance components calculated by REML and their relative importance (FALCONER 1989) are given in Table 5. These were used to calculate narrow sense heritabilities at family and individual levels (Table 6).

The variance component attributable to individual has the highest value, followed by those attributable to within-family, family and provenance, except for height where the provenance variance component is larger than the family component. Family heritability is much higher than individual heritability in all cases (Table 6).

Phenotypic and genetic correlations

All phenotypic correlations are statistically significant. The greatest correlation coefficient is found between height and diameter and the smallest is between DBH and crown width (Table 7).

All genetic correlation coefficients between pairs of traits are larger than corresponding phenotypic correlations, though the values rank in the same order. All genetic correlation coefficients are statistically significant at the $P < 0.01$ level. Height is most strongly

Table 3 Mean squares for height, DBH and crown width at 2.5 year after planting for a provenance/progeny trial of *Pinus caribaea* var. *bahamensis* in Guangxi, China

Source of variation	d.f.	M.S.		
		Height	DBH	Crown width
Replicate	5	3.58303	14.0541	3.60116
Region	3	5.74566**	11.7073**	0.98036*
Region/provenance	10	0.75234**	1.5149*	0.22884*
Region/provenance x family	107	0.21823**	0.9212**	0.09548**
Residual	600	0.07813	0.3266	0.04476
Total	**725**			

**: $P < 0.01$; *: $P < 0.05$

Table 4 Growth data (mean ± standard error) for each provenance at 2.5 year after planting in provenance/progeny trial of Pinus caribaea var. bahamensis in Guangxi, China (numbers in parentheses are ranks)

Region	Provenance	Height (m)	DBH (cm)	Crown width (m)
Abaco Island	Cedar Harbour	3.48±0.35 (4)	5.47±0.67 (4)	2.25±0.31 (1)
	Norman Castle	3.55±0.32 (3)	5.57±0.63 (1)	2.14±0.29 (2)
	Central Abaco	3.41±0.41 (6)	5.35±0.77 (6)	2.11±0.25 (3)
	Sandy Point	3.57±0.38 (2)	5.56±0.77 (2)	2.05±0.26 (4)
	Mean	***3.52±0.36***	***5.51±0.69***	***2.08±0.28***
Andros Island	San Andros	3.34±0.32 (9)	5.29±0.68 (8)	2.00±0.27 (8)
	Staniard Creek	3.20±0.38 (12)	5.00±0.73 (12)	1.91±0.27 (11)
	Roker Cay	3.01±0.41 (13)	4.83±0.92 (13)	1.85±0.27 (13)
	Kemps Bay	3.98±0.38 (14)	4.75±0.88 (14)	1.88±0.27 (12)
	Mean	***3.14±0.40***	***4.98±0.82***	***1.91±0.27***
Grand Bahama	Freeport	3.26±0.38 (11)	5.28±0.66 (9)	1.83±0.30 (14)
	South Riding	3.30±0.34 (10)	5.18±0.68 (11)	1.93±0.32 (10)
	Maclean's Town Cay	3.36±0.33 (8)	5.32±0.63 (7)	2.05±0.29 (5)
	Little Harbour Cay	3.43±0.29 (5)	5.26±0.58 (10)	2.05±0.26 (6)
	Mean	***3.36±0.33***	***5.34±0.65***	***1.99±0.29***
New Providence	Adelaide	3.59±0.32 (1)	5.53±0.68 (3)	1.99±0.30 (9)
	East New Providence	3.37±0.35 (7)	5.39±0.72 (5)	2.02±0.25 (7)
	Mean	***3.48±0.35***	***5.46±0.70***	***2.00±0.27***
	Grand Mean	**3.36±0.40**	**5.30±0.76**	**1.99±0.29**

correlated with DBH and then crown width. DBH and crown width are most weakly correlated (Table 7).

