Geographical variation and population biology in wild *Theobroma cacao*

John Benedict Allen

Doctor of Philosophy
University of Edinburgh
1988

GEOGRAPHICAL VARIATION AND POPULATION BIOLOGY IN WILD *Theobroma cacao*

ABSTRACT

Geographical variation in wild Upper Amazon populations of cocoa (*Theobroma cacao*) was studied with a view to improving germplasm collecting strategies for cocoa. Morphological descriptors, which covered pod, seed, flower, flush and tree habit characters, were assessed in 419 accessions from the Amazon Region of Ecuador, 136 accessions from the Río Caquetá in Colombia, and about 50 clones of wild origin from Peru. The Ecuadorian and Colombian material was collected in the course of the London Cocoa Trade Amazon Project (LCTAP), and data were acquired both at the time of collection and in the field genebank at San Carlos, Napo Province, Ecuador. The Peruvian material was examined in the genebank at Pichilingue, Ecuador. There were clear differences between the Ecuador, Caquetá and Peru accessions, but there was relatively little variation within Ecuador, so that the initial sampling strategy used to survey variation in Ecuador effectively captured most of the variation in that target region. Using cost data for LCTAP, it is argued that marginal costs for additional accessions are higher than is generally assumed, so that it is not cost-effective to acquire many similar accessions.

Field observations from LCTAP collecting trips as well as data from the San Carlos field genebank and from a study population of wild cocoa at San Carlos were used to build a picture of the population biology of *Theobroma cacao* in its natural habitat. Cocoa was consistently associated with certain soil and terrain types in Ecuador. Observed rates of seed and vegetative reproduction were low, but adult mortality was also low; cocoa could be characterized as a shade-bearing, K-adapted species, although short-term increases in seed production and growth rates resulted from shade reduction. Seed dispersal by mammals resulted in most seedlings germinating in groups close to parent trees, with high early mortality resulting from mammalian herbivory; after three months, mechanical damage was the most frequent cause of seedling mortality. Approximate tree ages were estimated from the relationship between basal area and growth rate in the study population over a two year period. The age distribution showed a peak in the age class 60-80 years, with fewer older or younger trees. Possible causes of disturbance in the past were reviewed.

Observations made in the course of this project, combined with published information on Upper Amazon ethnobotany, support the view that *Theobroma cacao* is a genuinely wild species whose present-day distribution in the Upper Amazon has been only marginally affected by human activities. It is argued that the Vavilovian centre of diversity-centre of origin model is not appropriate for cocoa, and that the Upper Amazon cocoa gene pool is valuable to breeders because it is distinct from the gene pool of cultivated cocoa.

I declare that this thesis has been composed by myself, and that the work presented in this thesis is my own. I acknowledge assistance with field work provided by INIAP staff working on LCTAP. Long-term production and flowering data referred to in Chapter 6 were collected by INIAP staff working under my supervision. No part of this thesis has been submitted to any other University.

Contents

Appendices

List of Figures

List of Tables

Glossary

CCCA	Cocoa, Chocolate and Confectionery Alliance (UK) (now BCCCA: Biscuit, Cake, Chocolate and Confectionery Alliance)
CATIE	Centro Agronómico Tropical de Investigacíon y Enseñanza (Turrialba, Costa Rica)
CEPLAC	Comissao Executiva do Plano de Lavoura Cacauaeira (Brazil)
CRU	Cocoa Research Unit, University of the West Indies (Trinidad)
EBC	Expedición Botánica Caquetá (code used for accessions from a collecting expedition on the Río Caquetá, Colombia, May-June 1984)
IBPGR	International Board for Plant Genetic Resources
ICA	Instituto Colombiano Agropecuario (Colombia)
INIAP	Instituto Nacional de Investigaciones Agropecuarias (Ecuador)
LCTAP	London Cocoa Trade Amazon Project
LCT-EEN	London Cocoa Trade - Estación Experimental Napo (code used for LCTAP accessions)
LCTMA	London Cocoa Terminal Market Association (UK) (now London Cocoa Terminal Market)
Pichilingue	INIAP (see above), Estación Experimental Tropical Pichilingue, Quevedo, Ecuador
San Carlos	INIAP (see above), Estación Experimental Napo-Payamino, Napo Province, Ecuador (the station has

	two sites, San Carlos, where the cocoa genebank is located, and Payamino, near Coca).
Turrialba	see CATIE
USDA	United States Department of Agriculture

Acknowledgements

I am greatly indebted to the many people who have have contributed to the London Cocoa Trade Amazon Project (LCTAP) and to the writing of this thesis, both in Ecuador and in the UK. Special thanks are due to the people and organizations listed here; my apologies are due to those who have not been mentioned.

Most fundamentally, LCTAP was conceived and carried to a successful conclusion through the expertise, imagination and enthusiasm of R.A.Lass, who acted as technical coordinator for the project, of E.T.Beauchamp, who was responsible for project administration, and of the London Cocoa Terminal Market, who funded the project on behalf of the London Cocoa Trade; the opportunity of working on this project is one for which I will always be extremely grateful. Although I would like this thesis to be considered as an additional personal contribution to LCTAP, I must emphasize that the writing of the thesis has not been funded by LCTAP, and responsibility for the contents of the thesis is mine alone.

In Ecuador, the project depended on the goodwill and facilities of the Instituto Nacional de Investigaciones Agropecuarias (INIAP), with the support of successive Director-Generals. Within INIAP, Dra Carmen Suarez and Ing Jaime Vera were closely involved with the planning of the project and ensured that needed resources were made available throughout a period of serious financial stringency for INIAP.

At the Napo-Payamino Experimental Station, and on collecting trips, I benefitted from the companionship and skilled technical assistance of Agr José Baquero. Two INIAP scientists, Ing Juvenal Cabrera and Ing Roger Mera, participated in project activities from 1984 onwards and are now responsible for management of the San Carlos genebank. Ing Victor Hugo Chalá, the Director of the Napo-Payamino Station from 1983 onwards, made my life easier by his enthusiastic support for the project and the smooth running of the station under his direction.

My participation in the collecting expedition on the Río Caquetá in Colombia was made possible by an invitation from the Corporación de Araracuara, with funding from the government of Colombia.

Many individuals helped by offering advice, practical assistance or timely criticism. I would like to record in particular the assistance of P.W.Allen, B.G.D.Bartley, G.Blaha, W.S.Chalmers, L.Cox, J.M.M.Engels, R.A.Ennos, H.C.Evans, D.Glendinning, C.Gonsalves, J.Grace, P.H.Gregory, W.Henderson, E.Imle, M.Jackson, J.-L. Jacob, A.J.Kennedy, A.Linton, G.Lockwood, L.Lopez, E.Mogrovejo, P.M.Mumford, E.Ng, A.F.Posnette, E.Potts, F.Rojas, E.A.Rosenquist, S.A.Rudgard, R.E.Schultes, N.W.Simmonds, W.Soegeng-Reksodihardjo, T.N.Sreenivasan, H.Toxopeus, R.L.Wastie, B.E.J.Wheeler, L.A.Withers and M.Wright. The use of cluster analysis to investigate geographical variation was suggested by D.J.Hand.

At Pichilingue, the ODA team provided me with five-star hospitality and contact with real science: Mo Hadfield, Bill Hadfield, John Hedger, Merylyn Hedger, Alan Maddison, John Maddison, Christine Orchard and John Orchard. At Finca Maria Zoyla, Denis and Marilda Hartley provided a refuge among the oil-palms. Nickie Irvine invited me to visit her field site at San José de Payamino, and stimulated me to look at Amazonian ecology from the perspective of its native inhabitants. Henning Reetz and Hilda Simon provided conversation and hospitality in Coca. Mary-Ellen Fieweger provided accomodation in Quito for my visit in 1986, and an inspiring flow of ideas on Ecuadorian history and culture. Away from the jungle, Diana Lewis encouraged me to see the rest of Ecuador.

In Edinburgh, my supervisors, D.C.Malcolm and M.G.H.Cannell, provided advice and support at critical stages in the planning and writing of this thesis; I am greatly indebted to them, and to Dr W.E.S.Mutch, Head of Department in the Department of Forestry and Natural Resources in 1983, for making it possible for me to undertake the thesis.

Finally, I would like to acknowledge the role of Catherine Barbour Allen; without her, I would probably never have left the Amazon.

Summary

(a) The overall objectives of this thesis were

to show how the pattern of geographical variation in wild cocoa might determine germplasm collecting strategies, and

to investigate the population biology of cocoa in its natural habitat,

using data acquired in the course of a major cocoa germplasm collecting project, the London Cocoa Trade Amazon Project (LCTAP);

(b) The planning of germplasm collection involves decisions about the choice of target region and about the sampling strategy to be used within the target region;

(c) Target regions have typically been selected using basic data on variation throughout the geographical range of cocoa, although practical and political considerations have often been of overriding importance. Most attention has been directed at the Upper Amazon region, which appears to be a *centre of diversity* for cocoa. I emphasize that this should not automatically be taken to mean that the Upper Amazon is also the *centre of origin* for cocoa;

(d) Within a target region, one widely accepted criterion for a sampling strategy is that collecting should capture at least 95% of alleles having frequency >.05. This presupposes that data are available to enable the delineation of populations within the target region;

(e) Cost-effective collecting also requires information on the allocation of variation within and between localities within the target region, and on the 'cost ratios', which are based on the relative costs of travelling to additional localities or collecting more samples at the same localities;

(f) Wild cocoa populations in the Amazon Region of Ecuador were systematically collected and characterized using a set of morphological descriptors. Partial data were also obtained for some Peruvian material, collected by Pound, and held in the Pichilingue genebank in Ecuador. The descriptor data were analyzed as frequency distributions for six areas, and by mapping the geographical distribution of groups of accessions generated by cluster analysis;

(g) Within Ecuador, variation was surprisingly limited, and each of the groups of accessions identified via cluster analysis was distributed throughout the region. In sampling strategy terms, it would be reasonable to treat the Amazon Region of Ecuador as a single population, occupying an area some 400 km (E-W) by 500 km (N-S). On this basis, the number of accessions successfully established at the San Carlos genebank (about 280) would mean that the probability of failing to capture four alleles with frequency .05 at a random locus would be less than 10^{-11}, far exceeding the criterion stated in (d) above;

(h) On the Río Caquetá in Colombia, in an area about 300 km (E-W) by 100 km (N-S), there were two distinct pod types with only limited geographical separation. The 34 accessions at San Carlos correspond to a probability of only 88.3% of capturing four alleles having frequency .05 at a random locus;

(i) The available Peruvian material was a small and unrepresentative sample of the original populations. The Pound accessions were most clearly distinguished from the Ecuadorian population on the seed, flush and flower pigmentation descriptors;

(j) Relationships between geographical accession groups were further explored by principal components analysis, which provided a concise summary of the similarities and differences noted above, and which generated a series of 'composite descriptors';

(k) The composite descriptors were used to estimate within- and between-area components of variation for two levels of geographic subdivision. The larger geographic subdivisions (Level 1 areas) covered areas of 30,000 km^2 or more, each corresponding to one or two entire provinces in Ecuador. The smaller geographic subdivisions (Level 2 areas) covered 10,000-30,000 km^2, and were based on groups of collecting localities within the Level 1 areas. It was concluded that Level 1 areas contributed about the same amount of variation as Level 2 areas, while Level 2 areas contributed about five times as much variation per tree as individual trees within Level 2 areas;

(l) The corresponding cost ratios were estimated from information on LCTAP expenditure. Cost ratios were of the order of 1, for sampling Level 2 areas within Level 1 areas, and 5, for trees within Level 2 areas; the latter ratio was reduced to about 2 if post-collection costs were included, and would be reduced still further by including quarantine and evaluation costs. It was concluded that relatively small sample sizes would be most cost-effective;

(m) The population biology of wild cocoa was investigated using observations on a study population of about 50 trees at San Carlos, as well as data from trees in the San Carlos genebank and data from collecting trips. In general, cocoa was a widespread component of the understory of undisturbed forest, with typical densities of around 5 trees/ha. It was not found in secondary forest;

(n) Observed rates of reproduction and mortality were low; it is suggested that cocoa is a very long-lived species which can maintain population numbers with minimal recruitment. It could be characterized as a shade-bearing, K-adapted species, although short-term seed production and growth rates were increased by shade removal;

(o) Seed dispersal by mammals resulted in most seeds germinating close to the parent tree and the resulting seedling clusters suffered high mortality within the first three months; subsequent seedling mortality was caused mainly by mechanical damage;

(p) Population age structures were deduced from short-term growth rate data from trees of varying sizes; problems with this method were discussed. The age distribution both for the study population and for trees encountered on collecting trips showed a peak in the age class 60-80 years, with fewer older or younger trees. Evidence relating to causes of disturbance which might have affected these populations was reviewed;

(q) The widespread occurrence of cocoa in the Upper Amazon forests, the evidence relating to its reproduction under forest conditions, and the fact that the indigenous inhabitants of the region do not treat cocoa as a crop species, make it very likely that *T. cacao* is a genuinely wild species, whose present-day distribution in the Upper Amazon has been only marginally affected by human activities;

(r) It appears most likely that cocoa is distributed as a wild species in both South and Central America. It is known that cocoa was first cultivated in Central America, but it does not appear that cocoa has evolved any characteristics related to cultivation by which it might be termed a domesticated species. As the practice of cultivation spread through the range of the wild species, it appears that local populations have been added to the set of cultivated types. On this model, the Upper Amazon cannot be described as the centre of origin, and, indeed, the concept of such a centre is not appropriate for cocoa. Rather, the potential value of Upper Amazon cocoa populations to cocoa breeders derives from the fact that, with the exception of the Pound collections, they comprise a wild gene pool not previously incorporated into cultivated cocoa populations. Adaptations for forest conditions, which appear to include very low rates of reproduction and mortality, and slow vegetative growth in low light levels, may explain why wild genotypes sometimes perform badly under plantation conditions.

1 Introduction

1.1 Background and objectives

Cocoa is a common neotropical rain-forest tree which has become an important crop plant. In the wild, cocoa (*Theobroma cacao* L.), together with other members of the genus *Theobroma*, occurs frequently in forests from Mexico south to Brazil and Bolivia, including the whole Amazon basin. Cocoa is cultivated as an export commodity throughout the tropics, the main centres of production being in South America, West Africa and Malaysia. The total value of world cocoa production in 1985 was about £2.25 x 10^9 (Wood and Lass 1986); it is a major source of foreign exchange for at least ten countries, and in most producing countries it is one of the principal sources of income for small farmers. Most of the voluminous literature on cocoa deals with its biology and management in cultivation, while there are relatively few references to cocoa as a wild species. Nonetheless, the wide distribution and frequent occurrence of cocoa would seem to justify study of this species as a significant component of the neotropical rain forest, the more so if such study can also contribute to more productive cocoa cultivation.

This thesis will emphasize one practical application of the study of cocoa as a wild species: the devising of more efficient strategies for the collection of wild cocoa germplasm for use in breeding programmes. Several such collections have been made in the past fifty years, including a major collecting programme in Ecuador from 1980 to 1985, for which I was responsible, and which provided the stimulus to undertake this thesis. This collecting programme, the London Cocoa Trade Amazon Project (LCTAP), was based at San Carlos, Napo Province, Ecuador, on part of the INIAP Napo-Payamino Experimental Station (Allen and Lass 1983; Allen 1988). As there is a continuing need for new wild germplasm in cocoa breeding programmes, and as no systematic collecting has taken place through much of the range of wild cocoa, it is likely that further collecting projects will be undertaken in the next few years. The results to be presented in this thesis should help to ensure that the resources available for such collecting are used to advantage.

Any study of cocoa as a wild species must address two fundamental questions. First, does cocoa have a 'centre of origin', in the sense used by Vavilov (1951), and, if so, is this centre of origin located in Amazonia or in Central America, or somewhere else? Second, are the wild cocoa populations being studied in the field truly wild? It is possible to formulate a wide variety of hypotheses about the history of cocoa both as a wild and as a cultivated species; it is hoped that the evidence presented in this thesis will enable the rejection of certain of these hypotheses and contribute to a definitive explanation of the known history and present status of cocoa.

The evidence to be presented in this thesis has been derived from two main sources. First, there are data on geographical variation in the wild cocoa populations of Ecuador and south-eastern Colombia. Some of these data were acquired in the field, in the course of a systematic programme of germplasm collecting. Further data were obtained from trees grown in a field genebank in Ecuador, these trees having been propagated from seed or budwood collected from wild populations. Some additional data were obtained from trees of wild origin in other genebanks. Limited published data are available from a few other wild populations, and these data have been used where appropriate.

The second source of information for this thesis is a body of data on the growth, reproduction and dispersal of wild cocoa. These data, which were obtained while I was working on the germplasm collecting project in Ecuador, are used in this thesis to build a picture of cocoa as a wild species in its natural habitat. Some of these data were acquired in the course of collecting expeditions or through study of the wild cocoa population at San Carlos. Additional data were obtained from accessions of wild origin growing in the field genebank at San Carlos.

This study was conducted within the constraints of a germplasm collecting programme. Management decisions, for example in relation to the collecting programme or the planting layout of the field genebank at San Carlos, were determined by the practical objectives of the programme, and not by the rigorous standards of scientific data collection. Nonetheless, the germplasm collecting programme created an unprecedented opportunity for studying cocoa in its natural habitat, and part of my intention in writing this thesis is to make available as much as possible of the information so acquired.

Structure of the thesis

Two threads run in parallel through this thesis. The first begins with a review of published ideas on germplasm collecting strategies (Section 1.2) and continues, via the description of the study region in Chapter 2, to the acquisition and analysis of data on geographical variation (Chapters 3 and 4); the results are discussed in Section 7.1. The second thread begins with a review of published information on wild cocoa (Section 1.3) and passes on, again via the description of the study region (Chapter 2), to methods and results relating to the population biology of wild cocoa (Chapters 5 and 6); these results are discussed in Section 7.2. Finally, the two threads of the thesis are brought together in the broader discussion of Section 7.3, and the overall conclusions of the thesis are presented briefly in Chapter 8.

1.2 Literature review: collecting strategies

The question of efficient strategies for sampling genetic diversity in crop species has been discussed by Allard (1970), Bennett (1970), Marshall and Brown (1975), and Bogyo *et al* (1980). The efficiency of the sampling strategy can be measured in terms of the 'cost', in money or time, of obtaining a given amount of useful genetic diversity. The components of the collecting strategy which are considered here are first, the choice of target region, and second, the number and distribution of samples within the target region. Other factors are assumed to be constant, although it is worth noting that radical improvements in collecting methodology would affect the relative costs of the strategies discussed here. Although their approaches vary widely, all these authors assume that useful genetic diversity can be measured in terms of the average number of alleles per locus, for a large number of loci. For practical purposes, where allele frequencies have not been measured, calculations are often based on frequencies of phenotypic character states, using the lists of standard descriptors which have been published for many crop species. These standard descriptors are based on characters for which phenotypic variation is essentially equivalent to genetic variation, with a minimal environmental component. Quantitative descriptors are also used, usually without any attempt to interpret their values in terms of allele combinations; hence useful variability for quantitative characters is sometimes equated with conventional statistical measures of variation such as standard deviations or coefficients of variation.

There are circumstances in which useful genetic diversity would not be assessed over a large number of loci (or descriptors) but would be determined by the presence of specified alleles at one or a few loci. For example, a collecting programme might seek genotypes having resistance to a particular disease. If the genotypes in question are detectable in the field, collecting strategies which assume random sampling are not relevant. Given, for the present, that screening for specific alleles in the field is rarely feasible, the consensus appears to be that breeding programmes benefit most from a general increase in genetic diversity, which is achieved by random sampling from related wild and primitive cultivated populations. This follows from the assumption that variation in 'useful' characters is likely to be correlated with variation in general.

The result of a sampling strategy can be expressed as the probability of obtaining alleles which have a frequency above some threshold value, in a specified target region. The question of choice of target region will be considered first. Ideally, the choice of target region would be based on an objective assessment of genetic diversity across the whole distribution of a crop species. Such an objective assessment can be achieved using conventional statistical measures of variability, such as coefficients of variation for quantitative characters. For example, Witcombe and Gilani (1979) compared wheat and barley collections from Nepal and Pakistan. Murphy and Witcombe (1981) extended the comparison for barley to include populations from Afghanistan, India and Tibet.

Another approach evaluates the distribution of variation between geographical regions, using the Shannon-Weaver index of diversity:

$$H = - p_i \log p_i \qquad (1.1)$$

where p_i = proportion of accessions in the ith class

The calculation of this index from qualitative descriptor data is straightforward; nonetheless, its use has been criticized (Marshall and Brown, *op cit*) on the ground that it measures the evenness of allele frequencies as well as the number of alleles in a population. If useful genetic diversity in a collection of germplasm is equated solely with the number of alleles per locus, then a measure which includes evenness may be misleading as a measure of comparative diversity. Despite the objection, this has become a recognized technique of which there are several published examples (Holcombe *et al* 1976, Jain *et al* 1975, Tolbert *et al* 1979; see also review by Chapman 1984). This approach was not used in the present project.

A further objective reason for selecting a target region might be the occurrence of conditions that would increase the likelihood of finding a particular category of useful alleles. For example, specific alleles for disease resistance would normally be sought in a region where the disease was prevalent.

It has often been tacitly assumed that intra-species genetic diversity is concentrated in the 'centres of diversity' described for many crops by Vavilov (1951). The studies referred to above show that this is often an over-simplification. Vavilov's analysis of diversity typically refers either to higher taxonomic levels (distributions of species within a genus), or to a relatively small range of characters compared to those used in more recent studies. High genetic diversity is often found in regions far removed from the centres of diversity identified by Vavilov. Also, the pattern of diversity for one group of characters may differ markedly from that for another group. There is little justification for choosing a target region solely on the basis that it is within the conventionally-defined centre of diversity for a species.

Despite the existence of methods for objectively comparing genetic diversity between regions, the choice of target regions for germplasm collecting projects has more often been governed by practical or political considerations. Practical considerations include the effectiveness of germplasm from a particular region in previous breeding programmes, as well as questions of ease of travel in collecting areas, and assessments of the rate of genetic erosion in different regions. Political considerations usually relate to the willingness of government agencies to cooperate with germplasm collecting projects. For the purposes of this discussion, it will be assumed that the choice of target region is not open to debate.

For random sampling of a target region in which all alleles are distributed at random, both geographically and between individuals (that is, assuming an outbreeding species with random mating), the probability of obtaining at least one copy of a given allele is a simple function of the frequency of the allele and the number of plants collected:

$$P = 1 - (1 - p)^{2n} \tag{1.2}$$

where P = probability of obtaining at least one copy of a specified allele
p = frequency of the specified allele
n = number of plants collected

Bogyo et al (1980) use the form

$$P = 1 - e^{-2np} \tag{1.3}$$

which is an approximation, based on the Poisson distribution, yielding virtually the same P for values of p less than 0.1. Values of P corresponding to realistic ranges of p (0.01 to 0.5) and n (10 to 1000) are given in Table 1.1; note that the probability of capturing an allele whose frequency is 0.05 lies in the range 0.92 to 0.99 if 50-100 samples are collected. For a species such as cocoa, this sort of sample size looks broadly feasible.

Table 1.1 *Probability (P) of obtaining at least 1 copy of an allele with frequency p in a sample of n individuals.*

	p = .5000	.1000	.0500	.0100
n = 10	.9990	.6513	.4013	.0956
20	1.0000	.8784	.6415	.1821
30	1.0000	.9576	.7854	.2603
40	1.0000	.9852	.8715	.3310
50	1.0000	.9948	.9231	.3950
100	1.0000	1.0000	.9941	.6340
500	1.0000	1.0000	1.0000	.9934
1000	1.0000	1.0000	1.0000	1.0000

Notes:
1. P calculated from Equation 1.2.

This basic equation can be developed in two ways. The first, and less interesting, development is calculation of the probability of collecting one allele at each of a number of loci. If all the specified alleles have the same frequency, then:

$$P = \{1 - (1 - p)^{2n}\}^{m} \qquad (1.4)$$

where P = probability of obtaining at least one copy of each the specified alleles (one per locus)
p = frequency of the specified alleles
m = number of loci
n = number of plants collected

Some consequences of this equation are illustrated by Table 1.2, which shows that the required sample size is more affected by extending the search to rarer alleles than by increasing the number of loci. This result is encouraging, but is difficult to use directly since the number of target loci is normally unknown.

Table 1.2 *Number of plants (n) which must be sampled to achieve a probability (P) of obtaining at least one copy of a specified allele (with frequency p) at each of m loci, given P = 0.95.*

	p = .50	.10	.05	.01
m = 10	8	50	103	525
20	9	57	116	594
30	9	60	124	634
40	10	63	130	663
50	10	65	134	685
100	11	72	148	754
500	13	87	179	914
1000	14	94	193	983

Notes:
1. n calculated the relationship of Equation 1.4.

The second and crucial development is calculation of the probability of collecting at least one copy of each of a number of alleles at a single locus. Both Marshall and Brown (1975) and Chapman (1984) present equations for the case of four alleles. Their equations are complex and awkward to work with, but both papers tabulate probabilities for various combinations of allele frequencies and sample sizes. It does appear, however, that a simpler expression can be used with no inaccuracy for realistic combinations of sample size and allele frequency:

$$P = \{1-(1-p_1)^{2n}\}\{1-(1-p_2)^{2n}\}\{1-(1-p_3)^{2n}\}\{1-(1-p_4)^{2n}\} \quad (1.5)$$

P = probability of obtaining at least one copy of each of four alleles at a single locus
p_1, p_2, p_3, p_4 = frequencies of the four alleles
n = number of plants collected

In other words, given that the occurrences of the four alleles are independent events, the probability of all four alleles being captured is simply the product of the probabilities for the individual alleles (it is not assumed that the four allele frequencies sum to unity). Similar expressions can be derived for larger numbers of alleles. These expressions have some useful properties. If $p_1 < p_2 < p_3 < \ldots p_m$, then it follows that

$$P > \{1 - (1-p_1)^{2n}\}^m \qquad (1.6)$$

where m = number of alleles
P,n as in Equation 1.5

This can be restated as

$$n < \{\log(1 - P^{1/m})\}/2\log(1 - p_1) \qquad (1.7)$$

In an extreme case, all m alleles will have the same frequency and the expression for P will be the same as that of Equation 1.4, except that m refers to the number of alleles rather than the number of loci. More realistically, if the target alleles have a range of frequencies, then the value of P will be affected principally by the frequency of the rarer alleles, since the value of

$$\{1 - (1-p_i)^{2n}\}$$

quickly approaches unity for $p_i > 0.1$. This generalization is also made by Chapman (1984), although it is not so easily deduced from his equation for P.

Before any use is made of Equation 1.5 for calculating minimum sample sizes, attention must be given to the question of how many alleles are likely to be present, and at what frequencies, at a typical locus. This question was discussed at length by Marshall and Brown, who concluded that a typical population would contain no more than four 'common' alleles (defined as $p_i > 0.05$), but many 'rare' alleles ($p_i < 0.05$). Their conclusion is supported both by population genetics theory and also by data from real plant populations.

The relationship between n and P for four alleles is demonstrated in Table 1.3, which shows that the probability of capture is affected most strongly by the frequency of the rarest allele. For example, when all four alleles have frequency 0.05, the sample size for P = 0.95 is 43 (case (e)). Cases (a)-(d) have one, two or three alleles at this frequency, and sample size varies only between 30 and 40. However, if the rarest allele has a frequency of only 0.01, then n rises to 218. Chapman (1984, Table 7.3) has tabulated n for various combinations of allele frequencies with at least one allele at a frequency of 0.01, showing that required sample sizes (for P =

0.95) range from 150 to 203 (after correcting for a different definition of allele frequency).

Table 1.3 *Probability of capturing four alleles at one locus; effect of varying allele frequencies and number of plants sampled.*

Case	p_1	p_2	p_3	p_4	n	P
(a)	.05	.1	.4	.4	10	.56
					20	.86
					30	.95
					40	.98
					50	.99
					60	1.00
(b)	.05	.1	.2	.6	30	.95
(c)	.05	.05	.4	.4	36	.95
(d)	.05	.05	.05	.8	40	.95
(e)	.05	.05	.05	.05	43	.95
(f)	.1	.1	.1	.1	21	.95
(g)	.01	.01	.01	.01	218	.95

Notes:
1. P calculated using Equation 1.5.
2. For cases (b)-(g), the value of n is the smallest sample size for which $P > 0.95$.

Marshall and Brown develop their argument by further categorizing alleles as 'widespread' or 'local'. Widespread alleles occur at similar frequencies in all the populations within the target region, while local alleles are restricted to one populations (or a few adjacent populations).

Widespread common alleles pose no difficulty for a collecting programme, and no special sampling strategy is indicated. Widespread rare alleles will be captured with probability determined by the total number of plants sampled from the whole target region. Use of Equation 1.2 or 1.3 will show that very large sample sizes are required for rarer alleles; for example, Bogyo *et al* (1980) calculated that a minimum sample of 30,000 plants is needed for 0.95 probability of capturing an allele with frequency .0001. No adjustments in collecting strategy would compensate for this basically unfavourable situation, which is even worse for local rare alleles. Marshall and Brown argue, first, that rare alleles probably lack adaptive value and are therefore inherently less interesting to breeders than common alleles and second, that rare alleles are more efficiently acquired by screening cultivated populations. Interest therefore centres on devising collecting strategies which are efficient for local common alleles, the aim of such strategies being 'the collection of at least one copy of each variant [allele] occurring in the target populations with frequency greater than .05' (Marshall and Brown 1975, p63). In practice, a strategy should be sought which captures '95% of the genes distributed in the population with a frequency of [at least] .05' (Oka 1969, cited by Marshall and Brown).

This recommendation seems to have been widely accepted, not least because it provides a basis for the calculation of minimum sample sizes which are realistic even for species which are expensive to collect; there would be little purpose in defining a criterion which could never in practice be achieved. Even so, the practical application of this criterion immediately raises a difficulty, which is the need to delineate the 'local populations' so that an appropriate number of samples can be collected in each one. If the local populations are delineated using geographical or administrative boundaries, no difficulty arises, but there is no guarantee that the resulting populations do not contain genetically distinct sub-populations; alternatively, the populations may be so similar genetically that full-scale sampling of each is wasteful. If the local populations are to be delineated using information on geographical variation in the target region, the process of obtaining such information may be more costly than the germplasm collecting which it is supposed to facilitate.

In practice, this difficulty can be avoided by treating germplasm collecting not as a single sampling operation but, at least potentially, as a sequence of stages, as suggested by Allard (1970) and Jain (1975), among others. The initial stage is planned on the basis of some broad assumptions about the distribution of variation in the target region, or any other available information. For subsequent stages, the collecting strategy is modified to take account of information on the pattern of variation obtained during the previous stage. In practice, a reasonable degree of caution applied to the first stage strategy will very likely mean that most of the variation can be captured in the first stage, unless the extent of local differentiation is much greater than expected; what is crucial is that a collecting programme is planned so that resources are available for a second stage if the results of the first stage show that distinctive local populations are being undersampled.

A cautious first stage collecting strategy depends on planning the distribution of samples. If the cost of collecting is ignored, then the most effective strategy is to collect small groups of samples throughout the target region, with the location of each group fixed on a regular grid. This will provide information on variation at a scale equal to or greater than the distances between grid points. The cost of collecting can be divided into two main elements, the cost of travelling to a collecting locality (plus any other preparation) and the cost of taking one sample at a given collecting locality. Since the 'travelling cost' is typically greater than the 'sampling cost', it is evident that travelling to large number of separate points and collecting a few samples at each is always more expensive than a strategy of collecting larger numbers of samples at fewer localities, even if the total number of samples is the same in both cases. In general, the more samples collected at each locality, the lower is the average per-sample cost. This is balanced by the fact that distinctive local populations which fall between the widely-spaced localities will not be detected by the cheaper strategy.

These strategies have deliberately been described in somewhat loose terms. In fact, there seems to be no consensus as to how the effects of these strategies can best be quantified, nor have cost data for real collecting programmes been published. For example, Allard (1970) used data on quantitative variation in physiological and morphological characters of *Avena fatua* (wild oats), in a 600 km x 200 km region of California, to deduce the distribution of samples that would be needed to represent the local populations in this region. He concluded that 500 local populations should be sampled; there was more small-scale variation in the east-west direction, across the mountain ranges, so that samples should be more closely spaced in the east-west direction (average spacing 2 km) than in the north-south direction (average spacing 200 km). Variation for the measured characters tended to be normally distributed; Allard stated that 'experience indicates that a sample of a few seeds (one panicle) from each of 200-300 plants would capture most of this variability'. According to Allard, this sample size would also result in the capture of alleles with frequencies greater than two or three per cent, although application of Equation 1.5 indicates that smaller samples would be adequate. In total, approximately one million seeds would be collected. Avena fatua typically has an outcrossing rate of less than five per cent (Jain 1975), but Allard suggests that similar collecting strategies would be required for highly cross-pollinated species such as *Lolium multiflorum* and *Collinsia heterophylla*, in the same region.

In general, the level of gene flow within a species does not seem greatly to restrict the development of genetically differentiated local populations, as demonstrated by studies of heavy-metal tolerance in plants (Antonovics *et al* 1971) and other cases of small-scale genetic differentiation in response to environmental mosaics (reviewed by Endler, 1986).

Although collecting a grass such as *Avena fatua* is likely to be quick and hence cheap, the total cost of acquiring, documenting, maintaining and evaluating such a large collection would certainly be prohibitive unless samples within each population were combined; for species where per-sample collecting costs are of the order of tens or hundreds of pounds, such rigorous delineation of local populations is obviously impracticable. Furthermore, to treat every local population as a separate target for germplasm collecting is to fail to recognize that the majority of alleles are probably common to most local populations. What are required, instead, are methods which quantify the proportion of total variation in a target region which is captured by different strategies. Marshall and Brown (1975), quoting Oka (1969), offer the equation:

$$G = 1 - \{(1 - P) + P(1 - p)^n\}^N \tag{1.6}$$

G = fraction of total variation captured
P = proportion of total variation per site
p = proportion of within-site variation per plant
N = number of sites sampled
n = number of plants sampled per site
Note: the terminology for Equations 1.6 and 1.7 has been slightly altered from Marshall and Brown.

The number of sites visited, and the number of plants collected at each site are controlled by the total resources available and the 'travelling' and 'sampling' costs referred to earlier:

$$E = n_1(c_1 + n_2c_2) \tag{1.7}$$

where E = total resources available
c_1 = 'travelling cost' = cost per site visited
c_2 = 'sampling cost' = marginal cost per plant sampled
n_1 = number of sites (= N in (1.6))
n_2 = number of samples per site (= n in (1.6))

It is then possible to obtain values of n_1 and n_2 which maximize G, for any given values of P, p, c_1 and c_2. The optimum value for n_2 is actually a function of the ratio of travelling cost to sampling cost, given P and p. Some representative solutions were tabulated by Marshall and Brown (Table 4.5), assuming cost ratios (c_1/c_2) of 25 or 100. At one extreme, for an outbreeding species in which there is much intra-site variation but populations at different sites are broadly similar ($P = 0.75$, $p = 0.05$), the optimum number of plants per site is 36-55. Making the reverse assumptions, for highly inbred cultivars where intra-site variation is minimal ($P = 0.01$, $p = 0.95$), it is not necessary to sample more than 1 or 2 plants per site.

A related formulation was used by Bogyo *et al* (1980), based explicitly on the theory of multistage sampling (Cochran 1963). The variance of an area mean can be partitioned into components for 'sites' and 'samples':

$$V = V_1/n_1 + V_2/n_1n_2 \tag{1.8}$$

V = variance of the overall mean for a character
V_1 = variance component for sites
V_2 = variance component for samples

Assuming that Equation 1.7 describes the cost constraints, then the optimum number of plants to sample at each site is again a function of the ratio of travelling to sampling costs:

$$n_2 = \sqrt{(c_1 V_2 / c_2 V_1)} \qquad (1.9)$$

Equations 1.7-1.9 can be extended to sampling at three levels, with cost ratios c_1/c_2 and c_2/c_3, and corresponding variance component ratios V_2/V_1 and V_3/V_2. Bogyo *et al* used data from a wheat collecting project in Sicily and obtained values for the ratio V_2/V_1 ranging from 3 to infinity, with an average of about 9. Given a cost ratio (c_1/c_2) of about 16, then the optimum number of samples per site would be 12; coincidentally, this was in fact the average number of samples collected in each of the ten areas into which Sicily was divided for collecting purposes. The advantage of Cochran's formulation is that V_1 and V_2 can be calculated readily from an analysis of variance (Som 1973). The form of Equation 1.9 also means that the optimum n_2 is insensitive to quite large changes in either the cost ratio or the ratio of the variance components; some representative values are given in Table 1.4.

Table 1.4 *The effect of cost ratio and variance component ratio on the optimum number of samples per site (n).*

	Cost ratio c_1/c_2						
	100	50	10	5	1	.5	.1
V_2/V_1	Optimum n						
100	100	71	32	22	10	7	3
50	71	50	22	16	7	5	2
10	32	22	10	7	3	2	1
5	22	16	7	5	2	2	1
1	10	7	3	2	1	1	1
.5	7	5	2	2	1	1	1
.1	3	2	1	1	1	1	1

Notes:

1. n calculated using Equation 1.9.
2. Bogyo et al (1980) give a similar table.

Another approach to the design of collecting strategies which ought to be fruitful is the use of clustering and other multivariate techniques.

Patterns of variation within a target region can either be summarized explicitly by classifying samples into clusters or they may be displayed graphically using ordination, factor analysis or related techniques (reviewed by Gould and Johnston, 1972). Most recent studies of geographic variation have used multivariate techniques to summarize data, but there are few published studies where the results have been used in the design of germplasm collecting strategies within target regions. Witcombe and Rao (1976) used reciprocal averaging ordination on data from wheat accessions collected at 91 sites in eastern Nepal; they were able to assign these sites to 5 areas in such a way that a sample of maximum variability could be obtained along a transect traversing four of these areas. The fifth area was shown by the ordination to overlap with three of the other areas. Accessions collected from distant sites tended to be further apart on the ordination diagram than those collected from adjacent sites, which would justify the use of collecting resources in travelling to more distant sites.

Bogyo *et al*, in the study described above, used a hierarchical cluster analysis to classify their ten collecting areas in Sicily into four groups. The groups contained geographically adjacent areas, again suggesting that genetic differentiation was a function of geographic separation. The implications of this result for germplasm collecting strategies were, surprisingly, not discussed.

Other studies have used multivariate analyses in the context of checking the identification of material held in genebanks (Yndgaard and Hoskuldsson 1985; Engels 1986).

1.3 Literature review: wild populations of cocoa

The taxonomic literature on *Theobroma*, which contains many brief references to observations of wild cocoa, was comprehensively reviewed by Cuatrecasas (1964). Soegeng-Reksodihardjo (1964) also revised the taxonomy of *Theobroma*, but although his monograph contains useful accounts, based on extensive fieldwork, of *Theobroma* species other than *T. cacao*, it provides little information on wild populations of cocoa. Cuatrecasas examined various hypotheses about the taxonomy and origins of cocoa, but concluded that it should be considered as a single species (*Theobroma cacao* L.) distributed as a wild species from southern Mexico to Bolivia and Brazil. Two subspecies were described, *T. cacao* subsp. *cacao* and *T. cacao* subsp. *leiocarpum*; these correspond broadly to the Criollo and Forastero groups of cultivated cocoa populations described by Cheeseman (1944). A third type, Trinitario, probably originated in cultivation, through hybridization between Criollo and Forastero populations. Cheeseman's classification is probably the most widely accepted way of describing cultivated cocoa populations and varieties, but his terminology is less useful when applied to wild populations, as wild trees in the Upper Amazon often combine characteristics from more than one type.

There is very little published information on wild cocoa populations in Central America (including Mexico), although cocoa was first cultivated there, by the Aztec and Maya Indians (Thompson 1956). On the other hand, there is a large body of information, much of it obtained in the course of germplasm collecting expeditions, on wild cocoa in South America, especially in the Amazon basin. Published descriptions of wild and primitive cultivated populations are summarized in Table 1.5. The approximate distributions of these populations are shown in Figure 1.1. Allen and Lass (1983) reviewed previous collecting expeditions; one recent report, on wild cocoa in French Guiana (Clément 1986), fills a gap noted in that review. Further information on the characteristics of wild cocoa populations in Ecuador and the Río Caquetá region of Colombia has also been presented in LCTAP reports (Allen and Lass 1983; Allen 1988).

Table 1.5 *Characteristics of some wild and primitive cultivated cocoa populations, from published descriptions.*

Population	Map Ref	Unripe Pod Colour	Pod Shape [d:l ratio]	Pod Apex	Pod Base Constriction	Pod Surface	Seed Colour	Flower Pigment	Flush Colour	References
Criollo	1	green or red	elongated [<0.5]	pointed	variable	very wrinkled 10 deep furrows	white or pale purple			Cuatrecasas 1964 Cheesman 1944
Western Ecuador 'Nacional'	2	green	oval	blunt point	slight	wrinkled deep furrows	pale purple			Cheesman 1944
Río Guaviare Colombia	3	green	[0.6]	rounded	absent	smooth	dark purple			Rojas-Peña 1951
Brazilian Amazon (and south-east Colombia)	4	green	[>0.5]	rounded or slight point	absent or slight	smooth or slightly wrinkled 10 shallow furrows	dark purple		red	Cuatrecasas 1964 Soria 1970
Río Nanay, Peru near Iquitos	5	green	long oval	rounded	absent	slightly wrinkled	dark purple	present		Pound 1938, 1945
Río Napo, Peru 'criollo de la montagne"	5	green or slightly red	long	pointed	present	wrinkled	pale purple or purple			Pound 1938, 1945
Parinari Río Marañon, Peru	5	green	long	pointed	present (variable)	smooth or wrinkled	dark purple	absent		Pound 1938, 1945
Río Marañon, Peru R Tigre to R Morona	5	green or slightly red	short oval			wrinkled 10 furrows	dark purple			Pound 1938, 1945
Río Huallaga, Peru	5	green	long variable	rounded or blunt point	absent	wrinkled	pale purple			Pound 1938, 1945
Río Ucayali, Peru	5	green	long	blunt point	absent	moderately wrinkled	dark purple		red	Pound 1938, 1945
Ecuador Amazon	6	green	[0.5-0.65]	rounded or moderate point	absent or slight	moderately or very wrinkled	white or pale to dark purple	absent	green	Chalmers 1969-1974
Alto Beni, Bolivia	7	green	elongated	small point		slightly wrinkled 10 shallow furrows	dark purple	absent		Soria 1965
Guianas	8	green	oval [0.6]	absent or slight point	slight	smooth or slightly wrinkled shallow furrows	dark purple		red	Myers 1930 Clément 1986

Notes:
1. These descriptions are simplified versions of those given in the original collectors' reports.
2. The collector's own terminology has been used as far as possible.
3. Flower and flush pigment data have been supplemented with information from Engels (1981) and LCTAP
4. LCTAP publications, and this thesis (Chapter 4) provide additional data on populations 2, 4, 5 and 6.
5. Map references refer to Figure 1.1.
6. This table is updated from Table 1 in Allen and Lass (1983).

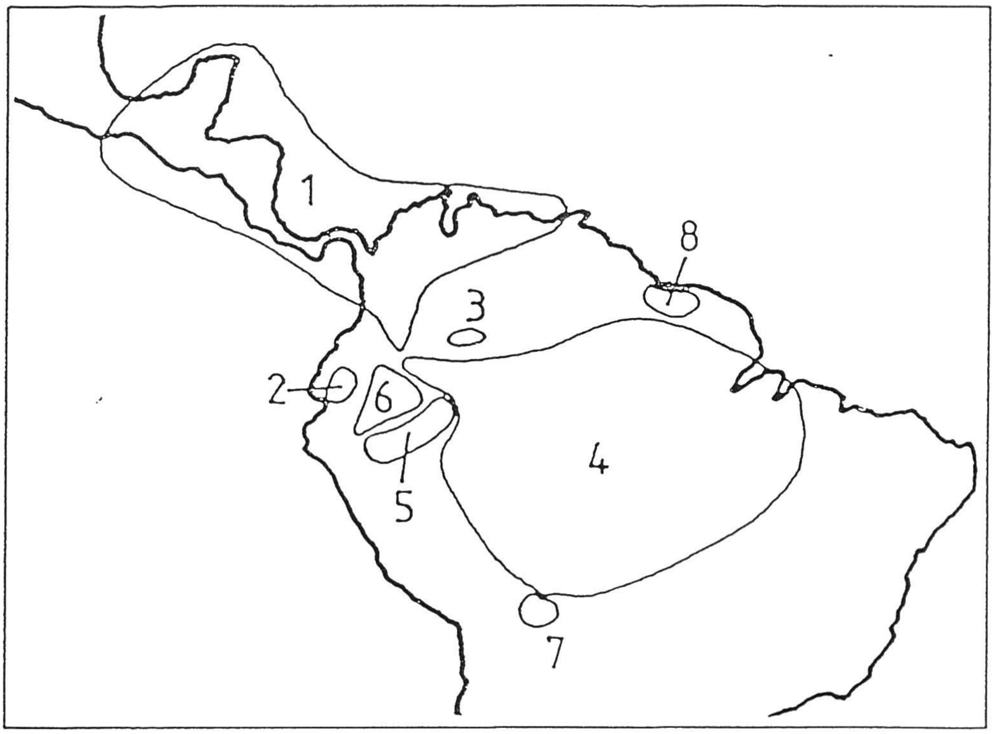


Figure 1.1 *Geographical distribution of the wild and primitive cultivated cocoa populations listed in Table 1.5.*

Cheeseman (1944), drawing in particular on the results of Pound's extensive explorations in 1937-38 and 1942-43 (Pound 1938, 1943, 1945) originated the idea that there is a 'centre of diversity', and hence a 'centre of origin' for cocoa in the Upper Amazon region. The existence of this centre is frequently assumed in standard accounts of cocoa as crop species (for example, Cope 1976). The relevance of centres of diversity and centres of origin to germplasm collecting strategies was mentioned earlier; application of these concepts to cocoa will be discussed in Chapter 7.

Only the most casual references to the ecology of wild cocoa can be found in the literature. A brief outline of generally accepted views on the ecology of cocoa provides the starting-point for Chapters 5 and 6. Cocoa is usually described as an understory species of lowland tropical rain forest, typically found along rivers (Richards 1952) in areas which are flooded for several months of the year (Wood 1975), although other observers found that it tended to grow above the regular flood level (Soegeng-Reksodihardjo (1964). From the far more extensive observations of cultivated cocoa (reviewed in Wood, 1975), it is known that cocoa produces large numbers of flowers (>10,000 flowers/tree/year) which are pollinated mainly by midges of the genus Forcipomyia (Diptera: Ceratopogonidae) (Young 1986). Virtually all Upper Amazon cocoa genotypes that have been tested are self-incompatible, but some cultivated populations are partly or entirely self-compatible. Despite the production of numerous flowers, even under the most favourable conditions only about 1% of flowers, and often only 0.1%, develop into mature fruits. The fruits remain attached to the trunk or main branches and seed dispersal depends on the pods being opened by mammals such as squirrels or monkeys. Under plantation conditions, pod production begins when the trees are 2-3 years old, and increases up to at least 8-10 years; apart from trees which are killed by disease or falling shade trees, plantation cocoa trees appear to live for hundreds of years.

Table 1.6 *The main cocoa collecting expeditions in South America, 1937-1985.*

Expedition	Date	Target Region	Map Ref	Number of Trees Collected	Main Genebanks	References
Pound (Trinidad)	1937-1938	Amazon and tributaries (Brazil, Peru, Colombia)	1	ca 50	Trinidad	Pound 1938, 1943, 1945 G.Lockwood and A.F.Posnette (personal communication)
Desrosiers and Von Buchwald (Ecuador)	1949	Río Napo, Ecuador	2	26	Pichilingue	Desrosiers and Von Buchwald 1950 Soria 1970
Anglo-Colombian (Colombia/Trinidad/UK)	1952-1953	Colombia	3	18	Trinidad	Baker et al 1954
Doak and Zambrano (Ecuador)	1961	Río Napo, Ecuador	4	46	Pichilingue	Soria 1970
CEPLAC/CEPEC/DEPEA (Brazil)	1965 onwards	Brazilian Amazon	5	>1000	Belém Manaus Santarém	Soria 1970 (see notes)
Chalmers et al (Ecuador/Trinidad)	1968-1973	Ecuadorian Amazon	6	ca 280	Pichilingue Trinidad	Chalmers 1969-1974
ICA (Colombia/IBPGR)	1978 onwards	Colombia	7	?	Palmira	L.Lopez (personal communication) (see notes)
LCTAP (Ecuador/UK)	1980-1985	Ecuadorian Amazon	8	420	San Carlos Pichilingue Trinidad	LCTAP publications
Expedicion Botánica (Colombia)	1984	Río Caquetá, Colombia	9	130	Palmira San Carlos Trinidad	Allen 1988
IRCC Guyane (France)	1984-1985	Guyane (French Guiana)	10	95	Sinnamary-Combi	Clément 1986

Notes:
1. Most expeditions have collected cultivated as well as wild cocoa. These are not consistently distinguished in reports.
2. Some short expeditions with less than 10 collections have been omitted.
3. The Brazilian (CEPLAC/CEPEC/DEPEA) collecting programme is the longest and most extensive mounted by any country.
   A detailed account of this programme was provided by B.G.D.Bartley (personal communication), but this could not be included here for reasons of space.
4. The Colombian (ICA) collecting programme has made several expeditions since 1978 for which no reports are available in the public domain.
5. Map references refer to Figure 1.2.
6. This table is updated from Table 2 in Allen and Lass (1983).

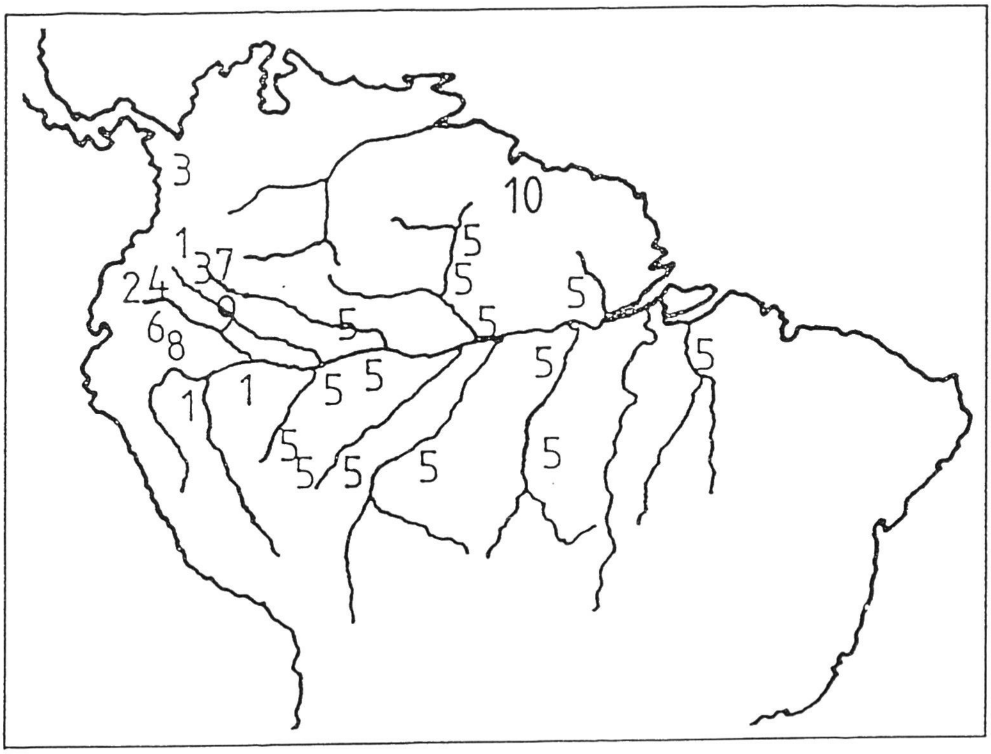


Figure 1.2 *Target regions for the cocoa collecting expeditions listed in Table 1.6.*

2 The study area

This chapter provides information on the physical environment and history of area in which the study took place, covering the Amazon Region of Ecuador and the Río Caquetá region in Colombia in general terms, with some additional information relating to the San Carlos genebank site.

2.1 The Amazon Region of Ecuador

Topography and soils

The Amazon Region of Ecuador (see Figure 2.1), also known as the *Oriente*, covers an area of about 120,000 km^2, about half the area of the UK. The majority of the region consists of a gently sloping plain, rising from about 200 m altitude in the east, to 300-400 m where it meets the edge of the Andean *sierra* in the west. Within the region, the mountains reach over 5000 m in several places, and there is a considerable area of land between 400 and 1200 m, including relatively level valley floors as well as rugged terrain with steeply-sloped hillsides. As an administrative entity, the region is divided into four provinces; from north to south, these are Napo, Pastaza, Morona-Santiago and Zamora-Chinchipe Provinces.

Collecting covered the area below about 1200 m altitude, which accounts for the greater part of the region. Several major rivers flow through this area, mainly from north-west to south-east, ultimately reaching the Amazon. All the rivers are fairly shallow; the more northerly rivers, including the Putumayo, San Miguel, Aguarico, Napo and Curaray, are navigable. Further south, the rivers, such as the Pastaza, Makuma, Upano and Santiago, fast-flowing with many rapids and are dangerous or impossible to navigate for much of their length within Ecuador.

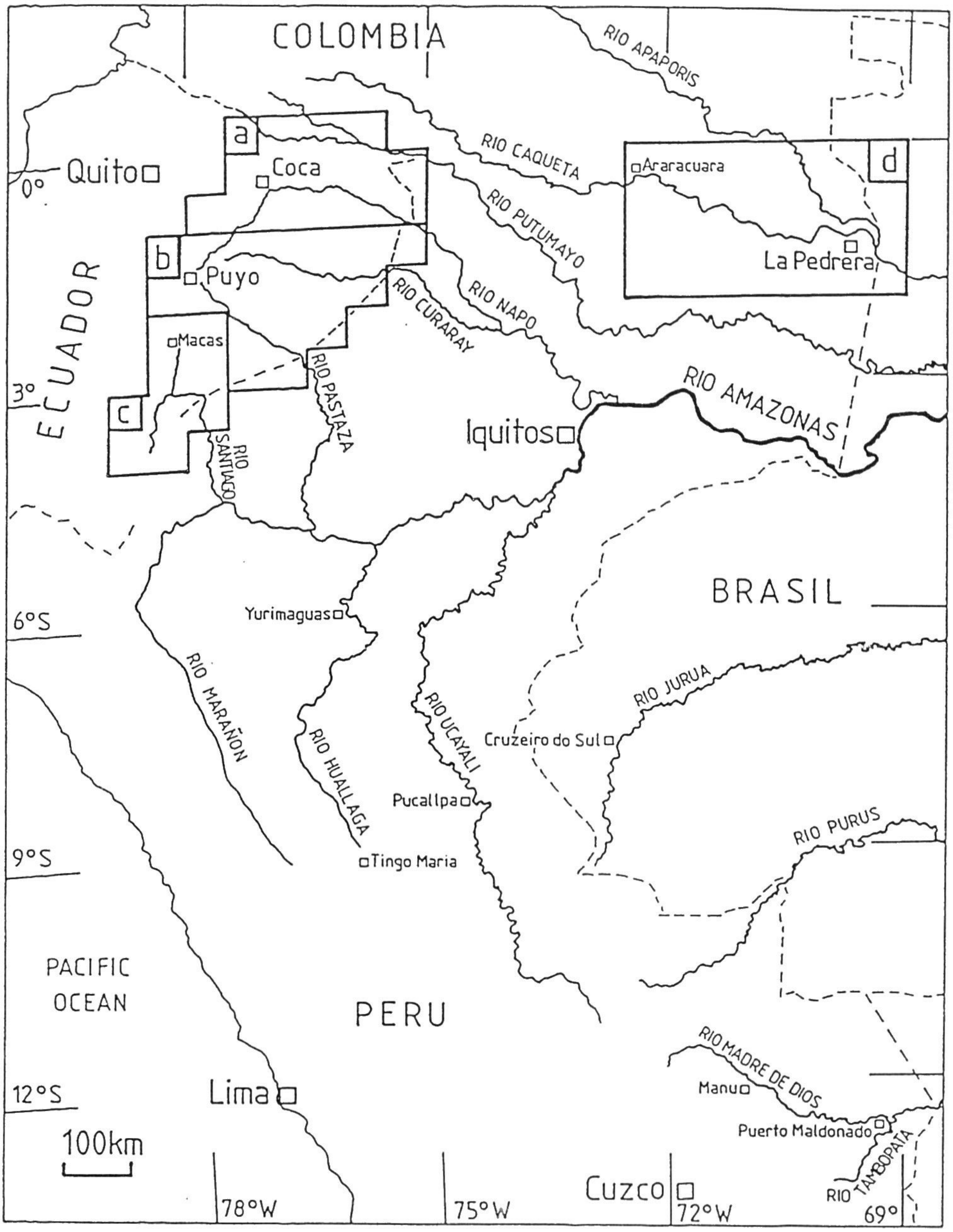
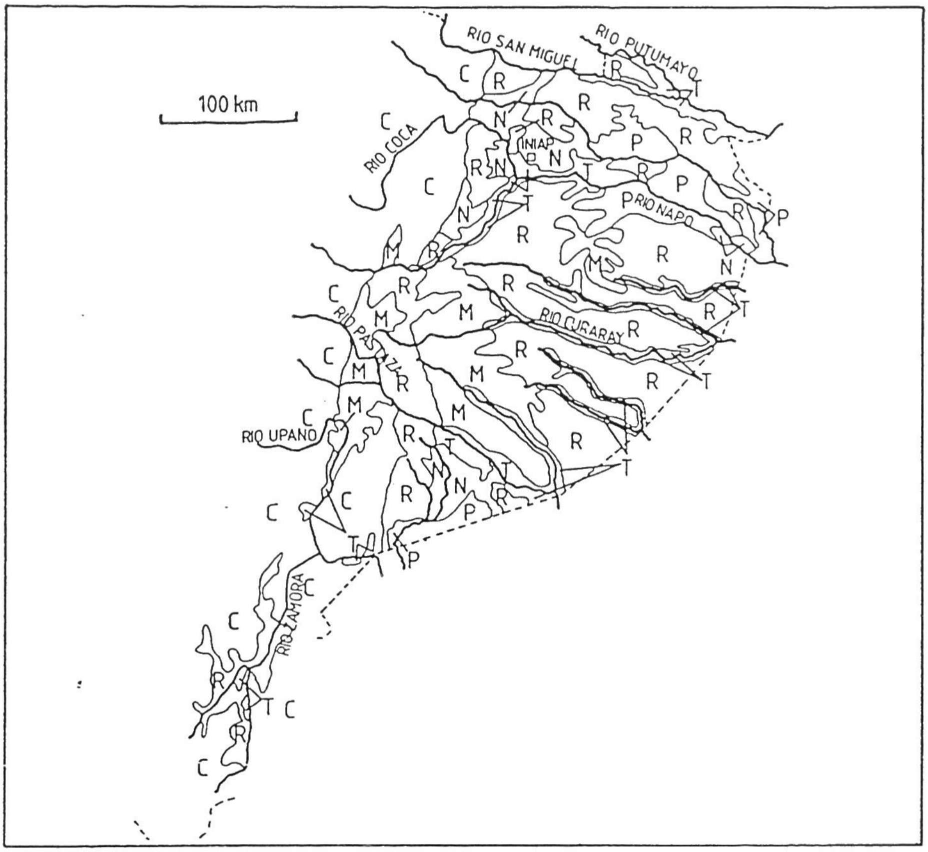


Figure 2.1 *Map of the Upper Amazon region, showing areas explored during LCTAP in Ecuador and Colombia. (Areas marked 'a', 'b', 'c', 'd' are shown in greater detail in Figures 4.1-4.4).*

In the flatter lowland areas, rivers generally follow a meandering course; river-channel erosion is probably the single most important determinant of the characteristic landforms of the Upper Amazon. A recent analysis of Landsat images for the whole of the Peruvian Amazon (Salo *et al* 1986) distinguished three typical landform categories: (1) present river floodplains, in which cut-off meanders leave a complex pattern of frequently-flooded depressions separated from the main river course by ridges of slightly higher ground which are flooded only exceptionally; (2) former floodplains, with similar topography, but beyond the present meander plains and (3) areas beyond the present and previous floodplains, with a characteristic undulating ('convex-concave') topography resulting from surface erosion of older deposits. In the Peruvian Amazon, the two floodplain formations accounted for 27% of the total area. Lateral erosion of the meander bends causes continuous changes in river courses. These may be surprising rapid, with movement of 100 m/year having been observed (Lathrap 1970), although Salo et al indicated a typical figure of 12 m/year, based on changes in the course of the Río Manú, Peru.

These landform categories can be observed in the Ecuadorian Amazon; they correspond to soil and terrain types which have been mapped in Ecuador. A simplified map of soil and terrain types, modified from Sourdat and Custode (1980) is presented in Figure 2.2. In relation to the distribution of wild cocoa, the most important feature of this map is the predominance in much of the region of undulating uplands corresponding to category (3) in the previous paragraph (Type R on map). This terrain type is derived from heavily weathered Tertiary sediments and is associated with relatively infertile, poorly drained red soils. Level terrain and the more fertile black soils coincide with present (Type T) and former (Type N) floodplains; extensive areas of the latter (corresponding to category (1)) are found in a few areas, notably between Lago Agrio and Coca (and including the San Carlos station), and these areas have been identified as suitable for intensive agricultural development. Present-day floodplains (category (2)) are found as patches or strips along all rivers, but the scale of the map does not in general allow them to be marked. The fertility of soils in the present-day floodplains is very variable; this may be partly accounted for by differences in the amount and composition of sediments deposited by different rivers.



Type	General Description	Soils	Soil classification
R	Undulating upland areas with small, rounded hills (originating in heavily weathered Tertiary sediments);	Shallow, poorly-drained red lateritic soils with high aluminium content; strongly leached, fertility low to moderate;	Oxic dystropepts;
N	Extensive flat areas, mainly former flood-plains;	Deep, well-drained, fertile black soils;	Eutropepts, distrandepts, vitrandepts.
T	River terraces;	Well-drained black soils, fertility variable;	Eutropepts, distrandepts, vitrandepts;
C	Mountainous areas above 1000m; steep slopes;	Mixed soils, heavily weathered; partial volcanic ash cover;	Various;
M	Dissected upland areas, rough terrain with abrupt changes of slope; originating in volcanic deposits;	Variable: includes areas with thick layers of boulders and sand of volcanic origin, also volcanic ash deposits, and areas of more weathered soils; variable fertility and drainage;	Various.
P	Permanently flooded or water-logged areas;	Accumulated organic deposits, sometimes fertile but permanently saturated with water;	Tropacuepts, tropofibrists;

Figure 2.2 Map of soil and terrain types in the Amazon Region of Ecuador, modified by J.B.Allen from Sourdat and Custode (1980).

In the southern part of the region, and along the edge of the *sierra*, there is a complex variety of terrain and soil types (Types M, C) still referable to landform category (3), but derived from older (Mesozoic) sediments overlaid in places by recent (Quaternary) deposits; soils in these areas vary widely in drainage and fertility. The geology of the region is reviewed in Baldock (1982).

Climate

In general terms, the climate of the Amazon Region of Ecuador, below 1200 m altitude, can be summarized as follows: mean daily maximum temperatures range from 26° to 30°; mean daily minimum temperatures range from 17° to 21°; annual rainfall totals vary from 2500 mm to over 6000 mm. The higher temperatures and lower rainfall figures are found in the lowland areas to the east, at 200-300 m altitude; travelling towards the mountains which mark the western edge of the region, rainfall increases, and temperatures decrease, with increasing altitude. However, the Zamora and Upano river valleys, which are enclosed within the mountains in the south of the region, receive less than 2000 mm of rain annually, because of a rain-shadow effect. Climatic data for some representative sites are given in Table 2.1. Some detailed observations on climate in the region were published by Grubb and Whitmore (1966, 1967).

Table 2.1 Representative climatic data for the Amazon Region of Ecuador

Station	Altitude	Annual Rainfall	Temperature Annual mean	Daily max	Daily min	Sunshine
	m	mm	oC	oC	oC	hour/day
San Carlos	265	3152		29.7	20.8	
Payamino	250	3402		29.8	21.2	4.2
Limoncocha	220	3074	24.6			
Tiputini	220	2337	25.3			
Putumayo	230	2917	25.3			
Curaray	300	2749	24.9			
Tena	527	6315	23.1			
Sucúa	910	1664	21.7			
Puyo	950	4461		26.2	16.7	2.7
Zamora	970	1907	21.2			
Shell	1043	4223	20.2			

Notes:

1. Sources as follows (years of records in brackets): San Carlos, LCTAP (2); Payamino, INIAP (2); Limoncocha, Summer Institute of Linguistics (19); Puyo, INERMHI (2); others, .i.Padilla 1977 ;(>4).

Seasonal differences in temperature are small, with monthly mean temperatures varying only by a degree or so through the year. The months of December, January and February are typically drier than the rest of the year, but this pattern is obscured in many years by much variation in the month to month distribution of rainfall. This is well illustrated by a long and reliable series of rainfall measurements from Limoncocha, 25 km east of San Carlos (Table 2.2). Atmospheric humidity is above 80% for much of the time, only dropping to around 50-60% for brief periods on sunny days.

Table 2.2 Rainfall data for Limoncocha, 1961-1979.

YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual Total
1961	132	282	247	320	270	246	167	144	91	365	248	198	2708
1962	192	207	179	217	381	285	352	238	222	288	175	189	2926
1963	55	183	291	401	268	334	269	192	132	188	278	172	2763
1964	30	91	225	449	311	419	282	267	242	235	132	262	2945
1965	214	186	445	285	216	211	219	232	194	338	257	269	3066
1966	229	99	150	239	370	256	417	119	329	139	253	281	2879
1967	235	72	238	312	434	226	315	178	155	405	248	80	2899
1968	325	129	258	462	361	333	213	183	341	344	153	192	3294
1969	206	250	127	274	145	331	191	354	100	135	281	210	2604
1970	273	215	225	385	463	297	238	249	211	261	234	202	3252
1971	358	245	346	282	311	288	220	192	234	352	388	124	3339
1972	372	160	549	270	245	368	343	223	241	428	267	265	3730
1973	195	310	415	162	311	247	311	249	370	184	326	180	3261
1974	211	215	348	197	279	340	326	351	389	300	142	121	3219
1975	309	231	406	341	343	411	255	270	196	277	338	231	3607
1976	339	177	280	289	401	286	274	227	397	300	274	199	3442
1977	45	249	316	389	352	329	283	221	212	263	317	299	3275
1978	125	353	161	379	221	454	311	182	213	226	153	102	2879
1979	41	113	459	249	269	240	260	155	53	159	207	88	2295
Mean	205	198	298	311	313	311	276	222	228	273	246	193	3073
Minimum	30	72	127	162	145	211	167	119	53	135	132	80	2295
Maximum	372	353	549	462	463	454	417	354	397	428	388	299	3730

(Source: Summer Institute of Linguistics) (all data in mm)

Notes:

1. Compiled by J.B.Allen from data supplied by the Summer Institute of Linguistics

The climate at San Carlos and Payamino may be taken as typical of the main lowland part of the region between 200 m and 300 m altitude. Temperature and rainfall data for San Carlos covering the period 1984-1986 are given in Table 2.3. Meteorological data collected by INIAP at the Payamino Station are available for the period September 1983 to June 1986. These data are summarized in Table 2.4. In general, the figures are very similar to those recorded at San Carlos.

Table 2.3 Meteorological data from INIAP Napo-Payamino Experimental Station, San Carlos site, 1984-1986

Year	Month	Rainfall mm	Temperature Daily max °C	 Daily min °C
1984	Jan (16.1-31.1)	98.3	29.9	20.5
	Feb	525.6	28.2	20.9
	Mar	281.5	29.7	21.4
	Apr	386.6	28.7	21.4
	May	267.7	30.1	21.3
	Jun	236.8	28.7	21.3
	Jul	174.2	28.8	20.2
	Aug	250.9	29.3	20.2
	Sep	265.1	30.2	20.6
	Oct	237.2	30.7	21.3
	Nov	409.5	30.3	21.0
	Dec	172.9	30.6	21.3
1985	Jan	44.5	32.0	21.3
	Feb	90.6	29.5	20.6
	Mar	119.1	30.1	21.9
	Apr	312.8	29.8	21.2
	May	85.9	29.4	20.8
	Jun	441.7	26.6	19.9
	Jul	334.0	27.4	19.8
	Aug	274.1	29.2	20.3
	Sep	299.7	30.4	20.4
	Oct	276.9	30.5	20.9
	Nov	312.9	31.2	21.1
	Dec	136.7	31.3	21.4
1986	Jan	187.9	29.8	21.0
	Feb	200.6	29.6	20.3
	Mar	379.5	29.5	20.3
	Apr	283.2	31.1	21.1
	May	384.7	30.3	20.8
	Jun	245.8	28.4	20.1
	Means	262.7 (excl 1.84)	29.7	20.8

Notes:

1. The San Carlos site is at 0°0'S, 76°50'W, altitude 265 m;
2. Data collected by LCTAP and INIAP.

Table 2.4 *Meteorological data from INIAP Napo-Payamino Experimental Station, Payamino site, 1983-1986.*

Year	Month	Rainfall	Temperature Daily max	Temperature Daily min	Humidity Daily min	Sunshine
		mm	°C	°C	%	hour/day
1983	Sep	171.4	29.4	19.8	58.8	4.6
	Oct	205.1	28.2	19.7	65.9	3.9
	Nov	426.5	29.0	20.2	64.2	4.4
	Dec	318.4	28.5	19.9	66.0	3.7
1984	Jan	218.0	29.1	19.8	61.9	4.5
	Feb	436.9	28.0	19.8	67.8	2.9
	Mar	402.2	28.6	25.5	67.0	3.2
	Apr	463.8	27.8	20.5	69.9	2.9
	May	506.8	28.2	20.2	65.6	4.8
	Jun	387.3	27.2	20.6	70.5	4.0
	Jul	224.5	27.2	19.7	62.3	4.2
	Aug	186.3	28.2	20.5	60.8	4.4
	Sep	355.3	31.6	23.0	59.2	5.3
	Oct	320.1	32.2	23.5	60.0	5.2
	Nov	326.2	31.0	23.3	64.8	4.6
	Dec	154.4	32.4	24.0	62.0	4.5
1985	Jan	32.2	34.0	23.5	50.0	6.2
	Feb	106.9	30.5	21.1	54.8	3.9
	Mar	165.4	30.5	21.9	56.1	3.6
	Apr	171.6	30.2	21.8	58.4	4.2
	May	117.2	29.6	20.8	58.7	4.0
	Jun	443.3	26.4	20.2	64.7	3.0
	Jul	363.8	28.6	19.1	66.4	3.6
	Aug	312.7	29.7	19.6	65.3	4.3
	Sep	276.6	31.5	20.1	60.2	4.5
	Oct	277.0	31.4	20.8	58.5	5.4
	Nov	244.7	31.3	21.2	59.1	6.3
	Dec	127.5	31.6	21.4	55.0	5.5
1986	Jan	253.4	30.5	21.4	62.2	3.8
	Feb	391.7	30.2	21.5	68.9	3.4
	Mar	354.9	29.8	21.1		3.1
	Apr	267.5	31.2	21.6		4.6
	May	384.1	30.3	21.4		4.3
	Jun	243.6	28.7	21.0		2.7
	Means	283.5	29.8	21.2	62.2	4.2

Notes:

1. The Payamino site is located at 0°27'S, 76°59'W, altitude 250 m.
2. The data, other than rainfall data, for the following months are incomplete: July 1984 (22 days only); August 1984 (5 days only).

The characteristics of the regional climate are well demonstrated by these data: for example, the minimal seasonal variation in temperature, and the marked year to year variation in the rainfall in any given month. The San Carlos figures for the period December 1984 to March 1985 are unusual; these four months were drier than any other four-month period during the nineteen years covered by the Limoncocha data (Table 2.2). During February and March 1985, the forest floor vegetation, which normally exists in a permanently moisture-saturated environment, was showing signs of moisture stress. Whitmore (1984, Section 4.5) has drawn attention to the importance of such 'rare climatic events' to rain-forest vegetation. Monthly rainfall fluctuations are often due to intense local storms. In this case, the dry period appears from the data to have ended in April 1985, when 313 mm of rain were recorded. However, about half this rain (157 mm) fell in a single four-hour storm on the night of 3-4 April, and the dry weather then resumed and continued through until the beginning of June. Local people attributed this unprecedented dry period to the effects of deforestation in the area since the oil company roads opened the area to colonization in the early 1970s. However, the Limoncocha data show that rainfall in the 1970s (mean annual rainfall 1970-1979: 3230 mm) was consistently higher than in the 1960s (mean annual rainfall 1961-1969: 2898 mm). This difference is statistically significant ($p < .01$).

Vegetation

The natural vegetation of the whole region is forest. There is little published information on the species composition of this forest, in Ecuador, or on vegetation types in relation to terrain and soil types, although important differences are evident. Some observations on structure and floristics for five sites were published by Grubb *et al* (1963). From my own observation, tree species diversity is marked; the most important families both in terms of number of species and number of individuals are the Moraceae and the Palmae. Stands dominated by a single tree species occur naturally only in permanently flooded areas, which are dominated by the palm *Mauritia flexuosa*, and on recent deposits along rivers, which carry dense stands of *Ochroma pyramidale* (balsa) or *Cecropia* species. Spontaneous successional forest occupies a significant proportion of the total land area in the floodplain zone of typical Upper Amazon rivers; a detailed description is provided by Salo *et al* (1986).

Quaternary history

In terms of the hypothesis of Pleistocene refuges in Amazonia, the whole of the Amazon Region of Ecuador has been identified as falling within one of the larger refuges, both on the grounds of present species distributions for several groups of plants and animals, and also present climate, with high rainfall distributed evenly through the year (Prance 1982). This hypothesis is usually taken to assert that a refuge area would have had a fairly stable climate through the last ice age and up to the present day. Detailed studies at several sites within Ecuador have provided evidence of significant climatic fluctuations within this period. Frost (1984) and Colinvaux *et al* (1985) studied sediments in four lakes in Napo Province, and found evidence of a major flooding episode between 1300 and 800 BP. P.A. Colinvaux (personal communication) believes that this would have inundated much of the area below 300 m altitude. Pollen spectra from a site near Puyo, dated at 26,000-33,000 BP, indicate that montane Andean forest may have extended into this region at the time of the last glaciation, implying mean annual temperatures in the Amazon lowlands at least 4.5° lower than at present (Liu and Colinvaux 1985). Such temperatures would have caused major changes in forest species composition. At Kumpak^{a} lake near Santiago (about 500 m altitude) pollen spectra showed that largely intact rain forest has persisted for at least 5,200 years (P.A. Colinvaux, personal communication), though with evidence of several shifts in species composition within this period.

Indigenous settlement

Over the greater part of the region, the majority of the inhabitants are indigenous Indians. In the north and central Oriente, these are mostly Quichua-speaking groups, with a small number of Cofanes, Secoyas, Sionas and Huaoranis (Aucas). In the southern Oriente, there is a large and flourishing population of Shuar (Jivaro). Along the navigable rivers in the north and central Oriente, the sites most favoured for present-day settlement are the non-flooding strips of older deposits along river banks.

However, in the recent past (up to as little as twenty years ago), most settlement and cultivation may have been located in the upland red soil areas away from the rivers (D. Irvine, personal communication). Archaeological evidence suggests that much of the Upper Amazon was more heavily populated before the Spanish conquest (Lathrap 1970), but there is little evidence on which one might reconstruct the pre-Columbian settlement pattern of the Ecuadorian Amazon. Recently, there has been a tendency for groups of families to live closer together in *comunas* (strictly speaking, *comuna* is an Ecuadorian legal term indicating a form of group land-holding).

The traditional form of indigenous agriculture, still widely practised today, involves the clearing of a patch of forest, typically around 1 ha, which is used to grow one or two crops of manioc (*Manihot esculenta*), maize (*Zea mays*) or, nowadays, bananas (*Musa* spp). This *chacra* (which is sometimes referred to in English as a 'garden') is then left fallow for perhaps 10 years before being cleared and planted again. Although the fallow *chacras* often appear to be given over to uncontrolled secondary forest, recent research (D. Irvine, personal communication) has shown that the species composition of these areas is carefully managed, and useful species may be protected and regularly harvested throughout the fallow period. Hunting and fishing provide a significant proportion of the regular diet. In the past twenty years, some *comunas* have established small coffee and cocoa plantations, and permanent pastures for beef production; these products are used to generate cash income and are hardly ever consumed within the *comunas.*

An aspect of indigenous culture specially relevant to this thesis is the planting of species other than the major food crops mentioned above: to give a few examples of the commonest species, *Inga* spp and *zapote* (*Quararibea cordata*), the fruits of which have a sweet edible pulp around the seeds; cotton (*Gossypium* spp); *achiote* (*Bixa orellana*), the source of an orange-red dye; chili peppers (*Capsicum* spp); the *chontaduro* palm (*Guilielma gasipaes*), of which the edible fruits are a major food item for a short period of the year. In the present context, it is helpful to divide these species into two categories. First, there are plants which are found only at settlement sites and may be presumed always to have been deliberately planted. Second, there are species which are found wild in the forest, while also occuring as cultivated or protected individuals near houses or in *chacras*; fruits of many forest species are harvested for food or other purposes, and the seeds then planted, so blurring the conventional distinction between wild and domesticated species.

European settlement

Parts of the Ecuadorian Amazon came under the nominal control of the Inca empire (Oberem 1980), but with no obvious influence on agricultural practices or settlement patterns. Furthermore, the Incas and other Andean civilizations, unlike their counterparts in Central America, apparently did not cultivate, trade in or use cocoa in any form (Mangelsdorf and Ree^{ve}s 1939). European and *mestizo* colonization of the region began immediately after the Spanish conquest of the Inca empire, with the founding, before 1600, of Baeza, Archidona, Avila and Macas. Although direct evidence is lacking, it is probable that cocoa was cultivated at these settlements as it was throughout the Spanish colonies (Barros 1981). However, the isolation of these settlements would probably have rendered unprofitable any large-scale planting and exportation of cocoa.

Documentary evidence relating to European settlement in the region in succeeding centuries is very sparse. Place names on maps, discussions with local people, and occasional references (for example, Schultes 1978) suggest that missions, farms and trading posts were established along the navigable rivers in the 18th, 19th and early 20th centuries. These settlements had easy communications downstream to Peru and Brazil, but travel from the Ecuadorian *sierra* was very arduous for Europeans. De la Condamine (1778) provides a graphic account of one such journey, on which nine people set out from Riobamba and lost their way in the forest along the Río Bobonaza; one by one, the travellers grew weak from hunger and died, except for a Frenchwoman called Mme Godin who wandered in the forest for nine days, eventually finding her way back to the river, where she was rescued by Indians. Mme Godin continued her journey down the Amazon and eventually rejoined her husband - the object of her journey - in Cayenne.

Some of these settlements were established during the Amazonian rubber boom of the late 19th and early 20th centuries (a recent account of this period is provided by Dean, 1987). Wild rubber trees (*Hevea brasiliensis* and other *Hevea* species) are now rare in Ecuador, possibly because of damage to trees caused by latex collection. By the middle of the 20th century, much of the land along the Río Napo was occupied by a series of farms, with some small plantings of cocoa. The indigenous population provided labour under conditions of near-slavery. To some extent, indigenous settlement patterns in the 20th century are probably the result of moving to less accessible areas to escape the demands of landowners and rubber traders. The power of the landowners on the Río Napo declined in the 1950s and 1960s following a strengthening of government control and the establishment of new Catholic missions at Coca and Nuevo Rocafuerte. By 1980, most land along the Río Napo was held by indigenous *comunas*, although many obstacles have been placed in the way of full legal ratification of these holdings.

Recent developments

The most recent wave of new settlement in the Ecuadorian Amazon has depended on the construction of roads into the region. The first of these was built by the Shell oil company to Puyo, in the 1940s, to facilitate the first oil exploration programme in the region. Shell also built airstrips, near Puyo and at Taisha and Villano. No commercial oilfield resulted from these explorations. Two more roads, to Zamora and Macas, were built later to encourage settlement in the southern Oriente. The first road into the lowland part of the region was built by the Texaco and Gulf oil companies in the early 1970s, to provide access to the oil fields in the Lago Agrio-Coca area and to service the trans-Andean pipeline, which runs alongside the road. The whole road network is continually being extended, both with new main roads (for example, from Baeza to Coca via Loreto, and from Puyo to Macas) and innumerable minor roads for which no up-to-date maps exist.

Colonization over more than 40 years along the roads in the Puyo-Tena area and in the Upano and Zamora valleys has resulted in nearly complete clearance of the forest, typically to 10 km or more away from the roads. Most of the land has been converted to pasture, usually of *gramalote* (*Axonopus* sp) for raising cattle. There are small plantings of many crops, including cocoa. The cocoa plantings are usually less than 1 ha and rarely more than 5 ha; management is usually restricted to weeding sufficiently to permit harvesting, and occasional pruning. Land holdings in these areas are irregular in size and shape.

Along the lowland part of the Quito-Lago Agrio-Coca road and its associated minor roads, colonization began in about 1970, and most individual colonists have been allocated 40-50 ha blocks of land. These farms, each with a road frontage of 200 or 250 m and 2 km deep, are arranged in a row along the main road, with ten or more rows of similar blocks beyond the first row. By 1986, the first two rows of farms (4 km) away from the road had typically been completely cleared of forest, and there was extensive clearance in the rows beyond these, particularly where side roads gave direct access. However, in the early 1980s it was still possible to find undisturbed forest close to roads.

The predominant uses of cleared land in the Lago Agrio-Coca area are for pasture (*Axonopus* sp (*gramalote), Pennisetum purpureum* (*elefante)* and *Brachiaria humidicola*) and coffee. Many colonists at first planted small areas (less than 1 ha) of cocoa, but these were badly affected by *Crinipellis* and *Moniliophthora* infections. Many cocoa plantings have now been abandoned or even cut down, and only occasional new plantings are found. Most cocoa planted by colonists in the last 40 years, in all parts of the Oriente, has been grown from seed brought from the *Costa* region; transport of seed (and diseases) has been facilitated by daily bus and truck services to the *Costa.* Seed from local wild trees has occasionally been planted. There are no reliable statistics on current cocoa planting and production in the region.

Other crops, in particular maize, bananas and manioc, are grown on a small scale by colonists, and two large oil palm plantations, totalling some 10,000 ha, are now in production. Plans exist for further large plantings of oil palm in Napo Province, but these have aroused strong opposition from Indians and colonists who fear the loss of existing or claimed land holdings.

In March 1987, a major earthquake and volcanic eruption destroyed some 50 km of the main Quito-Lago Agrio road and the trans-Andean pipeline. It is assumed that this will have affected the pace of development in Napo Province.

2.2 Río Caquetá, Colombia

Some collecting was also done in an area some 400 km east of the main target region in Ecuador, along the Río Caquetá and its tributaries in Colombia (Figures 2.1, 4.4). The section of the Caquetá which was explored, between Araracuara and La Pedrera, is a fast-flowing white-water river up to 2 km wide, with many large islands. There are major rapids at Araracuara, which block navigation upstream, and at the *Chorro de Cordoba* just above La Pedrera, which blocks the passage of larger boats downstream to Brazil. Away from the river, the terrain is characterized by an undulating landscape of eroded Tertiary sediments (altitude 100-200 m), with an isolated hill at La Pedrera rising to about 400 m, marking the frontier between the Upper Amazon landforms described here and the older, more resistant rocks of the Guyana Shield to the north and east.

Along the Caquetá itself, there is marginal strip of relatively dry ground typically 10-50 m wide, beyond which are found extensive areas of seasonally flooded depressions. Along its tributaries, many of which are major rivers in their own right, there is frequently no strip of dry ground to mark the banks, and the rivers are edged by flooded forest which may extend for hundreds of metres away from the main channel (a type of landscape which I did not encounter in Ecuador, except along the Río Yasuní). The flooded forest along the tributaries is often dominated by palms (typically, *Astrocaryum* and *Euterpe* spp); otherwise, the vegetation types encountered were broadly similar to those seen in Ecuador, though with significant differences in species composition. The tributaries of the Caquetá in this area are clear-water rivers, and it appears that the fertility of soils along the margins of these rivers is lower than along the Caquetá itself, which deposits nutrient-rich sediments from the *sierra.*

This area is populated almost entirely by indigenous groups: the Andoque, Yacuna-Matapi, Bora, Miraña, Nonuya and Macuna, notable for their system of linguistic exogamy. The Caquetá was also affected by the rubber boom, and the indigenous population along the main river appears to have been greatly reduced during the past hundred years by this or by other factors. To escape from these pressures, some Indians moved up to the headwaters of tributaries and even today have virtually no contact with outsiders. There is at present no government policy to promote colonization in this area, and there are no access roads from the rest of Colombia, so that extensive forest clearance is not occurring.

2.3 San Carlos

The field genebank occupies about 4 ha of land at the San Carlos site of the INIAP Napo-Payamino Experimental Station. San Carlos is located between Lago Agrio and Coca, Napo Province (0°20'S, 76°50'W) at 265 m altitude. The site was originally covered by mature forest, apparently undisturbed although frequent pottery fragments in the soil suggest that the forest was cleared at some time in the past, possibly several hundred years ago. The original forest to the south and east of the genebank remained undisturbed except for the removal of a few trees for timber. Most of the station consists of a level former floodplain, intersected by small streams which cut valleys 5-10 m below the general land level.

The results of a soil analysis for the San Carlos genebank site are given in Table 2.5, with some typical values for cocoa-growing soils for comparison. These results show that the soil at San Carlos is somewhat more acidic, and has a lower exchangeable calcium content, than most cocoa-growing soils. Nitrogen, magnesium and potassium levels are below average, but phosphorus and organic matter are high.

Table 2.5 *Soil analysis data for San Carlos genebank site, compared with typical cocoa-growing soils.*

			Organic	Total	Total		Exchangeable				
Site		pH	Matter	N	P	Al	Ca	Mg	K	Ca/Mg	(Ca+Mg)/K
			(%)	(%)	(ug/g)		(meq/100g)			(Ratios)	
San Carlos (0-10 cm)		4.2	9.4	0.30	702	1.31	5.19	1.60	0.48	3.24	14.1
Typical	Average	7.1	5.6	0.57	540		20.99	4.33	0.51	5.49	63.2
Cocoa-Growing	Maximum	7.7	15.4	1.60	550		36.80	10.20	0.89	13.09	110.3
Soils	Minimum	6.0	1.7	0.14	530		7.50	1.66	0.09	3.60	10.4

Notes:

1. The results for San Carlos are based on duplicate analyses of five samples taken from the genebank site and the adjacent forest. The analysis was done by A.Gray, Department of Forestry and Natural Resources, Edinburgh University.
2. Comparative data for typical cocoa growing soils were taken from Wood (1975), Table 4.1. They include soils from Ghana, Nigeria, Ivory Coast, Trinidad, Sabah and Western Samoa, sampled at depths between 0-4 cm and 0-20cm. The total phosphorus figures are from only two samples.

Part of the genebank site is shaded by trees remaining from the original forest, while the rest is shaded by planted *Inga edulis* at 9 m by 8 m spacing. The shade level is now around 20-40% of full sun throughout the site. Given the high rainfall throughout the year, and high average cloud cover, additional shade was not provided for newly-planted cocoa.

The plan of the genebank is indicated in Figure 2.3. Seedlings and budded plants were placed in separate blocks. For each collection of seed, a maximum of 10 seedlings were retained and planted in a single row. Where less than 10 seedlings were available, the gaps were left unplanted, and no attempt was made to fill these gaps by further propagation unless it appeared that there was a risk of losing all the seedlings through damage or disease. Budwood collections were planted as a row of up to 5 plants, and extra budded plants were propagated to fill gaps in these rows wherever possible. All trees were spaced at 4 m (between rows) by 3 m (within rows).

The collections were planted out in groups once or twice a year from January 1981 onwards.

Notes for Figure 2.3 (next page)

1. Updated to 31.7.86 (but excludes new Blocks 9 and 10).
2. All collections are LCT-EEN unless otherwise identified.
3. Other collection prefixes:
   - EBC = Expedicíon Botánica Caquetá
   - NAPO = NAPO (Chalmers)
   - AGU = AGUARICO (Chalmers)
4. Grouped collections (tree position in row in brackets):
   - [A] = 25 (2), 26 (3-6), 27 (7-10)
   - [B] = 28 (1), 29 (2-4)
   - [C] = 2 (1), 32 (3), 38 (5)
   - [D] = 103 (1), 146 (2), 137 (3)
   - [E] = 90, 120, 121, 124, 126
   - [F] = 145 (1), 158 (2)
   - [G] = NAPO 24 (1), AGUARICO 31 (2)

```
                           BLOCK 3A    BLOCK 3B               BLOCK 5             BLOCK 7A    BLOCK 7B
                               (CLONES)                   (SEEDLINGS) 1                (CLONES)
                           12345    54321              1234567890          12345    54321
    BLOCK 1                *****    *****              **********          *****    *****
(SEEDLINGS) 1              *****    *****              **********          *****    *****
       1234567890       30 ooooo    ooooo 373      348 oooooooooo      374 o----    ooooo 2
       _*********       31 ooooo    ooooo 371      347 oo-ooooooo      376 oo---    ooooo NAP024
       **********       33 ooooo    -oooo 370      346 oooooooooo      378 oo---    ooooo AGU31
     1 oooooooooo       36 ooooo    ooooo 369      345 ooooooo-o-      379 o----    -oooo 32
     4 oooooooooo       37 ooooo    -oooo 368      344 oooooooooo      382 oo---    ooo-o 38
     5 oooooooooo       46 ooooo    ooooo 367      343 ooo-oooooo   EBC114 o----    -oooo 97
     6 oooooooooo       49 ooooo    --ooo 366      342 ooooo-o-oo          -----    ooooo 103
     8 oo-ooooooo       52 ooooo    ooooo 364      341 oooooooooo   EBC118 o----    ooooo 120
    10 oooooooooo       57 ooooo    --ooo 362      340 oooooooooo   EBC122 o----    ooooo 126
    11 oooooooooo       60 ooooo    -oooo 142      339 oooooo-ooo   EBC135 -oooo    ooooo 128
    14 ooo-ooooo-       63 -oooo    --ooo 141      338 oooooooooo   EBC138 oo---    ooooo 137
    15 oooooooo--       65 ooooo    ooooo 133      337 oo-ooooooo   EBC141 o----    --ooo 145
    18 oooooooo--       68 ooooo    ooooo 333      336 ooooooooo-   EBC145 o----    ooooo 158
    19 oooooooo--       73 ooooo    ooooo 327      335 oooooooooo          -----    ooooo 391
    20 oooooooo-o       74 ooooo    --ooo 124      334 oooooo----        oooooooooo EBC148
    21 oooooooooo       80 ooooo    --ooo 297      332 oooooooooo        -------ooo EBC147
    22 oooooooooo       81 ooooo    ooooo 280      331 -ooo------        ---------o EBC142
    23 oooooooooo       91 ooooo    -oooo 121          ----------        oooooooooo EBC138
    24 oooooo-o-o      [C] o-o-o    ooooo 193      329 o---oo-o--        ---------o EBC135
   [A] -oo-oooooo       47 o----    ooooo 94       328 oooo-ooooo        --------oo EBC126
   [B] oooo------      [D] ooo--    ooooo 312      313 ooooooooo-        oooooooooo EBC125
    44 o-o-o-oooo      107 ooooo    ooooo 278      302 oooooooooo        --------oo EBC123
    48 ooooooooo-      108 ooooo    ooooo 195      300 oooooooooo        ----oooooo EBC121
    58 oooooooooo      122 ooooo    -oooo 189      326 oooooo--oo        --oooooooo EBC101
    59 oooooooooo      123 ooooo    --ooo 166      325 ooooo-o-oo        -----ooooo EBC100

       BLOCK 2             BLOCK 4A    BLOCK 4B                BLOCK 6        BLOCK 8
     (SEEDLINGS)               (CLONES)                    (SEEDLINGS)    (SEEDLINGS)
    61 o---oooooo      125 ooooo    -oooo 300      324 oooooooooo  -----ooooo EBC91
    62 oooooooooo      130 ooo--    ooooo 281      323 oooooooooo  oooooooooo EBC86
    64 oo--oo----      144 ooo--    oo-oo 269      321 oooooooooo  o-oooooooo EBC85
    65 ooooo-oooo      150 oo---    ooooo 267      320 oooooooooo  ------oooo EBC84
    66 oo-ooooooo      151 ooooo    ooooo 264      306 ooooooooo-  ------oooo EBC79
    67 oooooooooo      152 ooooo    -oooo 259      296 oooooooooo  -----ooooo EBC78
    68 oooooooooo      153 -oooo    ooooo 258      295 ooo-oooooo  oooooooooo EBC77
    69 -ooo------      154 -o-oo    ooooo 257      288 oooooooooo  -ooooooooo EBC76
    70 ooooooo-oo      155 ooooo    ooooo 255      134 ooooo-----  o-oooooooo EBC58
    71 ooooooo-oo      156 ooooo    ooooo 254      271 oo-ooo----  oooooooooo 393
    72 oooooooooo      157 ooooo    -oooo 253      248 -ooooooooo  ----------
    74 oo-ooooo--      [E] ooooo    -oooo 251      246 oooooooooo  --------oo EBC48
    76 oooo-oo---      [F] oo---    ooooo 250      244 ooo--ooooo  ------oooo EBC30
    77 ooooooo--o   NAP027 -oooo    ooooo 249      243 -o--------  -ooooooooo EBC29
    82 oooooooooo   NAP029 oooo-    ooooo 241      242 -oooooo-oo  oooooooooo EBC10
    79 oooooooooo   NAP051 ooooo    ooooo 238      241 ooo-------  oooooooooo EBC9
    78 oo-oooo-oo      [6] oo---    -----          212 oooooooooo  ooo-oooooo EBC6
    83 oooooooooo      282 ooooo    ooooo 283      215 ooooo-----  oooooooooo EBC5
    84 ooo-oooooo      146 o-ooo    ooooo 237      214 -oooo-----  o-o---oooo 372
    85 oooooooooo      148 o-ooo    -oooo 236      205 o---oooo-o  ooooo---oo 361
    86 oooooooooo      149 ooo--    ooooo 234      203 ooo-------  ooooo-oooo 360
    87 oooooooooo      162 ooo-o    ooooo 232      202 ooooo-oooo  ooo-ooo--o 359
    88 oooooooooo      163 ooooo    ooooo 231      201 o-o-oo-ooo  ooooooo-oo 358
    89 oooooooooo      164 ooooo    ooooo 228      199 oooooooooo  oooooo-ooo 357
    90 oooooooooo      165 ooo--    ---oo 227      196 o-oooooo--  ooo-oooooo 356
    92 oooooooooo      166 ooo--    ooooo 223      182 oooooooooo  ooooooo-o- 354
    93 oooooooooo      169 ooooo    ooooo 222      162 oooooooo-o  ---------o 355
   101 oooooooooo      172 ooooo    ooooo 221      136 oooooooooo  ----ooooo- 353
   109 oooooooooo      175 oooo-    oooo- 220      132 oooooooooo  o-oooooooo 352
   113 ooo-------      178 oooo-    ooooo 219      160 o-oooooooo  oo--oooooo 351
   115 oooooooo-o      179 ooooo    ooooo 218      159 --ooooooo-  -o---ooooo 350
   127 oooooooooo      181 ooooo    ooooo 217      145 o-oooooooo  ooo-oooooo 349
   128 -o--------      185 ooooo    o-ooo 216      142 ooo--oo---  **********
   130 oooooooooo      187 ooo--    -----          131 o--o-oo-oo  **********
       **********          *****    *****              **********  ******____
       **********          *****    *****              **********  ******____
       1234567891          12345    54321              1234567891  1987654321
                0                                               0  0
```

Figure 2.3 *Plan of the field genebank at San Carlos (see notes on previous page).*

3 Geographical variation: methods

This chapter describes the techniques that were used to acquire and analyze data relating to geographical variation in the wild cocoa populations being studied. Since this work was done within a germplasm collecting project, it is necessary to outline here the methods used in collecting, since the collecting process determined the sample of trees from which descriptor data were obtained.