Selection response

When comparing the relative responses to family and individual selections, selection intensity is assumed to be constant. The selection response is therefore determined by the magnitude of the heritability and the phenotypic variance at each level. The relative responses can be obtained by substituting the corresponding standard deviation and heritability into equation [7]. Suppose ⅓ of the total population of the trial is to be selected, i.e. standardised selection intensity $i = 1.091$ (COTTERILL & DEAN 1990). The genetic gains from different selection methods are shown in table 6. Family selection will yield over four times the genetic gain anticipated under individual selection (mass selection)

Table 5 Estimated variance components (± standard error) and the relative contribution of provenance, family and within family to each trait at 2.5 year after planting

Trait	Height		DBH		Crown width	
Source of variation	Variance component	Relative (%)	Variance component	Relative (%)	Variance component	Relative (%)
Provenance	0.03133±0.01	18.5	0.0565±0.03	9.0	0.00602±0.003	10.4
Family	0.02296±0.01	13.6	0.0948±0.02	15.2	0.00862±0.002	14.9
Individual	0.06888±0.01	40.7	0.2844±0.02	45.5	0.02586±0.002	44.8
Within family	0.04592±0.01	27.2	0.1896±0.02	30.3	0.01724±0.002	29.9

Table 6 Estimated heritabilities (± standard error) and anticipated selection responses (genetic gains) under family selection and individual selection supposing ⅓ of the population is selected (selectionintensity = 1.091) for three traits at 2.5 year after planting in a provenance/progeny trial of *Pinus cariabea* var. *bahamensis* in Guangxi, China

Trait	Height		DBH		Crown width	
	Heritability	Genetic gain	Heritability	Genetic gain	Heritability	Genetic gain
Family	0.64±0.11	18.94%	0.64±0.05	24.23 %	0.57±0.06	18.14 %
Individual	0.25±0.10	4.46 %	0.28±0.06	6.37 %	0.22±0.05	4.60 %

intensity i = 1.091 (COTTERILL & DEAN 1990). The genetic gains from different selection methods are shown in table 6. Family selection will yield over four times the genetic gain anticipated under individual selection (mass selection) within provenance. Genetic gains in diameter are larger than those for height and crown width.

Age to age stability

Genetic correlations between traits measured at different ages are given in table 8. Most correlations are statistically significant. Correlations involving height are generally higher than those involving diameter and crown width. Correlations between character values at planting age and other ages tend to be the weakest.

Table 7 Phenotypic (above diagonal) and genetic (below diagonal) correlation coefficients among traits for a provenance/progeny trial of *Pinus caribaea* var. *bahamensis* in Guangxi, China

	Height	DBH	Crown width
Height	–	0.8175**	0.5479**
DBH	0.8580**	–	0.4808**
Crown width	0.6281**	0.6083**	–

Table 8 Age to age genetic correlations for DBH, height and crown width in a provenance/progeny trial of *Pinus cariabaea* var. *bahamensis* in Gunagxi, China

	Dec–93	Dec–92	Nov–91	Apr–91	Dec–92	Nov–91
	DBH				Crown width	
Dec–93	1	0.4358**	0.4649**	0.0302	0.4287**	0.4715**
Dec–92	0.5225**	1	0.5086**	0.0094		0.3718**
Nov–91	0.5529**	0.3868**	1	0.3282**		
Apr–1	0.2171*	0.1593	0.7459**	1		
	Height					

**: $P < 0.01$; * $P < 0.05$

Values of the correlation coefficients generally tend to increase with age.

DISCUSSION

Genetic variation and implication for provenance selection

Success in the establishment and productivity of forest tree plantation is determined largely by the species used and the source of seed within species (ZOBEL & TALBERT 1984). The results of the trial show that for *P. caribaea* var. *bahamensis* significant genetic variation exists both among the regions (islands) and among the provenances within regions. The growth rates at the experimental site are probably related to adaptation of provenances to rainfall in their natural distribution. Superior performance is shown by provenance coming from northern areas of the natural range where rainfall is greater.