3.1 Collecting

A full description of collecting methods was given in Allen and Lass (1983). These methods may be summarized as follows:

(a) material was collected in the course of short trips, lasting from two days up to three weeks;

(b) pods were collected wherever possible, and budwood was collected if no pods were available;

(c) seeds were stored in unopened pods, or in damp charcoal; budwood was stored in damp newspaper, after sealing the ends of the budsticks with wax;

(d) generally, collected material was taken to San Carlos for propagation and establishment in the field genebank. Some of the Río Caquetá accessions were propagated only in Colombia;

(e) seed was sown in plastic bags in the nursery, with germination rates approaching 100%;

(f) budwood was used either for patch budding of individual buds, or for side-grafting of entire budsticks (Soderholm and Shaw 1965). The latter technique was preferable;

Accessions were identified by the prefixes LCT-EEN (*L*ondon *C*ocoa *T*rade - *E*stacíon *E*xperimental *N*apo), for the main LCTAP collection, and EBC (*E*xpedicíon *B*otánica *C*aquetá) for the Río Caquetá material. Individual seedlings derived from pod collections were given the main accession number with a suffix /s1, /s2, /s3, etc.

In general, no attempt was made to select trees for collection on the basis of characteristics such as freedom from disease symptoms. This policy, which marked a significant change from most earlier collecting expeditions, was adopted after lengthy discussion with various people involved in cocoa collecting and breeding. There was general agreement that the apparent performance of a cocoa tree in the wild is a poor predictor of performance in cultivation. At least two factors explain this general rule. First, the plantation environment is quite different from the forest environment, both in terms of microclimate and in terms of the proximity of other individuals of the species. Second, there is genetic variation among fungal pathogens of cocoa, and pathogen races in cocoa-growing regions may be distinct from those to which a wild population is exposed. Both factors apply to *Crinipellis* infection, where the expression of symptoms is enhanced by increased light levels, and the existence of pathogenic races has been demonstrated by Wheeler and Mepsted (1982).

So far as was practicable, therefore, trees from which material was collected were intended to be a random sample of the population. Some constraints were imposed by the practical logistics of collecting. The maximum number of trees which could be collected in a day varied between 8 and 15, depending on the time spent travelling and the type of material collected, budwood taking longer to collect than pods (measurements on pods and seeds were very time-consuming, but were normally done after returning to San Carlos). In areas where wild cocoa was scarce, it was feasible to collect material from every tree located. If it seemed likely that more than 8-10 trees would be encountered in the course of a particular day, material would be taken from only a proportion of trees, spaced out along the day's route (for example, by stopping every 20 minutes along a river). When a group of trees was located, material was normally collected from two or three trees. If pods were available, they were always collected (because of the much higher success rate in propagation); otherwise, the trees which seemed to have the best budwood were collected.

3.2 Descriptors

The descriptors used follow general IBPGR recommendations for constructing descriptor lists (Hobbs 1982) as well as specific recommendations for cocoa from IBPGR (IBPGR 1981) and from other workers active in this field (Engels *et al* 1980), and the catalogue of the Turrialba cocoa genebank (Engels 1981). Where these recommendations conflict, I have followed the Turrialba catalogue, on the grounds that published descriptions of the LCTAP accessions should be compatible with the only other major genebank catalogue for cocoa to have been published to date. Some additional descriptors have been devised specifically for the LCTAP collections.

The descriptor data were grouped into six data sets. Data set 1 consists of the *passport data,* basically identification codes, species and location data, with some information on the status of the original tree. Data set 2 consists of the main *pod and seed* descriptors. Data set 3 contains three descriptors which, taken together, describe the *habit* of trees grown from seed; these descriptors were devised for this project and have not so far been applied in other genebanks. Data set 4 consists of a single descriptor referring to *flush colour,* and data set 5 consists of *flower* descriptors. Data set 6 comprises two *production descriptors* derived from Data set 2. The descriptor data were treated in this way for two reasons: first, because some data sets only existed for part of the collection and second, because the data sets could not all be analyzed in the same way.

The complete descriptor list is given in Appendix 1. The methodology of recording follows Engels (1981), with the following differences:

(a) pod and seed data were recorded from a sample of 1-5 pods per collection, rather than the recommended sample of 10-20 pods. Many of these data were obtained on collecting trips, at the time of collection, when few wild trees yielded more than two or three healthy pods. For collections made as budwood, it was necessary to wait until pods were produced at San Carlos, but time did not permit measurements to be made on large samples of pods. Nonetheless, many descriptors showed little variation within a collection, so that the data do adequately describe the characteristics of the accessions. This is particularly true of qualitative descriptors such as pod surface texture and pod apex form, and of ratios between measurements, such as the ratio of pod diameter to pod length.

(b) Turrialba used separate three pod pigment descriptors referring to anthocyanin intensity in: (1) unripe pod ridges; (2) unripe pod furrows; (3) ripe pod ridges and furrows. I used a single descriptor for general anthocyanin intensity in unripe pods, most closely comparable with (1).

Particular care has to be exercised in the interpretation of the bean colour descriptors. For pods collected in the wild, these data reflect the gene frequencies of the wild populations; these are of much more interest than the bean colour variation which occurs as a result of cross-pollination within the genebank. However, it can be deduced that a clone which produces only white beans in the genebank is homozygous for at least one recessive allele controlling bean pigmentation.

Descriptor data were obtained from the LCT-EEN and EBC series of accessions, both at the time of collection and in the San Carlos genebank. Some additional data were acquired from 52 clones in the cocoa genebank at the INIAP Pichilingue station. The Pichilingue data were from clonal material, mainly Pound's Peruvian expeditions in 1937-38 and 1942-43, and also from three clones representing the Ecuadorian *Nacional* variety (see Section 1.2). Table 3.1 shows which data sets were available for each group of accessions, and indicates which methods of data analysis were applied to each data set.

All descriptor data used in this thesis were collected by myself in the course of the project, with two exceptions. First, for comparison, I have used published means and coefficients of variation for the Turrialba genebank (Table 4.1). Second, to supplement the Pichilingue data, I took data on nine descriptors for 11 Pound clones also held at Turrialba from Engels (1981). The descriptors were: seed weight; seed number; seed length; seed width; seed thickness; presence of white, light purple, dark purple seeds; flush colour.

Table 3.1 *Descriptor data sets and analysis methods for each group of accessions.*

Data Set	Descriptors		ACCESSION SERIES			DATA ANALYSIS			
			LCT-EEN	EBC	Pichilingue	Descriptive statistics	Cluster analysis	Principal Components	Analysis of variance
1	Passport data								
		complete	all	all		-	-	-	-
		partial			all	-	-	-	-
2	Pod and seed								
		complete	part	part		yes	yes	yes	yes
		partial			all	yes	yes	yes	no
3	Habit		part	part	none	yes	yes	yes	no
4	Flush colour		most	part	part	yes	yes	no	no
5	Flower		part	few	most	yes	yes	yes	no
6	Production		part	part	none	yes	no	no	no

3.3 Analysis of descriptor data

Computing

All data handling and analysis was done using an IBM PC-compatible microcomputer running under PC-DOS. Basic data entry and manipulation was performed using dBase II and dBase III Plus (Ashton-Tate), while statistical and multivariate analyses were performed using NCSS Version 5.1 (NCSS, Kaysville, Utah).

Geographical subdivision

At several stages in the data analysis, it was convenient to divide accessions into geographically defined groups, as shown in Table 3.2. The larger geographic subdivisions (*Level 1* areas) covered areas of 30,000 km^2 or more, each corresponding to one or two entire provinces in Ecuador. The smaller geographic subdivisions (*Level 2* areas) covered 10,000-30,000 km^2, and were based on groups of collecting localities within the Level 1 areas, associated with geographical features such as rivers.

Table 3.2 Geographically defined sub-populations.

Level 1	Level 2
(1) Napo Province, Ecuador (and adjacent areas of Colombia)	(1) San Carlos (INIAP station only) (2) Río Putumayo and Río San Miguel (3) Río Napo (including Coca area)
(2) Pastaza Province, Ecuador	(4) Rios Curaray, Villano, Bobonaza, etc
(3) Morona-Santiago and Zamora-Chinchipe Provinces, Ecuador	(5) Río Upano (including Taisha) and Río Santiago (6) Río Zamora
(4) Río Caquetá region of Colombia	(7) Río Caquetá (upper part of explored section) (8) Río Caquetá (middle part of explored section) (9) Río Caquetá (lower part of explored section)
(5) Coastal Region, Ecuador (some Nacional accessions at Pichilingue)	(10) Coastal Region, Ecuador
(6) Perú (Pound clones at Pichilingue)	(11) MORONA series: Río Morona (12) NANAY series: Río Nanay, near Iquitos, Río Marañon (13) PARINARI series: Parinari, Río Marañon (14) SCAVINA series: original locality unknown (15) IMC series: near Iquitos, Río Marañon (16) POUND series: Río Nanay, near Iquitos, Río Marañon

Notes:
1. The designations of these areas were chosen as convenient ways of labelling collecting areas, and do not imply that collecting took place mainly along rivers.

Descriptive statistics

Means and coefficients of variation were calculated for all except binary descriptors, for each of the main series of accessions (LCT-EEN, EBC). Frequency distributions for most descriptors in the six Level 1 areas are presented as histograms.

Cluster analysis

An exploration of the relationship between geographical distribution and genetic differentiation was achieved by grouping the accessions into clusters based on the main pod and seed descriptors. The k-means clustering algorithm was used, minimizing the within-cluster sum of squares (Hand 1981). The data were standardized to zero mean and unit standard deviation within the clustering program. Accessions belonging to each cluster were plotted on geographical (latitude-longitude) coordinates. As there were typically several accessions at each location, the distribution of each cluster was plotted separately.

It was hypothesized that a correlation between geographical separation and genetic distance would produce a clear separation of the clusters when plotted on geographical coordinates. This is a conventional use of cluster analysis (Wishart 1978), which does not seem to have been applied previously in the context of germplasm collection strategies.

Principal components analysis

Principal components analysis was used, first, to make broad comparisons between areas and, second, to generate 'composite descriptors' for the analysis of variance (see below). All data were standardized to zero mean and unit standard deviation before running the analysis.

Analysis of variance and resource allocation

An analysis of variance for pod and seed descriptors (Data set 2, LCT-EEN and EBC accessions only) was used to explore the optimum allocation of resources to exploration of Level 1 and Level 2 areas. Initially, this analysis was performed separately for every descriptor, but for simplicity it was decided to use the first three axes from the principal components analysis as composite descriptors.

Using the approach of Bogyo *et al* (1980) (with calculations as outlined by Som (1973, Chapter 17)), it was possible to estimate the ratio of within-area to between-area components of variation, which, together with the 'cost ratio' (defined in Section 1.2) enabled calculation of the optimum number of samples per area.

4 Geographical variation: results

Chapter 4 presents the main results of the thesis in relation to the geographical pattern of variation in wild cocoa. The initial section (Section 4.1) provides information on the numbers of accessions available from the collecting programme. These accessions are described first in Section 4.2, which gives a summary of the main pod and seed descriptors, and then in greater detail in Section 4.3, with a breakdown by Level 1 areas for all descriptors. Two complementary approaches to summarizing geographical variation using multivariate analysis and mapping techniques are presented in Sections 4.4 and 4.5. Finally, Section 4.6 presents estimates of the partition of variation between Level 1 and Level 2 areas; these estimates will be used in Section 7.1 to explore the optimum allocation of resources in collecting.

4.1 Collected material available for study at San Carlos

The collecting trips were described in detail in Allen and Lass (1983) and Allen (1987). Collecting covered the whole of the Amazon Region of Ecuador as well as an area along the Río Caquetá in Colombia. The localities for all accessions are shown on the maps in Figures 4.1-4.4. The numbers of accessions collected and established at San Carlos are summarized in Table 4.1.

Table 4.1 *Numbers of accessions collected and established in the San Carlos genebank.*

Accession series	LCT-EEN	EBC
Total collected	419	136
Total in San Carlos genebank	281	34
of which: as clone	154	7
as seedling progenies	137	27

Notes:
1. A few accessions are held both as clones and as seedling progenies;
2. Not all the EBC collections were taken to San Carlos. Other accessions are held in the ICA genebank at Palmira, Colombia.

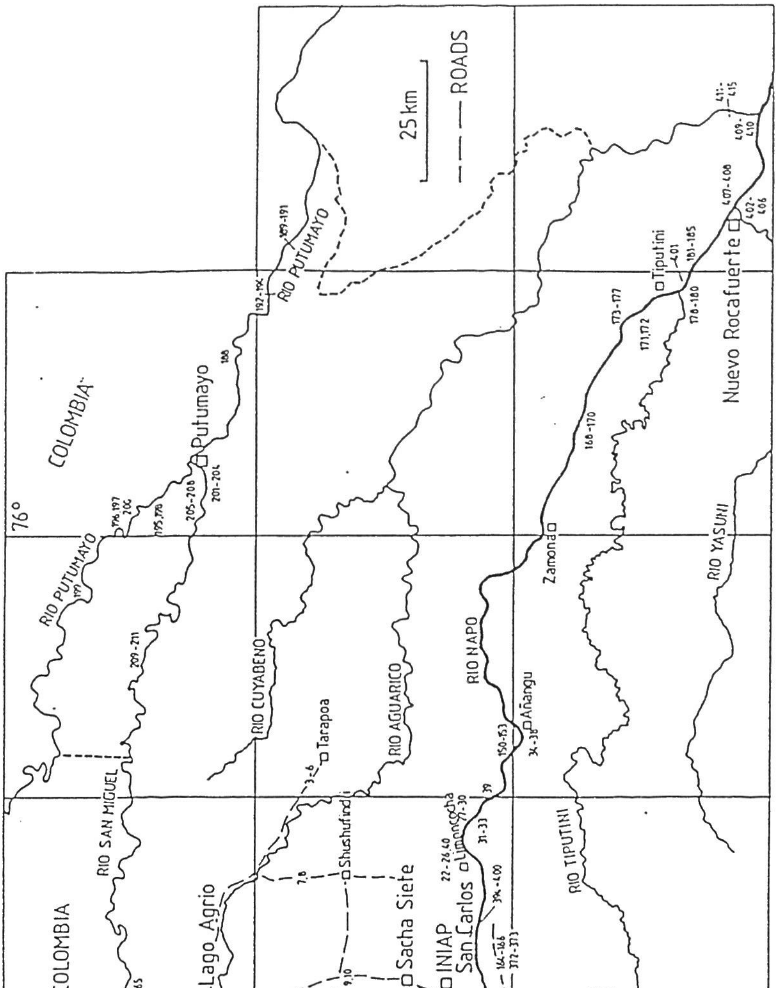


Figure 4.1 Map of Amazon Region of Ecuador, showing locations of LCT-EEN collections: northern part (area 'a' of Figure 3.1)

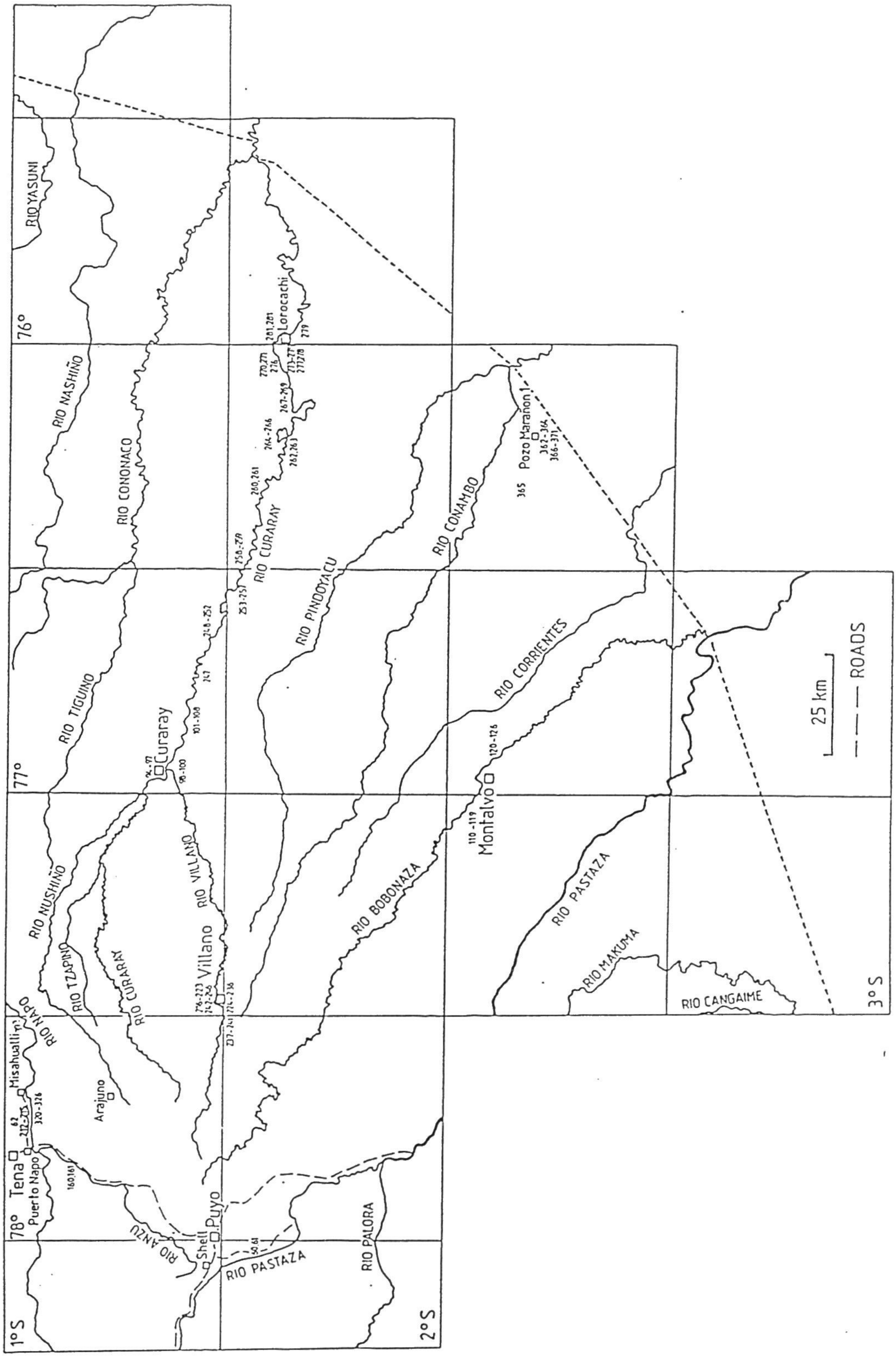


Figure 4.2 *Map of Amazon Region of Ecuador, showing locations of LCT-EEN collections: central part (area 'b' of Figure 3.1)*

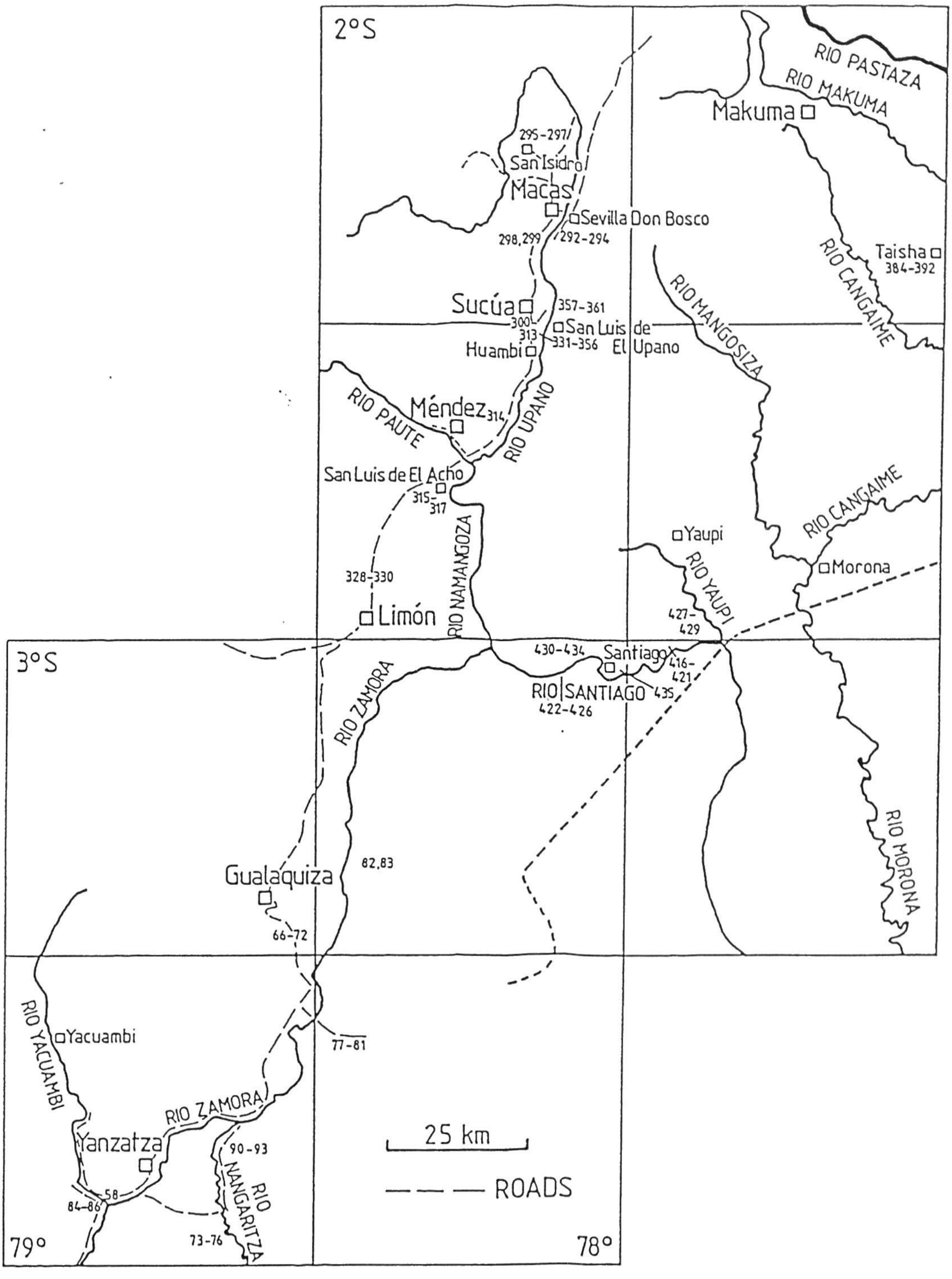


Figure 4.3 *Map of Amazon Region of Ecuador, showing locations of LCT-EEN collections: southern part (area 'c' of Figure 3.1)*

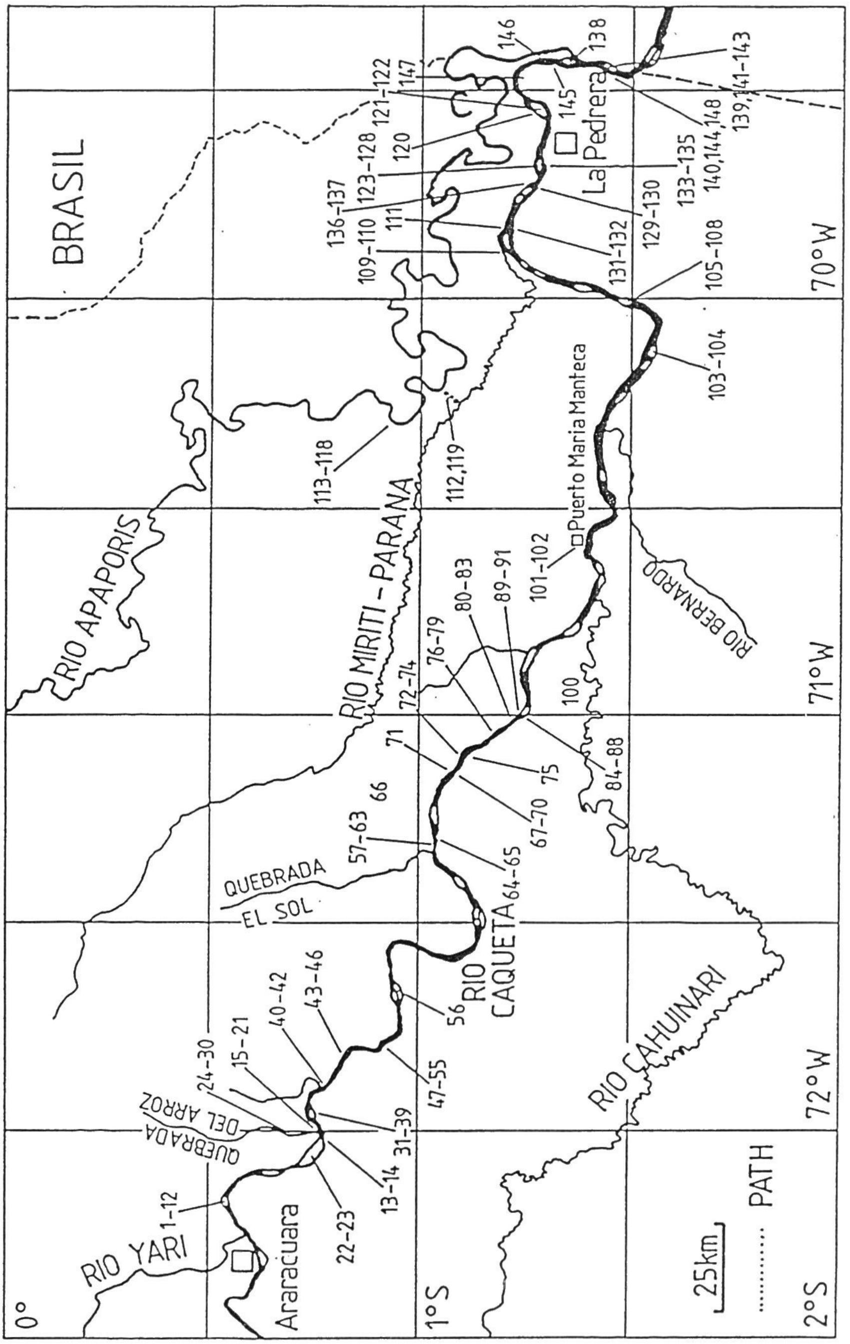


Figure 4.4 *Map of Río Caquetá region in Colombia, showing locations of EBC collections (area 'd' of Figure 3.1)*

4.2 Pod and seed descriptors: descriptive summary

The complete descriptor data are presented in Appendices 2-4. The main pod and seed descriptors (Data set 2) are summarized in Table 4.2, with published data from the germplasm collection at Turrialba for comparison.

The descriptors in Table 4.2 are listed with their mean values and coefficients of variation (CV), or frequencies, for three groups of collections: (1) LCT-EEN collections at San Carlos, (2) EBC collections at San Carlos and (3) all described collections at Turrialba.

The Turrialba genebank includes a fair cross-section of cultivated cocoa clones, with a few accessions of known wild origin (these are not clearly distinguished in the catalogue). In comparing the mean descriptor values for the LCT-EEN collections and those for the Turrialba clones, the similarities are more striking than the differences; pod size and weight, and seed size and weight, are nearly identical in the two groups. However, the LCT-EEN pods have higher mean values for apex form (more pointed pods), surface rugosity, primary furrow depth, and diameter:length ratio, although there is considerable overlap for all these characters. Considering the qualitative descriptors, pigmented pods occur in 35% of the Turrialba clones but only 5% of the LCT-EEN collections, mostly of cultivated origin. White seeds are very common (65%) in the LCT-EEN collections, but were recorded in only 17% of the Turrialba clones, although the Turrialba catalogue does point out that their seed colour percentages reflect cross-pollination within the genebank.

The EBC accessions should be compared with the LCT-EEN series. The EBC accessions evidently represent a distinct population or populations, with narrower, smoother pods and smaller seeds, which are never white, although light purple seeds are common. High CV values for several pod descriptors, notably pod wall thickness, result from the inclusion of a small group of rather distinct collections, having very large pods with unusually thick, hard walls.

Table 4.2 Pod and seed descriptors: summary.

DESCRIPTOR	LCT-EEN		EBC		Turrialba	
	mean	CV(%)	mean	CV(%)	mean	CV(%)
POD						
Basal constriction	1.3	138	2.8	61	1.8	100
Apex form	5.2	48	5.0	34	4.6	39
Surface rugosity	5.4	30	2.2	114	3.8	44
Furrow depth	5.2	19	3.6	25	3.9	41
Length (mm)	166	18	176	27	174	15
Diameter (mm)	93	15	76	20	86	8
Diameter:length ratio	0.56	16	0.44	16	0.51	16
Wall thickness(mm)	18	27	13	40	-	-
Weight (g)	574	44	475	74	570	8
SEED						
Fresh peeled weight (g)	2.02	28	1.44	41	1.95	28
Number	28	39	36	28	34	12
Length (mm)	24.7	11	23.1	13	24.4	9
Width (mm)	13.4	12	11.0	15	13.1	13
Thickness (mm)	9.3	15	8.4	12	9.0	13
Width:length ratio	0.54	8	0.47	9	0.54	8
Seed index	1.74	37	1.05	42	-	-
	f	%	f	%	f	%
Anthocyanin pigment in unripe pods (scale)						
0	188	95	33	100	192	65
>0	10	5	0	0	102	35
White seeds						
Present	131	66	0	0	49	17
Absent	67	34	30	100	245	83
Light purple seeds						
Present	74	37	25	83	186	63
Absent	124	63	5	17	108	37
Dark purple seeds						
Present	86	43	15	50	180	61
Absent	112	57	15	50	114	39

Notes:
1. Turrialba data are from Engels (1981), Appendix D.
2. Codes and symbols:

CV = coefficient of variation
f = frequency
"-" = no comparable descriptor

4.3 Geographical variation

The frequency distributions in Figures 4.5-4.10 provide a more complete picture of the geographical pattern of variation, including the Pound and *Nacional* accessions at Pichilingue. The *Nacional* data are very limited and are included for interest rather than detailed analysis. The geographical breakdown of the data corresponds to the Level 1 areas in Table 3.2. Several observations may be noted here:

(a) for most descriptors, the three areas in Ecuador (Napo, Pastaza, Zamora) were broadly similar;

(b) the Zamora accessions had a higher proportion of pods with moderate surface rugosity (5), compared to Napo, where very rough-surfaced pods (7) were most frequent. Deep furrows were also less frequent in Zamora than in Napo. Pod walls in the Zamora accessions were typically thinner than in the Napo group. Most accessions from Zamora had white seeds, whereas for the Napo group wholly purple-seeded pods were fairly common. Zamora flowers were typically smaller (shorter sepals) than Napo flowers;

(c) the Pastaza accessions appeared more similar to the Napo group than to the Zamora group;

(d) as mentioned in Section 4.2, there were bimodal distributions for several descriptors in the Caquetá accessions from Colombia: pod surface rugosity; pod length; pod wall thickness;pod weight; seed number; seedweight; seed length. These correspond to two distinct types of cocoa which were noted in the field. One type (small smooth thin-walled pods, small seeds) was found throughout the Caquetá area. This type was very different from the majority of accessions from the Napo, Pastaza or Zamora areas, although a few similar trees occurred in the Napo group, originating along the Río Putumayo. Curiously, they were actually collected on the Colombian side of the river. The other Caquetá type (large rough thick-walled pods, large seeds) was found in the eastern half of the area explored, and bore some resemblance to the typical Ecuador accessions;

(e) on most pod and seed descriptors, the Pound accessions were not clearly distinct from the Ecuador accessions. The pod diameter:length ratio was typically lower, and seeds were smaller and lighter. The Pound accessions were clearly distinct in terms of seed colour, flush colour and flower pigmentation (see Section 4.4);

(f) in general, the three Nacional accessions appeared slightly closer to the Pound material than to the Ecuadorian Amazon accessions.

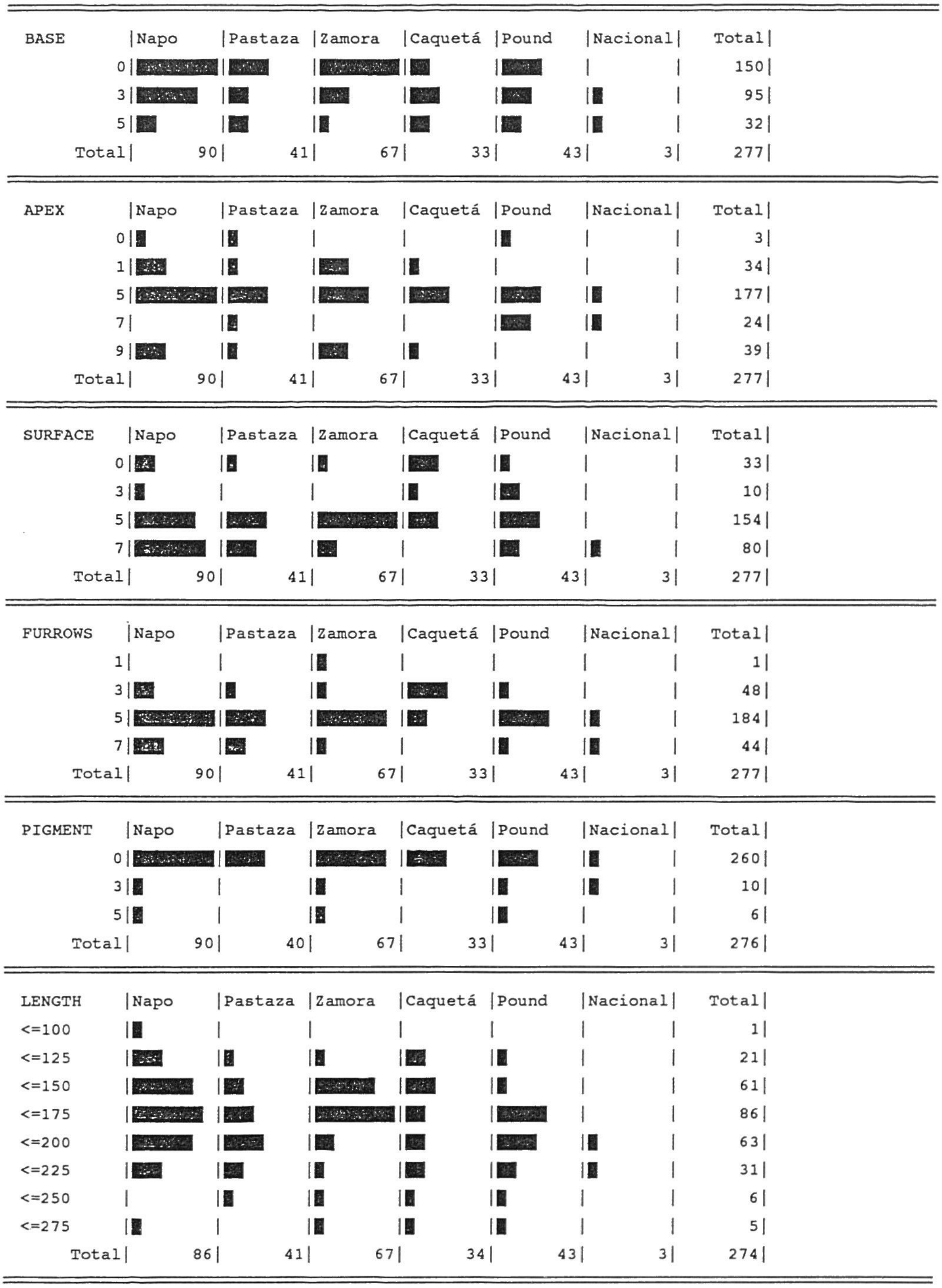
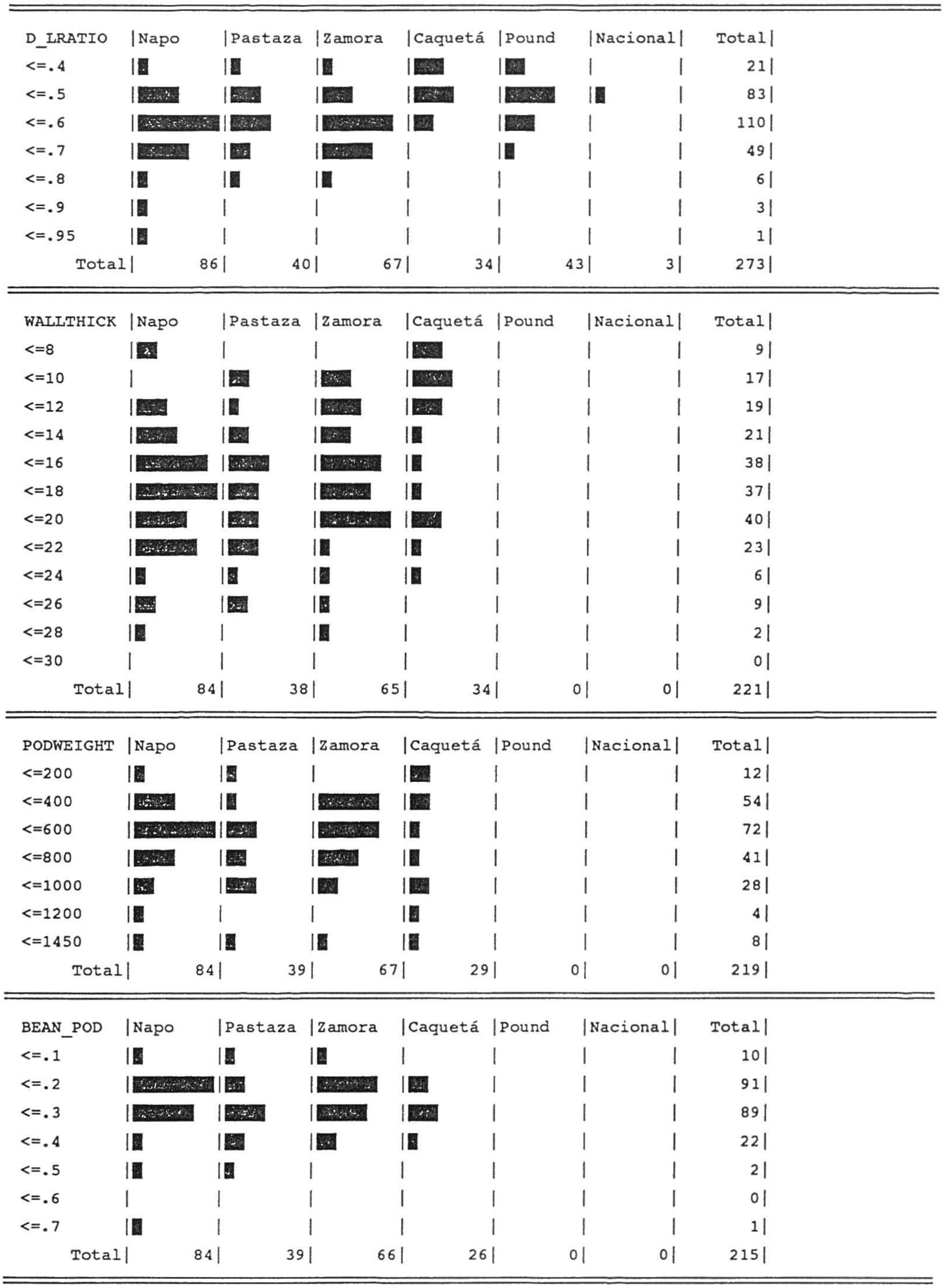


Figure 4.5 *Frequency distributions for pod descriptors (Data set 2) in six Level 1 areas (continues on next page). See Appendix 1 for details of descriptors.*



D_LRATIO	Napo	Pastaza	Zamora	Caquetá	Pound	Nacional	Total
<=.4	█	█	█	█	█		21
<=.5	█	█	█	█	█	█	83
<=.6	█	█	█	█	█		110
<=.7	█	█	█		█		49
<=.8	█	█	█				6
<=.9	█						3
<=.95	█						1
Total	86	40	67	34	43	3	273

WALLTHICK	Napo	Pastaza	Zamora	Caquetá	Pound	Nacional	Total
<=8	█			█			9
<=10		█	█	█			17
<=12	█	█	█	█			19
<=14	█	█	█	█			21
<=16	█	█	█	█			38
<=18	█	█	█	█			37
<=20	█	█	█	█			40
<=22	█	█	█	█			23
<=24	█	█	█	█			6
<=26	█	█	█				9
<=28	█		█				2
<=30							0
Total	84	38	65	34	0	0	221

PODWEIGHT	Napo	Pastaza	Zamora	Caquetá	Pound	Nacional	Total
<=200	█	█		█			12
<=400	█	█	█	█			54
<=600	█	█	█	█			72
<=800	█	█	█	█			41
<=1000	█	█	█	█			28
<=1200	█			█			4
<=1450	█	█	█	█			8
Total	84	39	67	29	0	0	219

BEAN_POD	Napo	Pastaza	Zamora	Caquetá	Pound	Nacional	Total
<=.1	█	█	█				10
<=.2	█	█	█	█			91
<=.3	█	█	█	█			89
<=.4	█	█	█	█			22
<=.5	█	█					2
<=.6							0
<=.7	█						1
Total	84	39	66	26	0	0	215

Figure 4.5 (continued from previous page).

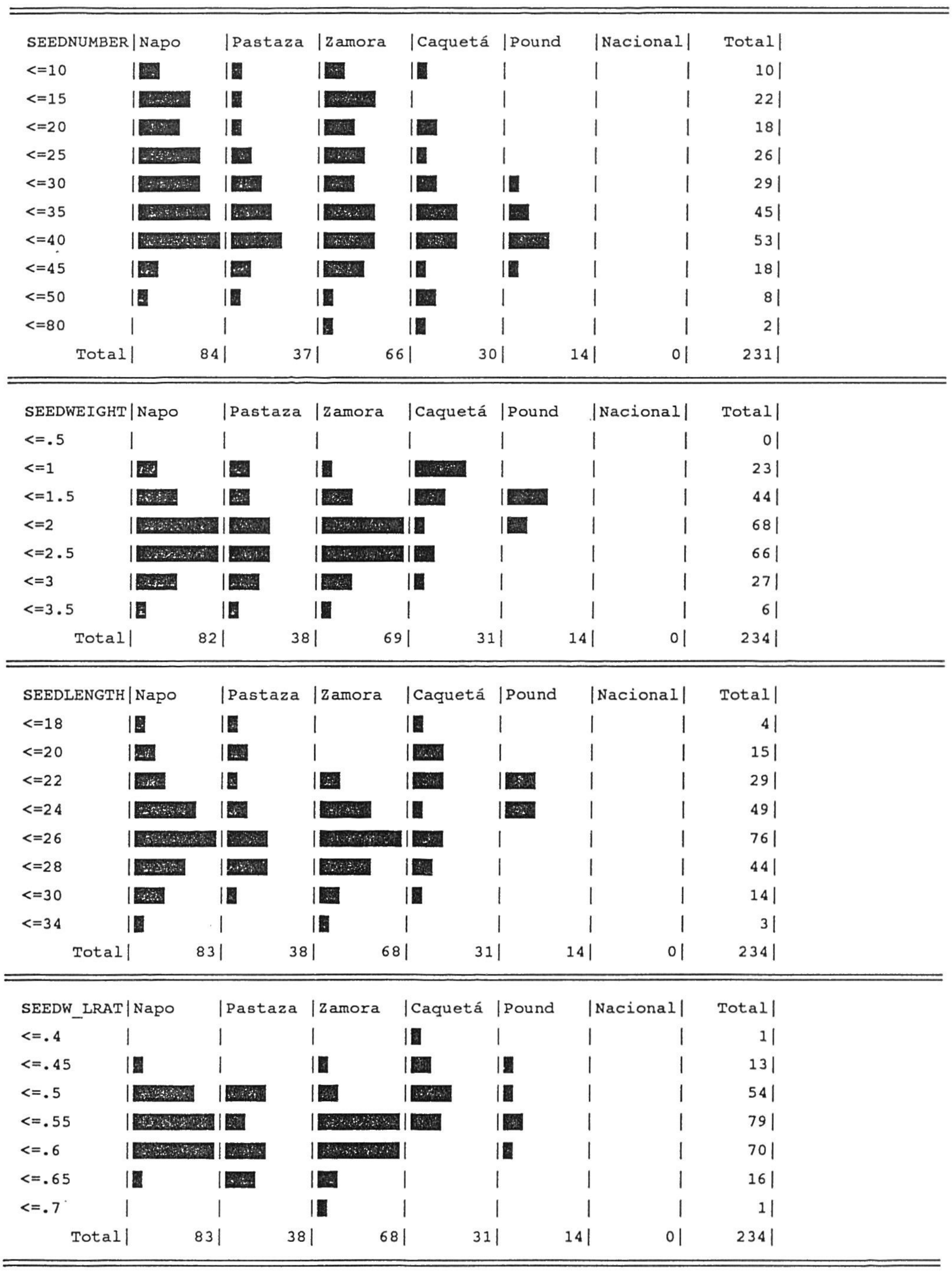


Figure 4.6 *Frequency distributions for seed descriptors (Data set 2) in six Level 1 areas (continues on next page). See Appendix 1 for details of descriptors.*

```
================================================================================
SEEDWHITE |Napo    |Pastaza |Zamora  |Caquetá |Pound   |Nacional|   Total|
         0|██████  |███     |██      |█████   |██      |        |     110|
         1|████████|████    |████████|        |        |        |     131|
     Total|      89|      39|      69|      30|      14|       0|     241|
================================================================================
SEEDLTPURP|Napo    |Pastaza |Zamora  |Caquetá |Pound   |Nacional|   Total|
         0|████████|███     |████████|█       |██      |        |     135|
         1|██████  |███     |███     |████    |██      |        |     106|
     Total|      89|      39|      69|      30|      14|       0|     241|
================================================================================
SEEDDKPURP|Napo    |Pastaza |Zamora  |Caquetá |Pound   |Nacional|   Total|
         0|████████|████    |████████|███     |█       |        |     128|
         1|████████|████    |████    |███     |███     |        |     114|
     Total|      89|      40|      69|      30|      14|       0|     242|
================================================================================
```

Figure 4.6 (continued from previous page).