The most practical way of breeding *P. caribaea* var. *bahamensis* is to start at the highest possible level. Selection of provenances from the more productive islands will give the most genetic gain for the least expense. The immediate application of the results is to guide the seed importing for the national afforestation programme in which *P. caribaea* has been taken as a major fast-growing species for industrial purposes. Clearly the Abaco Island and New Providence Islands are the desirable sources for establishment at the test site in China. However, it should be noted that the prospective area for planting *P. caribaea* var. *bahamensis* stretches from Hainan Island to south-eastern Fujian province (PAN 1991). There is large environmental variation within this region and, therefore, testing across a wider range of environmental conditions is necessary to establish the generality of this result.

Experimental design and analysis for combined tests

The value of the data derived from provenance/progeny trials for tree improvement purpose is closely related to the experimental design used. Three features of the current experiment need to be borne in mind when interpreting the results that have been obtained; the experiment is a range-wide test of variation, it has been conducted at only one site, and the experiment is unbalanced with unequal numbers of families per provenance.

The consequence of using a range-wide test is that poorly adapted provenances are likely to be included in the test. This will decrease the test mean while inflating the estimated genetic variance and anticipated genetic gains in comparison with a trial that includes only a narrow range of adapted provenances. Range-wide tests may therefore not be as suitable as limited-range tests for selection and breeding, particularly for species with wide distribution (LAND *et al.* 1987). If results from limited range tests are not available, it may be most appropriate to estimate genetic parameters from a subset of the data, including only provenances selected for use in the appropriate breeding zone.

A second feature of the present experiment is that it has been conducted at only one site. If genotype × environment interaction (G × E) occurs across the range of sites to be planted in China, the single trial will overestimate genetic gains anticipated from selection. G × E interaction has been documented in the analysis of previous trials of Carribean pine (GIBSON 1982) although its importance for *P. caribaea* var. *bahamensis* is not yet known.

The final feature of this experiment is its unbalanced design. The imbalance caused by use of unequal sized samples occurs commonly in genetic studies (SNEDECOR & COCHRAN 1967). It is often believed that such unbalanced design will cause difficulties in statistical analysis and that the best way is to avoid it (KANOWSKI & NIKLES 1989; LAND *et al.* 1987). Practically, however, such imbalance may be unavoidable in forestry trials. With the rapid development of both computer hardware and software for the analysis of variance components by the techniques such as REML, this problem can be overcome relatively easily.

Age to age stability and implication for early selection

Although it seems that family performance is stabilising as the trees age, it is too early to draw meaningful conclusions as to the optimum timing of selection. Selection in tropical pines is commonly made around 6 years (KANOWSKI & NIKLES 1989a), although longer intervals of not less than half rotation age have been advocated (ZOBEL & TALBERT 1984). Further investigation of age-age correlations for *P. caribaea* var. *bahamensis* in southern China is necessary to define the optimum selection age.

Genetic parameters and implication for breeding strategy

The substantial variation which exists between regions and between provenances within regions suggests that rapid improvement can be realised by selection of trees within the best families of the superior provenances of the Abaco and New Providence Islands. However, breeding strategy needs to accommodate maintenance

of genetic diversity (ERIKSSON *et al.* 1993); given the limited testing thus far of *P. caribaea* var. *bahamensis* in southern China, this objective will be particularly important. The dual objectives of improvement and conservation can best be realised within the MPBS (BARNES 1984; ERIKSSON *et al.*, 1993). The Multiple Population Breeding System (MPBS) is implemented by considering each trial as a separate sub-population and managing it as a BSO (BARNES & MULLIN 1989). Seed is collected initially from a large number of individuals representing the genetic base of the trial (eg by using only within family selection). The trial is subsequently thinned on the basis of provenance, family within provenance, and individual information to deliver improved seed for operational use. Establishing structures such as a nucleus population (COTTERILL *et al.* 1989) within MPBS offers one means of realising both short and longer term gains.