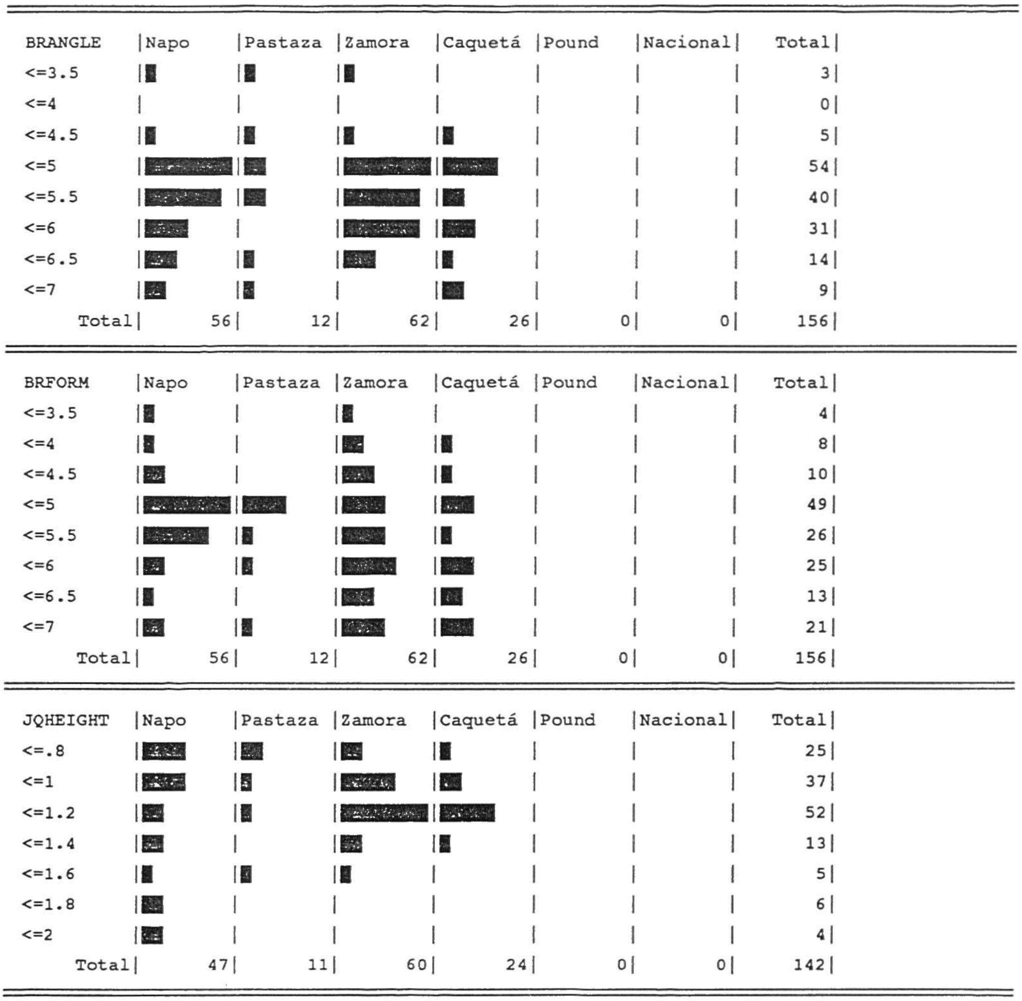


Figure 4.7 *Frequency distributions for tree habit descriptors (Data set 3) in six Level 1 areas. See Appendix 1 for details of descriptors.*

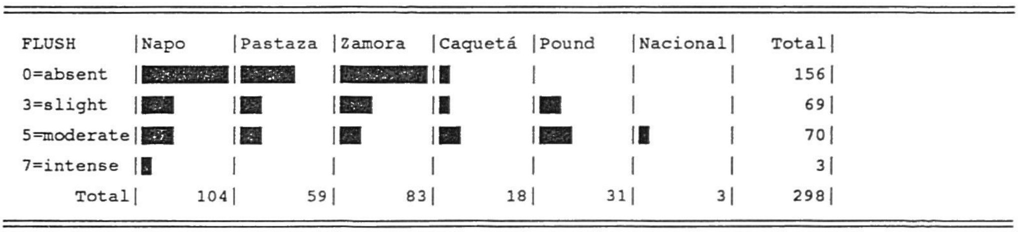
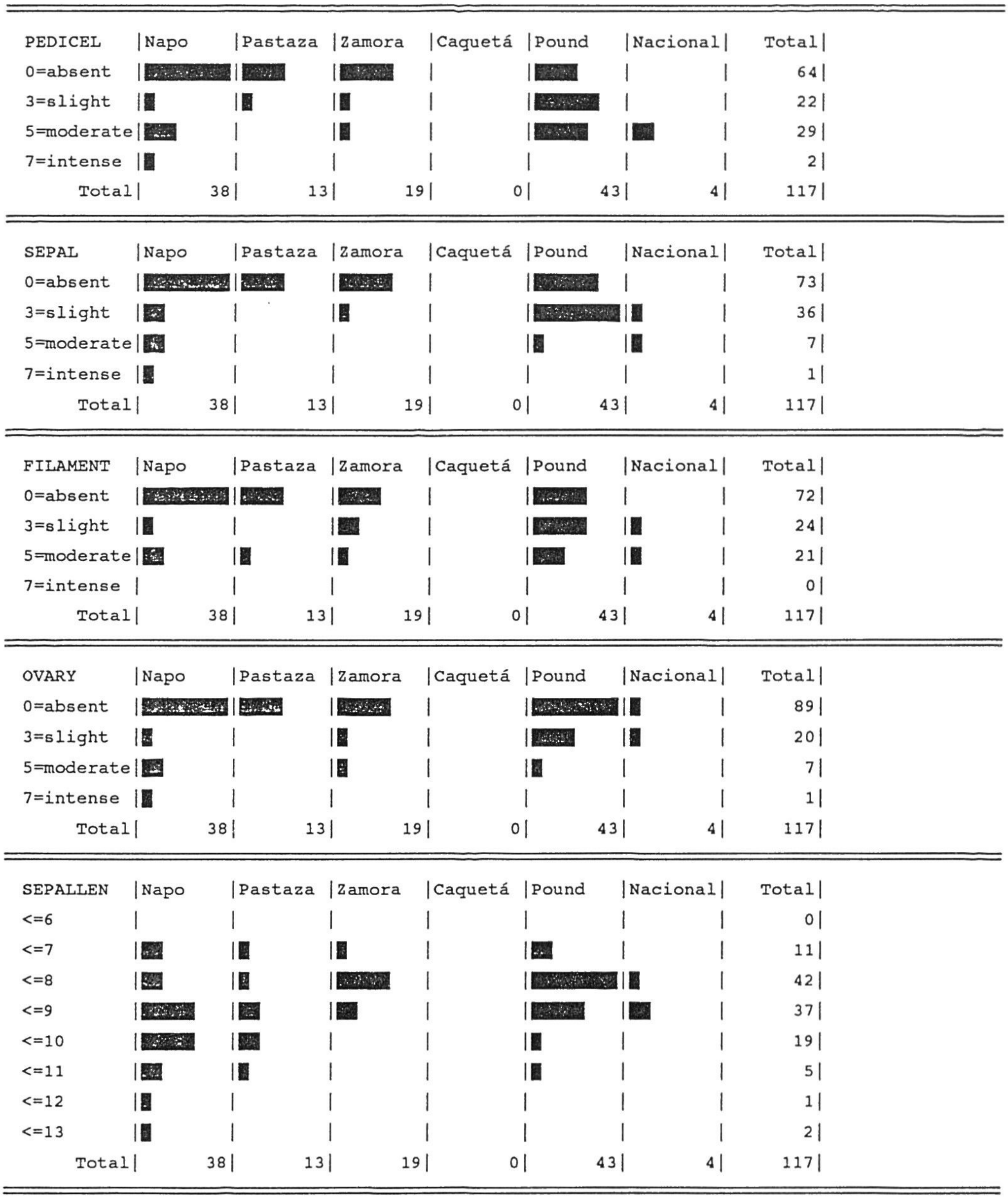


Figure 4.8 *Frequency distribution for flush colour descriptor (Data set 4) in six Level 1 areas. See Appendix 1 for details of descriptors.*



```
PEDICEL   |Napo     |Pastaza |Zamora  |Caquetá |Pound    |Nacional|   Total|
0=absent  |         |        |        |        |         |        |      64|
3=slight  |         |        |        |        |         |        |      22|
5=moderate|         |        |        |        |         |        |      29|
7=intense |         |        |        |        |         |        |       2|
     Total|       38|      13|      19|       0|       43|       4|     117|

SEPAL     |Napo     |Pastaza |Zamora  |Caquetá |Pound    |Nacional|   Total|
0=absent  |         |        |        |        |         |        |      73|
3=slight  |         |        |        |        |         |        |      36|
5=moderate|         |        |        |        |         |        |       7|
7=intense |         |        |        |        |         |        |       1|
     Total|       38|      13|      19|       0|       43|       4|     117|

FILAMENT  |Napo     |Pastaza |Zamora  |Caquetá |Pound    |Nacional|   Total|
0=absent  |         |        |        |        |         |        |      72|
3=slight  |         |        |        |        |         |        |      24|
5=moderate|         |        |        |        |         |        |      21|
7=intense |         |        |        |        |         |        |       0|
     Total|       38|      13|      19|       0|       43|       4|     117|

OVARY     |Napo     |Pastaza |Zamora  |Caquetá |Pound    |Nacional|   Total|
0=absent  |         |        |        |        |         |        |      89|
3=slight  |         |        |        |        |         |        |      20|
5=moderate|         |        |        |        |         |        |       7|
7=intense |         |        |        |        |         |        |       1|
     Total|       38|      13|      19|       0|       43|       4|     117|

SEPALLEN  |Napo     |Pastaza |Zamora  |Caquetá |Pound    |Nacional|   Total|
<=6       |         |        |        |        |         |        |       0|
<=7       |         |        |        |        |         |        |      11|
<=8       |         |        |        |        |         |        |      42|
<=9       |         |        |        |        |         |        |      37|
<=10      |         |        |        |        |         |        |      19|
<=11      |         |        |        |        |         |        |       5|
<=12      |         |        |        |        |         |        |       1|
<=13      |         |        |        |        |         |        |       2|
     Total|       38|      13|      19|       0|       43|       4|     117|
```

Figure 4.9 *Frequency distributions for flower descriptors (Data set 5) in six Level 1 areas. See Appendix 1 for details of descriptors.*

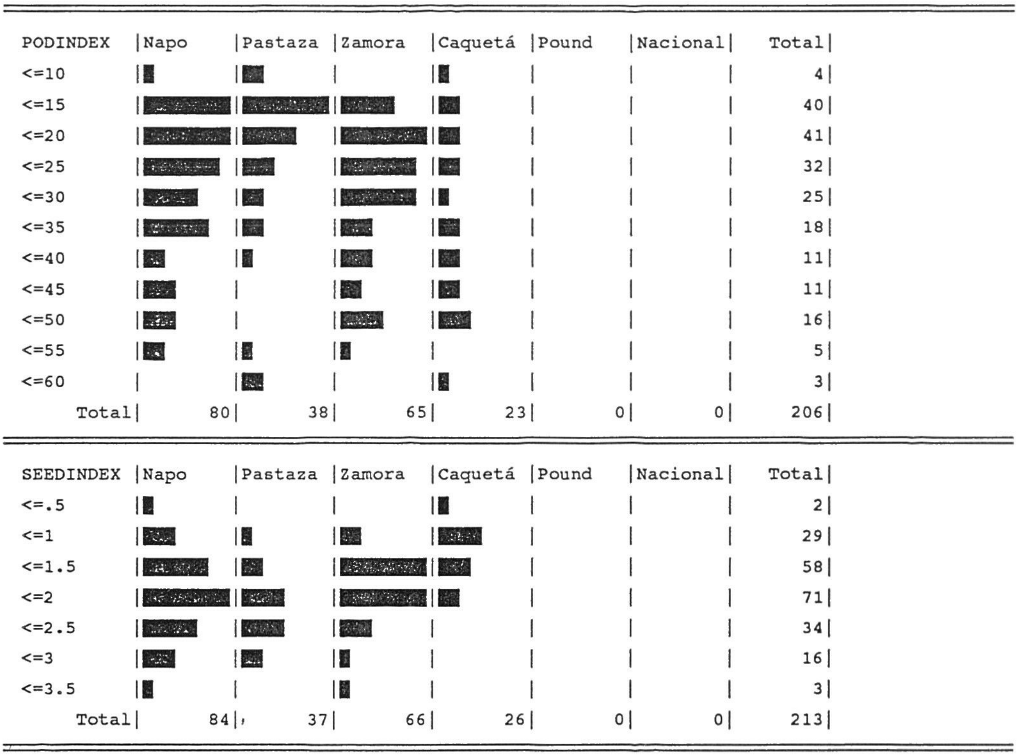


Figure 4.10 *Frequency distributions for production descriptors (Data set 6) in six Level 1 areas. See Appendix 1 for details of descriptors.*

4.4 Cluster mapping

The results of the cluster analysis are presented in Figures 4.12-4.16. The cluster maps are intended (a) to reveal, within the LCT-EEN and EBC accession groups, evidence of morphologically distinct sub-populations occupying distinct geographical areas and (b) to explore the relationships between all the accession groups. Because the Pound and Nacional groups consisted of rather few clones of uncertain origin, the cluster analysis was not intended to elucidate the pattern of variation within these groups.

The validity of the clustering was explored by plotting within-cluster variation, as a percentage of total variation, against the number of clusters. This was done both for the real pod and seed data and for four artificial data sets constructed from random normal deviates, following the approach of Doran and Hodson (1975). It was hypothesized that, as the number of clusters increases from two to ten, within-cluster variation should reduce more rapidly for the real data set than for the artificial data, in which any clustering is the result of chance. The results of this exercise, presented in Figure 4.11, showed that clustering did result in a larger reduction in variation when applied to the real data than when applied to artificial data in which any structure exists only through chance effects. This confirms that the clustering process did identify genuinely distinct groups of accessions, although substantial intra-cluster variation remained.

However, even with ten clusters, only slightly more than half the total variation was accounted for by the clustering. Additional clusters beyond about seven or eight produced only small reductions in within-cluster variation, so that for clarity of presentation it was decided to present the results for eight clusters (except for the habit and flower descriptors). I have verified that the spatial distributions of clusters overlapped in a broadly similar way for up to twenty clusters.

The following paragraphs present the results of the cluster mapping for each set of descriptors.

Pod descriptors; all accession groups (Figure 4.12): eight clusters; clustering accounted for 57% of total variation.
LCT-EEN: mainly in clusters 1-6; no evident spatial separation;
EBC: clusters 1 and 7; no spatial separation;
Pound accessions: all clusters except 7. Cluster 8 mainly Pound material with a few LCT-EEN accessions;
Nacional: one accession in each of clusters 1, 2 and 8.

Seed descriptors; all accession groups except Nacional (Figure 4.13): eight clusters; clustering accounted for 61% of total variation.
LCT-EEN: clusters 1-6 and 8; no evident spatial separation;
EBC: clusters 1, 2, 7 and 8, with cluster 7 containing only a few LCT-EEN accessions; no evident spatial separation;
Pound accessions: clusters 2, 7 and 8; grouped with LCT-EEN and EBC accessions in 2 and 8, grouped mainly with EBC material in 7.

Pod and seed descriptors; LCT-EEN and EBC accessions (Figure 4.14):

eight clusters; clustering accounted for 48% of total variation.

LCT-EEN: clusters 1-6, with a few accessions in clusters 7 and 8; no evident spatial separation for clusters 1-6;

EBC: mainly cluster 7, grouped with a few LCT-EEN accessions from the north-east Oriente; also clusters 1, 2 and 5; some spatial separation, with cluster 7 accessions located in the western part of the Río Caquetá collecting area.

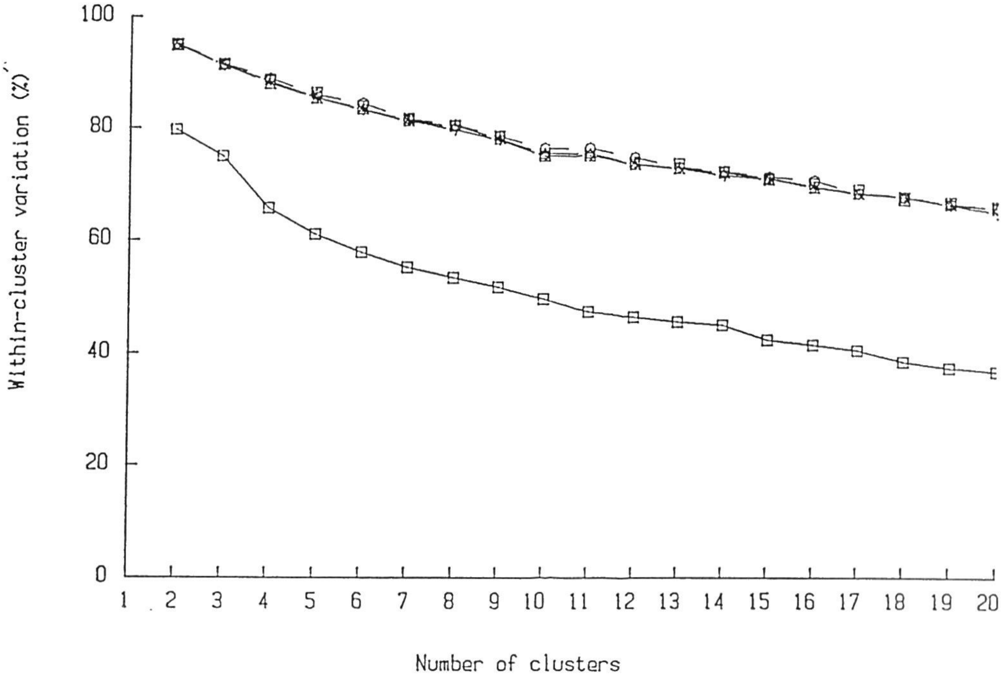


Figure 4.11 *Comparison of clustering applied to real and artificial data sets. The real data set (lower, solid line) consisted of 21 descriptors from Data set 2 applied to 234 LCT-EEN and EBC accessions (as in Figure 4.14). The artificial data sets (upper, dashed lines) contained randomized normal data of identical size.*

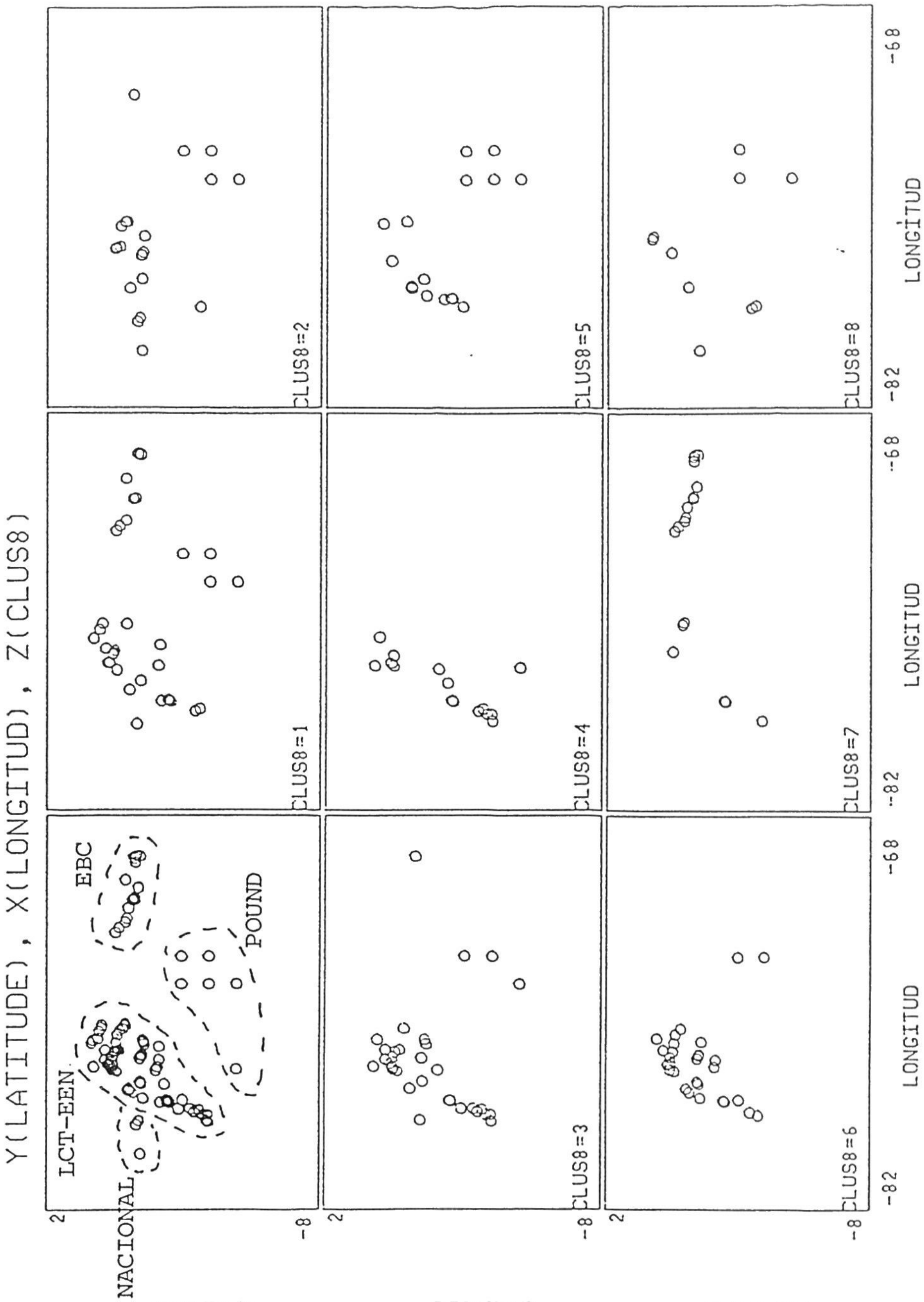


Figure 4.12 *Cluster maps: clustering based on pod descriptors (Data set 2; all accession groups).*

Notes for Figures 4.12-4.18:

1. Each of these figures shows the locations of individual accessions plotted on latitude-longitude coordinates. The area covered by each of the small plots corresponds to the upper part of the map in Figure 2.1 (the map of the Upper Amazon region).
2. The first plot shows all the accessions for a particular cluster analysis.
3. The remaining plots show accession locations for each cluster, labelled (for example) CLUS8=1, CLUS8=2...CLUS8=8.
4. An arbitrary location was assigned for each of the MORONA, NANAY, PARINARI and SCAVINA groups of the Pound accessions from Peru, since their exact locations are not known.

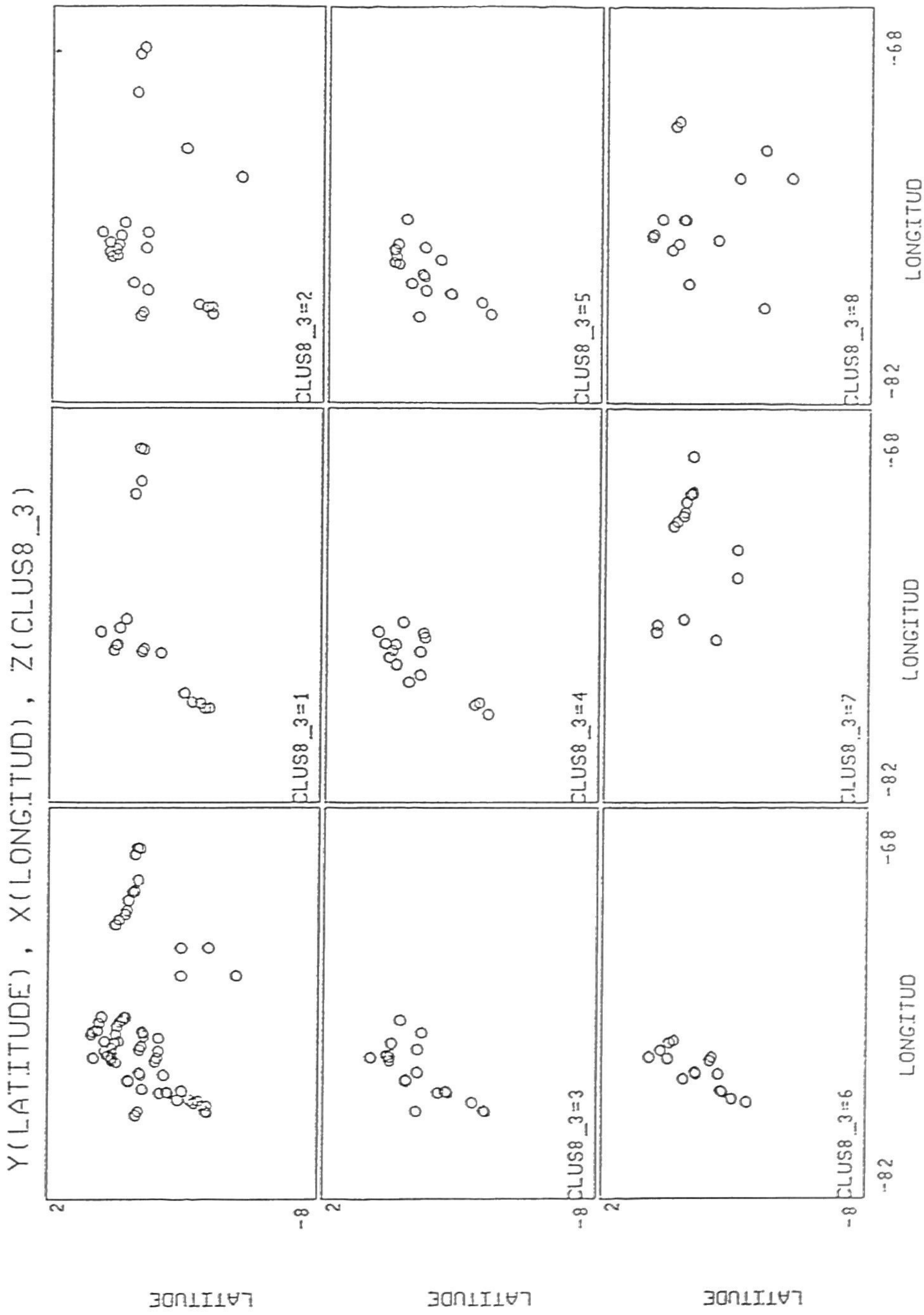


Figure 4.13 *Cluster maps: clustering based on seed descriptors (Data set 2; all accession groups). For notes, see Figure 4.12.*

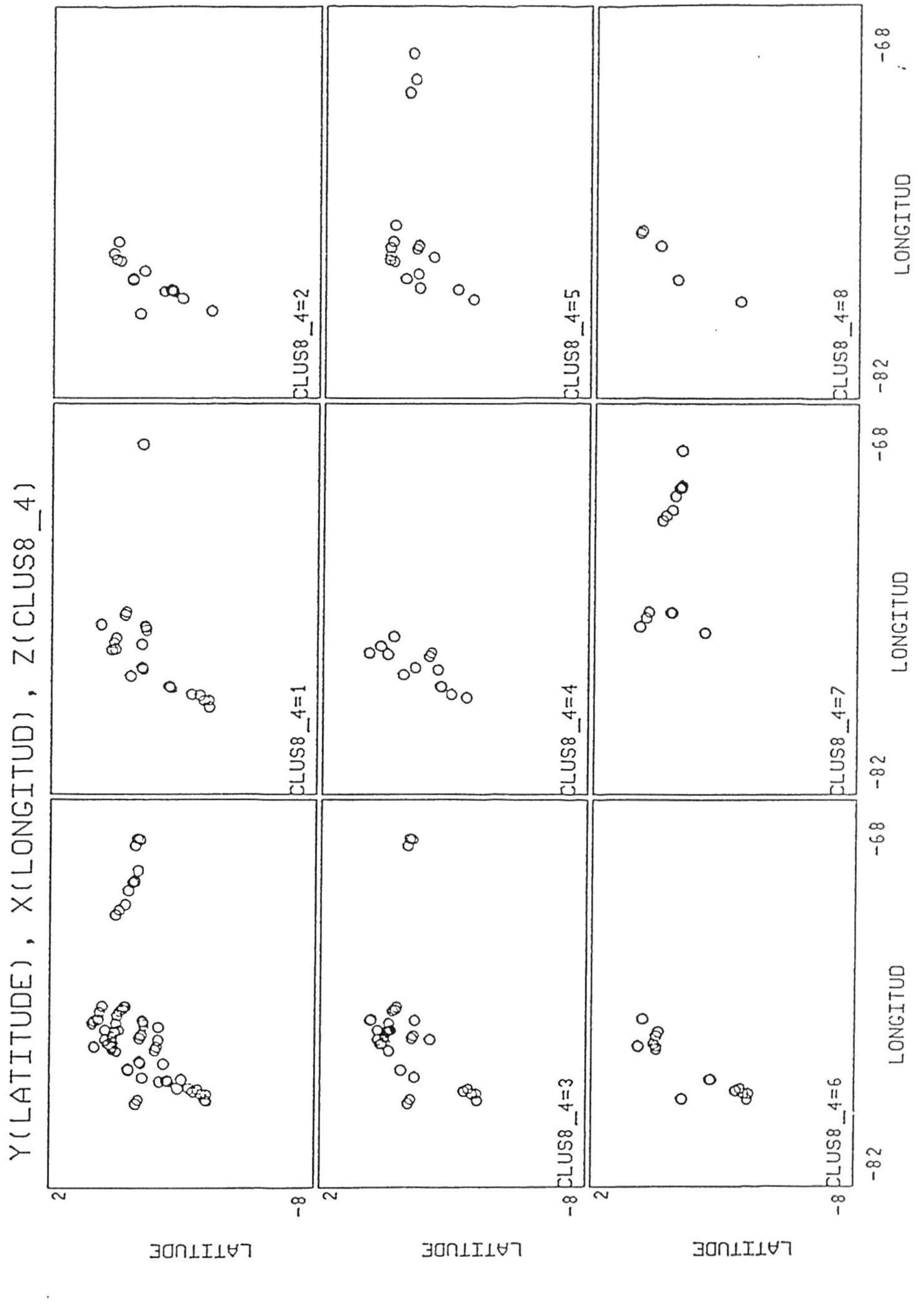


Figure 4.14 *Cluster maps: clustering based on pod and seed descriptors (Data set 2; LCT-EEN and EBC accessions). For notes, see Figure 4.12.*

Data sets 3 and 5 comprised fewer descriptors, so that smaller numbers of clusters were appropriate, and it was possible to characterize the clusters.

Habit descriptors; LCT-EEN and EBC seed accessions (Figure 4.15): this data set comprised only three descriptors, so that 67% of total variation was accounted for by grouping the accessions into five clusters. Clusters 1-3 all had jorquette heights around 1 m, with cluster 2 being distinguished by high branch angle at the jorquette, and cluster 3 by upward-curved branch form. Clusters 1-3 were all widely distributed in Ecuador. Cluster 4 was characterized by low branch angle and downward-curved branch form; this type was widespread in Ecuador, but essentially absent (one accession only) from the Caquetá accessions. Cluster 5 comprised a small group of accessions from Napo Province in Ecuador, with first jorquette heights around 2 m; this cluster may be an anomaly resulting from recording jorquette heights in relatively old trees.

Flower descriptors; LCT-EEN, Pound and Nacional accessions: this data set comprised only five descriptors, four of which described flower pigmentation and were strongly correlated, so that two clusters (Figure 4.16(a)) accounted for 43% of total variation. Cluster 1 consisted of accessions with little or no anthocyanin pigment in pedicels, sepals, stamen filaments or ovaries; cluster 2 accessions had more or less strongly pigmented flowers. The majority of LCT-EEN accessions belonged to cluster 1, whereas the Pound accessions were evenly divided between the two clusters. All three Nacional accessions were in cluster 2. Further subdivision, to five clusters (Figure 4.16(b)) accounted for 68% of total variation and accessions appeared to be grouped on the basis of sepal length as well as flower pigmentation. For example, cluster 4 consisted of a few accessions with minimal flower pigment but mean sepal length 11.6 mm (overall mean sepal length was 8.5 mm). In general, there appeared to be no spatial separation of the clusters within Ecuador.

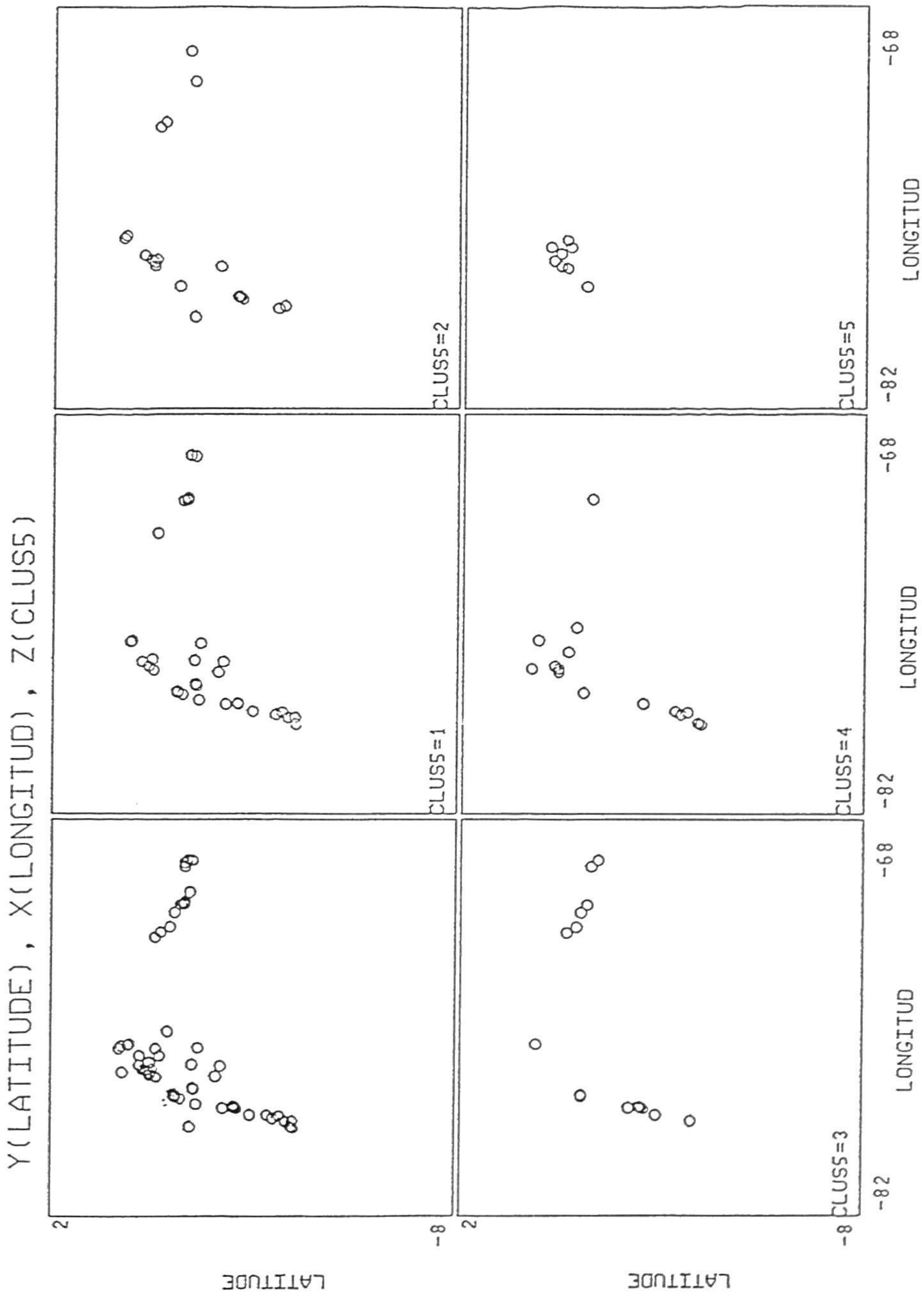
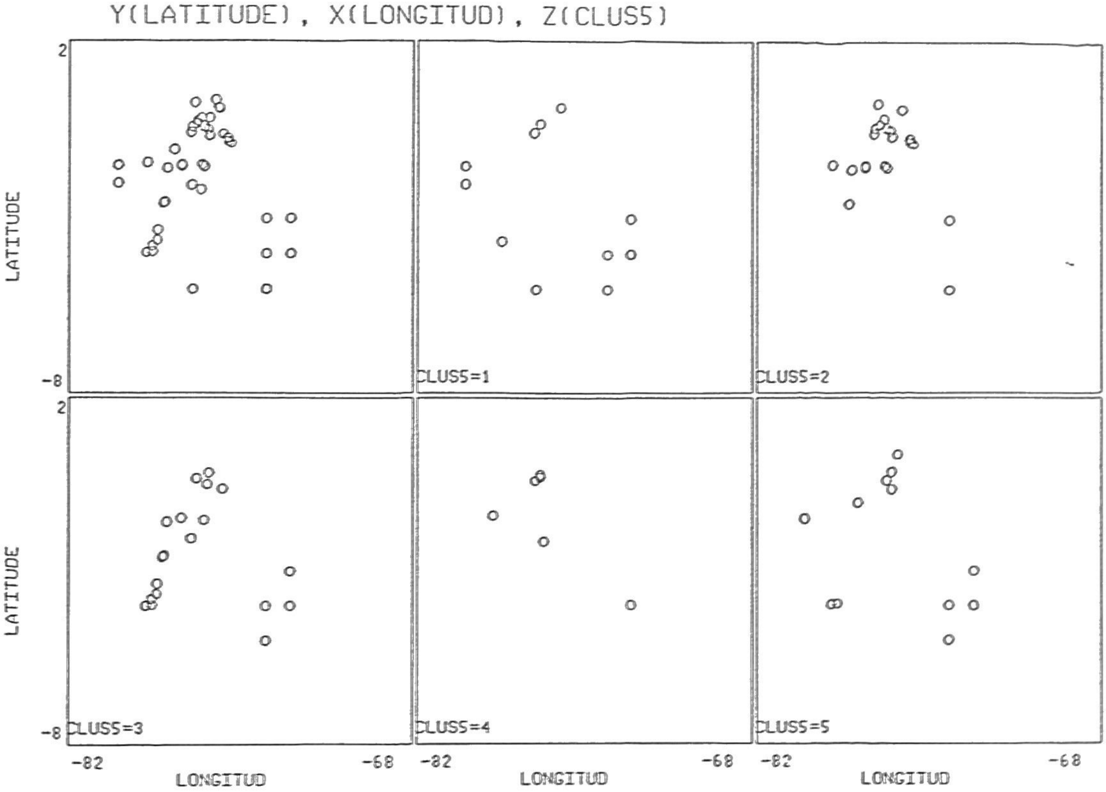


Figure 4.15 *Cluster maps: clustering based on habit descriptors (Data set 3; LCT-EEN and EBC accessions). For notes, see Figure 4.12.*



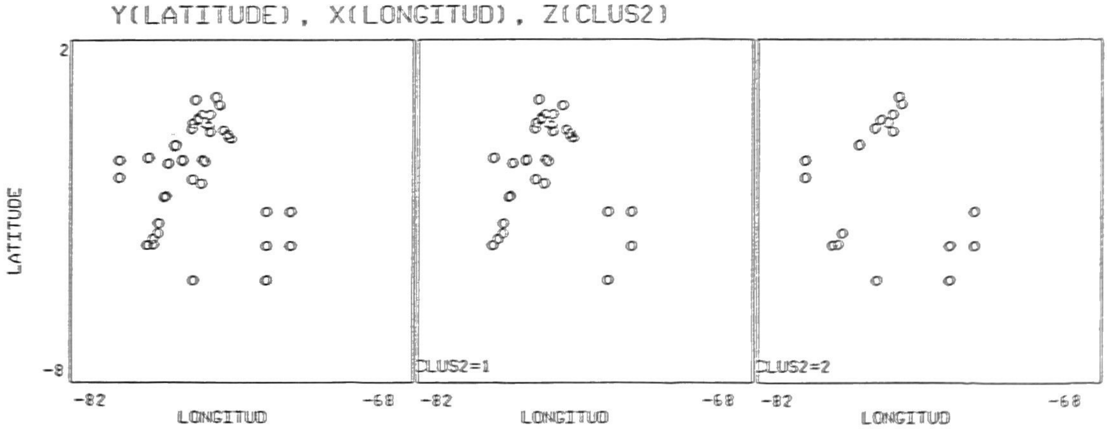


Figure 4.16 *Cluster maps: clustering based on flower descriptors (Data set 5; all accession groups). For notes, see Figure 4.12.*

Flush colour (Data set 4) was not used in the cluster analysis since it consisted of a single descriptor. However, it was informative to map the distribution of flush colour (Figure 4.17). The majority of LCT-EEN accessions had no anthocyanin pigment in the flush (flush = 0), whereas this type was not found among the Pound material. Low and moderate levels of pigmentation (flush = 3 or 5) were widespread in all accession groups. Two accessions with heavily pigmented flushes (flush = 7) were from planted trees of unknown origin, and are not relevant to this discussion.

The distribution of white seeds, one of the descriptors from Data set 2, was also of interest (Figure 4.18); white seeds were *never* found outside Ecuador, although they were found in the majority (66%) of the LCT-EEN accessions.

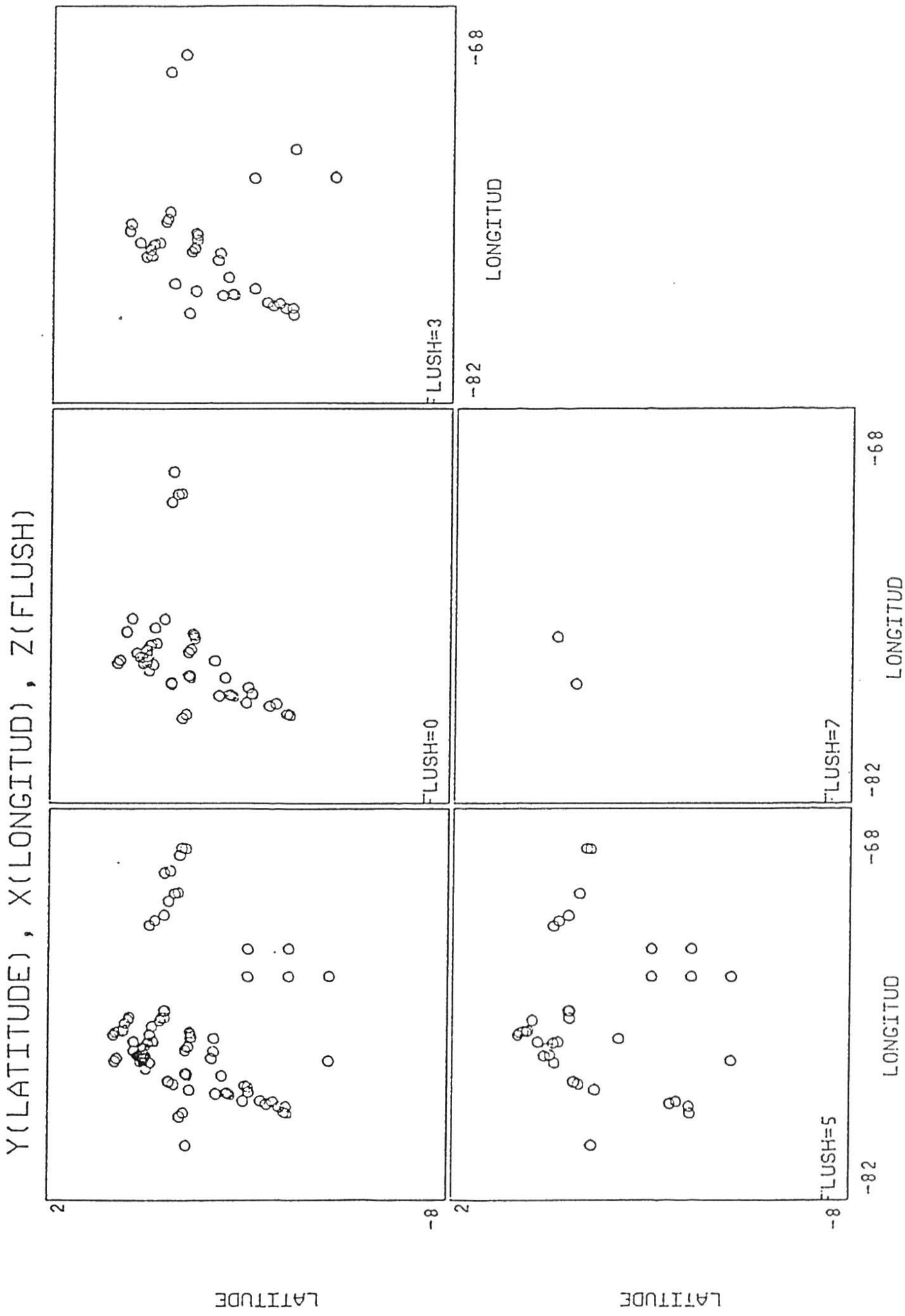


Figure 4.17 *Descriptor maps: flush colour (Data set 4; all accession groups). For notes, see Figure 4.12.*

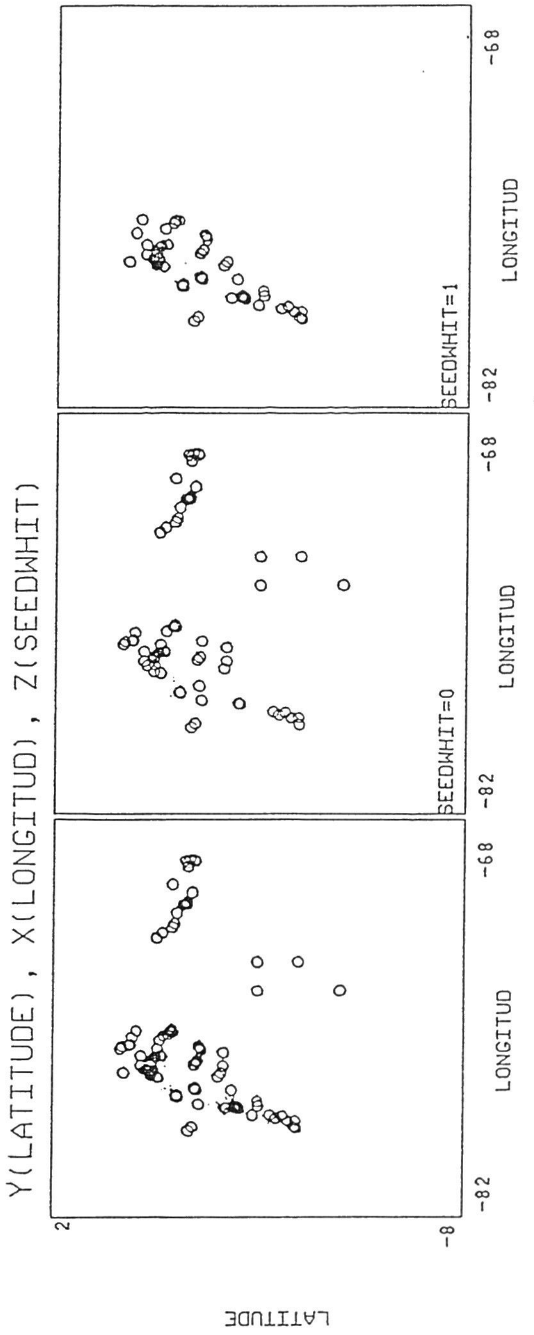


Figure 4.18 *Descriptor maps: presence of white seeds (Data set 2; all accession groups). For notes, see Figure 4.12.*

4.5 Principal components analysis

Figures 4.19-4.21 show the mean scores for Level 2 areas plotted on the first three axes derived from the principal components analysis of pod, seed and flower descriptors. The axis loadings and the variation accounted for by each axis are given in Table 4.3. It should be noted that for most areas there was a large range of values and hence considerable overlap between accessions from different areas; plotting the mean scores is convenient for summarizing the overall pattern of variation, but no undue significance should be attached to the precise locations of the plotted means.

Pod and seed descriptors; LCT-EEN and EBC accessions (Figure 4.19): the EBC accessions (areas 7-9) appeared somewhat distinct from the LCT-EEN material (areas 1-6). The three EBC points were widely spaced, reflecting the divergence noted in Section 4.3; on Axis 1, area 9 was closest to the Ecuador areas. Among the latter, area 2 (Putumayo) was associated with areas 7 and 8 on Axis 1, reflecting obvious affinities in terms of thin pod walls and small seeds.

Pod descriptors; all accession groups (Figure 4.20): in general, the range of variation in the Nacional and Pound accessions (areas A-G) was similar to that seen in the LCT-EEN material (areas 1-6). One point for area B referred to a single accession and should be ignored. The EBC accessions (areas 7-9) again appeared somewhat distinct on Axis 1.

Flower descriptors; LCT-EEN, Nacional and Pound accessions (Figure 4.21): in general, the Nacional and Pound accessions (areas A-G) were well separated from the LCT-EEN material (areas 1-6) on Axes 1 and 3, which reflected overall flower pigmentation. Axis 2 values were mainly determined by flower size; it appeared that the LCT-EEN accessions had slightly larger flowers than other areas.

Although the principal components analysis provides a concise overview of the descriptor data, its primary purpose was to generate 'composite descriptors' for the analysis of variance, as described in the following section.

Table 4.3 Principal components analysis: descriptors used and axis loadings for axes 1-3.

(a) pod and seed descriptors (Figure 4.19)

Descriptor	AXIS 1	AXIS 2	AXIS 3
% total variation	31	17	9
BASE	0.17	0.14	-0.05
APEX	0.04	0.08	0.33
SURFACE	-0.20	-0.01	0.09
FURROWS	-0.19	0.08	-0.06
PIGMENT	0.12	-0.03	-0.05
LENGTH	-0.19	0.41	-0.06
DIAMETER	-0.35	0.12	-0.07
D_LRATIO	-0.14	-0.36	-0.04
WALLTHIC	-0.31	0.05	-0.36
PODWEIGH	-0.29	0.29	-0.19
TOTBEANW	-0.20	0.39	0.22
BEAN_POD	0.16	0.19	0.51
SEEDNUMB	0.06	0.44	0.22
SEEDWEIG	-0.34	-0.07	0.11
SEEDLENG	-0.30	0.03	0.09
SEEDWIDT	-0.34	-0.08	0.24
SEEDTHIC	-0.28	-0.18	-0.06
SEEDW_LR	-0.15	-0.16	0.25
SEEDWHIT	-0.17	-0.23	0.30
SEEDLTPU	0.06	0.17	-0.13
SEEDDKPU	0.09	0.18	-0.29

(b) pod descriptors (Figure 4.20)

Descriptor	AXIS 1	AXIS 2	AXIS 3
% total variation	28	25	14
BASE	-0.26	0.40	0.29
APEX	0.05	0.41	0.07
SURFACE	0.47	0.14	0.48
FURROWS	0.48	0.17	0.43
PIGMENT	-0.18	0.09	0.41
LENGTH	0.34	0.47	-0.45
DIAMETER	0.56	-0.12	-0.28
D_LRATIO	0.16	0.62	0.23

(c) flower descriptors (Figure 4.21)

Descriptor	AXIS 1	AXIS 2	AXIS 3
% total variation	52	20	16
PEDICEL	0.58	0.06	-0.07
SEPAL	0.56	0.08	0.08
FILAMENT	0.40	-0.19	-0.73
OVARY	0.42	0.22	0.61
SEPALLEN	-0.10	0.95	-0.29

Notes:
1. Axis loadings > 0.3 are underlined.

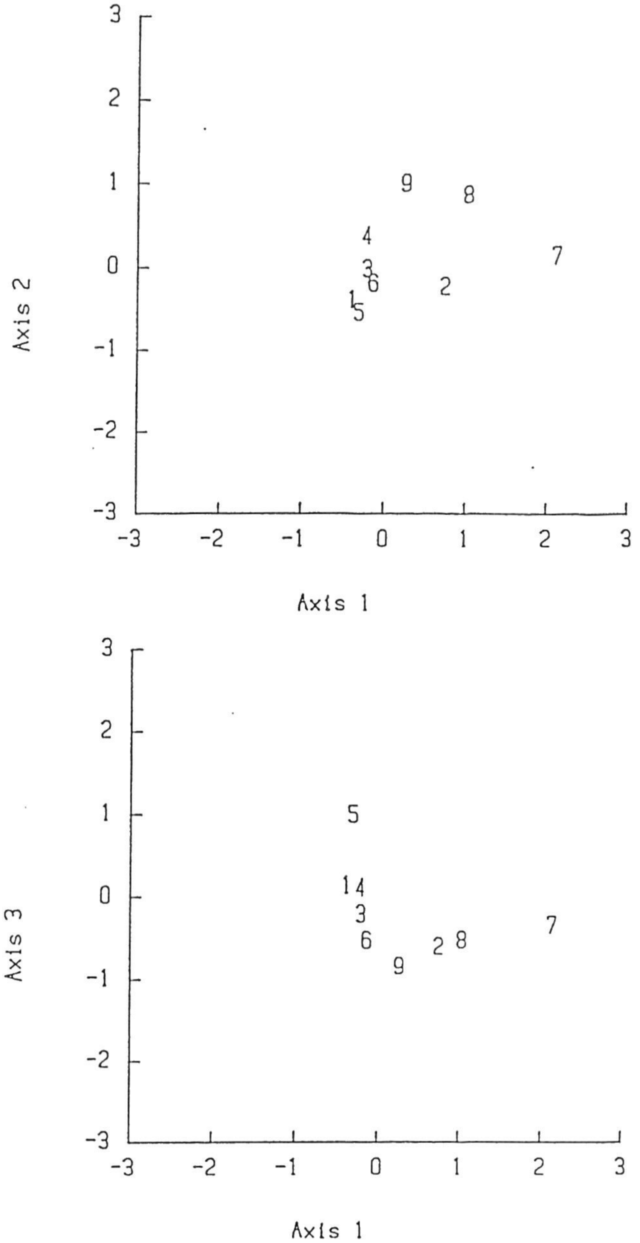


Figure 4.19 *Principal component means for Level 2 areas plotted against axes 1-3: pod and seed descriptors (Data set 2; LCT-EEN and EBC accessions).*

Notes for Figures 4.19-4.21:
1. Level 2 areas (see also Table 3.2) coded as 1 = San Carlos, 2 = Río Putumayo, 3 = Río Napo, 4 = Río Curaray etc, 5 = Río Upano etc, 6 = Río Zamora, 7,8,9 = Río Caquetá, A = Ecuador Coastal Region, B = MORONA, C = NANAY, D = PARINARI, E = SCAVINA, F = IMC, G = POUND.

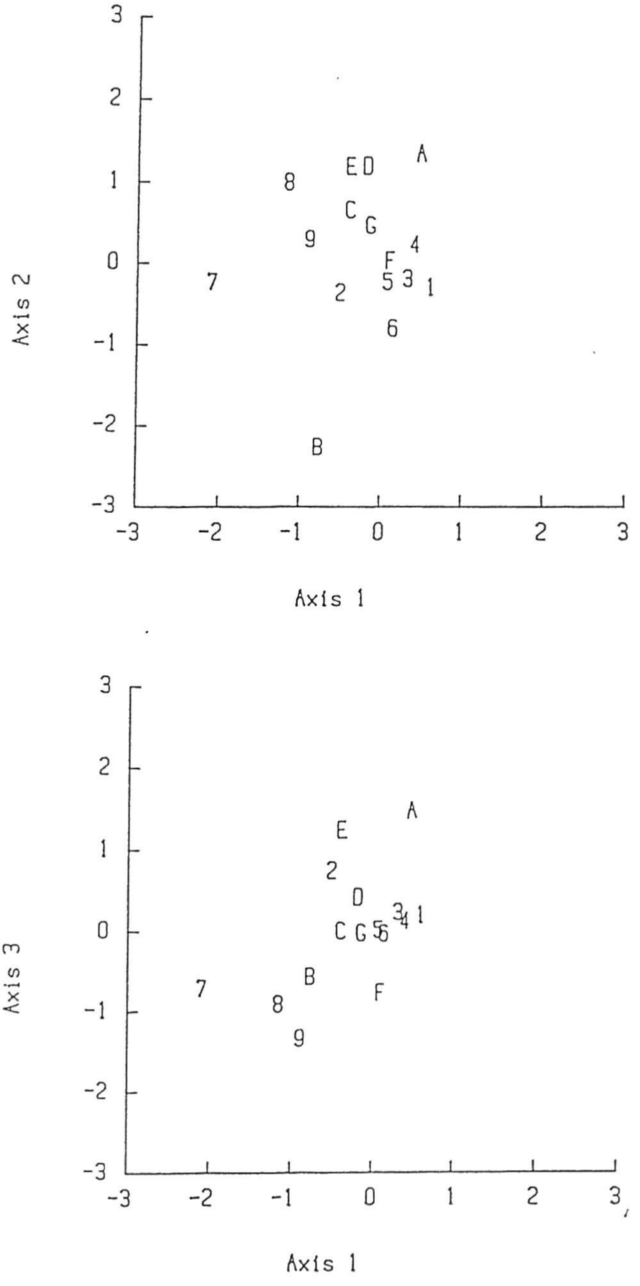


Figure 4.20 *Principal component means for Level 2 areas plotted against axes 1-3: pod descriptors (Data set 2; all accession groups).*

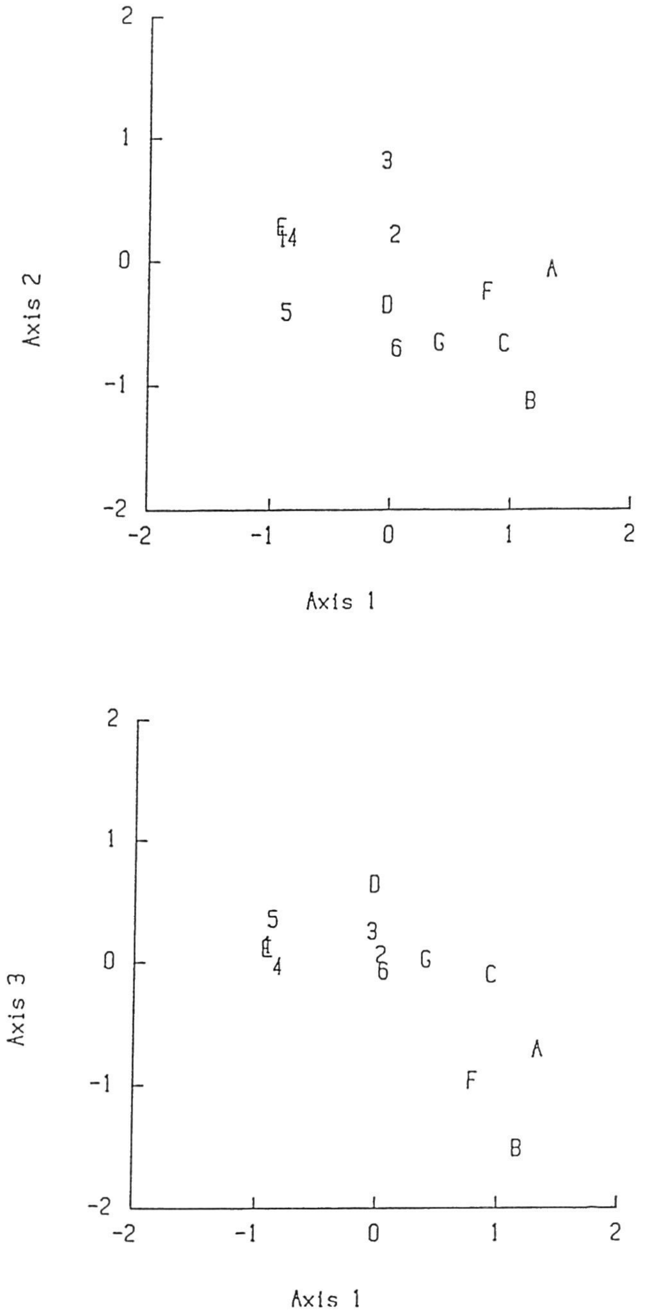


Figure 4.21 *Principal component means for Level 2 areas plotted against axes 1-3: flower descriptors (Data set 5; LCT-EEN, Nacional and Pound accessions).*

4.6 Analysis of variance

Starting from the principal components analyses described above, individual accession scores on Axes 1-3 were used as variables for an analysis of variance, the immediate object of which was to estimate within- and between-area components of variation for Level 1 and Level 2 areas. The resulting analysis of variance, and the calculated V_i values and ratios are given in Tables 4.4 and 4.5. In general, mean squares between Level 1 areas were higher than mean squares between Level 2 areas, and these in turn were higher than mean squares within Level 2 areas; where this was not the case, variance components were calculated from pooled sums of squares.

Values for the V_2/V_1 ratio varied from 0.12 to 2.76, with a median of 0.9. For V_3/V_2, values ranged from 2.45 to 46.3, with a median of 5.4. This ratio in effect measures the relative amount of variation captured per accession for two sizes of sampling unit. Thus, broadly, it can be said that Level 1 areas contribute about the same amount of variation per tree as Level 2 areas, and that Level 2 areas contribute about five times as much variation per tree as do individual trees within Level 2 areas. The practical significance of these ratios depends on the sampling cost ratios defined in Section 1.2; discussion will be deferred until Chapter 7.

Table 4.4 *Analysis of variance using composite descriptors derived from principle components analysis, for two levels of geographic subdivision (continues on next page).*

(a) Pod and seed descriptors (as Figure 4.19)

AXIS 1

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	3	38.4994	19	12.8331	0.1661
Between Level 2 within Level 1	5	21.3977	10	4.2795	0.1547
Within Level 2 areas	197	145.1031	71	0.7366	0.7366
TOTAL	205	205.0002	100		

AXIS 2

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	3	27.8904	14	9.2968	0.1554
Between Level 2 within Level 1	5	6.4686	3	1.2937	0.0187
Within Level 2 areas	197	170.6412	83	0.8662	0.8662
TOTAL	205	205.2002	100		

AXIS 3

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	3	20.9080	10	6.9693	-
Between Level 2 within Level 1	5	39.6910	19	7.9381	0.2988
Within Level 2 areas	197	144.4017	70	0.7330	0.7330
TOTAL	205	205.0002	100		

(b) Pod descriptors (as Figure 4.20)

AXIS 1

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	5	76.9564	29	15.3910	0.3007
Between Level 2 within Level 1	10	18.0970	7	1.8097	0.0663
Within Level 2 areas	255	174.9467	65	0.6861	0.6861
TOTAL	270	270.0001	100		

AXIS 2

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	5	43.6464	16	8.7293	0.1211
Between Level 2 within Level 1	10	32.6130	12	3.2613	0.1477
Within Level 2 areas	255	193.7413	72	0.7598	0.7598
TOTAL	270	270.0001	100		

AXIS 3

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	5	46.9576	17	9.3915	0.1665
Between Level 2 within Level 1	10	18.7180	7	1.8718	0.0632
Within Level 2 areas	255	204.3242	76	0.8013	0.8013
TOTAL	270	270.0001	100		

(c) Flower descriptors (as Figure 4.21)

AXIS 1

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	4	26.4277	23	6.6069	0.1882
Between Level 2 within Level 1	8	17.6250	15	2.2031	0.1679
Within Level 2 areas	104	71.9475	62	0.6918	0.6918
TOTAL	116	116.0000	100		

AXIS 2

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	4	29.3490	25	7.3373	0.2811
Between Level 2 within Level 1	8	6.0858	5	0.7607	0.7737
Within Level 2 areas	104	80.5653	69	0.7747	-
TOTAL	116	116.0000	100		

AXIS 3

SOURCE	DF	SS	%	MS	V_i
Between Level 1 areas	4	5.2683	5	1.3171	-
Between Level 2 within Level 1	8	21.2630	18	2.6579	0.1501
Within Level 2 areas	104	89.4683	77	0.8603	0.8603
TOTAL	116	116.0000	100		

Table 4.5 *Components of variance for sampling at two levels of geographic subdivision.*

(a) pod and seed descriptors (as Figure 4.19)

	V_1	V_2	V_3	V_2/V_1	V_3/V_2
AXIS 1	0.1661	0.1547	0.7366	0.93	4.76
AXIS 2	0.1554	0.0187	0.8662	0.12	46.32
AXIS 3	-	0.2988	0.7330	-	2.45

(b) pod descriptors (as Figure 4.20)

	V_1	V_2	V_3	V_2/V_1	V_3/V_2
AXIS 1	0.3007	0.0663	0.6861	0.22	10.35
AXIS 2	0.1211	0.1477	0.7598	1.22	5.14
AXIS 3	0.1665	0.0632	0.8013	0.38	12.68

(c) flower descriptors (as Figure 4.21)

	V_1	V_2	V_3	V_2/V_1	V_3/V_2
AXIS 1	0.1882	0.1679	0.6918	0.89	4.12
AXIS 2	0.2811	0.7737	-	2.76	-
AXIS 3	-	0.1501	0.8603	5.73	5.73

4.7 Other descriptors

Some effort was devoted to looking at potentially interesting descriptors other than those for which results have been presented.

One possible descriptor for cocoa, mentioned by Posnette (1945), is the degree of separation of the pulvini on the petioles. According to Posnette, Pound's NANAY clones have pulvini which are not clearly separated. This character was not very distinct in the four NANAY clones at Pichilingue, and a complete survey of the accessions held at San Carlos revealed none in which the pulvini were not clearly separated. For the purposes of this project, therefore, no use was made of this descriptor, although it might provide an additional measure of variability in other cocoa populations.

Another character of interest is the distribution of flower cushions on the tree. Most general descriptions of cocoa refer to the flower cushions occurring on the trunk and main branches of the tree, and pictures of cocoa plantations typically show pods hanging from trunks. Among the LCT-EEN accessions, however, a comprehensive survey of seedling trees in the genebank in 1986 showed no flower cushions at all on the trunks of trees of wild origin. The 'trunks' of budded trees are plagiotropic branches, and do [illegible] have flower cushions. In general, pods on wild trees in Ecuador, whether in the genebank or in the wild, were borne on the branches. On the other hand, wild trees along the Río Caquetá in Colombia were seen to have a high density of flower cushions all the way up the trunk; this characteristic could not be checked in the genebank because the Caquetá trees were too young in 1986, but it is probable that there is a striking difference between the two areas in this character. The Pound accessions at Pichilingue were all budded trees and no assessment of this character was possible.

5 Population biology: methods

The purpose of Chapters 5 and 6 is to present a series of observations which, taken as a whole, provide a picture of *Theobroma cacao* as a wild species in its natural habitat. The methodology for these observations is outlined in Chapter 5. Some data were acquired in the course of collecting trips, for which Chapter 2 provide background descriptions; the remaining data relate to a population of wild cocoa at the INIAP Napo-Payamino Experimental Station at San Carlos, where LCTAP was based. This population will be referred to here as the 'study population'.

5.1 The study population

A population of wild cocoa, apparently of natural origin, existed at San Carlos in the area of forest chosen for the genebank site (see Section 2.3). The existence of this population made possible the study of a wild population of cocoa in parallel with the normal activities of LCTAP. Although part of the site was used for planting the field genebank, cocoa trees were carefully preserved during clearance operations, and larger forest trees were retained for shade over most of the genebank area. Although the genebank plantings clearly resulted in considerable disturbance to part of the study population, they also facilitated the mapping of the wild population and some aspects of data collection.

The locations of trees in the study population were mapped, to the nearest 1 m; in undisturbed areas, careful searches were made for small trees and recent seedlings. The spatial distribution of the trees was tested for randomness by comparing the observed mean nearest-neighbour distance to that expected for a random distribution, with a correction for boundary length in a finite area (Diggle 1983).

Population structure was investigated by measuring the girths of every stem at least once, at 1.3 m above ground level, for stems having girth >100 mm; individual trees had up to 28 such stems. For a sample of trees representing the size range found in the study area, individual stems were marked (using plastic labels stapled to the stem at the measurement height) and measured three times in the period 1984-1986, to provide information on growth rates. Repeated girth measurements were generally within the range ±1 mm of the original measurement. Observations of flowering and pod production were also made, at irregular intervals.

Data on the study population were supplemented by data from trees in the genebank. Although this was not strictly a 'wild population', it did consist in the main of wild genotypes from the same region, growing in conditions identical to those of the study population.

5.2 Flowering and pollination

The immediate objects of collecting data on flowering were to determine, first, the pattern of flowering through the year and, second, the age at which young trees begin to flower.

The pattern of flowering through the year

Changes in flower numbers through the year were assessed by counting fallen flowers on 1 m by 1 m squares of plastic mesh placed beneath the canopies of a sample of trees. Counts were made at 3-4 day intervals during periods of high flower production, or otherwise at 7 day intervals. The mesh squares were cleaned of flowers and other debris after each count. Cocoa flowers which are not pollinated (typically more than 99% of all flowers) fall from the tree within 1-3 days after opening, but remain easily recognizable on the mesh squares for at least 10 days, so that this method provides a complete count of flowers produced by 1 m^2 of canopy. At each count, the trunk of each tree was also checked for flowers (up to the first jorquette), but none were ever found. This technique for flower counting has been used routinely at Pichilingue for several years, and is regarded as reliable (A.C.Maddison, personal communication).

Flower counts were made on 18 trees in the genebank and on 5 wild trees. The genebank trees were a random sample from the oldest blocks (3 accessions taken at random from each of Block 1 and Block 2; 3 trees taken at random within each accession). The wild trees were chosen for their accessibility. Counting started on 5 April 1984 and continued (with assistance from INIAP staff) up to 30 June 1986.

The age at which young trees begin to flower

The age of first flowering under plantation conditions was assessed by checking all trees in the genebank for the presence of flowers approximately once a month between January 1983 and October 1985. As this check was made quickly, it is likely that some trees with few flowers were recorded as not in flower, but this should not affect the overall conclusions.

Pollination and compatibility

It is generally assumed that wild cocoa is self-incompatible, as is virtually all Upper Amazon material so far tested (Engels 1981). To check self-incompatibility, stamens were removed from freshly opened flowers using tweezers, and examined to see if the anthers were releasing pollen. If so, the anthers were brushed lightly over the stigmas of other freshly-opened flowers, which were then marked with pins pushed through pieces cut from orange plastic plant labels. The flowers were checked after about five days for signs of ovary enlargement, which signified a successful pollination. Self-pollination was attempted on 5-16 flowers of each clone. Flowers were not protected against natural pollination, as this was expected to occur in less than 1% of flowers. I determined in advance that any clone which produced more than one apparently successful self-pollination would be rechecked, as spurious positives might have been a problem if the level of natural pollination was higher than expected. In the event, only one apparently successful self-pollination occurred, out of 141 attempted self-pollinations, so that natural pollination did not complicate the interpretation of either self- or cross-pollinations.

Shortage of time, and scarcity of flowers during periods when time was available, caused the abandonment of plans for a fuller study of cross-compatibility.

5.3 Seed production

Observations of pod production were made in the study population at irregular intervals between February 1982 and July 1986 of pod production in the wild population at San Carlos. Additional data were obtained during collecting trips, when pods on wild trees were being actively sought for collection. The field record sheet for each collection indicates whether the original tree had no pods, few pods (approximately 1-10 pods), or numerous pods (more than 10 pods).

The interpretation of these data depends to some extent on the flowering pattern of the trees, and knowledge of the period during which a pod would be counted as such. Typically, a pod takes about five months to develop to maturity, and is large enough to be readily counted for around four months. After ripening, a pod which is not removed from the tree turns black within about one month, but remains more or less intact for a further 4-8 months. This means that a fair estimate of a year's pod production can be obtained by a single count of pods (including pods at all stages of development, as well as old, decaying pods), if there is a marked peak of flowering in the course of the year and pods are counted between 4 and 8 months after the peak flowering period.

Seed production was calculated from pod numbers using the average number of seeds per pod, which for Ecuadorian wild populations was taken to be 28 (from Table 4.2). From collecting experience, it was known that seed viability was virtually 100%.

5.4 Seed dispersal

The role of mammals

It is generally believed that wild cocoa seeds are dispersed by mammals, including monkeys, squirrels and rats. In cocoa plantations, all these animals have been reported as pests (Thorold 1975), in that they open pods and remove the seeds, presumably attracted by the sweet pulp surrounding the seeds. Direct observation of this behaviour in the wild was not feasible, due to the small numbers of pods on accessible wild trees and to the fact that pods seemed to be opened mainly at night. However, some direct observations were made of monkeys opening cocoa pods while visiting the Anaconda Hotel, a jungle lodge on an island in the Río Napo, near Misahuallí, in April 1985.

Seed shadows

Further evidence relating to seed dispersal was obtained from seed shadow maps. Seedlings were mapped around two wild cocoa trees at San Carlos and one wild *Theobroma subincanum* at San José de Payamino. Seedlings were located by careful searching of the forest floor, starting at the parent tree and working outwards. Most seedlings were fairly close to the parent tree; after locating the most distant seedling in each case, searching continued in a circular zone, centred on the parent tree, to at least 10 m beyond the furthest seedling. The position of each individual was worked out using the distance to a fixed origin point near the base of the tree, and a compass bearing from this origin to the seedling. For large trees, several origin points were used and their positions relative to the parent tree noted.

5.5 Seedling survival

Three simple experiments were set up to obtain data on seedling survival under forest conditions. In all three experiments, seeds were planted just below the surface litter, at 100 mm intervals along 2 m transects. The seeds were not peeled. The ends of the transects were marked with posts, so that the locations of seeds could be found by measurement. The transects were placed in the forest in areas not subject to any regular human disturbance. In each experiment, the transects were placed parallel to each other, 5 m apart. After germination, seedlings were individually identified using plastic plant labels. The experiments were recorded at irregular intervals, when the stem diameters and heights of surviving seedlings were measured. Where seedlings had died, the cause of death, if evident, was noted. Details of the three experiments are listed in Table 5.1.

Table 5.1 Seedling survival experiments

Experiment	Site	Transect	Seeds /transect	Peeled weight (g)	Date planted	Seed source
1	1	A-D	20	0.97	21.10.84	LCT-EEN 46
	2	E-H				
2	1	I,K,M	5	1.92	12.2.85	LCT-EEN 73 (light purple)
		J,L,N				LCT-EEN 73 (dark purple)
3	1	P,R,T,V	10	2.41	5.5.85	LCT-EEN 288/s1, 288/s3
		Q,S,U,W		0.68		LCT-EEN 199/s2, 199/s6, 199/s7

Notes:

1. Sites: 1 = forest to S of genebank 2 = forest to E of oilwell Sacha-75

5.6 Adult survival, vegetative reproduction and age-structure

The adult survival rate and the related rate of vegetative reproduction were deduced from events observed in the study population and among other wild trees at San Carlos. Absolute age determination was not possible, but estimates of relative ages were obtained by calculating rates of basal area increase for trees of different sizes, over the period 1984-1986.

6 Population biology: results

The first section of this chapter presents some general observations relating to wild cocoa in Ecuador and Colombia; succeeding sections present the results corresponding to the methodology outlined in Chapter 5 for the study of cocoa as a wild species.

6.1 General observations relating to wild cocoa

In general, *Theobroma cacao* is a common understorey tree in mature forest throughout much of the region, at altitudes below 1000 m. The density of individuals varies widely; in some areas, a day's exploration of the forest yielded only one or two trees, while many sites were found with 5-10 individuals/ha.

In Ecuador, the distribution of wild cocoa showed a clear relationship with the soil and terrain types listed in Section 2.1. Cocoa was hardly ever found in the undulating uplands (Figure 2.2, Type R), which typically have relatively infertile red lateritic soils with a high aluminium content. Two other *Theobroma* species, *T. glaucum* and *T. subincanum* (using the nomenclature of Cuatrecasas (1964)) were, however, preferentially associated with this terrain type, and were not usually found in association with *T. cacao*. It was particularly common on the level terrain associated with former floodplains (Type N), with relatively fertile black soils. In the present-day floodplain complexes (Type T), cocoa was frequent on the higher, rarely-flooded river terraces, and absent from low-lying, regularly flooded areas. It was completely absent from permanently flooded or swampy areas (Type Q). In the mountainous terrain close to the *sierra* (Type M), cocoa was found both on level sites and on slopes of 60° or more, for example on the cliffs along the valleys of the Río Upano and the Río Santiago. The occurrence of wild cocoa in such sites had not previously been reported, and came as something of a surprise.

Along the Río Caquetá in Colombia, extensive stands of cocoa were found on the marginal strip of relatively dry ground (see Section 2.2) along the main river. No cocoa was found in the flooded areas further away from the river, nor was it found along the minor clear-water tributaries of the Caquetá, possibly because of the typically lower fertility of the soils along these tributaries.

Typical wild cocoa trees throughout the region were tall (up to 20 m) with 2-20 or more trunks and a vigorous growth of chupons (orthotropic shoots) from the base of the established trunks. The largest trees seen had 40-50 trunks, and the bases of these trees were 2-3 m across (see Appendix 5 for photographs of typical trees).

The morphological characteristics of wild cocoa have been described in detail in Chapter 3. In brief, wild trees in Ecuador were generally distinctive (in relation to the species in general) by virtue of their strongly wrinkled pods, the high frequency of white seeds, the absence of red pigmentation in the flush, and the absence of flower-cushions on the trunk. Photographs showing the range of pod types found in Ecuador and Colombia are provided in Appendix 5.

6.2 The study population

The mapped study population is shown in Figure 6.1. In an area of approximately 12 ha, 56 trees were located. The observed density of 4.7 trees/ha was similar to that at other sites in the region where wild cocoa was found.

Nearest-neighbour distances were measured, from the map, for 45 trees in an area 320 m x 240 m; the mean distance was 19.44 m. The expected mean nearest-neighbour distance (± 1 standard error), assuming the trees to be randomly distributed, was calculated to be 22.08±1.87 m. The difference between the observed and expected values is tested in terms of the normal deviate:

$$z = (19.44 - 22.08)/1.87 = -1.41 \qquad (6.1)$$

This difference is not significant ($p = 0.16$); there is no reason to question the random spatial distribution of the trees in the study population.

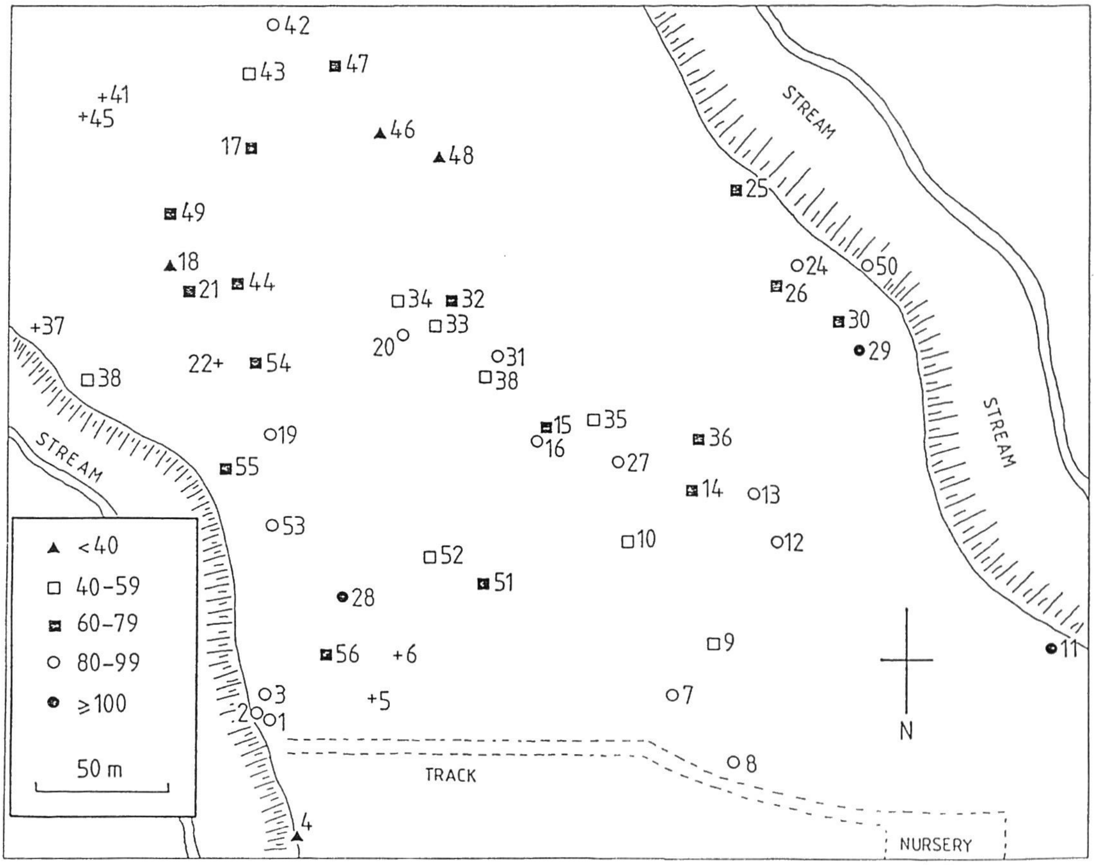


Figure 6.1 *Map of the study population at San Carlos. The site is at 265 m above sea level, and is approximately level, bounded by steep-sided stream valleys, 5-10 m deep, to the east and west. Symbols on the map indicate locations of trees and are coded to indicate the estimated age of the tree (see Section 6.8 and key on map). The numbers are tree numbers in the series W-1 to W-56 (no direct reference is made in the thesis to these tree numbers, but they have been included for completeness).*

6.3 Flowering and pollination

The pattern of flowering through the year

Flower counts are presented in Figure 6.2. For the trees recorded, flowering during 1984-1986 was concentrated in the period from September to March, with the main peak in October-November and a lower peak in February-March. Flowering rates were 5-10 times higher in 1984-1985 than in 1985-1986, possibly because of the unusually hot, dry weather in the former period (see Section 2.1). Between April and August, flower production virtually ceased in all three years.

Extensive production data are available for the genebank (Figure 6.3). These show that there are peaks in pod production corresponding to the flowering peaks, but they are far less marked. Appreciable numbers of pods appear to be initiated even in periods of minimal flower production.

The age at which young trees begin to flower

It seems to be generally accepted that cocoa trees start to flower and bear fruit when they reach a certain size. Glendinning (1960) stated that bearing commenced at a stem diameter of 6 cm, corresponding to a girth of about 200 mm. Flowering data from 24 accessions in Block 1 of the San Carlos genebank support the application of this generalization to wild genotypes. These seedlings were planted in January 1981, and started to flower around January 1983. As average girth increased from 170 mm (January 1983) to 351 mm (May 1985), the proportion of trees flowering increased from 8% to 97% (Figure 6.4). For individual accessions in Block 1, there was a clear relationship between mean girth and the proportion of trees flowering at a particular date (Figure 6.5).

Under plantation conditions (as in the genebank) there is substantial variation in the time required to reach flowering size. Nonetheless, wild genotypes in the San Carlos genebank generally reached flowering size in 24-36 months from field planting (or 30-48 months from germination). Seedlings in the forest typically exist under much heavier shade and growth rates are far lower (see Section 6.5).

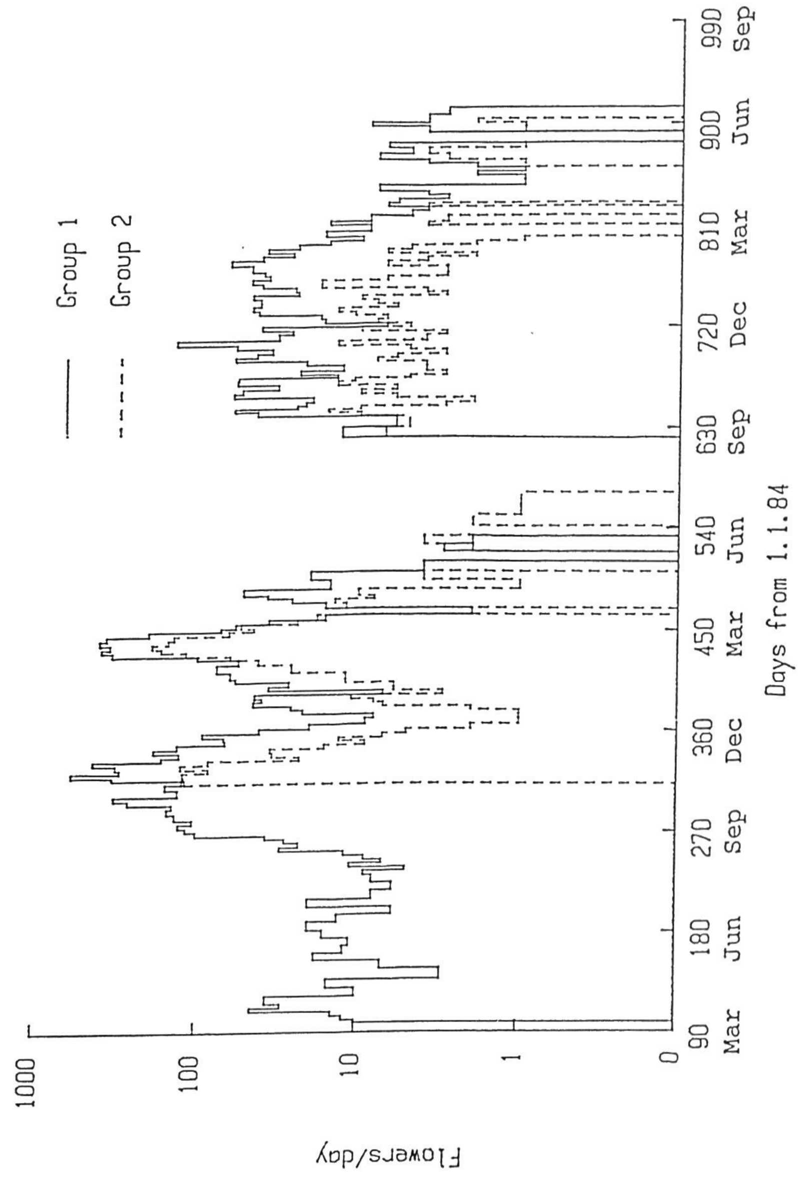


Figure 6.2 *Flower counts 1984-1986. Group 1: counts from 18 accessions in the genebank (18 m²) plus one wild tree (2 m²); Group 2: counts from four wild trees (4 m²). Data cover the period 5.4.84 to 30.6.86; from October 1985 to May 1986, data were collected by J.Baquero.*

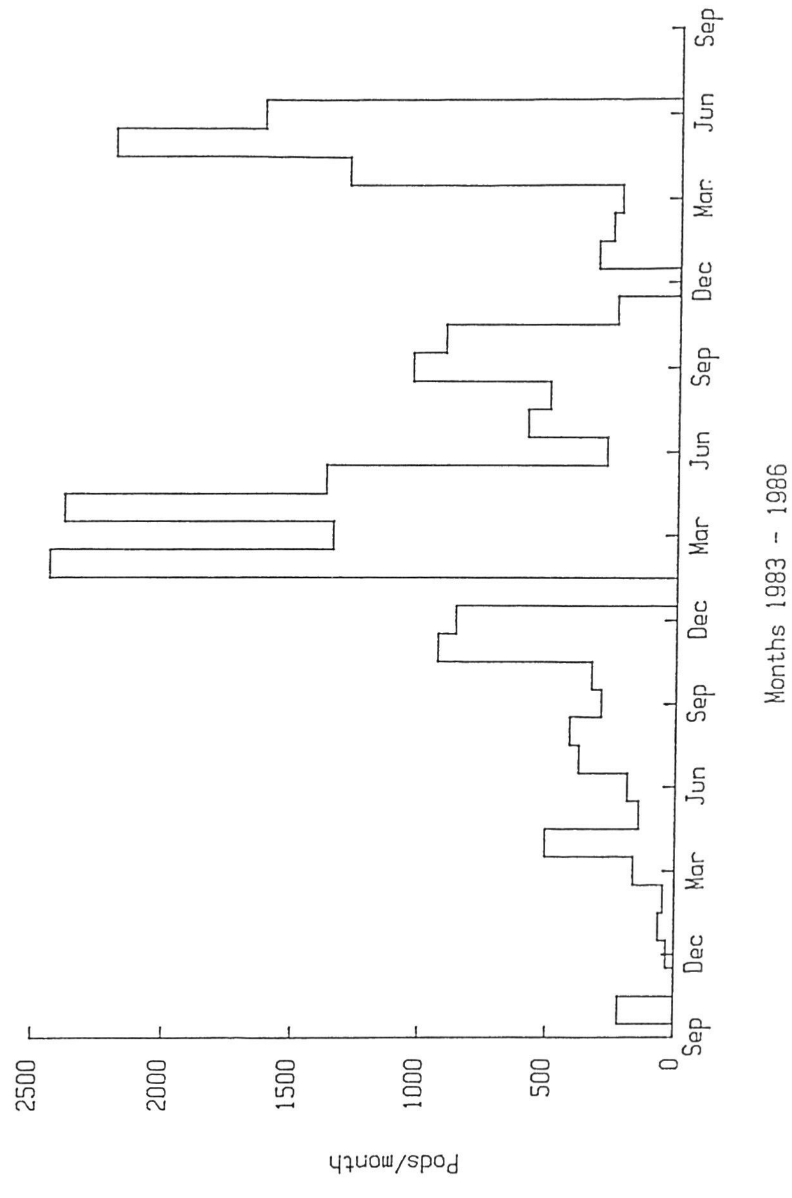


Figure 6.3 *Pod production 1983-1986. Total number of pods (including diseased pods) harvested in the San Carlos genebank each month, October 1983 - June 1986. From October 1985 to May 1986, data were collected by J.Baquero.*

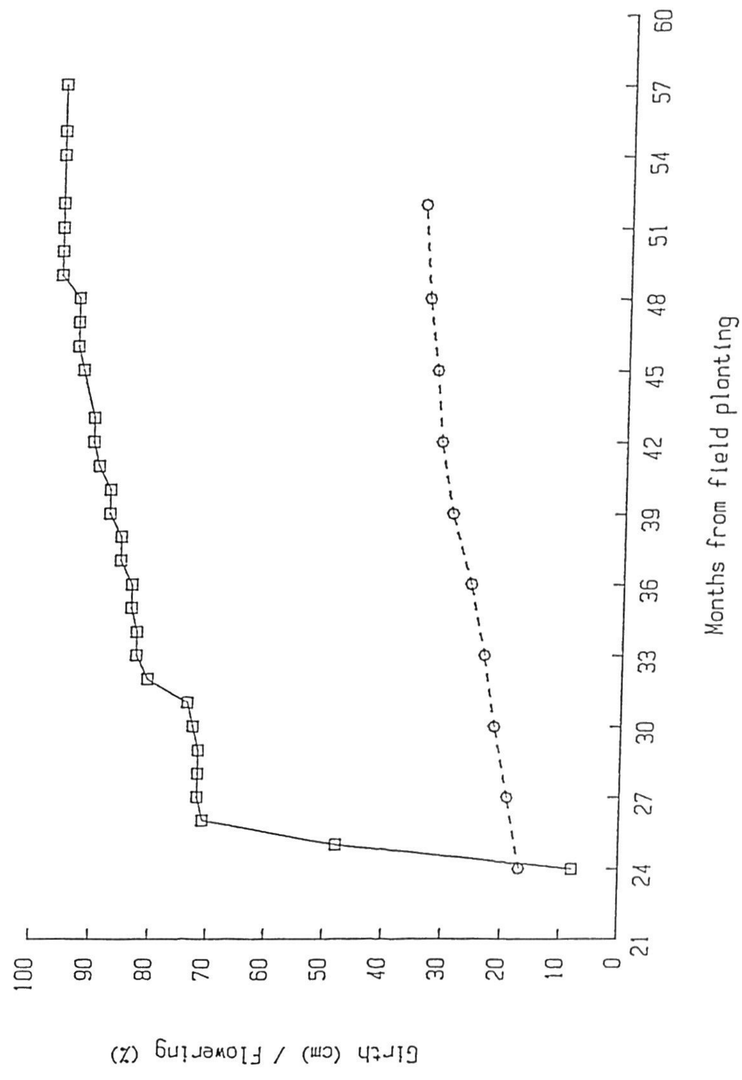


Figure 6.4 Tree size and initiation of flowering; genebank Block 1 means. Upper curve: percentage of trees which have been recorded as flowering at least once in current or earlier months; lower curve: mean girth of all trees in block. Total number of trees = 205 (1983); planted January 1981.

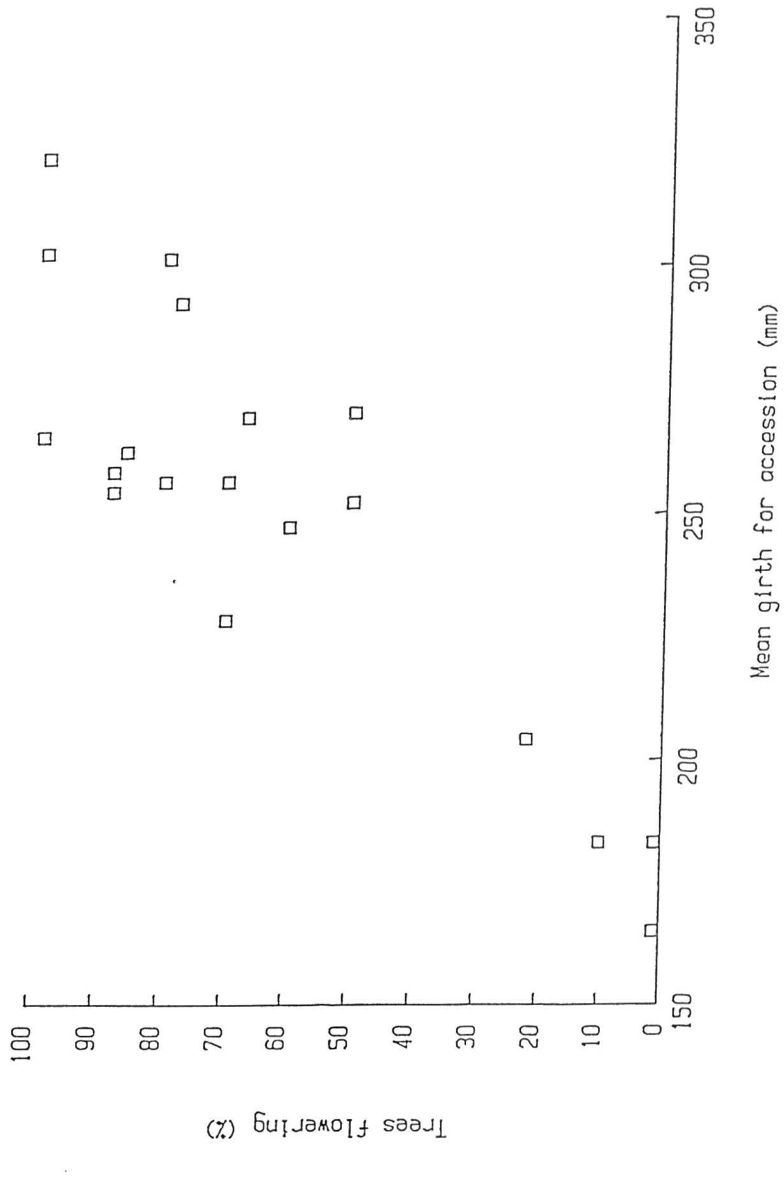


Figure 6.5 *Relationship between girth and proportion of trees flowering for 19 accessions in genebank Block 1, in January 1984. Number of trees per accession: 7-10; planted January 1981.*

Pollination and compatibility

The results of the controlled pollinations are shown in Table 6.1. They support the hypothesis that wild cocoa is self-incompatible.