Since the trial is still too young to produce seeds, it would be wise to develop a separate clonal seed orchard (CSO) of the best individuals in the best families of the best provenances of Abaco and New Providence for early seed production for sites typical of the test site. Establishing a CSO will meet short-term production needs and retain flexibility. It is necessary to establish provenance/progeny tests on contrasting sites throughout the planting region while monitoring the development of this trial to determine its best use on the basis of longer term information from this site and information from other sites.

ACKNOWLEDGEMENTS

Thanks are given to the Oxford Forestry Institute for providing the seeds for the experiment. This research is co-funded by the Overseas Development Administration of British government and the Chinese Academy of Forestry. The first author thanks Dr. P. Kanowski at the OFI for his advice on analysis of the experiment and very valuable comments on the manuscript. An anonymous reviewer is acknowledged for comments.

REFERENCES

BARNES, R.D., 1984: A multiple population breeding strategy for Zimbabwe. *In:* Provenance and Genetic Improvement Strategies in Tropical Forest Trees. (Eds. Barnes, R.D. & Gibson, G.L.). pp. 619–632. Commonwealth Forestry Institute, Oxford and Forest Research Centre, Harare.

BARNES, R.D. & MULLIN, L.J., 1989: The multiple population breeding strategy in Zimbabwe – five year results. *In:* Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin, A.R. & Matheson, A.C.). pp. 148–158. Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA.

BAYLIS, W.H.B. & BARNES, R.D., 1989: International provenance trial of *Pinus caribaea* var. *bahamensis*. *In:* Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin, A.R. & Matheson, A.C.). pp. 363–365. Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA.

BRIDGWATER, F., 1992: Mating design. *In:* Handbook of Quantitative Forest Genetics (eds. Fins, L., Friedman, S.T. & Brotschol, J.V.). pp. 69–95. Kluwer Academic Publishers, Dordrecht, Boston and London.

COTTERILL, P. & DEAN ,C.A., 1990: Successful Tree Breeding with Index Selection. CSIRO, Australia. 80 pp.

COTTERIL, P.P., DEAN, C.A., CAMERON, J. & BRINDBERGS, M., 1989: Mucleus breeding: A new strategy for rapid improvement under clonal forestry. *In:* Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin, A.R. & Matheson, A.C.). pp.39-52. Oxford Forestry Institute, Oxford, UK and Winrock International, Arlington Virginia, USA, 1989.

CROCKFORD, K.J., BIRKS, J.S. & BARNES, R.D., 1989: Family within provenance trials of *Pinus caribaea, P. oocarpa* and *P. patula* ssp. *tecunumanii* – early results and implications for breeding strategies. *In:* Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin, A.R. & Matheson, A.C.). pp. 175–186. Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA.

ERIKSSON, G., NAMKOONG, G. & ROBERDS, J.H., 1993: Dynamic gene conservation for uncertain futures. *Forest Ecology and Management* **62**: 15–37.

FALCONERD.S., 1989: Introduction to quantitative genetics. 3rd ed. Longman Scientific & Technical and John Wiley & Sons, Inc., New York, 438 pp.

GENSTAT 5 COMMITTEE, 1990: Genstat 5 Reference Manual. Clarendon Press, Oxford, 749 pp.

GIBSON, G.L., 1982: Genotype-environment interaction in *Pinus caribaea*. Unit of Tropical Silviculture, Interim Report, Commonwealth Forestry Institute, Oxford, 112 pp.

GREAVES, A., 1978: Description of seed sources and collections for provenances of *Pinus caribaea*. CFI, Tropical Forestry Paper No. 12, 98 pp.

HARVILLE, D.A. & CALLANAN T.P., 1990: Computational aspects of likelihood-based inference for variance components. *In*: Advances in Statistical Methods for Genetic Improvement of Livestock (eds. D. Gianola & K. Hammond). Springer-Verlag. Berlin and London, 534 pp.

KANOWSKI, P.J. & NIKLES, D.G., 1989A: A summary of plans for continuing genetic improvement of *Pinus caribaea* var. *hondurensis* in Queensland. *In:* Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin, A.R. & Matheson, A.C.). pp. 68–78.

Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

KANOWSKI, P.J. & NIKLES, D.G., 1989B: Family structure in provenance trials, conservation stands and genetic base populations. *In:* Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin, A.R. & Matheson, A.C.). pp. 236–249. Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

KENNEDY, W.J. JR. & GENTLE, J.E., 1980: Statistical Computing. Marcel Dekker Inc., New York and Basel, 591 pp.

LAMB, A.F.A., 1973: *Pinus caribaea.* Vol. 1, Fast Growing Timber Trees of the Lowland Tropics (Now "Tropical Forestry Papers". No. 6. University of Oxford, 254 pp.

LAMPRECHT, H., 1989: Silviculture in the tropics: Tropical forest ecosystems and their tree species – possibilities and methods for their long term utilisation. Dt. Ges. für Tech. Zusammenarbeit (GTZ) GmbH, Eschborn. [Transl. by: John Brose et al.] – Rossdorf: TZ-Verl.-Ges. 1990].

LAND, S.B. JR., BONGARTEN, B.C. & TOLIVER, J.R., 1987: Genetic parameters and selection indices from provenance/progeny tests. *In:* Statistical considerations in genetic testing of forest trees. pp. 59–74. Southern Cooperative Series Bulletin No. 324.

NAMKOONG, G., 1979: Introduction to Quantitative Genetics in Forestry. Forest Service, United States Department of Agriculture, Washington D.C., 342 pp.

NAMKOONG, G., SNYDER , E.B. & STONECYPHER, R.W., 1966: Heritability and gain concepts for evaluating breeding system such as seedling orchards. *Silvae Genet.* **15**:76–84.

NIKLES, D.G., 1989: Early growth and potential value of *P. caribaea* var. *hondurensis* x *P. tecunumanii* and x *P. oocarpa* F_1 hybrids in Australia. *In:* Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. (eds. Gibson, G.L., Griffin, A.R. & Matheson, A.C.). pp. 412–414. Oxford Forestry Institute, Oxford, United Kingdom and Winrock International, Arlington, Virginia, USA, 1989.

PAN, Z.–G., 1991: Growing exotic trees in China. China Forestry Press, Beijing. 194 pp. [in Chinese]

SNEDECOR G.W. & COCHRAN, W.G., 1967: Statistical Methods. 6th edition. Iowa State University Press. Ames, Iowa, USA, 593 pp.

WANG, H.–R., ZHENG, Y.–Q. & WANG, W.–H., 1992: Growing Performance of *Casuarina junghuniana* and other 6 species of *Casuarinaceae* in Southeast China. *Scientia Silvae Sinicae* **28**(4): 343–347. (In Chinese).

WANG, H.–R. & ZHENG, Y.–Q, 1993: Species and Provenance Selection of Eucalyptus in Subtropical China. Final report of Sino-Australian co-operative research project. Chinese Academy of Forestry, Beijing, 60 pp. [In Chinese].

WESTFALL, R.D., 1992: Developing Seed Transfer Zones. *In:* Handbook of Quantitative Forest Genetics (eds. Fins, L., Friedman, S.T. & Brotschol, J.V.). pp. 140–194. Kluwer Academic Publishers, Dordrecht, Boston and London.

WHITE, T.L., 1992: Advanced generation breeding populations: size and structure. *In:* Breeding tropical trees. Proc. IUFRO S.02.08, Cartagena and Cali, Colombia. 9–12 October 1992, 208–222.

ZHENG, Y.–Q. & WANG, H.–R., 1993: Species/provenance trial of *Eucalyptus smithii* R. Baker and other 23 temperate eucalypt species in China. *In:* Australian Broad-leaved Tree Species Research in China. (ed. A.G. Brown). ACIAR Proceedings **48**:116–122.

ZOBEL, B. & TALBERT, J., 1984: Applied forest tree improvement. John Wiley & Sons, New York, Chichester, Brisbane, Toronto and Singapore, 505 pp.