Table 6.1 Results of controlled pollinations on LCT-EEN accessions.

(a) attempted self-pollinations (all unsuccessful)

1	74	120	150	223
46	80	125	151	228
60	91	132	156	214/s4
73	107	148	175	214/s5

(b) attempted cross-pollinations (female parent first)

Successful	Unsuccessful
44/s3 x 61/s4	46 x 91
60/s3 x 61/s4	46 x 223
91 x 228	46 x 228
120 x 46	60 x 46
132 x 133	60 x 80
152 x 46	60 x 91
	91 x 46
	133 x 1

6.4 Seed production

The numbers of pods seen at the time of collection on trees from which LCT-EEN accessions were derived are summarized in Table 6.2. Shade level is clearly a major determinant of pod production. This can be quantified by assuming figures of 5 pods/tree for 'few pods' and 15 pods/tree for 'many pods'. On this basis, trees growing in full sun had an average of 7 pods, trees in light shade had 4 pods and trees in dense shade had only 1 pod.

Table 6.2 *Pod numbers at time of collection on trees from which LCT-EEN accessions were derived.*

	No shade	Light shade	Dense shade
No pods	52	81	60
Few pods (1-10)	42	29	8
Many pods (>10)	51	22	2
TOTAL	145	132	70

Notes:
1. Data for LCT-EEN 1-373.

Pod production data for the study population are given in Table 6.3. There was considerable year-to-year variation, but the effect of shade level was consistent, with trees in heavy shade producing an average of 3 pods/year, and trees in light shade or full sun having 10 pods/year. It should be noted that these figures have not been adjusted for pods which might have developed and disappeared between recordings; given that dead pods persist for long periods, this should result in only a slight underestimate of pod production. Seedlings were not recorded at all before 1985, and could not be recorded at all in the area used for the genebank, but it appeared that there were very few seedlings, and no seedlings were found which had survived from previous years.

Table 6.3 Pod production in the study population, 1982-1986

(a) trees under forest canopy (heavy shade)	1982	1984	1985	1986
Pods/tree	6.7	0.0	3.9	0.3
Seedlings/tree	-	-	12.2	0.4
(b) trees exposed to sun (light shade or full sun)				
Pods/tree	10.0	10.5	16.8	3.2

Notes:
1. Pod and seedling numbers were recorded in April 1982, October 1984, May-June 1985 and July 1986.

6.5 Seed dispersal

The role of mammals

On Isla Anaconda, three species of monkey, spider monkey (*Ateles belzebuth*), howler monkey (*Alouatta seniculus*) and capuchin monkey (*Cebus albifrons*), were tame enough to be observed at close quarters. Most of the vegetation on the island was secondary forest, with some semi-abandoned cocoa plantings. Ripe pods were obtained from one of these plantations and nailed to tree trunks at about 1.5 m above ground level. The pods were of a typical cultivated type, having some red pigmentation and small to medium sized purple seeds.

All three species reacted immediately to the offered pods, wrenching them off the tree trunks and carrying them away. One capuchin monkey carried a pod up into the branches of a tree about 10 m away; holding the pod at both ends, the central part of the pod was battered against a branch until the pod broke into two pieces. The monkey then ate the seeds, during which process a few seeds fell to the ground. A howler monkey also carried a pod up into a nearby tree, but opened the pod by chewing at one end. Again, most of the seeds were eaten, but a few fell to the ground. All three species were also seen to open pods by hitting them on the ground, the seeds then being eaten.

These observations confirm, first, that cocoa pods are very attractive to monkeys, and second, that once removed from the tree, pods are likely to be taken some distance away before being opened. However, I had expected that the monkeys would eat the sweet pulp around the seeds and reject the seeds themselves, which are bitter; the observation that at least three-quarters of the seeds were actually consumed casts doubt on the role of monkeys in seed dispersal. Circumstances did not allow any check on the state of seeds passed through the monkeys' digestive tracts. Cocoa seeds are easily damaged, and even if swallowed intact it seems likely that they would not emerge in a viable condition.

From these observations, and from pods seen occasionally in the field, it appeared that pods opened by monkeys were completely disintegrated, as well as being removed from the tree. A much commoner form of damage was the opening of a hole, around 40-70 mm in diameter, in the side of a pod attached to a tree, with part or all of the contents being removed. In one small cocoa plantation near San Carlos, around 30% of the mature pods were damaged in this way (M. Wright and A. Linton, personal communication). The general belief among local people there and elsewhere was that this damage was caused by squirrels, which are common in the region; this could not be confirmed by observation, but there was no reason to doubt the explanation. Wright and Linton also found evidence, from measurements of toothmarks, that some pods growing near the ground had been opened by 'guatusa' (usually identified as *Dasyprocta punctata*, but possibly another species of *Dasyprocta* (D.Irvine, personal communication)), a caviomorph rodent common in the region; this would probably not occur in wild trees, which rarely carry pods near the ground.

It has occasionally been suggested that rivers might play some role in cocoa seed dispersal. Although I established, using a number of pods of varying origin, that intact detached cocoa pods do float, the likelihood is that pods which are removed from trees by monkeys will be broken open, and the seeds do not float. If a pod did float down a river, it would still have to be opened; although seeds do germinate in intact pods, the seedlings cannot develop unless the pod is opened within a week or two. Dispersal along rivers is not impossible, but it would require a highly improbable sequence of events to occur.

Seed shadows

The seed shadow maps are presented in Figure 6.6. The three maps show the positions of seedlings around 1-2 months after germination. In the *Theobroma subincanum* map, about two-thirds of the seedlings are under the canopy of the parent tree. Of the remaining seedlings, none are more than 8 m from the parent tree. In the two *T. cacao* maps, more than 95% of the seedlings are under the canopies of the parent trees and the remainder are no more than 1-2 m further away. In all three cases, ample time was available to search for seedlings and I am confident that no other seedlings were growing within a circle centred on the parent tree and extending at least 10 m beyond the most distant mapped seedling.

From these maps, it is reasonable to conclude that only a small proportion of seeds are dispersed more than 5-10 m from the parent tree. This conclusion is supported by experience on collecting trips; in six years of collecting, I never found a cocoa seedling in a place where there were no cocoa trees. Seedlings under the canopies of parent trees were found many times; only once (Río Curaray, 28.8.81) was a seedling found as much as 25 m away from a putative parent tree. Walking through the forest, it was not uncommon to see cocoa seedlings, then look up and realise that one was standing underneath a large cocoa tree. In the study population area at San Carlos, which was intensively explored over several years, only one group of seedlings was found more than a metre or two away from a parent tree, and these were less than 15 m from a possible parent tree. These observations are consistent with the hypothesis that squirrels are commonly responsible for opening cocoa pods; studies of seed dispersal in temperate nut-bearing trees suggest that squirrels usually carry nuts only 15-30 m before burying them (Lanner 1985).

At the trees where the seed shadows were mapped, there were no seedlings surviving from previous years. However, of 57 seedlings mapped around tree W-26 on 29.5.85, and marked with plastic plant labels, 5 were still alive on 25.7.86.

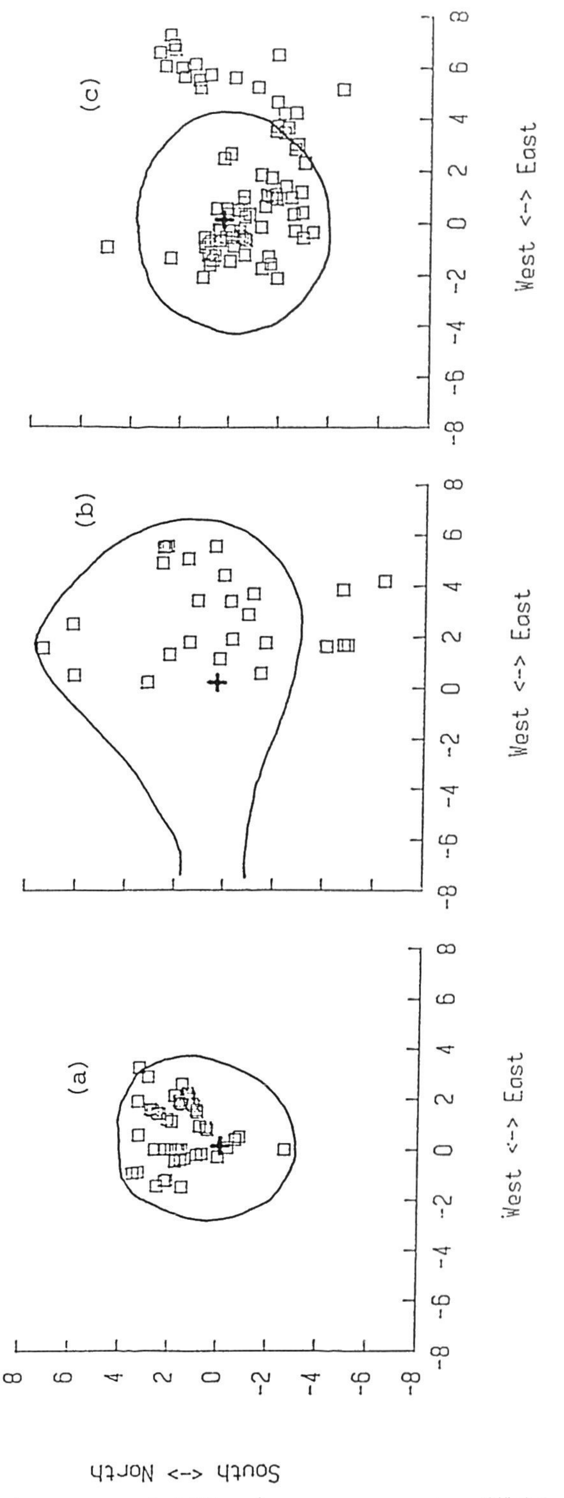


Figure 6.6 *Seed shadow maps. (a) Theobroma cacao, tree W-26, San Carlos, 29.2.85; (b) T. cacao, tree W-29, San Carlos, 1.8.82; (c) T. subincanum, San José de Payamino, 25.2.84. In each map, the parent tree is marked (+); the outline indicates the extent of the parent tree's canopy.* Scales in metres.

6.6 Seedling survival

In all three seedling survival experiments, differences in survival rates between 'treatments' *within* experiments were small and statistically insignificant. Further interpretation requires caution, since there are many differences in starting conditions *between* experiments. The survivorship curves for Experiment 1 are at first sight quite different to those for Experiments 2 and 3 (Figure 6.7). Rather than seeking to explain this difference, I would draw attention to the similarity in mortality rates after the first 90 days of the experiments (Figure 6.8). This highly variable early mortality is partly explained by an analysis of the causes of death across all three experiments (Table 6.4).

Table 6.4 Causes of death recorded for seedlings in seedling survival experiments

Probable cause of death	Number of seedlings	(%)
Mammal damage:		
stem severed	38	(32)
cotyledons eaten	5	(4)
pulled out of ground	3	(3)
Insect damage:		
defoliation	6	(5)
Fungal attack		
and general attrition	8	(7)
Crinipellis perniciosa	1	(1)
Mechanical damage		
hit by falling Cecropia leaf (1 also with Crinipellis)	3	(3)
hit by falling Iriartea leaf	2	(2)
hit by unidentified leaf	1	(1)
entangled in roots of shrub	1	(1)
No cause of death ascertained	49	(42)

The greatest single cause of death was damage by mammals, most often by severing the stem of the seedling with a single cut below the cotyledons. This damage was also seen in the genebank nursery; from the type of damage, it was thought that the animals responsible were probably rabbits (*Sylvilagus brasiliensis*) which were common in the area. The motivation for this behaviour is unknown, since no part of the seedlings was actually consumed, the severed part being left on the ground nearby. The incidence of mammal damage in the transects was somewhat variable, with some transects escaping altogether and others being almost wiped out. After 2-3 months, seedlings apparently no longer attract this type of attack, and are then subject to steady attrition by insects, fungal pathogens and large falling leaves (for example, much damage is caused by the leaves of *Iriartea* palms, which are 3-4 m long and weigh more than 20 kg).

Seedling growth rates, from stem diameter measurements, were generally similar across all transects (Figure 6.9). All the transects were in dense shade, and these seedlings grew very slowly by comparison with seedlings grown in the nursery and genebank under light shade (compare Figures 6.4 and 6.9). On the basis of the growth rates in Figure 6.9, many years of growth would be required to reach flowering size (as stated in Section 6.2, flowering does not begin until a stem diameter of about 60 mm is attained).

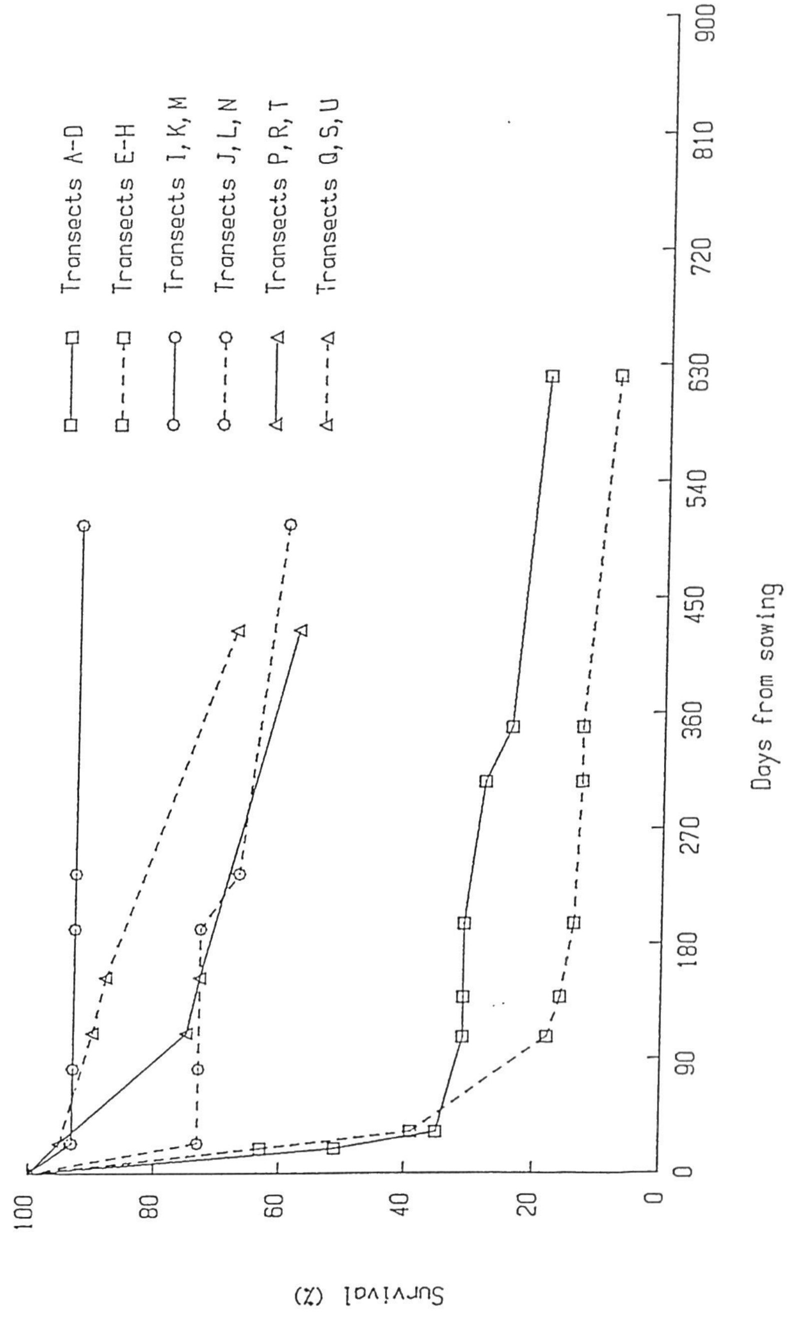


Figure 6.7 *Seedling survival; the proportion of seedlings surviving in each group of transects for the experiments listed in Table 5.1, at irregular intervals 0-630 days from sowing.*

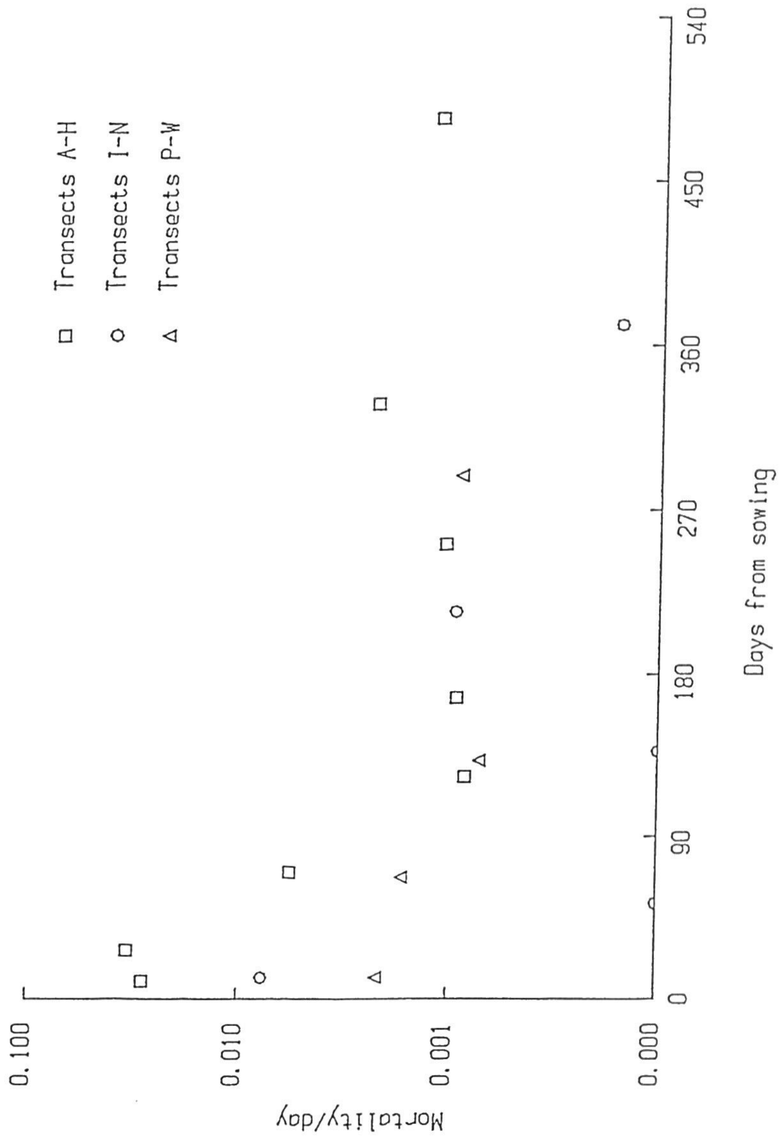


Figure 6.8 *Seedling mortality; the mean daily mortality rate for each group of transects for the experiments listed in Table 5.1, calculated from the survival curves (Figure 6.7).*

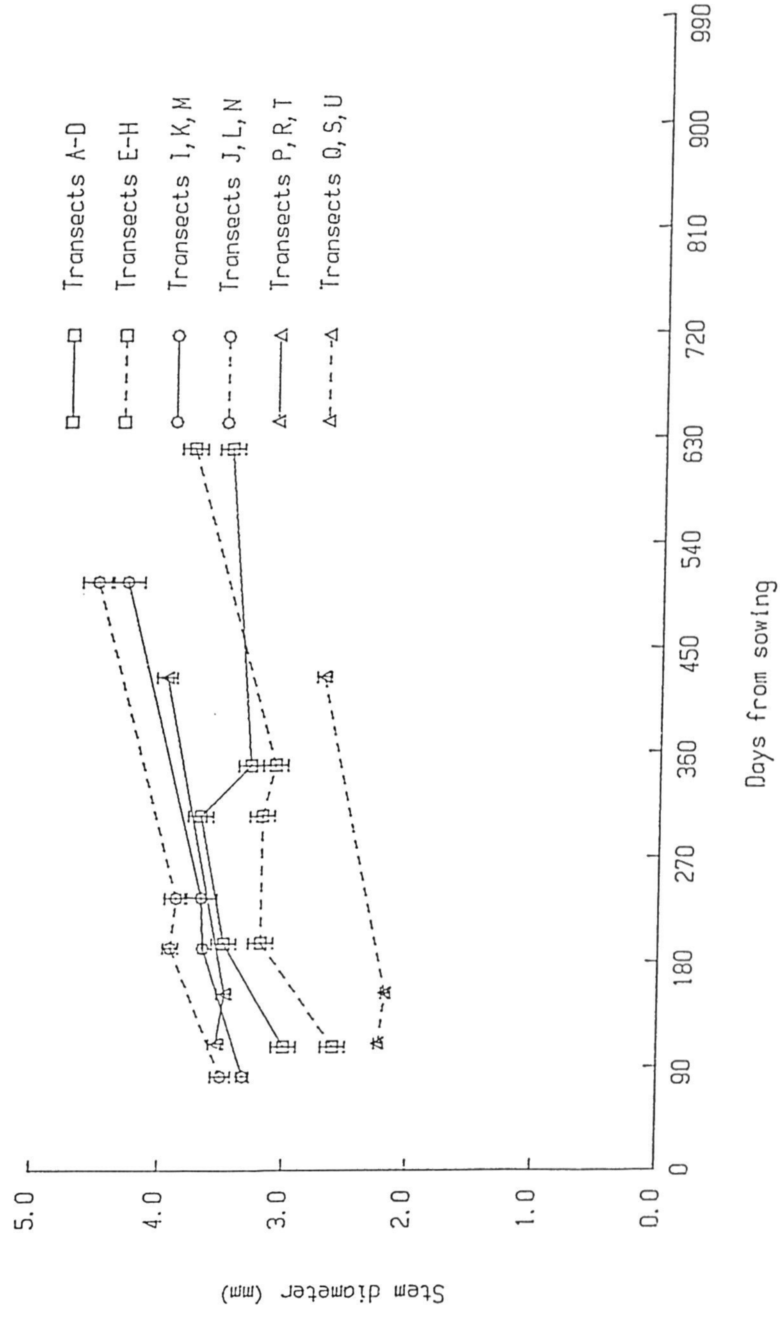


Figure 6.9 *Seedling growth rates; mean stem diameters (± 1 standard error) in each group of transects for the experiments listed in Table 5.1, at irregular intervals 0-630 days from sowing.*

6.7 Adult survival and vegetative reproduction

No deaths were observed in the study population over the period 1983-1986, when the population was subject to close monitoring. Furthermore, there was no evidence of tree deaths in the period 1980-1983, although some parts of the study area were not explored until the later part of this period. It is clear that adult tree mortality is low; this statement can be quantified by calculating the probability of observing zero deaths given various levels of average long-term mortality. The probability, P, of observing at least one death is given by

$$P = 1 - (1 - m)^{nt} \qquad (6.2)$$

where m = mean annual mortality (proportion of trees which die each year)
n = number of individuals in study population
t = period in years

For $P = .5$, $n = 50$ and $t = 3$, m works out at .0046; in other words, with mortality at this level, there would be an even chance of observing at least one tree death during the study period. With mortality above this level, one would expect to have recorded deaths in the study population. This provides a rough upper limit to adult mortality; the true mortality may well be much lower.

This low estimate of adult mortality in a wild population accords with the observable fact that cocoa trees are not easily killed. When a cocoa tree is cut down, the stump immediately starts to sprout, and several new trunks form within a year or two. Similarly, when part or all of a cocoa tree falls over, shoots appear at intervals along the fallen trunks, and these root adventitiously. The result may be regrowth of the original tree, or the development of a clump of two or more trees 5-10 metres apart. This type of vegetative reproduction may be very common. One tree in the study population (W-7) fell over in 1980 and by 1985 had formed a vigorous clump with 17 stems of 100 mm girth or greater (one stem had a girth of 908 mm); one pair of trees (W-15 and W-16), about 5 m apart, appeared to be joined at ground level by a decaying trunk when first recorded.

In the genebank, some mortality (less than 0.5% per year) was caused by fungal infections affecting the roots and the lower part of the trunk, but this type of infection was only seen once in a wild tree (Villano, 17.7.81). The ubiquitous *Crinipellis perniciosa* infection did not cause the death of infected trees. A few trees in open sites were covered in 'mistletoes' (probably *Loranthus* spp), but there was no sign of trees dying as a result; in any event, mistletoe infestation was not found in trees growing in their natural habitat.

This evidence of the longevity of wild cocoa trees accords with the generally accepted view that cultivated cocoa does not have a finite lifespan, although old plantations normally have very low yields. Cocoa plantations have survived since the eighteenth century or earlier near Belém (Brazil) and along the Río Guaviare in Colombia (B.G.D.Bartley, personal communication; Rojas-Peña 1951; Barros 1981).

6.8 Population age-structure

The girth data were used to establish a relationship between growth rate and tree size. The total basal area for each tree was calculated for each measurement date and used to calculate the basal area increments in 1984-85 and 1985-86. These data were then fitted to the logistic growth equation:

$$dx/dt = rx(1 - x/K) \qquad (6.3)$$

where x = basal area (cm^2)
dx/dt = basal area growth rate (cm^2/year)
r = rate constant
K = maximum (asymptotic) for basal area

Within each year, the fit was good, although the estimated K values for the two years were quite different; the asymptote cannot be reliably estimated without data from more large trees. The fitted r and K values are given in Table 6.5, and the resulting age-basal area curves are plotted in Figure 6.10. Given that the purpose of this exercise was to obtain approximate estimates of relative age, I decided to use the 'combined data' curve, based on mean values for r and K, to calculate ages. On this graph, zero age refers to the age at which the basal area was equal to 10 cm^2, so that these ages do not include the period required to reach this size. This curve was used to calculate ages (a) for 48 wild trees at San Carlos and (b) for 30 wild trees from which LCT-EEN accessions were collected, at a variety of sites in Napo Province (girth data were not generally obtained on collecting trips because measurement on large multi-trunked trees was too time-consuming). From Figure 6.11, it is clear that the two age distributions are broadly similar, with few young or very old trees. The median age for both groups was 75 years.

In the absence of corroborative data, the calculation of ages from basal areas must be treated sceptically. There did not appear, unfortunately, to be any other way of estimating ages; the growth rings seen in cocoa trunks appeared to bear no relationship to age, which was not surprising in view of the lack of seasonality in the region's climate. In even-aged single-species stands, size distributions typically reflect the effect of competition on the growth of individuals, giving the illusion of a population of variable age. In the present case, intra-specific competition was generally minimal and it is reasonable to suppose that the size distribution does reflect the presence of trees of different ages in the population. Although the absolute ages estimated by this method should be treated as approximate, the lack of younger and older trees in the population does not appear to be an artefact. The similarity of the age/size distributions for the two groups of wild trees is notable, and provides evidence that processes observed in the study population are typical of wild cocoa populations in the region.

Table 6.5 *Parameter estimates for tree growth equation*

Period	r day^{-1}	K cm^2
October 1984 - June 1985	.0431	7329.9
June 1985 - July 1986	.0642	1075.4
Mean values as used	.0537	4202.7

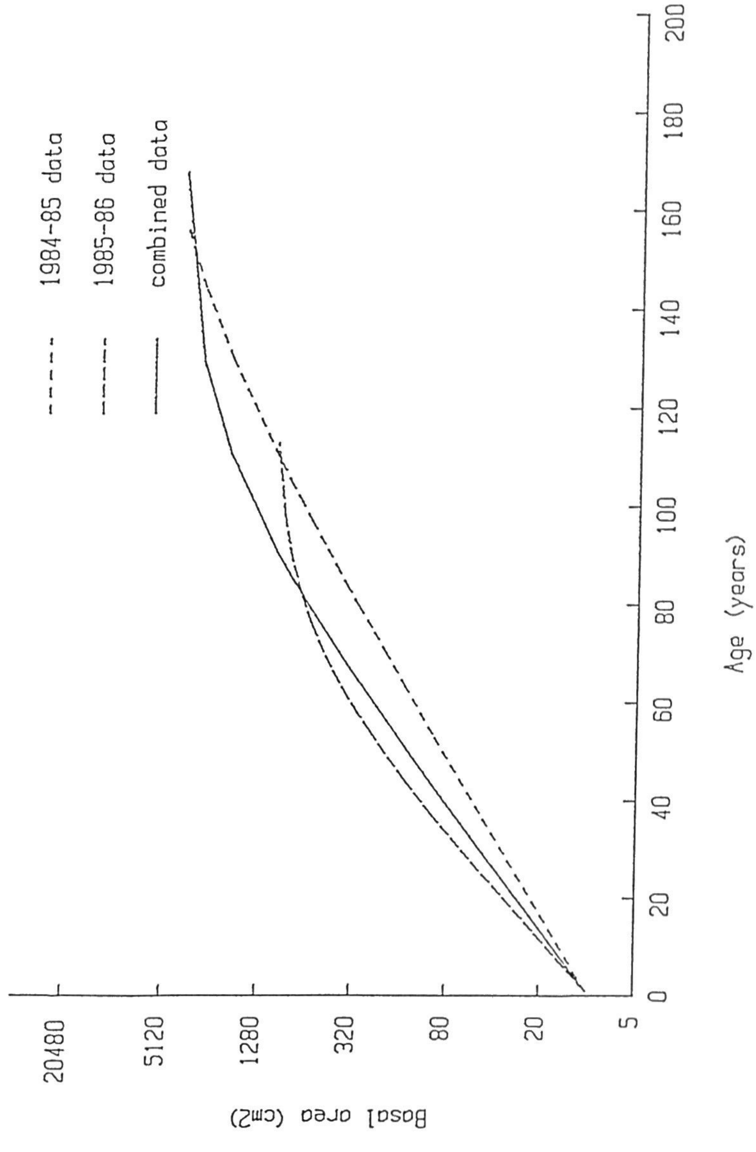


Figure 6.10 Calculated relationship between age and basal area for trees in the study population, using Equation 6.3 and the parameter estimates of Table 6.5.

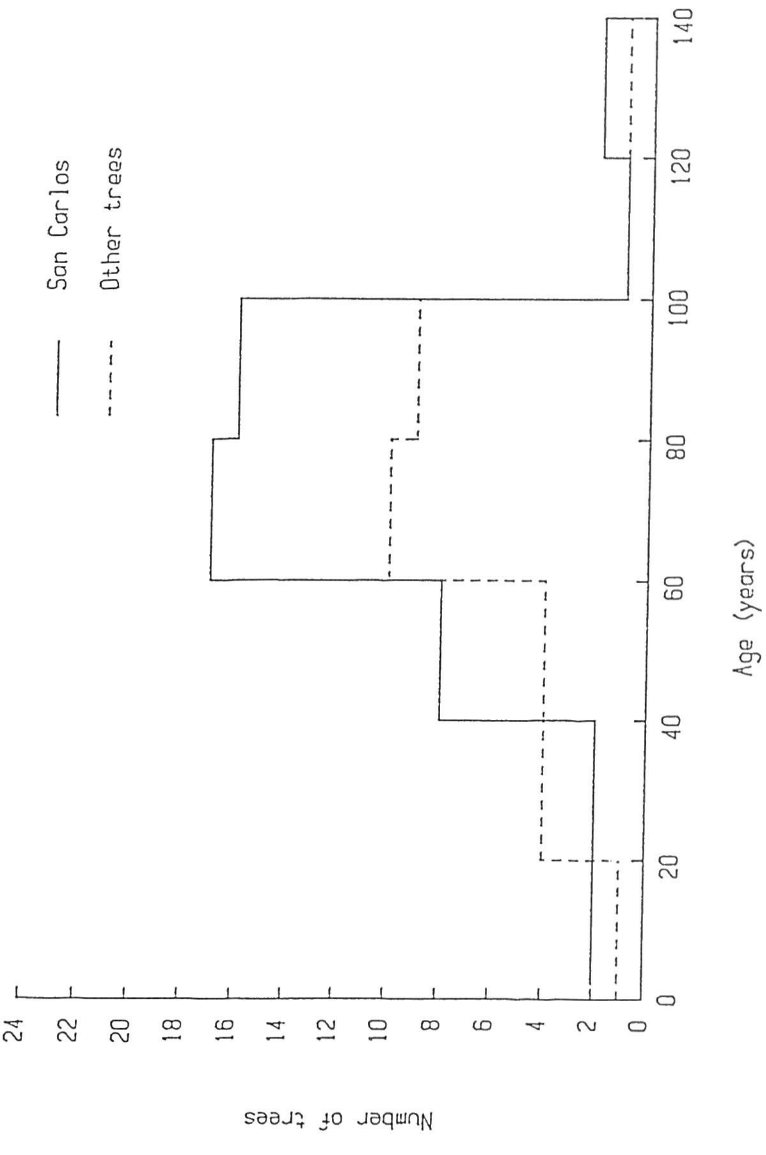


Figure 6.11 *Calculated age distributions, based on the basal area-age relationship of Figure 6.10 ('combined data') for (a) 48 wild trees at San Carlos and (b) 30 other wild trees in Napo Province (original trees for LCT-EEN accessions).*

7 Discussion

In this chapter, the results are discussed under three main headings. First, I consider the implications of the results of Chapter 4 in relation to questions about collecting strategy for cocoa germplasm. In the second part, I make use of the results of Chapter 6 to draw some general conclusions about reproduction and population structure in wild cocoa populations. Finally, I go on to discuss some broad issues arising from this study, including the question of the wild status of cocoa and the possible existence of a centre of origin.

7.1 Collecting strategy

The collecting strategy used in LCTAP involved subdivision of the target region at a scale of about 50 km, corresponding to one to three collecting trips to each of the Level 2 areas listed in Table 3.2. This strategy was determined by previous collectors' descriptions of wild cocoa populations, notably the observations of Pound (1938, 1943, 1945), who drew attention to the occurrence of variant types of cocoa he found in a succession of Upper Amazon tributaries above Iquitos. The proposed strategy was also designed to accord with the funding and timescale for LCTAP. A preliminary data analysis at the end of the first three years of the project (Allen and Lass 1983) suggested, in general terms, that there was little variation across the target region in Ecuador. On this basis, it appeared that this first-stage collecting provided an adequate sample of variation in the target region, so that the additional collecting during LCTAP Phase 2 did not involve exploration of more closely-spaced localities.

The results presented in Chapter 4 enable a fundamental reappraisal of the collecting strategy used during LCTAP. Since LCTAP Phase 2 also involved some collecting in Colombia, this reappraisal covers variation over a much larger area than the preliminary data analysis referred to above, and more descriptor data are now available.

Choice of target region

The Amazon Region of Ecuador was selected as the target region for LCTAP for three main reasons. First, the national agricultural research institute in Ecuador, INIAP, had a long history of involvement in cocoa research, and wished to participate in germplasm collecting in the Amazon. Second, with the construction of new roads into the Ecuadorian Amazon in the 1970s, it was considered necessary to collect cocoa germplasm in areas where extensive forest clearance was taking place; improved access to the region also made collecting easier, particularly away from the rivers along which earlier collectors had travelled.

Third, it was known from experience in West Africa and elsewhere (Toxopeus 1972; Wood and Lass 1986) that Upper Amazon cocoa germplasm was likely to be useful in breeding programmes; it was also believed, from Pound's expedition reports (cited above) and from collections made by Chalmers (1969-1974), that there existed in Ecuador a population of wild cocoa quite distinct from the Upper Amazon populations collected by Pound. It was hypothesized that germplasm from Ecuador would contribute additional variability to cocoa breeding programmes, with increased likelihood of finding genotypes having desirable characteristics, and new possibilities for exploiting heterosis via inter-population crosses. It should be said here that this hypothesis cannot be tested until the LCTAP material is actually used in breeding programmes, and, with cocoa germplasm, this is likely to take several decades; meanwhile, discussions of germplasm collecting strategy have to assume that the end product will ultimately be of value.

The Río Caquetá target region was selected by the Colombian agencies which mounted the 1984 international collecting expedition there. The Colombian national agricultural research institute (ICA) planned to collect cocoa germplasm systematically throughout Colombia, and the Río Caquetá region provided an appropriate field of activity for a large-scale collecting expedition, when Colombian government funding for such an expedition became available. The comments made in the previous paragraph about the potential value of Upper Amazon germplasm are also relevant here.

Delineation of target populations

Given the approach to collecting strategy outlined in Section 1.2, application of Marshall and Brown's (1975) criterion for collectable 'common' alleles (those of frequency >.05) required delineation of populations within the target region. Two distinct approaches were used here. For an initial view of the descriptor data, the assumption was that populations existed which coincided with geographic subdivisions of the target region. Frequency distributions for individual descriptors in six geographically-defined groups of accessions were presented in Section 4.3. They demonstrated the broad similarity of accessions from different areas within Ecuador and the existence of two distinct populations within the EBC accessions from the Río Caquetá in Colombia. The frequency distributions also showed a substantial resemblance between the Pound accessions from Peru and the Ecuador material, while identifying some descriptors on which these groups were clearly distinguished.

The complementary technique of cluster mapping did not start by assuming the existence of distinct subpopulations in different areas; instead, it divided accessions into morphologically distinct groups, using cluster analysis, and then tested the hypothesis that these groups were spatially separated within the target region. The cluster maps drew attention to the broad differences between the Ecuador, Caquetá and Pound accession groups, while producing very little evidence of geographically-related variation within Ecuador. Within the Caquetá material, although the frequency distributions showed two distinct types, the cluster maps showed that the distributions of the two types overlapped.

Either approach can be used to justify the view that the Ecuador accessions represent a single population for sampling purposes. This population is found throughout an area of some 120,000 km^2, extending about 400 km in an east-west direction and 500 km in a north-south direction. Although the frequency distributions indicate some differences between areas within Ecuador, the cluster mapping suggests that morphological variation is not geographically related. In practice, this means that a sample of a suitably large number of trees from, say, Napo Province will contain the same alleles as a sample from Zamora Province.

Differences in allele frequencies between the two areas are not relevant to discussions of collecting strategy, so long as the same alleles are present above the .05 frequency threshold in both areas.

Among the accessions from the Río Caquetá, the overlapping distributions of the two pod types mean that this area should also be treated for sampling purposes as a single population, albeit a strongly polymorphic one. This population extends over an area of at least 300 km (east-west) by 100 km (north-south), although it should be noted that trees appeared to be found only on narrow strips of land along the Caquetá and Apaporis rivers.

The Pound accessions were not collected within LCTAP and were not intended as a representative sample of wild cocoa within a defined target region. The data available from the Pound accessions cannot reliably be used to delineate populations, but they do indicate a clear discontinuity between the Ecuador population and a population (or populations) along the Río Marañon and its immediate tributaries in Peru.

Both the frequency distributions for individual descriptors and the cluster mapping provide evidence that the Ecuador, Caquetá and Pound populations should not be lumped together for the purpose of evaluating sampling strategy. The Caquetá target region is about 350 km to the east of Ecuador; a few accessions collected from the Río Putumayo, along the Ecuador-Colombia frontier, resemble the most widespread Caquetá type, so it is tempting to speculate that trees resembling the Caquetá population are found across a much wider area than was explored during the Caquetá expedition. On the other hand, Pound (cited above) described trees similar to the typical Ecuador accessions on the Río Orteguaza, a tributary of the Río Caquetá.

It is safer to accept that this approach to delineating populations within a target region does not lead to any conclusions about the distribution of the identified populations outside the target region. On the other hand, I suggest that it is reasonable to extrapolate from the target regions considered here to other potential target regions of comparable size, topography and history. This leads to the working hypothesis that target regions of comparable size to those considered here (say, around

100,000 km2), in the Upper Amazon, should be treated as single populations for sampling purposes.

Numbers of trees sampled

The calculation of minimum sample sizes was reviewed in Section 1.2. From Equation 1.7, it follows that even in the 'worst case' of a locus with four alleles at frequency .05 (Table 1.3, case (e)), a sample size of 43 trees is sufficient to achieve 95% probability of capturing all four alleles. This sample size was comfortably exceeded for the Ecuadorian population, with 281 accessions established at San Carlos (see Table 4.1). For the Caquetá population, an adequate number of trees was collected (136 accessions), but only part of the material was taken to San Carlos. The 34 accessions which were established there do not by themselves constitute an adequate sample of the population; however, further accessions were established at the ICA genebank at Palmira, Colombia. Reversing the calculation (that is, using Equation 1.6), one may say that for the actual number of LCT-EEN accessions at San Carlos, the probability of failing to capture the hypothetical four alleles of the 'worst case' for the Ecuador population is less than 10^{-11}, so that the 'capture probability' exceeds 99.9999%. For the Caquetá population, the capture probability works out at 88.3%.

If, for the sake of argument, the three Level 1 areas in Ecuador had each been treated as a separate population for sampling purposes, then the numbers of accessions established (Napo: 126, Pastaza: 64, Zamora: 91) would correspond to capture probabilities of at least 99.999%, 99.4% and 99.96%.

Some reservations

Although the results of these calculations appear satisfactory, several reservations should be noted. First, Marshall and Brown's criterion for collectable 'common' alleles was to some extent fixed on the basis that it led to feasible sample sizes. If sample sizes in the present project appear adequate in relation to this criterion, then the LCTAP collecting programme may be judged effective in relation to conventional germplasm collecting practice. However, one must consider the possibility that in practice the cocoa alleles of value to breeders may fall below the .05 frequency threshold. In this context, it is encouraging to note that for the Ecuador population (the principal target of LCTAP), there is a large safety margin in the number of samples collected; the actual sample of 281 accessions exceeds the calculated sample size (218) required for capture of four alleles even at frequency .01 (Table 1.3, case (g)).

Second, the capture of a single copy of each of the 'common' alleles may be a valid minimum target, but it may not be adequate for long-term conservation of the genetic diversity of a wild population. For short-lived species which have to be raised from seed at regular intervals to maintain a germplasm collection, the loss of variability in a small captive population may be an acute problem (Gale and Lawrence 1984). For a long-lived perennial species which can be propagated vegetatively, such as cocoa, it would probably be more cost-effective to replicate the collection at two or more sites than to collect more material with the aim of capturing additional copies of alleles. The replication of a cocoa collection by conventional methods (budding or grafting) is in principle perfectly feasible, although it requires experienced staff and close cooperation between genebanks over several years, which has in practice been difficult to achieve.

The third problem concerns the lack of information on allele frequencies in cocoa populations. Isoenzyme data from genotypes of wild origin would provide estimates of the mean number of alleles per locus and of heterozygosity. The calculations of minimum sample size assume that there are four or more alleles to be captured at each locus, yet some surveys of enzyme polymorphism (for example, Jain 1975b) have suggested typical values, for a range of species, of only 1.2-2.6 alleles per locus. The calculations also assume that each accession provides two alleles for each locus which are drawn at random from all alleles present in the population for that locus; again, it is possible that this assumption is too pessimistic. Because of the unusual features of the life history and population dynamics of cocoa, discussed later in this chapter, application of typical values may be inappropriate. The isoenzyme data for cocoa published to date (Amefia *et al* 1984; Atkinson *et al* 1986) are limited, as techniques for cocoa are still under development.

In the absence of specific gene frequency data, an objective assessment of variation in wild cocoa populations must depend primarily, as here, on morphological descriptors. These have the advantage that populations described in terms of these descriptors can be related to numerous published accounts of wild and cultivated populations (although descriptions found in the literature are often vague about crucial characters such as flush colour). Some of these descriptors (for example, seed size) are also of direct agronomic significance, which is unlikely to be true of any isoenzyme character. Set against these advantages is the risk that some conspicuous, easily assessed characters may be over-emphasized as measures of overall similarity or difference; this possibility was discussed by Witcombe and Gilani (1979), who concluded that for cereal germplasm collecting it was preferable to explore outside the Vavilovian centres of diversity, which had been delineated on the basis of a small set of qualitative characters controlled by major genes. For cocoa, there would appear to be a need for additional quantitative descriptors related to agronomic performance, with the proviso that such descriptors must be capable of assessment in a cost-effective way.

Allocation of resources

Calculating required numbers of samples for defined populations leaves a range of options for planning collecting expeditions. In particular, the question arises as to whether available resources should be used to maximise the number of areas or 'sites', or to increase the number of plants collected per site. Where data on costs and variability are available, Equations 1.7 and 1.9 enable calculation of an optimal solution. The conventional view (see Section 1.2) is that the marginal cost of additional accessions from a site will be considerably less than the initial cost of visiting a site; in other words, the cost ratio c_1/c_2 is much greater than 1. This should more than offset the fact that additional accessions from the same site would normally contribute less variability than accessions from additional sites; on balance, applying Equation 1.9, it will be most cost-effective to collect a number of samples at every site visited. In some cases, the optimum number of accessions per site will be quite large (Table 1.4).

Extending the method to cover two levels of sampling (Level 1 areas, Level 2 areas, trees), the appropriate variance component ratios were calculated in Section 4.6. The corresponding cost ratios can be estimated from a breakdown of LCTAP expenditure. The total direct cost of LCTAP was about £130,000. This includes UK funding (£110,000) and additional staff costs funded by INIAP (£20,000), although it does not include additional administrative costs incurred both in the UK and in Ecuador, which would be very difficult to quantify. For present purposes, it is assumed that exploration, collection, establishment of collected material and documentation accounted for half of total LCTAP expenditure. The remaining expenditure covered 'post-collection' costs, including genebank maintenance, transfer of plants or budwood to quarantine, and the recording of yield and pod infection for evaluation purposes. The full cost breakdown is given in Table 7.1.

For sampling Level 2 areas within Level 1 areas, the cost ratio (c_1/c_2) and the variance component ratio (V_2/V_1) are approximately 1, so that the optimum n_2 is also 1 (from Equation 1.9); in other words, there is no economic advantage to sampling from more than one Level 2 area within each Level 1 area. This result is not especially remarkable given the observed pattern of geographical variation and the crude estimation of collecting costs for Level 1 and Level 2 areas.

For sampling individual trees within Level 2 areas, the median variance component ratio (V_3/V_2) was about 5. If collecting costs only are considered, then the cost ratio (c_2/c_3) is also about 5 and the optimum number of samples (n_3) is 5 trees per Level 2 area. Experience suggests, however, that it is quite misleading to ignore 'post-collection' costs, which for LCTAP, as noted above, constituted around half the total budget for the collecting project. When post-collection costs are included, the cost ratio reduces to 2.3 and the optimum n_3 is reduced to 3.5 trees per Level 2 area.

These figures should be regarded as approximations, from which two firm conclusions may be drawn. First, from a theoretical economic viewpoint, it is clear that sample sizes could have been much smaller and the spare resources used to collect over a wider area. In practice, other factors entered into the concentration of resources on collecting within

Ecuador. The project was based in Ecuador and costs would have been much higher if extensive collecting had taken place outside Ecuador. More fundamentally, intensive, systematic coverage of the original target region was necessary for the delineation of populations (discussed earlier in this section). The analysis presented here would not have been feasible if sample sizes had, in advance, been reduced to the bare minimum. Nonetheless, if the principle of extrapolation to other areas is accepted, future cocoa germplasm collecting projects may be able to deploy their resources over relatively larger areas.

Secondly, the inclusion of post-collection costs radically alters the calculation of optimum sample sizes. Post-collection cost data for LCTAP cover a period of six years only for the earliest accessions (collected in 1980), and do not include any of the very heavy costs associated with the construction and running of quarantine facilities; for example, a new quarantine house was built in Barbados specifically to receive LCTAP material, with a full-time technician. In addition, full-scale evaluation of accessions at breeding centres may involve five to ten or more years of data recording. I am aware of no published figures for long-term post-collection costs, but it is obvious that the further ahead one looks, the smaller will be the cost ratio c_2/c_3, as investment in individual accessions accumulates over time. In my opinion, this casts serious doubt on the conventional use of cost ratio calculations to justify the acquisition of large numbers of relatively similar accessions.

Table 7.1 Estimation of cost ratios for LCTAP

ITEM	COST	CALCULATED AS
(a) total direct LCTAP expenditure	130,000	
(b) 'collection' costs: expenditure on collecting and establishment of 280 LCT-EEN accessions in genebank	65,000	50% of (a)
(c) 'post-collection' costs: maintenance, transfer to quarantine, evaluation and related activities 1981-1985 total	65,000	50% of (a)
(d) per accession	232	65000/280
(e) collection cost per Level 1 area	1,000	estimated cost of including Level 1 area
(f) collection cost per Level 2 area	1,000	estimated cost of additional Level 2 area
(g) marginal collection cost per accession	200	{(b) - (3x(e) + 6x(f))}/280
COST RATIOS	RATIO	
(h) c_1/c_2	1.0	(e)/(f)
(i) collection costs only c_2/c_3	5.0	(f)/(g)
(j) including post-collection costs c_2/c_3	2.3	(f)/{(g) + (d)}

Notes:

1. Collection costs and the number of accessions do not include the expedition to the Río Caquetá, which was funded by the Colombian government.
2. Post-collection costs have not been adjusted for the small number of EBC accessions.

7.2 Population biology of wild cocoa

Reproductive strategy

The observations presented in Chapter 6 provide a clear picture of the reproductive 'strategy' of *Theobroma cacao* in its natural habitat. On the one hand, the pod production data indicate that trees growing under forest conditions may produce as few as 1-3 pods per year. At 28 seeds per pod (Table 4.2), this corresponds to only 30-90 seeds per tree. Cocoa seeds are 'recalcitrant', remaining viable for only 3-4 weeks even under the most favourable conditions, so that there is no soil seed bank. The seeds have to be removed from their pods by animals at the correct time for germination; it appears that both monkeys and squirrels may open pods, but that monkeys, at least, may destroy a high proportion of seeds in the process. Observations on seed shadows and seedling survival further indicate that most seeds germinate within less than 10 m of the parent tree, and that these seedlings, often growing in groups, are particularly susceptible to destruction at an early stage by rabbits. The result of all these factors is an extremely low rate of reproduction by seed.

Vegetative reproduction was observed, but this was also a rare event, with only a single case seen in a population of about fifty trees over a six year period.

These observations might suggest the hypothesis that cocoa is a maladapted species, incapable of maintaining itself in this habitat and likely gradually to disappear without human intervention. However, observations on adult mortality and on the longevity of cocoa trees suggest that the death of an established tree is also an extremely rare event. The multiple-stemmed habit of cocoa trees, combined with the ability to produce adventitious shoots from cut or fallen stems, makes them resistant to most kinds of mechanical damage, while the fungal root pathogens which affect trees in plantations appear to be rare in wild populations. This means that the observed rate of recruitment into the study population, though low (4 trees less than 40 years old), is still adequate to maintain or even increase the population over the long term.

It may be helpful to categorize *T. cacao* as an example of a K-adapted species, in the sense used by Southwood (1981). Such species are found in stable habitats, and produce small numbers of offspring with a relatively large parental investment in each. The adults are typically long-lived, so that the low fecundity is balanced by low mortality. For cocoa, this description seems accurate although it is not actually known what proportion of net annual production, under forest conditions, is accounted for by pod and seed production. Whitmore (1984, Sections 6 and 7) discusses reproductive strategies in rain forest trees without recourse to the concept of r and K-adapted species, by distinguishing between 'pioneer', light-demanding species and shade-bearing species. Pioneer species, which colonize large gaps, are characterized by high rates of seed production, efficient seed dispersal, the existence of soil seed banks, and very rapid height growth in young seedlings. By contrast, shade-bearing species germinate and grow, albeit slowly, under forest floor conditions; they include (but are not limited to) species whose seeds germinate immediately, often in the vicinity of the parent tree. Most shade-bearing species do require increased light to develop into large trees; the point is that they can survive for many years while waiting for suitable conditions.

Cocoa is most definitely not a pioneer species; wild cocoa was never found in association with trees such as *Cecropia* or *Ochroma lagopus* (balsa), and slow-growing cocoa seedlings would be hopelessly outcompeted by seedlings of such pioneer species. Nonetheless, it seems likely that gaps are important to cocoa in two ways. First, higher light levels result in much increased pod production (Tables 6.2, 6.3). Where wild cocoa trees had been left uncut in land cleared for pasture, it was evident that the trees would continue for several years to produce ten or more times as many pods as nearby trees in forest. This could also occur when a tree in forest was exposed by river-bank erosion, although such trees would usually fall into the river as the bank continued to erode. Heavily-fruiting trees on river banks were often very conspicuous, which may partly account for published reports that wild cocoa is typically found along rivers. Second, measurements of seedling growth (Figure 6.9) under forest floor conditions imply that several decades of growth under these conditions would be required to reach flowering size. A modest increase in light levels, as might be produced by a small canopy gap, would increase rates of growth to levels seen in the genebank (Figure 6.4), where flowering size was reached in less than four years from germination.

Established cocoa trees growing in sites where competition is minimized (for example, in pastures) appeared to adapt to the removal of all shade. Such trees suffered an initial die-back, followed by the development of a tight crown with shorter internodes and smaller leaves than before; as stated earlier, pod production was far higher in such trees. Whitmore emphasizes the variety of strategies included in the shade-bearing category, and cocoa, although clearly a shade-bearing species, is tolerant of a wide range of light regimes. This accords with the earlier assertion that *T. cacao* is a K-strategist, emphasizing long-term survival rather than reproductive effort.

Population structure and habitat stability

Any further interpretation of the observed size or age distribution in the study population requires a reassessment of the stability of the habitat. This can be based partly on information from published sources on the geomorphology and quaternary history of the region, and partly by direct observation. It turns out that the superficially stable habitats of this region have been subjected to a whole series of perturbations with different time-scales (see Section 2.1 for references).

First, it is generally accepted that the Amazon basin was much drier during the Pleistocene; although rainfall in the Ecuadorian Amazon was probably sufficient to support forest throughout this period, recent evidence suggests that areas adjacent to the Andes may have been significantly cooler during this period, so that shifts in species composition may have occurred. These broad-scale climatic changes are relevant to the present-day distribution of species, but should probably not be invoked to explain population structures.

Second, recent studies of lake sediments in the Ecuadorian Amazon have provided evidence of a prolonged flooding episode as recently as 800 years ago. This event probably affected much of Napo Province below the 300 m contour, and would have eliminated cocoa from the flooded areas, since cocoa is intolerant of waterlogged soils. The present-day wild cocoa population must therefore have spread through this area within the last 800 or so years.

Third, the majority of wild cocoa trees occur in present or former river floodplains, and these are (or have been) affected by the erosion-deposition processes taking place along all rivers in the region. For example, over a period of several hundred years, it is likely that much of the forest in the San Carlos region has been destroyed more than once by changes in the course of the Río Napo. Again, this means that present-day cocoa populations are the result of dispersal, possibly over quite short periods. As suggested above, this type of disturbance may stimulate dramatic short-term increases in seed production by cocoa.

Fourth, those parts of the Amazon basin close to the Andes have been, and continue to be, affected by volcanic activity. Lakes of volcanic origin have been identified in the Ecuadorian Amazon (Colinvaux *et al* 1985) and the volcano Reventador, about 125 km north-west of San Carlos, is continuously active, with major eruptions in 1960 (Grubb *et al* 1963) and 1987. No evidence is presently available as to the effects of past volcanic activity on the San Carlos area, but it is certain that cocoa populations in other parts of the region will have been affected at various times.

Finally, it is evident that the forest is subject to continual disturbance by the falling of large trees, typically blown down by strong winds associated with convectional rainstorms, which create gaps of up to 0.04 ha (Whitmore 1984, Section 7). Large cocoa trees are unlikely to be killed by such windthrown trees; again, the additional exposure to light will certainly increase pod production and seedling growth rates in areas adjacent to the gap.

It may therefore be concluded that despite the impression of stability conveyed by the Amazon rain forest, there is reason to believe that the study population is not the result of a steady approach to equilibrium over several thousand years, but is rather the result of a combination of influences over the past 800 or so years, with the likelihood of partial or complete elimination of the population at least once in this period, as well as the periodic stimulation of seed production by increases in light levels. Such events could readily explain the observed age and size distribution.

The implications for cocoa breeding

This account of cocoa in its natural habitat may have interesting implications for cocoa breeding. It has long been known that shade is required for the establishment of planted cocoa in most areas, and that careful management is required if shade reduction in plantations is to result in sustained yield increases. What is probably not generally realized, however, is the extreme lack of productivity, in agronomic terms, of wild cocoa in its natural habitat. In physiological terms, this low agronomic productivity may have two explanations. It may simply be a reflection of low net primary production, or it may indicate that dry matter partitioning is weighted away from seed production (to use the terminology of Cannell (1985), the 'harvest increment' is low. Hard data on partitioning in cocoa do not appear to exist, although some estimated ranges for different classes of plantations were given by Ng (1982). It has been observed (G.Lockwood, personal communication) that some wild clones from Ecuador, when grown under good plantation conditions in West Africa, showed vigorous vegetative growth but produced few pods; many accessions in the San Carlos genebank appear to be behaving similarly. There is a strong implication that at least some wild cocoa has an inherently low harvest increment which is not modifiable by environmental factors. However, as mentioned in Section 7.2, some wild trees seen on collecting trips had apparently responded to the removal of forest shade by a large and apparently sustained increase in pod production. I would suggest that there is a need for a study of partitioning in cocoa which attempts to quantify the allocation of dry matter production under different environments, and which seeks to identify variation in this character among genotypes of wild origin.

7.3 Broad issues

Are these populations really wild?

Throughout this thesis, and in particular in Section 7.2, there has been an implicit assumption that the Ecuadorian and Caquetá populations in general, and the study population in particular, are genuine wild populations in which the present-day distribution of trees has only been marginally affected by human activities. It has sometimes been suggested, however, that *T. cacao* was originally native to Central America, and was introduced to the Amazon basin as a cultivated plant; Cuatrecasas (1964) summarizes the views of several authors on this subject. Members of the Anglo-Colombian collecting expedition (Baker *et al* 1954) were certainly not convinced of the wild status of cocoa along most of the rivers they explored in eastern Colombia.

Although this question may be always be an intriguing subject for debate, I suggest that the weight of evidence is now firmly in favour of the wild status of cocoa populations in eastern Ecuador and south-eastern

Colombia, and probably elsewhere in the Upper Amazon. The evidence comes from two sources. First, careful reading of a wide range of sources - and my personal observations - have failed to produce any evidence that the indigenous inhabitants of the Upper Amazon have traditionally regarded cocoa as a crop, in the sense of a species which is systematically cultivated. As with many forest fruits, casual harvesting and consumption (in the case of cocoa, of the pulp) are customary, but authors from De la Condamine (1778) onwards concur in the view that cocoa was of no greater importance than a wide range of other forest species, several authors omitting it entirely from their ethnobotanical inventories (Harner 1972; Lathrap 1970; Mangelsdorf and Reeves 1939; Oberem 1980; Rumazo Gonzales 1982; Vickers and Plowman 1984; Davis and Yost 1983).

Second, observations in the course of the present project have made it clear that cocoa is widespread and sometimes very common as a forest species, with a consistent pattern of geographical variation. There is ample evidence that cocoa reproduces under forest conditions. To sustain the counter-hypothesis, it would be necessary to postulate that cocoa was very widely cultivated in the past and that the tradition of cultivation was completely lost. This explanation seems unreasonable, since other aspects of Amazon agriculture, such as the cultivation of manioc (*Manihot esculenta*) have remained stable over wide areas for very long periods. Such changes as have been recorded generally relate to increases in the number of cultivated species, such as the spread of plantains (*Musa* spp).

One *Theobroma* species, *T. bicolor*, is traditionally cultivated as a food plant in the region and its distribution is quite different to that of *T. cacao*. Trees of *bicolor* are typically found in groups close to existing or recent indigenous settlements; I have never found this species other than in association with settlements, and there are no well-documented accounts of wild *T. bicolor* populations. If cocoa had been introduced to the Upper Amazon as a cultivated species, I would expect its distribution to resemble that of *T. bicolor*.

Acceptance of the wild status of cocoa does not exclude the possibility that the wild cocoa population in some areas may have been augmented by planting seeds or protecting naturally-dispersed seedlings.

D. Irvine (personal communication) has shown that such practices are common even for species which are never considered as crops; in fact, large areas of Amazonian forest may have been subject to such management in the past, when the indigenous populations were more numerous than at present.

A centre of origin for cocoa?

The hypothesis that the Upper Amazon region contains both a centre of diversity for *T. cacao*, and its centre of origin, originated with Cheeseman (1944), and has often been treated as a simple fact (for example, by Cope, 1976). As indicated in Section 1.2, the concept of Vavilovian centres of diversity (Vavilov 1951) has been modified in recent years as surveys of variation in crop species have been published. Furthermore, the Vavilov model of domestication and evolution is almost certainly inapplicable to cocoa, since, in my view, *T. cacao* is not a 'domesticated' species; there is no evolutionary step between the wild species, as seen in the forests of the Upper Amazon (and possibly elsewhere), and cultivated populations or varieties. Making due allowance for incompatibility mechanisms within populations, all known wild and cultivated populations appear to be fully interfertile, and there is no single morphological character, or group of characters, which is peculiar to the wild or cultivated populations.

This project has provided no new information on the wild status of cocoa in Central America, and it is not possible to exclude the hypothesis that cocoa originated in the Amazon and was taken to Central America. Nonetheless, there are well-documented wild populations outside the Amazon basin in the Guianas (see Table 1.5), and consistent reports of wild cocoa at a range of localities in Central America (Cuatrecasas 1964). It is also generally believed that there were wild cocoa trees in the original forests of the coastal region of Ecuador, in the Guayas basin, although caution is necessary here since this was an area of early agricultural activity (Parsons and Schlemon 1982) and cultural links with the Amazon basin have been postulated (Lathrap 1970). If one accepts the simplest hypothesis, namely that *T. cacao* is distributed as a wild species in both Central and South America, then it reasonable to postulate that the earliest systematic cultivation of cocoa, in Mexico and Guatemala (Thompson 1956) was based on the wild populations of that area. As the cultivation of cocoa spread in the post-Colombian period, other wild populations were added to the cultivated population, as documented by Cheesman (1944), giving rise to a complex set of cultivated types in which the characteristics of several wild populations appear in new combinations.

As the practice of establishing cocoa plantations spread through the natural range of *T. cacao*, the question of seed sources would have arisen repeatedly. Where large plantations were established in remote areas, as with the Jesuit missions on the Río Guaviare in Colombia (Rojas-Peña 1951), it is likely that locally available seed of wild origin was used. In other cases, missionaries and travellers may have carried small numbers of pods over long distances; it is tempting to suggest that this may account for the exceptional variability found along rivers with a long history of colonial settlement (for example, the Río Marañon, as explored by Pound, and to some extent the Río Napo).

The broad conclusion is therefore that the present-day wild cocoa populations of the Upper Amazon do not represent the centre of origin for cocoa, and that such a centre of origin, in the classical sense, does not exist; rather, the Upper Amazon populations, such as those described in this thesis, should be considered both as a source of genetic variation which has not yet been incorporated into the cultivated populations of cocoa, and as a laboratory in which to study the adaptation to its original habitat of *Theobroma cacao*.

8 Conclusions

The conclusions of this thesis are built on several complementary lines of reasoning which take as their starting points the objectives laid out in Chapter 1.

The first main objective of this thesis was to show how information on the geographical pattern of variation might be used to devise an efficient strategy for cocoa germplasm collecting in Upper Amazon cocoa populations. Geographical variation was described in terms of the morphological descriptors conventionally used for cocoa; the thesis presents the first systematic application of these descriptors to wild populations of cocoa. The main data for the study were obtained from about 600 accessions collected in the course of the London Cocoa Trade Amazon Project (LCTAP) in Ecuador and along the Río Caquetá in Colombia; about half of these were successfully established at a genebank at San Carlos, Napo Province, Ecuador. Additional data were obtained for about 50 accessions collected by Pound in Peru in 1937-42 and held in the Pichilingue genebank in Ecuador. The collecting strategy originally devised for LCTAP in Ecuador was based on the working hypothesis that the target region needed to be sampled intensively, with collecting sites at intervals of 50 km or less, with the expectation that this would capture more variability than a coarser sampling grid.

The results presented here show that this working hypothesis was too pessimistic, in the sense that, using cluster analysis, it was not possible to identify morphologically distinct subpopulations associated with particular areas within Ecuador (with the exception of a few accessions from the Río Putumayo). The characteristics which separated the Ecuadorian accessions from the Colombian and Peruvian groups (white seeds, rough-surfaced pods, unpigmented flush and flowers, and, probably, the absence of flower cushions on the trunk) were found throughout the Ecuadorian Amazon. For sampling purposes, the whole of this region could be treated as a single population, and the number of accessions established at San Carlos was far in excess of the minimum sample size according to the most widely accepted criterion: a probability of 95% of capturing all alleles with frequency >.05 at a random locus. Nonetheless, the information needed to draw this conclusion would not have been available if collecting had not taken place systematically thoughout the target region in Ecuador.

The Colombian and Peruvian populations were distinct from each other and from the Ecuadorian material. Within the accessions from the Río Caquetá, two distinct types were evident on the basis of pod and seed characters, with only a limited degree of geographical separation. Even if the Río Caquetá target region were treated as containing a single population, the number of accessions held at the San Carlos genebank would fall below the desirable minimum sample size (additional accessions from the Río Caquetá expedition are held in genebanks in Colombia). The small number of Peruvian accessions available for study did not justify any attempt to delineate separate populations within Peru; indeed, the Pound material in general, despite its crucial importance to cocoa breeding programmes, is probably too small a sample to be used to describe geographical variation in Peruvian wild cocoa.

Using principal components axes as 'composite descriptors', an analysis of variance was used to explore the partition of variation between areas of different sizes. Conventionally, additional samples from sites are cheap, relative to travelling and other fixed costs, thereby justifying the acquisition of additional genetically similar acqusitions. Using cost data from LCTAP, it was possible to show that taking into account only collecting costs, the marginal cost of additional accessions from sites was high enough to render uneconomic the acquisition of large numbers of accessions from the same sites; when post-collection costs are considered, the optimum number of samples per site tends to decline as the period over which costs are calculated is increased. If these results could be extrapolated to other areas, they suggest that future cocoa germplasm collecting projects may be able to deploy their resources to cover larger target regions.

The second main objective of the thesis was to use observations on wild cocoa, acquired during the germplasm collecting project, to arrive at an overall view of the population biology of cocoa in its natural habitat. Several broad conclusions emerge from this part of the thesis. Cocoa is a long-lived species, which by virtue of its multi-trunked habit and capacity for adventitious shoot and root formation, is capable of indefinite growth and of vegetative reproduction. This longevity offsets the low rate of reproduction observed, with trees under forest conditions producing only 30-90 seeds per year. Dispersal requires the intervention of mammals to open the pods; this could not be observed directly, but two distinct types of pod damage seen in the field were probably caused by squirrels and monkeys. Most seeds germinated in groups under the canopy of the parent tree and were susceptible to early damage by mammals, probably rabbits; this would account for the observed random distribution of established trees. Although cocoa is clearly a shade-bearing species, increased light levels did result in increased rates of seed production and seedling growth. The overall age-structure of the study population at San Carlos, and of a sample of wild trees from Napo Province, suggested that there had been increased recruitment in the past, which could be explained by increased light levels resulting from any of several listed causes of disturbance. However, in completely unshaded sites, cocoa was adversely affected by competition from epiphytes and other light-adapted species, and cocoa was never found in association with the pioneer tree species found in large forest clearings or on recent alluvial deposits along meandering rivers. Published references to wild cocoa had described it being typically found along rivers; in Ecuador, cocoa was not associated with river banks, but with specific soil and terrain types. Along the Río Caquetá, most wild cocoa was found close to the river, but this was determined by the existence of extensive floodable areas beyond the raised strip of dry land along the river; cocoa was never found in flooded or swampy sites.

This thesis provides the first detailed account of geographical variation and population biology in any wild population of cocoa. As such, it justifies some conclusions relating to broader issues. First, I suggest that the reproductive strategy of cocoa in its natural habitat will cause difficulties when wild germplasm is brought into breeding programmes; it may be helpful to quantify changes in dry-matter partioning which occur when wild trees are exposed to increased light, and to look for variation in this character. Second, I suggest that there can longer be any doubt of the wild status of cocoa in much of the Upper Amazon. Its observed distribution as a forest species cannot be accounted for by the counter-hypothesis of dispersal as a cultivated species, given the evidence that cocoa is not a traditional crop species for Amazonian Indians. Finally, I suggest that the wild status of cocoa in the Upper Amazon, and the variability of the Upper Amazon populations, do not justify the frequent assertion that this region is the 'centre of origin'; indeed, the concept of a centre of origin does not appear appropriate for cocoa. As a wild species, cocoa is probably very widely distributed and there are no characters which consistently distinguish cultivated from wild cocoa; it seems likely that the spread of cocoa cultivation in South America in the post-Colombian period made use of seed from local wild populations. The value of Upper Amazon cocoa populations to cocoa breeders is not that these populations are ancestral to cultivated cocoa, but rather that they are a source of variation not yet incorporated into the cultivated gene pool.

9 References

LCTAP Reports and publications

Allen, J.B. (1980-1985). London Cocoa Trade Amazon Project, Reports 1-10. Cocoa, Chocolate and Confectionery Alliance, London.

Allen, J.B. (1982). London Cocoa Trade Amazon Project. Cocoa Growers' Bulletin, 33, 5-10.

Allen, J.B. (1982). Collecting wild cocoa in its centre of diversity. Proceedings 8th International Cocoa Research Conference, October 1981, Cartagena, Colombia, 665-662.

Allen, J.B. (1983). In search of the perfect cocoa bean. New Scientist, 97(1343), 293-296.

Allen, J.B. (1984). Strategies and methods for collecting Theobroma. Plant Genetic Resources Newsletter 57, 8-14. International Board for Plant Genetic Resources, Rome.

Allen, J.B. (1985). Evaluation of early vigour of wild cocoa from the Amazon Region of Ecuador. Proceedings International Conference on Cocoa and Coconuts, Kuala Lumpur, October 1984. Incorporated Society of Planters, Kuala Lumpur, Malaysia.

Allen, J.B. (1988). London Cocoa Trade Amazon Project: Final Report, Phase Two. Cocoa Growers' Bulletin. Cadbury Schweppes PLC, Birmingham, UK.

Allen, J.B. and Lass, R.A. (1983). London Cocoa Trade Amazon Project: Final Report, Phase 1. Cocoa Growers' Bulletin, 34, i-iii, 1-71. Cadbury Schweppes, Birmingham, UK.

Potts, G. (1984). An evaluation of methods to improve the transportation of cocoa budwood and a study of growing conditions at the Napo-Payamino Experimental Station (San Carlos), Ecuador. Ghana Cocoa Growing Research Association, CCCA, London.

Wright, M. (1984). A note on some insects associated with cocoa in Ecuador. Tropical Pest Management, 30, 29-31.

References cited in this thesis

Allard, R.W. (1970). Population structure and sampling methods. In: Frankel, O.H. and Bennett, E. (eds). Genetic resources in plants - their exploration and conservation. Blackwell Scientific Publications, Oxford, UK.

Amefia, Y.K., Cilas, C., Djiekpor, E.K., and Partiot, M. (1984). Etude du polymorphisme enzymatique chez le cacaoyer. I. Mise au pointd'une methode d'extraction et mise en evidence d'un locus specifiant une esterase. Cafe, Cacao, The 28(2) 89-94.

Antonovics, J., Bradshaw, A.D. and Turner, R.G. (1971). Heavy metal tolerance in plants. Advances in Ecological Research, 7, 1-85.

Atkinson, M.D., Withers, L.A. and Simpson, M.J.A (1986). Characterizationof cacao germplasm using isoenzyme markers. 1. A preliminary study of diversity using starch gel electrophoresis and standardization of the procedure. Euphytica 35, 741-750.

Baker, R.E.D., Cope, F.W., Holliday, P.C., Bartley, B.G.D. and Taylor, D.J. (1954). The Anglo-Colombian Cocoa Collecting Expedition. Report on Cacao Research, 1953. Imperial College of Tropical Agriculture, Trinidad.

Baldock, J.W. (1982). Geology of Ecuador: explanatory bulletin of the National Geological map of the Republic of Ecuador, 1:1,000,000 scale. Direccion General de Geologia y Minas.

Barros N., O. (1981). Cacao. Manual de Asistencia Tecnica 23. Instituto Colombiano Agropecuario, Bogota, Colombia.

Bennett, E. (1970). Tactics of plant exploration. In: Frankel, O.H. and Bennett, E. (eds). Genetic resources in plants - their exploration and conservation. Blackwell Scientific Publications, Oxford, UK.

Bogyo, T.P., Porceddu, E. and Perrino, P. (1980). Analysis of sampling strategies for collecting genetic material. Economic Botany, 34, 160-174.

Buck, A.L. (1981). New equations for computing vapor pressure and enhancement factor. Journal of Applied Meteorology, 20, 1527-1532.

Cannell, M.G.R. (1985). Dry matter partitioning in tree crops. In: M.G.R.Cannell and J.E.Jackson (eds). Attributes of trees as crop plants, 160-193. Institute of Terrestrial Ecology, Huntingdon, UK.

Chalmers, W.S. (1969-1974). Cacao germplasm collecting in the Oriente Region of Ecuador. Annual Reports on Cocoa Research, Trinidad, 1968, 1969, 1970, 1971, 1972, 1973. Imperial College of Tropical Agriculture/University of the West Indies.

Chapman, C.G.D. (1984). On the size of a genebank and the genetic variation it contains. In: Holden, J.H.W. and Williams, J.T. (eds). Crop genetic resources: conservation and evaluation. George Allen and Unwin, London.

Cheesman, E.E. (1944). Notes on the nomenclature, classification and possible relationships of cacao populations. Tropical Agriculture (Trinidad), 21(8), 144-159. (reprinted in: Archives of Cocoa Research, 1).

Clement, D. (1986). Cacaoyers de Guyane: prospections. Cafe Cacao Thé, 30(1) 11-36.

Cochran, W.G. (1963). Sampling techniques (2nd Edition). Wiley, New York.

Colinvaux, P.A., Miller, M.C., Liu, K., Steinitz-Kannan, M. and Frost, I. (1985). Discovery of permanent Amazon lakes and hydraulic disturbance in the upper Amazon Basin. Nature 313, 42-45.

Cope, F.W. (1976). Cacao. in N.W.Simmonds (ed) Evolution of crop plants, 285-289. Longman, London.

Cuatrecasas, J. (1964). Cacao and its allies: a taxonomic revision of the genus Theobroma. Contributions from the U.S. National Herbarium, 35(6), 375-614. Smithsonian Institution, Washington D.C..

Dash, J.S. (1929). Agricultural Journal of British Guiana, June 1929.

Davis, E.W. and Yost, J.A. (1983). The ethnobotany of the Waorani of eastern Ecuador. Harvard University Botanical Museum Leaflets, 29(3), 159-211.

De la Condamine (1778). Relation abregee d'un voyage fait dans l'interieur de l'Amerique Meriodionale [...]. Maestricht.

Dean, W (1987). Brazil and the struggle for rubber. Cambridge University Press, Cambridge, UK.

Desrosiers, R. and Von Buchwald, A. (1950). Report of a trip to the Napo River. Estacion Experimental Tropical Pichilingue, Ecuador.

Diggle, P.J. (1983). Statistical analysis of spatial point patterns. Academic Press, London.

Doran, J.E. and Hodson, F.R. (1975). Mathematics and computers in archaeology. Edinburgh University Press, Edinburgh, UK.

Endler, J.A. (1986). Natural selection in the wild. Princeton University Press, Princeton, NJ.

Engels, J.M.M. (1981). Genetic resources of cacao: a catalogue of the CATIE collection. CATIE Technical Bulletin 7. CATIE, Turrialba, Costa Rica.

Engels, J.M.M. (1986). The systematic description of cacao clones and its significance for taxonomy and plant breeding. PhD Thesis, Agricultural University, Wageningen. 125 pp.

Engels, J.M.M., Bartley, B.G.D. and Enriquez, G.A. (1980). Cacao descriptors, their states and modus operandi. Turrialba, 30, 209-218.

Frost, I.G. (1984). A paleolimnological and palynological investigation in the Ecuadorian rainforest: evidence of regional flooding and paleohydrological disturbance in the Amazon rainforest. MSc Thesis, Ohio State University.

Gale, J.S. and Lawrence, M.J. (1984). The decay of variability. In: Holden, J.H.W and Williams, J.T. (eds). Crop genetic resources: conservation and evaluation.. George Allen and Unwin, London, UK.

Gould, S.J and Johnston, R.F. (1972). Geographic variation. Annual Review of Ecology and Systematics, 3, 457-498.

Grubb, P.J. and Whitmore, T.C. (1966). A comparison of montane and lowland forest in Ecuador. II. The climate and its effects on the distribution and physiognomy of the forests. Journal of Ecology, 54, 303-333.

Grubb, P.J. and Whitmore, T.C. (1967). A comparison of montane and lowland forest in Ecuador. III. The light reaching the ground vegetation. Journal of Ecology, 55, 33-57.

Grubb, P.J., Lloyd, J.R., Pennington, T.D. and Whitmore, T.C. (1963). A comparison of montane and lowland forest in Ecuador. I. The forest structure, physiognomy and floristics. Journal of Ecology, 51, 567-601.

Hand, D.J. (1981). Discrimination and classification. Wiley, Chichester, UK.

Harner, M.J. (1973). The Jivaro: people of the sacred waterfalls. Anchor Press/Doubleday, New York.

Hobbs, R. (1982). IBPGR standardized format for descriptor lists. AGP:IBPGR/82/46. International Board for Plant Genetic Resources, Rome.

Holcomb, J., Tolbert, D.M. and Jain, S.K. (1976). A diversity analysis of genetic resources in rice. Euphytica, 26, 441-450.

Holcomb, J., Tolbert, D.M., and Jain, S.K. (1976). A diversity analysis of genetic resources in rice.. Euphytica, 26, 441-450.

IBPGR (1981). Genetic resources of cocoa. International Board for Plant Genetic Resources, Rome.

Jain, S.K. (1975a). Population structure and the effects of breeding systems.. In Frankel, O.H. and Hawkes, J.G. (eds). Crop genetic resources for today and tomorrow, 15-36.

Jain, S.K. (1975b). Genetic reserves. In Frankel, O.H. and Hawkes, J.G. (eds). Crop genetic resources for today and tomorrow, 379-396

Jain, S.K., Qualset, C.O., Bhatt, C.M and Wu, K.K. (1975). Geographical patterns of phenotypic diversity in a world collection of durum wheats. Crop Science, 15, 700-704.

Kornerup, A. and Wanscher, J.H. (1978). Methuen handbook of colour. Methuen, London, UK.

Lanner, R.M. (1985). Some attributes of nut-bearing trees of temperate forest origin. In: M.G.R.Cannell and J.E.Jackson (eds). Attributes of trees as crop plants, 426-437. Institute of Terrestrial Ecology, Huntingdon, UK.

Lathrap, D.W. (1970). The Upper Amazon. Thomas and Hudson, London.

Lent, Roy (1966). The origin of the cauliflorous inflorescence of Theobroma cacao. Turrialba, 16, 352-358.

Liu, K. and Colinvaux, P.A. (1985). Forest changes in the Amazon Basin during the last glacial maximum. Nature 318, 556-557.

Mangelsdorf, P.C. and Reeves, R.G. (1939). The origin of Indian corn and its relatives. Texas Agricultural Experiment Station Bulletin, 574. Agricultural and Mechanical College, Texas.

Marshall, D.R. and Brown, A.H.D. (1975). Optimum sampling strategies in genetic conservation. In Frankel,O.H. and Hawkes, J.G. (eds). Crop genetic resources for today and tomorrow, 53-80. Cambridge University Press, Cambridge, UK.

Murphy, P.H. and Witcombe, J.R. (1981). Variation in Himalayan barley and the concept of centres of diversity. Proceedings Fourth International Barley Genetics Symposium, Edinburgh, July 1981, 26-36.

Myers, J.G. (1930). Notes on wild cacao in Surinam and in British Guiana. Kew Bulletin, 1, 1-10.

Ng, E. (1982). Potential cocoa photosynthetic productivity. Proceedings 8th International Cocoa Research Conference, October 1981, Cartagena, Colombia, 235-244.

Oberem, Udo (1980). Los Quijos. Instituto Otavaleno de Antropologia, Otavalo, Ecuador.

Oka, H. (1969). A note on the design of germplasm presentation work in grain crops. SABRAO Newsletter 1, 127-134.

Padilla G., W.A. (1977). Breve diagnostico agro-socio-economico de la Region Oriental, para la ubicacion de un centro experimental agropecuario de INIAP. Boletin Tecnico, 23. INIAP, Quito, Ecuador.

Parsons, J.J. and Schlemon, R (1982). Nuevo informe sobre los campos elevados prehistoricos de la Cuenca de Guayas, Ecuador. Miscelanea antropologica ecuatoriana, Boletin de los Museos del Banco Central del Ecuador, 2(2) 31-37. Banco Central, Quito, Ecuador.

Posnette, A.F. (1945). Incompatibility in Amazon cocoa. Tropical Agriculture (Trinidad), 22, 184-187.

Potts, G.E. (1984). An evaluation of methods to improve the transportation of cocoa budwood and a study of growing conditions at the Napo-Payamino Experimental Station (San Carlos), Ecuador. Ghana Cocoa Growing Research Association, CCCA, London.

Pound, F.J. (1938). Cacao and witchbroom disease of South America. Trinidad and Tobago.

Pound, F.J. (1943). Cacao and witches' broom disease. Government Printer, Trinidad and Tobago.

Pound, F.J. (1945). A note on the cocoa population of South America. Report and Proceedings of the Cocoa Research Conference held at the Colonial Office, May-June 1945, 131-133. His Majesty's Stationery Office, London.

Prance, G.T. (ed) (1982). Biological diversification in the tropics. Columbia University Press, New York, NY.

Richards, P.W. (1952). The tropical rain forest. Cambridge University Press, Cambridge, UK.

Rojas-Pena, Enrique (1951). Los cacaotales del Rio Guaviare. Agricultura Tropical (Bogota), 7, 27-32 and 49-56.

Rumazo G., J. (1982). La region amazonica del Ecuador en el siglo XVI. Banco Central, Quito, Ecuador.

Salo, J. Kalliola, R., Hakkinen, I, Makinen, Y., Niemela, P., Puhakka, M., and Coley, P.D. (1986). River dynamics and the diversity of Amazon lowland forest. Nature 322, 254-258.

Schultes, R.E. (1978). Richard Spruce and the potential for European settlement of the Amazon: an unpublished letter. Botanical Journal of the Linnaean Society, 77, 131-139.

Soderholm, P.K. and Shaw, E.W. (1965). A modified side graft technique for use in a cacao virus indexing program. Proceedings American Society for Horticultural Science, Caribbean Region, 9, 25-29.

Soegeng-Reksodihardjo (1964). The species of the genus Theobroma. PhD thesis, Harvard University, Cambridge, Massachusetts.

Som, R.K. (1973). A manual of sampling techniques. Heinemann, London, UK.

Soria V., J. (1965). Notes on the native cacaos in the vicinity of Iquitos (Peru) and of Alto Beni (Bolivia). Cacao, Turrialba, 10(4), 14-16.

Soria V., J. (1970). The latest cocoa expeditions to the Amazon basin. Cacao, Turrialba, 15(1), 5-15.

Sourdat, M. and Custode, E. (1980). La problematica del manejo integral y el estudio morfopedologico de la Region Amazonica Ecuatoriana. Ministerio de Agricultura y Ganaderia, Quito, Ecuador.

Thompson, J. Eric S. (1956). Notes on the use of cacao in Middle America. Notes on Middle American Archaeology and Ethnology, No. 128. Carnegie Institution of Washington, Department of Archaeology.

Thorold, C.A. (1975). Diseases of cocoa. Oxford University Press, Oxford, UK.

Tolbert, D.M., Qualset, C.D., Jain, S.K. and Craddock, J.C. (1979). Diversity analysis of a world collection of barley. Crop Science, 19, 789-794.

Toxopeus, H. (1972). Cocoa breeding: a consequence of mating system, heterosis and population structure. Proceedings Conference on Cocoa and Coconuts in Malaysia, 1972, 3-12. Incorporated Society of Planters, Kuala Lumpur, Malaysia.

Vavilov, N.I. (1951). Phytogeographic basis of plant breeding. The origin, variation, immunity and breeding of cultivated plants [translation from Russian]. Chronica Botanica, 13, 1-366.

Vello, F. and Madeiros, A.G. (1965). Botanical expedition to Brazilian Amazon. CEPLAC/CEPEC, Itabuna, Bahia, Brazil.

Vickers, W.T. and Plowman, T. (1984). Useful plants of the Siona and Secoya Indians of eastern Ecuador. Fieldiana Botany (page reference not available). Field Museum of Natural History, Chicago, Illinois.

Wheeler, B.E.J. and Mepsted, R. (1982). Pathogenic races of Crinipellis perniciosa. Proceedings 8th International Cocoa Research Conference, Cartagena, Colombia, October 1981, 365-370.

Whitmore, T.C. (1984). Tropical rain forests of the Far East, Second Edition. Tropical rain forests of the Far East. Second Edition. Clarendon Press, Oxford, UK.

Wishart, D. (1978). CLUSTAN user manual. Edinburgh University Computing Service.

Witcombe, J.R. and Gilani, M.M. (1979). Variation in cereals from the Himalayas and the optimum strategy for sampling plant germplasm. Journal of Applied Ecology, 16, 633-640.

Witcombe, J.R., and Rao, A.R. (1976). The genecology of wheat in a Nepalese centre of diversity. Journal of Applied Ecology, 13, 915-924.

Wood, G.A.R. (1975). Cocoa. Longman, London.

Wood, G.A.R. and Lass, R.A. (eds) (in press). Cocoa (4th Edition). Longman, London.

Yndgaard, F. and Hoskuldsson, A. (1985). Electrophoresis: a tool for genebanks. Plant Genetic Resources Newsletter, 63, 34-40.

Young, A.M. (1986). Cocoa pollination. Cocoa Growers' Bulletin, 37, 5-23.

APPENDIX 1 DESCRIPTOR LIST

DATA SET 1: PASSPORT DATA

REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES
1.1	ACCESSION	Access	Accession code	Unique code used to identify an accession within the database
1.2	COLLPREFIX	Collpref	Collection prefix	LCT-EEN = London Cocoa Trade Estación Experimental Napo EBC = Expedición Botánica Caquetá
1.3	COLLNUM	Num	Collection number	for LCT-EEN, 1..435 for EBC, 1..148
1.4	SPECIES	Spe	Species	Theobroma c = cacao b = bicolor g = glaucum m = microcarpum o = obovatum r = grandiflorum s = subincanum h = Herrania spp
1.5	DATE	Date	Collection date	day/month/year
1.6		Latitud	Latitude	
1.6.1	LATDEG			degrees
1.6.2	LATMIN			minutes
1.6.3	NS			north/south
1.6.4	LAT_DECIM			degrees/minutes as decimal
1.7		Longitu	Longitude	
1.7.1	LONGDEG			degrees
1.7.2	LONGMIN			minutes
1.7.3	EW			east/west
1.7.4	LONG_DECIM			degrees/minutes as decimal

APPENDIX 1 DESCRIPTOR LIST

DATA SET 1: PASSPORT DATA

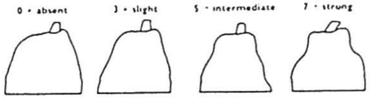
REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES
1.8	COUNTRY	Country	Country	
1.9	PROVINCE	Province	Province	
1.10	LOCALITY1	Area	Locality 1	general area
1.11	LOCALITY2	Locality1	Locality 2	specific location
1.12	ALTITUDE	Alti	Altitude	altitude (m)
1.13	STATUS	Stat	Status	1 = wild 2 = weedy 3 = primitive cultivar 4 = advanced cultivar
1.14	MATERIAL	Mate	Material	1 = budwood or cuttings 2 = seed 3 = both
1.15	SOBREVIV	Surv	Survival	1 = established in San Carlos genebank 0 = not established at San Carlos
1.16	NOTES		Notes	not printed in Appendix 2
1.17	POPLEVEL1	Pop1	Level 1 Population	see Table 3.2
1.18	POPLEVEL2	Pop2	Level 2 Population	see Table 3.3
1.19	POPLEVEL3		Level 3 Population	not used

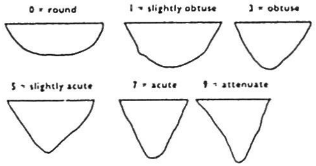
APPENDIX 1 DESCRIPTOR LIST

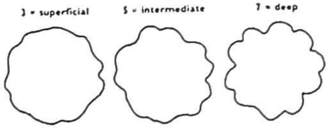
DATA SET 1: PASSPORT DATA

REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES
(POD CHARACTERS)				
2.1	BASE	Bas	Basal constriction	0 = absent 3 = slight 5 = intermediate 7 = strong
2.2	APEX	Apx	Apex form	0 = round 1 = slightly obtuse 3 = obtuse 5 = slightly acute 7 = acute 9 = attenuate
2.3	SURFACE	Sur	Surface rugosity	0 = absent 3 = slight 5 = intermediate 7 = intense
2.4	FURROWS	Fur	Furrow depth	primary furrow depth 3 = superficial 5 = intermediate 7 = deep
2.5	PAIRED	Pai	Furrows paired	1 = yes (= 5 pairs of furrows) 0 = no (= 10 unpaired furrows)

(drawings modified by J.B.Allen from Engels 1981)







DATA SET 2: POD AND SEED DESCRIPTORS

REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES
2.6	PIGMENT	Pig	Anthocyanin pigment	in unripe pod 0 = absent 3 = slight 5 = intermediate 7 = intense
2.7	LENGTH	Len	Length	in mm
2.8	DIAMETER	Diam	Diameter	in mm
2.9	D_LRATIO	D:LRa	Diameter:length ratio	
2.10	WALLTHICK	Wall	Wall thickness	at thickest point, in mm (including endocarp)
2.11	PODWEIGHT	Podw	Weight	fresh weight of pod with beans, in g
2.12	TOTBEANWT		Total bean weight	fresh weight of beans, in g (used for calculation of Pod index and seed index)
2.13	BEAN_POD		Ratio of total bean	weight to pod weight (not printed in Appendix 3)
(SEED CHARACTERS)				
2.14	SEEDINDEX	Sindx	Seed index	see Data Set 6
2.15	SEEDNUMBER	Snum	Number	number of fully-developed seeds per pod
2.16	SEEDWEIGHT	Sweig	Weight	fresh peeled weight (g)
2.17	SEEDLENGTH	Slen	Length	in mm

DATA SET 2: POD AND SEED DESCRIPTORS

REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES
2.18	SEEDWIDTH	Swid	Width	in mm
2.19	SEEDTHICK	Sthi	Thickness	in mm
2.20	SEEDW_LRAT	Swlra	Width:length ratio	
2.21	SEEDWHITE	Swhi	White beans	1 = present 0 = absent
2.22	SEEDLTPURP	Sltp	Light purple beans	1 = present 0 = absent
2.23	SEEDDKPURP	Sdkp	Dark purple beans	1 = present 0 = absent
2.24	DATAFROM		Source of data	for pod and seed data (not printed in Appendix 3) 1 = data recorded at time of collection 2 = data recorded in genebank

APPENDIX 1 DESCRIPTOR LIST

DATA SET 2: POD AND SEED DESCRIPTORS

REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES
3.1	BRANGLE	Ban	Branch angle	based on jorquette branches in seedling trees 3 = low: < 15 degrees above horizontal 5 = intermediate: 15-30 degrees above horizontal 7 = high: >30 degrees above horizontal
3.2	BRFORM	Bfo	Branch form	based on jorquette branches in seedling trees 3 = curved downwards 5 = straight 7 = curved upwards
3.3	JQHEIGHT	Jhe	Jorquette height	height of first jorquette (m)

APPENDIX 1 DESCRIPTOR LIST

DATA SET 3: HABIT DESCRIPTORS

REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES	
4.1	FLUSH	Flush	Flush colour	anthocyanin intensity in new flush	
					Methuen colour codes (Kornerup and Wanscher 1978)
				0 = absent	30D...1C-E...2E-F...
				3 = slight	...3D-F...
				5 = moderate	...4E...5E...6D-E...7D-E
				7 = intense	...8D...9E

DATA SET 5: FLOWER DESCRIPTORS

REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES
5.1	PEDICEL	Ped	Pedicel pigment	anthocyanin intensity in pedicel 0 = absent 3 = slight 5 = moderate 7 = intense
5.2	SEPAL	Sep	Sepal pigment	anthocyanin intensity in sepal (coded as for 5.1)
5.3	FILAMENT	Fil	Filament pigment	anthocyanin intensity in stamen filament (coded as for 5.1)
5.4	OVARY	Ova	Ovary pigment	anthocyanin intensity in upper part of ovary (coded as for 5.1)
5.5	SEPAL_LEN	Spl	Sepal length	mean sepal length (mm, rounded to 0.5 mm)

APPENDIX 1 DESCRIPTOR LIST

DATA SET 4: FLUSH COLOUR DESCRIPTOR

REFERENCE	FIELD	HEADING	DESCRIPTOR	DESCRIPTOR STATES
6.1		Pindx	Pod index	Estimated number of pods required to produce 1 kg of dry beans calculated as 1000/(0.4 x TOTBEANWT)
6.2	SEEDINDEX	Sindx	Seed index	Estimated mean weight of a single dry bean calculated as (TOTBEANWT x 0.4)/SEEDNUMBER

Notes:

1. The list of descriptors is based on the following sources (see Chapter 3)

(a) Hobbs (1982). IBPGR standardized format for descriptor lists. (AGP:IBPGR/82/46)
(b) IBPGR (1981). Genetic resources of cocoa. IBPGR document (AGP:IBPGR/80/56)
(c) Engels, Bartley and Enriquez (1980). Cacao descriptors, their states and modus operandi. Turrialba 30(2),209-218
(d) Engels (1981). Genetic resources of cacao: a catalogue of the Turrialba collection. CATIE Technical Bulletin No.7;

2. REFERENCE numbers for descriptors are used in the headings for Appendices 2-4;
3. FIELD refers to the dBase field names, which are given here for reference;
4. HEADING refers to the headings used in Appendices 2-4;
5. Missing data in numerically coded fields are represented by -1;
6. Some Pound accession numbers at represented by two or more possibly distinct clones at Pichilingue;
   the accessions involved are: PARINARI-121 (2 clones), POUND-7 (2 clones), POUND-12 (3 clones).

Note on descriptors

Ideally, descriptors should refer to characters which are (1) heritable and (2) distinct. Practicality also dictates (3) that descriptors should be capable of assessment with minimum effort and (4) that a descriptor list should include descriptors used in past publications.

Most of the descriptors used here can be justified in terms of criteria (2), (3) and (4). The heritability of the characters used in this descriptor list is generally accepted by breeders although it is rarely possible to cite experimental data to support this assertion. Toxopeus (1969) listed only three characters (apart from pollen compatibility) for which genetic data were available, and these were all pigment characters. *Axil spot* is associated with intense red pigmentation in pods, and appears to be controlled by two complementary dominant alleles (Pound 1932). This character was absent from the wild populations studied here, and was therefore not included in this descriptor list. *Albinism* is determined by a single recessive allele; some albino seedlings were found in the progeny of a wild tree at San Carlos, but this condition was too rare to be used as a descriptor. *Seed colour* may be controlled by one (Toxopeus 1969) or possibly two (Pound 1932) loci; it appears to be controlled independently of pod and flush pigmentation. The pattern of inheritance of pigment characters led Engels (1984) to propose that cocoa is a tetraploid species, although the genetic data could also be interpreted in terms of the action of multiple alleles.

The inheritance of pod shape descriptors does not appear to have been formally investigated, but there is no doubt, from experience in the San Carlos genebank, that the pod surface, apex form and furrow depth are expressed consistently within accessions; the pod base descriptor was somewhat variable within accessions.

Glendinning (1963) provided data on the inheritance of bean size and number, and pod size; the first two characters were highly heritable, whereas pod size was strongly influenced by environmental and physiological factors. However, experience in the San Carlos genebank suggested that extremes in pod size (for example, the small pods of LCT-EEN 46 and LCT-EEN 199) were expressed consistently. In general, quantitative pod and seed characters appear to be heritable, but large sample sizes are required for assessment of individual accessions (Engels *et al* 1980); the procedures used here (samples of 1-3 pods per accession) are adequate for characterization of populations.

In a few cases, seedling progenies could be compared with budded material from the same accession; the pod, seed and flush descriptors were consistent.

Additional references

Engels, J.M.M. (1984). A systematic description of cacao clones. IV. Some evidence of tetraploid inheritance. Cafe, Cacao, The 28, 95-102 [this reference forms one chapter of Engels (1986) in main bibliography]

Glendinning, D.R. (1963). The inheritance of bean size, pod size and number of beans per pod in cocoa (Theobroma cacao L.) with a note on bean shape. Euphytica 12, 311-322.

Pound, F.J. (1932). A preliminary survey of pigment factors in cacao. 2nd Annual report on Caco Research, Trinidad, 11-15.

Toxopeus, H. (1969). Cocoa in F.P.Ferweda and F.Wit (eds), Outlines of perennial crop breeding in the tropics. Wageningen Miscellaneous Papers 4, 79-109.

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
10010	LCT-EEN	1	c	30/04/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	2	1	1	1
10020	LCT-EEN	2	c	30/04/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	San Carlos	260	1	1	1	1	3
10030	LCT-EEN	3	c	03/05/80	0 10 S	76 20 W	Ecuador	Napo	Lago Agrio	camino Tarapoa - Fanny km2	200	1	1	0	1	3
10040	LCT-EEN	4	c	03/05/80	0 10 S	76 20 W	Ecuador	Napo	Lago Agrio	Via Tarapoa - Lago Agrio km4	200	1	2	1	1	3
10050	LCT-EEN	5	c	04/05/80	0 10 S	76 20 W	Ecuador	Napo	Lago Agrio	Via Lago Agrio - Tarapoa km43	200	1	2	1	1	3
10060	LCT-EEN	6	c	04/05/80	0 10 S	76 20 W	Ecuador	Napo	Lago Agrio	Via Lago Agrio - Tarapoa km43	200	1	2	1	1	3
10070	LCT-EEN	7	c	06/05/80	0 10 S	76 40 W	Ecuador	Napo	Coca - Lago Agrio	Via Shushufindi - Rio Aguarico	250	1	1	0	1	3
10080	LCT-EEN	8	c	06/05/80	0 10 S	76 40 W	Ecuador	Napo	Coca - Lago Agrio	Via Shushufindi - Rio Aguarico	250	1	2	1	1	3
10090	LCT-EEN	9	c	17/05/80	0 15 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	Via Lago Agrio - Coca km33	250	1	1	0	1	3
10100	LCT-EEN	10	c	17/05/80	0 15 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	Via lago Agrio - Coca km32	250	1	2	1	1	3
10110	LCT-EEN	11	c	19/05/80	0 25 S	76 55 W	Ecuador	Napo	Coca	Via Auca km2	200	1	2	1	1	3
10120	LCT-EEN	12	c	19/05/80	0 25 S	76 55 W	Ecuador	Napo	Coca	Via Auca	200	1	1	0	1	3
10130	LCT-EEN	13	c	19/05/80	0 25 S	76 55 W	Ecuador	Napo	Coca	Via Auca	200	1	1	0	1	3
10140	LCT-EEN	14	c	19/05/80	0 25 S	76 55 W	Ecuador	Napo	Coca	Via Auca km14	200	1	2	1	1	3
10150	LCT-EEN	15	c	20/05/80	0 35 S	77 5 W	Ecuador	Napo	Coca	S de Rio Napo camino al O km4	200	1	1	1	1	3
10160	LCT-EEN	16	c	20/05/80	0 35 S	77 5 W	Ecuador	Napo	Coca	S de Rio Napo camino al O km4	200	1	1	0	1	3
10170	LCT-EEN	17	c	20/05/80	0 35 S	77 5 W	Ecuador	Napo	Coca	S de Rio Napo camino al O km11	200	1	1	0	1	3
10180	LCT-EEN	18	c	20/05/80	0 35 S	77 5 W	Ecuador	Napo	Coca	S de Rio Napo camino al O km11	200	1	2	1	1	3
10190	LCT-EEN	19	c	20/05/80	0 35 S	77 5 W	Ecuador	Napo	Coca	S de Rio Napo camino al O	200	1	2	1	1	3
10200	LCT-EEN	20	c	20/05/80	0 35 S	77 5 W	Ecuador	Napo	Coca	S de Rio Napo camino al O	200	1	2	1	1	3
10210	LCT-EEN	21	c	20/05/80	0 35 S	77 5 W	Ecuador	Napo	Coca	S de Rio Napo	200	1	2	1	1	3
10220	LCT-EEN	22	c	03/06/80	0 35 S	76 5 W	Ecuador	Napo	Rio Napo	Limoncocha	220	1	2	1	1	3
10230	LCT-EEN	23	c	03/06/80	0 25 S	76 35 W	Ecuador	Napo	Rio Napo	Limoncocha	220	1	2	1	1	3
10240	LCT-EEN	24	c	03/06/80	0 25 S	76 35 W	Ecuador	Napo	Rio Napo	Limoncocha	220	1	2	1	1	3
10250	LCT-EEN	25	c	03/06/80	0 25 S	76 35 W	Ecuador	Napo	Rio Napo	Limoncocha	220	1	2	1	1	3
10260	LCT-EEN	26	c	03/06/80	0 25 S	76 35 W	Ecuador	Napo	Rio Napo	Limoncocha	220	1	2	1	1	3
10270	LCT-EEN	27	c	04/06/80	0 25 S	76 33 W	Ecuador	Napo	Rio Napo	Itaya (Rio Capocuy)	220	3	3	1	1	3
10280	LCT-EEN	28	c	04/06/80	0 25 S	76 33 W	Ecuador	Napo	Rio Napo	Itaya (Rio Capocuy)	220	1	2	1	1	3
10290	LCT-EEN	29	c	04/06/80	0 25 S	76 33 W	Ecuador	Napo	Rio Napo	Itaya (Rio Capocuy)	220	1	2	1	1	3
10300	LCT-EEN	30	c	04/06/80	0 25 S	76 33 W	Ecuador	Napo	Rio Napo	Itaya (Rio Capocuy)	220	1	1	1	1	3

DATA SET 1: PASSPORT DATA

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
10310	LCT-EEN	31	c	04/06/80	0 27 S	76 35 W	Ecuador	Napo	Rio Napo	Rio Indillama	220	1	1	1	1	3
10320	LCT-EEN	32	c	04/06/80	0 27 S	76 35 W	Ecuador	Napo	Rio Napo	Rio Indillama	220	1	1	1	1	3
10330	LCT-EEN	33	c	04/06/80	0 27 S	76 35 W	Ecuador	Napo	Rio Napo	Rio Indillama	220	1	1	1	1	3
10340	LCT-EEN	34	c	05/06/80	0 33 S	76 22 W	Ecuador	Napo	Rio Napo	Rio Anangu	220	1	1	0	1	3
10350	LCT-EEN	35	c	05/06/80	0 33 S	76 22 W	Ecuador	Napo	Rio Napo	Rio Anangu	220	1	1	0	1	3
10360	LCT-EEN	36	c	05/06/80	0 33 S	76 22 W	Ecuador	Napo	Rio Napo	Rio Anangu	220	1	2	1	1	3
10370	LCT-EEN	37	c	05/06/80	0 33 S	76 22 W	Ecuador	Napo	Rio Napo	Rio Anangu	220	1	3	1	1	3
10380	LCT-EEN	38	c	05/06/80	0 33 S	76 22 W	Ecuador	Napo	Rio Napo	Rio Anangu	220	1	3	1	1	3
10390	LCT-EEN	39	c	05/06/80	0 27 S	76 29 W	Ecuador	Napo	Rio Napo	Providencia	220	1	2	0	1	3
10400	LCT-EEN	40	h	05/06/80	0 0 S	76 0 W	Ecuador	Napo	Rio Napo	Limoncocha	220	1	2	1	1	3
10410	LCT-EEN	41	c	10/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10420	LCT-EEN	42	c	10/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10430	LCT-EEN	43	c	10/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10440	LCT-EEN	44	c	11/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	2	1	1	3
10450	LCT-EEN	45	c	11/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10460	LCT-EEN	46	c	11/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	3	1	1	1	3
10470	LCT-EEN	47	c	11/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	1	1	3
10480	LCT-EEN	48	c	11/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	1	1	3
10490	LCT-EEN	49	c	11/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	1	1	3
10500	LCT-EEN	50	c	11/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10510	LCT-EEN	51	c	12/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10520	LCT-EEN	52	c	12/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	1	1	3
10530	LCT-EEN	53	c	12/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10540	LCT-EEN	54	c	12/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10550	LCT-EEN	55	c	12/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10560	LCT-EEN	56	c	12/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	0	1	3
10570	LCT-EEN	57	c	12/06/80	0 40 S	76 20 W	Ecuador	Napo	Loreto		450	1	1	1	1	3
10580	LCT-EEN	58	c	16/06/80	3 54 S	78 49 W	Ecuador	Zamora-Chinchipe	Rio Zamora	Punguinza	800	1	2	1	3	6
10590	LCT-EEN	59	c	03/07/80	0 25 S	77 2 W	Ecuador	Napo	Coca	INIAP Payamino	250	1	2	1	1	3
10600	LCT-EEN	60	c	10/07/80	1 35 S	78 2 W	Ecuador	Pastaza	Puyo	Via Palora km7	900	3	1	1	2	4

DATA SET 1: PASSPORT DATA

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
10610	LCT-EEN	61	c	10/07/80	1 35 S	78 2 W	Ecuador	Pastaza	Puyo	Via Palora km7	900	3	2	1	2	4
10620	LCT-EEN	62	c	11/07/80	1 4 S	77 45 W	Ecuador	Napo	Rio Napo	Via Misahualli-Puerto Napo	400	3	2	1	1	3
10630	LCT-EEN	63	c	22/07/80	0 15 N	76 55 W	Ecuador	Napo	Rio San Miguel	Puerto San Miguel	300	1	2	1	1	2
10640	LCT-EEN	64	c	22/07/80	0 15 N	76 55 W	Ecuador	Napo	Rio San Miguel	Puerto San Miguel	300	1	2	1	1	2
10650	LCT-EEN	65	c	22/07/80	0 15 N	76 55 W	Ecuador	Napo	Rio San Miguel	Puerto San Miguel	300	1	3	1	1	2
10660	LCT-EEN	66	c	08/08/80	3 28 S	78 32 W	Ecuador	Morona-Santiago	Rio Zamora	Gualaquiza-Zamora km5	800	1	2	1	3	6
10670	LCT-EEN	67	c	08/08/80	3 28 S	78 32 W	Ecuador	Morona-Santiago	Rio Zamora	Gualaquiza-Zamora km5	800	3	2	1	3	6
10680	LCT-EEN	68	c	08/08/80	3 28 S	78 32 W	Ecuador	Morona-Santiago	Rio Zamora	Gualaquiza-Zamora km5	800	1	3	1	3	6
10690	LCT-EEN	69	c	08/08/80	3 28 S	78 32 W	Ecuador	Morona-Santiago	Rio Zamora	Gualaquiza-Zamora km11	800	1	2	1	3	6
10700	LCT-EEN	70	c	08/08/80	3 28 S	78 32 W	Ecuador	Morona-Santiago	Rio Zamora	Gualaquiza-Zamora km11	800	1	2	1	3	6
10710	LCT-EEN	71	c	08/08/80	3 28 S	78 32 W	Ecuador	Morona-Santiago	Rio Zamora	Via Chuchumbleza km3	800	1	2	1	3	6
10720	LCT-EEN	72	c	08/08/80	3 28 S	78 32 W	Ecuador	Morona-Santiago	Rio Zamora	Via Chuchumbleza km3	800	1	2	1	3	6
10730	LCT-EEN	73	c	09/08/80	3 57 S	78 38 W	Ecuador	Zamora-Chinchipe	Rio Nangaritza	La Guinza	900	1	1	1	3	6
10740	LCT-EEN	74	c	09/08/80	3 57 S	78 38 W	Ecuador	Zamora-Chinchipe	Rio Nangaritza	La Guinza	900	1	3	1	3	6
10750	LCT-EEN	75	c	09/08/80	3 57 S	78 38 W	Ecuador	Zamora-Chinchipe	Rio Nangaritza	La Guinza	900	1	1	0	3	6
10760	LCT-EEN	76	c	09/08/80	3 57 S	78 38 W	Ecuador	Zamora-Chinchipe	Rio Nangaritza	La Guinza	900	1	2	1	3	6
10770	LCT-EEN	77	c	11/08/80	3 38 S	78 27 W	Ecuador	Morona-Santiago	Rio Zamora	Chuchumbleza	800	1	2	1	3	6
10780	LCT-EEN	78	c	11/08/80	3 38 S	78 27 W	Ecuador	Morona-Santiago	Rio Zamora	Chuchumbleza	700	1	2	1	3	6
10790	LCT-EEN	79	c	11/08/80	3 38 S	78 27 W	Ecuador	Morona-Santiago	Rio Zamora	Rio Zamora-Tundaime km2	750	1	2	1	3	6
10800	LCT-EEN	80	c	11/08/80	3 38 S	78 27 W	Ecuador	Morona-Santiago	Rio Zamora	Rio Zamora-Tundaime km1	750	1	1	1	3	6
10810	LCT-EEN	81	c	11/08/80	3 38 S	78 27 W	Ecuador	Morona-Santiago	Rio Zamora	Rio Zamora-Tandaime km1	750	1	1	1	3	6
10820	LCT-EEN	82	c	12/08/80	3 38 S	78 27 W	Ecuador	Morona-Santiago	Rio Zamora	Proveduria	700	1	2	1	3	6
10830	LCT-EEN	83	c	12/08/80	3 20 S	78 25 W	Ecuador	Morona-Santiago	Rio Zamora	Arenal (2km N de Proveduria)	700	1	2	1	3	6
10840	LCT-EEN	84	c	13/08/80	3 58 S	78 52 W	Ecuador	Zamora-Chinchipe	Rio Yacuambi	Via Guadalupe km7	800	1	2	1	3	6
10850	LCT-EEN	85	c	13/08/80	3 58 S	78 52 W	Ecuador	Zamora-Chinchipe	Rio Yacuambi	Via Guadalupe km7	800	1	2	1	3	6
10860	LCT-EEN	86	c	13/08/80	3 58 S	78 52 W	Ecuador	Zamora-Chinchipe	Rio Yacuambi	Via Guadalupe km5	800	1	2	1	3	6
10870	LCT-EEN	87	c	13/08/80	3 58 S	78 52 W	Ecuador	Zamora-Chinchipe	Rio Yacuambi	Via Guadalupe km 2	800	1	2	1	3	6
10880	LCT-EEN	88	c	13/08/80	3 58 S	78 52 W	Ecuador	Zamora-Chinchipe	Rio Yacuambi	Via orilla oeste km1	800	1	2	1	3	6
10890	LCT-EEN	89	c	13/08/80	3 58 S	78 52 W	Ecuador	Zamora-Chinchipe	Rio Yacuambi	Via orilla oeste km1	800	1	2	1	3	6
10900	LCT-EEN	90	c	14/08/80	3 47 S	78 38 W	Ecuador	Zamora-Chinchipe	Rio Nangaritza	Los Encuentros-Paquisha km16	800	1	3	1	3	6

APPENDIX 2 DATA SET 1: PASSPORT DATA

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
10910	LCT-EEN	91	c	14/08/80	3 47 S	78 38 W	Ecuador	Zamora-Chinchipe	Rio Nangaritza	Los Encuentros-Paquisha km16	800	1	1	1	3	6
10920	LCT-EEN	92	c	14/08/80	3 47 S	78 38 W	Ecuador	Zamora-Chinchipe	Rio Nangaritza	Los Encuentros-Paquisha km1	800	1	2	1	3	6
10930	LCT-EEN	93	c	14/08/80	3 47 S	78 38 W	Ecuador	Zamora-Chinchipe	Rio Zamora	Los Encuentros orilla E, 1km N	800	1	2	1	3	6
10940	LCT-EEN	94	c	30/09/80	1 20 S	78 58 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	1	2	4
10950	LCT-EEN	95	c	30/09/80	1 20 S	78 58 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
10960	LCT-EEN	96	c	30/09/80	1 20 S	78 58 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
10970	LCT-EEN	97	c	30/09/80	1 20 S	78 58 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
10980	LCT-EEN	98	c	01/10/80	1 25 S	78 28 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
10990	LCT-EEN	99	c	01/10/80	1 25 S	78 28 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
11000	LCT-EEN	100	c	01/10/80	1 25 S	78 28 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
11010	LCT-EEN	101	c	02/10/80	1 25 S	78 50 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	3	1	2	4
11020	LCT-EEN	102	c	02/10/80	1 25 S	78 50 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
11030	LCT-EEN	103	c	02/10/80	1 25 S	78 50 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	1	2	4
11040	LCT-EEN	104	c	02/10/80	1 25 S	78 50 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
11050	LCT-EEN	105	c	02/10/80	1 25 S	78 50 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
11060	LCT-EEN	106	c	02/10/80	1 25 S	78 50 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	0	2	4
11070	LCT-EEN	107	c	02/10/80	1 25 S	78 50 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	1	2	4
11080	LCT-EEN	108	c	02/10/80	1 25 S	78 50 W	Ecuador	Pastaza	Rio Curaray	Curaray	200	1	1	1	2	4
11090	LCT-EEN	109	c	16/10/80	0 25 S	77 2 W	Ecuador	Napo	Coca	INIAP Payamino	250	1	2	1	1	3
11100	LCT-EEN	110	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	0	2	4
11110	LCT-EEN	111	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	0	2	4
11120	LCT-EEN	112	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	0	2	4
11130	LCT-EEN	113	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	2	1	2	4
11140	LCT-EEN	114	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	0	2	4
11150	LCT-EEN	115	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	2	1	2	4
11160	LCT-EEN	116	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	0	2	4
11170	LCT-EEN	117	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	0	2	4
11180	LCT-EEN	118	h	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	2	0	2	4
11190	LCT-EEN	119	c	22/10/80	2 3 S	77 3 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	0	2	4
11200	LCT-EEN	120	c	23/10/80	2 7 S	76 55 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	1	2	4

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
11210	LCT-EEN	121	c	23/10/80	2 7 S	76 55 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	1	2	4
11220	LCT-EEN	122	c	23/10/80	2 7 S	76 55 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	1	2	4
11230	LCT-EEN	123	c	23/10/80	2 7 S	76 55 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	1	2	4
11240	LCT-EEN	124	c	23/10/80	2 7 S	76 55 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	1	2	4
11250	LCT-EEN	125	c	23/10/80	2 7 S	76 55 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	1	2	4
11260	LCT-EEN	126	c	23/10/80	2 7 S	76 55 W	Ecuador	Pastaza	Rio Bobonaza	Montalvo	260	1	1	1	2	4
11270	LCT-EEN	127	c	22/10/80	2 10 S	76 40 W	Ecuador	Napo	Coca - Lago Agrio	Shushufindi: km6 al S	250	1	2	1	1	3
11280	LCT-EEN	128	c	22/10/80	0 10 S	76 40 W	Ecuador	Napo	Coca - Lago Agrio	km7 al S	250	1	2	1	1	3
11290	LCT-EEN	129	c	22/10/80	0 10 S	76 40 W	Ecuador	Napo	Coca - Lago Agrio	km11 al S	250	1	2	0	1	3
11300	LCT-EEN	130	c	22/10/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	3	1	1	1
11310	LCT-EEN	131	c	22/10/80	0 10 S	76 40 W	Ecuador	Napo	Coca - Lago Agrio	Lago Agrio - Coca km33	250	1	2	1	1	3
11320	LCT-EEN	132	c	02/10/81	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	2	1	1	1
11330	LCT-EEN	133	c	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
11340	LCT-EEN	134	c	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	2	1	1	1
11350	LCT-EEN	135	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
11360	LCT-EEN	136	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	2	1	1	1
11370	LCT-EEN	137	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
11380	LCT-EEN	138	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
11390	LCT-EEN	139	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
11400	LCT-EEN	140	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
11410	LCT-EEN	141	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
11420	LCT-EEN	142	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	3	1	1	1
11430	LCT-EEN	143	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
11440	LCT-EEN	144	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
11450	LCT-EEN	145	c	12/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
11460	LCT-EEN	146	c	10/11/80	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	3	1	1	1
11470	LCT-EEN	147	c	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
11480	LCT-EEN	148	c	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
11490	LCT-EEN	149	c	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
11500	LCT-EEN	150	c	25/11/80	0 30 S	76 24 W	Ecuador	Napo	Rio Napo	Anangu	250	1	1	1	1	3

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
11510	LCT-EEN	151	c	25/11/80	0 30 S	76 24 W	Ecuador	Napo	Rio Napo	Anangu	250	1	1	1	1	3
11520	LCT-EEN	152	c	26/11/80	0 30 S	76 24 W	Ecuador	Napo	Rio Napo	Anangu	250	1	1	1	1	3
11530	LCT-EEN	153	c	26/11/80	0 30 S	76 24 W	Ecuador	Napo	Rio Napo	Itaya	250	1	1	1	1	3
11540	LCT-EEN	154	c	27/11/80	0 30 S	76 24 W	Ecuador	Napo	Rio Napo	Paroto Yacu	250	1	1	1	1	3
11550	LCT-EEN	155	c	27/11/80	0 30 S	76 24 W	Ecuador	Napo	Rio Napo	Paroto Yacu	250	1	1	1	1	3
11560	LCT-EEN	156	c	28/11/80	0 30 S	76 24 W	Ecuador	Napo	Rio Napo	El Descanso	250	1	1	1	1	3
11570	LCT-EEN	157	c	28/11/80	0 30 S	76 24 W	Ecuador	Napo	Rio Napo	El Descanso	250	1	1	1	1	3
11580	LCT-EEN	158	c	28/11/80	0 27 S	76 53 W	Ecuador	Napo	Rio Napo	Comuna San Carlos	250	1	1	1	1	3
11590	LCT-EEN	159	c	29/11/80	0 27 S	76 58 W	Ecuador	Napo	Rio Napo	Comuna San Jose	250	1	2	1	1	3
11600	LCT-EEN	160	c	11/12/80	1 10 S	77 50 W	Ecuador	Pastaza	Rio Anzu	Via Puyo-Tena km46	550	3	2	1	2	3
11610	LCT-EEN	161	b	11/12/80	1 10 S	77 50 W	Ecuador	Pastaza	Rio Anzu	Via Puyo-Tena km44	550	1	2	1	2	3
11620	LCT-EEN	162	c	18/02/81	0 25 S	77 2 W	Ecuador	Napo	Coca	INIAP Payamino	250	1	3	1	1	3
11630	LCT-EEN	163	c	18/02/81	0 25 S	77 2 W	Ecuador	Napo	Coca	INIAP Payamino	250	1	1	1	1	3
11640	LCT-EEN	164	c	18/02/81	0 28 S	76 48 W	Ecuador	Napo	Rio Napo	Via Coca-Yuca km12	250	1	1	1	1	3
11650	LCT-EEN	165	c	18/02/81	0 28 S	76 48 W	Ecuador	Napo	Rio Napo	Via Coca-Yuca km13	250	1	1	1	1	3
11660	LCT-EEN	166	c	18/02/81	0 28 S	76 48 W	Ecuador	Napo	Yuca	Via Coca-Yuca km21	250	1	1	1	1	3
11670	LCT-EEN	167	g	30/03/81	-1 -1	-1 -1 W	Ecuador	Napo	Rio Tzapino		300	1	2	1	1	3
11680	LCT-EEN	168	c	11/04/81	0 38 S	75 47 W	Ecuador	Napo	Rio Napo	Chiru Isla	220	1	1	0	1	3
11690	LCT-EEN	169	c	11/04/81	0 38 S	75 47 W	Ecuador	Napo	Rio Napo	Chiru Isla	220	1	1	1	1	3
11700	LCT-EEN	170	c	11/04/81	0 38 S	75 47 W	Ecuador	Napo	Rio Napo	Chiru Isla	220	1	1	0	1	3
11710	LCT-EEN	171	c	12/04/81	0 45 S	75 35 W	Ecuador	Napo	Rio Napo	Sinchi Chicta	220	1	1	0	1	3
11720	LCT-EEN	172	c	12/04/81	0 45 S	75 35 W	Ecuador	Napo	Rio Napo	Sinchi Chicta	220	1	1	1	1	3
11730	LCT-EEN	173	c	13/04/81	0 43 S	75 37 W	Ecuador	Napo	Rio Napo	Sinchi Chicta	220	1	1	0	1	3
11740	LCT-EEN	174	c	13/04/81	0 43 S	75 37 W	Ecuador	Napo	Rio Napo	Sinchi Chicta	220	1	1	0	1	3
11750	LCT-EEN	175	c	13/04/81	0 43 S	75 37 W	Ecuador	Napo	Rio Napo	Sinchi Chicta	220	1	1	1	1	3
11760	LCT-EEN	176	c	13/04/81	0 43 S	75 37 W	Ecuador	Napo	Rio Napo	Sinchi Chicta	220	1	1	0	1	3
11770	LCT-EEN	177	c	14/04/81	0 43 S	75 37 W	Ecuador	Napo	Rio Napo	Puerto Quinche	220	1	1	0	1	3
11780	LCT-EEN	178	c	15/04/81	0 50 S	75 35 W	Ecuador	Napo	Rio Napo	Boca Tiputini	220	1	1	1	1	3
11790	LCT-EEN	179	c	15/04/81	0 50 S	75 35 W	Ecuador	Napo	Rio Napo	Boca Tiputini	220	1	1	1	1	3
11800	LCT-EEN	180	c	15/04/81	0 50 S	75 35 W	Ecuador	Napo	Rio Napo	Boca Tiputini	220	1	1	0	1	3

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
11810	LCT-EEN	181	c	16/04/81	0 52 S	75 28 W	Ecuador	Napo	Rio Napo	Santa Rosa	220	1	1	1	1	3
11820	LCT-EEN	182	c	16/04/81	0 52 S	75 28 W	Ecuador	Napo	Rio Napo	Santa Rosa	220	1	2	1	1	3
11830	LCT-EEN	183	c	16/04/81	0 52 S	75 28 W	Ecuador	Napo	Rio Napo	Santa Rosa	220	1	1	0	1	3
11840	LCT-EEN	184	s	16/04/81	0 52 S	75 28 W	Ecuador	Napo	Rio Napo	Santa Rosa	220	1	1	0	1	3
11850	LCT-EEN	185	c	16/04/81	0 52 S	75 28 W	Ecuador	Napo	Rio Napo	Santa Rosa	220	1	1	1	1	3
11860	LCT-EEN	186	b	18/04/81	0 38 S	75 47 W	Ecuador	Napo	Rio Napo	Chiru Isla	220	1	2	0	1	3
11870	LCT-EEN	187	c	29/04/81	0 20 S	77 3 W	Ecuador	Napo	Coca	Palmoriente (Huashito)	220	1	2	1	1	3
11880	LCT-EEN	188	c	16/05/81	0 3 N	75 40 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	1	1	1	2
11890	LCT-EEN	189	c	16/05/81	0 3 S	75 27 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	1	1	1	2
11900	LCT-EEN	190	c	16/05/81	0 3 S	75 27 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	1	0	1	2
11910	LCT-EEN	191	c	16/05/81	0 3 S	75 27 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	1	0	1	2
11920	LCT-EEN	192	c	16/05/81	0 2 S	75 33 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	1	0	1	2
11930	LCT-EEN	193	c	16/05/81	0 2 S	75 33 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	1	1	1	2
11940	LCT-EEN	194	c	16/05/81	0 2 S	75 33 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	2	0	1	2
11950	LCT-EEN	195	c	17/05/81	0 12 N	75 57 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	1	1	1	2
11960	LCT-EEN	196	c	17/05/81	0 17 N	75 59 W	Colombia	Putumayo	Rio Putumayo	Puerto El Carmen	200	1	2	1	1	2
11970	LCT-EEN	197	c	17/05/81	0 17 N	75 59 W	Colombia	Putumayo	Rio Putumayo	Puerto El Carmen	200	1	2	0	1	2
11980	LCT-EEN	198	c	17/05/81	0 12 N	75 57 W	Ecuador	Napo	Rio Putumayo	Puerto El Carmen	200	1	1	0	1	2
11990	LCT-EEN	199	c	17/05/81	0 20 N	76 5 W	Colombia	Putumayo	Rio Putumayo	Puerto El Carmen	200	1	2	1	1	2
12000	LCT-EEN	200	c	17/05/81	0 15 N	75 58 W	Colombia	Putumayo	Rio Putumayo	Puerto El Carmen	200	1	1	0	1	2
12010	LCT-EEN	201	c	18/05/81	0 5 N	75 55 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	2	1	1	2
12020	LCT-EEN	202	c	18/05/81	0 5 N	75 55 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	2	1	1	2
12030	LCT-EEN	203	c	18/05/81	0 5 N	75 55 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	2	1	1	2
12040	LCT-EEN	204	c	18/05/81	0 5 N	75 55 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	1	0	1	2
12050	LCT-EEN	205	c	18/05/81	0 7 N	75 58 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	2	1	1	2
12060	LCT-EEN	206	c	18/05/81	0 7 N	75 58 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	1	0	1	2
12070	LCT-EEN	207	c	18/05/81	0 7 N	75 58 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	1	0	1	2
12080	LCT-EEN	208	c	18/05/81	0 7 N	75 58 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	1	0	1	2
12090	LCT-EEN	209	c	18/05/81	0 13 N	76 15 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	1	0	1	2
12100	LCT-EEN	210	c	18/05/81	0 13 N	76 15 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	1	0	1	2

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
12110	LCT-EEN	211	c	18/05/81	0 13 N	76 15 W	Ecuador	Napo	Rio San Miguel	Puerto El Carmen	200	1	1	0	1	2
12120	LCT-EEN	212	c	15/05/81	1 3 S	77 45 W	Ecuador	Napo	Rio Napo	Via Misahualli-Puerto Napo	400	3	2	1	1	3
12130	LCT-EEN	213	c	15/07/81	1 3 S	77 45 W	Ecuador	Napo	Rio Napo	Via Misahualli-Puerto Napo	400	3	2	0	1	3
12140	LCT-EEN	214	c	15/07/81	1 3 S	77 45 W	Ecuador	Napo	Rio Napo	Via Misahualli-Puerto Napo	400	3	2	1	1	3
12150	LCT-EEN	215	c	15/07/81	1 3 S	77 45 W	Ecuador	Napo	Rio Napo	Via Misahualli-Puerto Napo	400	3	2	1	1	3
12160	LCT-EEN	216	c	17/07/81	1 28 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12170	LCT-EEN	217	c	17/07/81	1 28 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12180	LCT-EEN	218	c	17/07/81	1 28 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12190	LCT-EEN	219	c	17/07/81	1 28 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12200	LCT-EEN	220	c	17/07/81	1 28 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12210	LCT-EEN	221	c	17/07/81	1 28 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12220	LCT-EEN	222	c	17/07/81	1 28 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12230	LCT-EEN	223	c	17/07/81	1 28 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12240	LCT-EEN	224	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	0	2	4
12250	LCT-EEN	225	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	0	2	4
12260	LCT-EEN	226	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	0	2	4
12270	LCT-EEN	227	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12280	LCT-EEN	228	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12290	LCT-EEN	229	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	0	2	4
12300	LCT-EEN	230	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	0	2	4
12310	LCT-EEN	231	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12320	LCT-EEN	232	c	18/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12330	LCT-EEN	233	c	19/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	0	2	4
12340	LCT-EEN	234	c	19/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12350	LCT-EEN	235	c	19/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12360	LCT-EEN	236	c	19/07/81	1 31 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12370	LCT-EEN	237	c	19/07/81	1 31 S	77 32 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12380	LCT-EEN	238	c	19/07/81	1 31 S	77 32 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	1	2	4
12390	LCT-EEN	239	c	19/07/81	1 31 S	77 32 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	0	2	4
12400	LCT-EEN	240	c	19/07/81	1 31 S	77 32 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	1	0	2	4

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
12410	LCT-EEN	241	c	19/07/81	1 31 S	77 32 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	3	1	2	4
12420	LCT-EEN	242	c	20/07/81	1 30 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	2	1	2	4
12430	LCT-EEN	243	c	20/07/81	1 30 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	2	1	2	4
12440	LCT-EEN	244	c	20/07/81	1 30 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	2	1	2	4
12450	LCT-EEN	245	c	20/07/81	1 30 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	2	0	2	4
12460	LCT-EEN	246	c	20/07/81	1 30 S	77 27 W	Ecuador	Pastaza	Rio Villano	Villano	330	1	2	1	2	4
12470	LCT-EEN	247	c	27/08/81	1 27 S	76 44 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12480	LCT-EEN	248	c	27/08/81	1 28 S	76 38 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12490	LCT-EEN	249	c	27/08/81	1 28 S	76 38 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12500	LCT-EEN	250	c	27/08/81	1 28 S	76 38 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12510	LCT-EEN	251	c	27/08/81	1 28 S	76 38 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12520	LCT-EEN	252	c	27/08/81	1 28 S	76 38 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12530	LCT-EEN	253	c	27/08/81	1 32 S	76 32 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12540	LCT-EEN	254	c	28/08/81	1 32 S	76 32 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12550	LCT-EEN	255	c	28/08/81	1 32 S	76 32 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12560	LCT-EEN	256	c	28/08/81	1 32 S	76 32 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12570	LCT-EEN	257	c	28/08/81	1 32 S	76 32 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12580	LCT-EEN	258	c	28/08/81	1 33 S	76 28 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12590	LCT-EEN	259	c	28/08/81	1 33 S	76 28 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12600	LCT-EEN	260	c	29/08/81	1 34 S	76 20 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12610	LCT-EEN	261	c	29/08/81	1 34 S	76 20 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12620	LCT-EEN	262	c	29/08/81	1 32 S	76 13 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12630	LCT-EEN	263	c	29/08/81	1 32 S	76 13 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12640	LCT-EEN	264	c	29/08/81	1 36 S	76 12 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12650	LCT-EEN	265	c	29/08/81	1 36 S	76 12 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12660	LCT-EEN	266	c	29/08/81	1 36 S	76 12 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12670	LCT-EEN	267	c	29/08/81	1 38 S	76 8 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12680	LCT-EEN	268	c	29/08/81	1 38 S	76 8 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12690	LCT-EEN	269	c	29/08/81	1 38 S	76 8 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12700	LCT-EEN	270	c	29/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
12710	LCT-EEN	271	c	29/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	2	1	2	4
12720	LCT-EEN	272	c	29/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12730	LCT-EEN	273	c	29/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12740	LCT-EEN	274	c	30/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12750	LCT-EEN	275	c	30/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12760	LCT-EEN	276	c	30/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12770	LCT-EEN	277	c	30/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12780	LCT-EEN	278	c	30/08/81	1 37 S	76 2 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12790	LCT-EEN	279	c	30/08/81	1 38 S	75 58 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12800	LCT-EEN	280	c	30/08/81	1 35 S	75 58 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	0	2	4
12810	LCT-EEN	281	c	30/08/81	1 35 S	75 58 W	Ecuador	Pastaza	Rio Curaray	entre Curaray y Lorocachi	200	1	1	1	2	4
12820	LCT-EEN	282	c	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
12830	LCT-EEN	283	c	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
12840	LCT-EEN	284	h	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	2	0	1	1
12850	LCT-EEN	285	c	24/12/81	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
12860	LCT-EEN	286	c	24/12/81	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
12870	LCT-EEN	287	c	10/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	0	0	1	1
12880	LCT-EEN	288	c	20/03/82	0 20 S	76 50 W	Ecuador	Napo	Coca - Lago Agrio	INIAP San Carlos	260	1	1	1	1	1
12890	LCT-EEN	289	a	28/08/82	-1 -1	-1 -1 W	Ecuador	Los Rios	Quevedo	INIAP Pichilingue	-1	3	2	0	-1	-1
12900	LCT-EEN	290	c	04/11/81	0 27 S	77 7 W	Ecuador	Napo	Rio Payamino	1hr rio arriba de INIAP	250	1	1	0	1	3
12910	LCT-EEN	291	c	04/11/81	0 27 S	77 7 W	Ecuador	Napo	Rio Payamino	1hr rio arriba de INIAP	250	1	1	0	1	3
12920	LCT-EEN	292	c	20/11/81	2 22 S	78 7 W	Ecuador	Morona-Santiago	Rio Upano	Sevilla Don Bosco	950	1	1	0	3	5
12930	LCT-EEN	293	c	20/11/81	2 22 S	78 7 W	Ecuador	Morona-Santiago	Rio Upano	Sevilla Don Bosco	950	1	1	0	3	5
12940	LCT-EEN	294	c	20/11/81	2 22 S	78 7 W	Ecuador	Morona-Santiago	Rio Upano	Sevilla Don Bosco	950	1	1	0	3	5
12950	LCT-EEN	295	c	21/11/81	2 14 S	78 10 W	Ecuador	Morona-Santiago	Rio Upano	San Isidro	1120	3	2	1	3	5
12960	LCT-EEN	296	c	21/11/81	2 14 S	78 10 W	Ecuador	Morona-Santiago	Rio Upano	San Isidro	1120	3	2	1	3	5
12970	LCT-EEN	297	c	21/11/81	2 14 S	78 10 W	Ecuador	Morona-Santiago	Rio Upano	San Isidro	1120	3	1	1	3	5
12980	LCT-EEN	298	c	23/11/81	2 22 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Macas	910	1	1	0	3	5
12990	LCT-EEN	299	c	23/11/81	2 22 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Macas	910	1	1	0	3	5
13000	LCT-EEN	300	c	23/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi	760	1	3	1	3	5

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
13010	LCT-EEN	301	c	23/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi	760	1	1	0	3	5
13020	LCT-EEN	302	c	23/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi	760	1	2	1	3	5
13030	LCT-EEN	303	c	23/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi	760	1	1	0	3	5
13040	LCT-EEN	304	c	23/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	700	1	1	0	3	5
13050	LCT-EEN	305	c	23/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	800	1	1	1	3	5
13060	LCT-EEN	306	c	24/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	850	1	2	1	3	5
13070	LCT-EEN	307	c	24/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	800	1	1	0	3	5
13080	LCT-EEN	308	c	24/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	800	1	1	0	3	5
13090	LCT-EEN	309	c	24/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	800	1	1	0	3	5
13100	LCT-EEN	310	c	24/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	800	1	1	0	3	5
13110	LCT-EEN	311	c	24/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	750	1	1	0	3	5
13120	LCT-EEN	312	c	24/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	750	1	1	1	3	5
13130	LCT-EEN	313	c	24/11/81	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	800	1	1	1	3	5
13140	LCT-EEN	314	c	25/11/81	2 38 S	78 12 W	Ecuador	Morona-Santiago	Rio Upano	Sucua	600	1	1	0	3	5
13150	LCT-EEN	315	c	26/11/81	2 47 S	78 18 W	Ecuador	Morona-Santiago	Rio Upano	Mendez	480	1	1	0	3	5
13160	LCT-EEN	316	c	26/11/81	2 47 S	78 18 W	Ecuador	Morona-Santiago	Rio Upano	Mendez	480	1	1	0	3	5
13170	LCT-EEN	317	c	26/11/81	2 47 S	78 18 W	Ecuador	Morona-Santiago	Rio Upano	Mendez	480	1	1	0	3	5
13180	LCT-EEN	318	b	20/02/82	0 58 S	77 28 W	Ecuador	Napo	Rio Napo	Campanacocha	350	1	2	0	1	3
13190	LCT-EEN	319	c	21/02/82	0 58 S	77 28 W	Ecuador	Napo	Rio Napo	Santa Rosa	350	1	1	0	1	3
13200	LCT-EEN	320	c	22/02/82	1 3 S	77 42 W	Ecuador	Napo	Rio Napo	Misahualli-Puerto Napo km5	400	1	2	1	1	3
13210	LCT-EEN	321	c	22/02/82	1 3 S	77 42 W	Ecuador	Napo	Rio Napo	Misahualli-Puerto Napo km5	400	1	2	1	1	3
13220	LCT-EEN	322	c	23/02/82	1 3 S	77 46 W	Ecuador	Napo	Rio Napo	Misahualli-Puerto Napo km16	400	3	2	0	1	3
13230	LCT-EEN	323	c	23/02/82	1 3 S	77 46 W	Ecuador	Napo	Rio Napo	Misahualli-Puerto Napo km16	400	3	2	1	1	3
13240	LCT-EEN	324	c	23/02/82	1 3 S	77 46 W	Ecuador	Napo	Rio Napo	Misahualli-Puerto Napo km16	400	3	2	1	1	3
13250	LCT-EEN	325	c	23/02/82	1 3 S	77 46 W	Ecuador	Napo	Rio Napo	Misahualli-Puerto Napo km12	400	3	2	1	1	3
13260	LCT-EEN	326	c	23/02/82	1 3 S	77 42 W	Ecuador	Napo	Rio Napo	Misahualli-Puerto Napo km6	400	1	2	1	1	3
13270	LCT-EEN	327	c	31/03/82	0 18 N	77 2 W	Colombia	Putumayo	Rio San Miguel	2hr rio arriba de Puerto Colon	290	3	1	1	4	2
13280	LCT-EEN	328	c	15/07/82	2 54 S	78 25 W	Ecuador	Morona-Santiago	Rio Upano	Via Limon-Mendez km8	950	1	2	1	3	5
13290	LCT-EEN	329	c	15/07/82	2 54 S	78 25 W	Ecuador	Morona-Santiago	Rio Upano	Via Limon-Mendez km8	950	1	2	1	3	5
13300	LCT-EEN	330	c	15/07/82	2 54 S	78 25 W	Ecuador	Morona-Santiago	Rio Upano	Via Limon-Mendez km8	950	1	2	0	3	5

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
13310	LCT-EEN	331	c	17/07/82	2 35 S	78 12 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: camino a Bellavista	800	1	2	1	3	5
13320	LCT-EEN	332	c	17/07/82	2 35 S	78 12 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: camino a Bellavista	800	1	2	1	3	5
13330	LCT-EEN	333	c	17/07/82	2 35 S	78 12 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: camino a Bellavista	800	1	1	1	3	5
13340	LCT-EEN	334	c	17/07/82	2 35 S	78 12 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: camino a Bellavista	800	1	2	1	3	5
13350	LCT-EEN	335	c	18/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km1 al N	800	1	2	1	3	5
13360	LCT-EEN	336	c	18/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	750	1	2	1	3	5
13370	LCT-EEN	337	c	18/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	750	1	2	1	3	5
13380	LCT-EEN	338	c	18/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	750	1	2	1	3	5
13390	LCT-EEN	339	c	18/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	750	1	2	1	3	5
13400	LCT-EEN	340	c	18/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	750	1	2	1	3	5
13410	LCT-EEN	341	c	20/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	750	1	2	1	3	5
13420	LCT-EEN	342	c	20/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	800	1	2	1	3	5
13430	LCT-EEN	343	c	20/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	800	1	2	1	3	5
13440	LCT-EEN	344	c	20/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	800	1	2	1	3	5
13450	LCT-EEN	345	c	20/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	800	1	2	1	3	5
13460	LCT-EEN	346	c	20/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	800	1	2	1	3	5
13470	LCT-EEN	347	c	20/07/82	2 32 S	78 9 W	Ecuador	Morona-Santiago	Rio Upano	Huambi: km2 al N	800	1	2	1	3	5
13480	LCT-EEN	348	c	20/07/82	2 28 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	Sucua: hacia Rio Upano	800	1	2	1	3	5
13490	LCT-EEN	349	c	20/07/82	2 28 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	Sucua: hacia Rio Upano	800	1	2	1	3	5
13500	LCT-EEN	350	c	20/07/82	2 28 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	Sucua: hacia Rio Upano	800	1	2	1	3	5
13510	LCT-EEN	351	c	21/07/82	2 32 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	San Luis de El Upano	700	1	2	1	3	5
13520	LCT-EEN	352	c	21/07/82	2 32 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	camino a San Luis de El Upano	700	1	2	1	3	5
13530	LCT-EEN	353	c	21/07/82	2 32 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	camino a San Luis de El Upano	700	1	2	1	3	5
13540	LCT-EEN	354	c	21/07/82	2 32 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	camino a San Luis de El Upano	700	1	2	1	3	5
13550	LCT-EEN	355	c	21/07/82	2 32 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	camino a San Luis de El Upano	700	1	2	1	3	5
13560	LCT-EEN	356	c	21/07/82	2 32 S	78 8 W	Ecuador	Morona-Santiago	Rio Upano	camino a San Luis de El Upano	700	1	2	1	3	5
13570	LCT-EEN	357	c	21/07/82	2 28 S	78 7 W	Ecuador	Morona-Santiago	Rio Upano	San Luis de El Upano	750	1	2	1	3	5
13580	LCT-EEN	358	c	21/07/82	2 28 S	78 7 W	Ecuador	Morona-Santiago	Rio Upano	San Luis de El Upano	750	1	2	1	3	5
13590	LCT-EEN	359	c	21/07/82	2 28 S	78 7 W	Ecuador	Morona-Santiago	Rio Upano	San Luis de El Upano	750	1	2	1	3	5
13600	LCT-EEN	360	c	21/07/82	2 28 S	78 7 W	Ecuador	Morona-Santiago	Rio Upano	San Luis de El Upano	750	1	2	1	3	5

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
13610	LCT-EEN	361	c	21/07/82	2 28 S	78 7 W	Ecuador	Morona-Santiago	Rio Upano	San Luis de El Upano	750	1	2	1	3	5
13620	LCT-EEN	362	c	10/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	1	2	4
13630	LCT-EEN	363	c	10/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	0	2	4
13640	LCT-EEN	364	c	11/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	1	2	4
13650	LCT-EEN	365	c	12/12/82	2 9 S	76 19 W	Ecuador	Pastaza	Rio Conambo	Pozo Huito 1	200	1	1	0	2	4
13660	LCT-EEN	366	c	12/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	1	2	4
13670	LCT-EEN	367	c	12/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	1	2	4
13680	LCT-EEN	368	c	12/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	1	2	4
13690	LCT-EEN	369	c	13/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	1	2	4
13700	LCT-EEN	370	c	13/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	1	2	4
13710	LCT-EEN	371	c	13/12/82	2 11 S	76 12 W	Ecuador	Pastaza	Rio Conambo	Pozo Maranon 1	200	1	1	1	2	4
13720	LCT-EEN	372	c	16/02/83	0 28 S	76 48 W	Ecuador	Napo	Rio Napo	Yuca: camino al E km2	250	1	2	1	1	3
13730	LCT-EEN	373	c	16/02/83	0 28 S	76 48 W	Ecuador	Napo	Rio Napo	Yuca: camino al E km2	250	1	1	1	1	3
13740	LCT-EEN	374	c	08/11/83	0 8 S	76 53 W	Ecuador	Napo	Coca - Lago Agrio	Lago Agrio-Coca km32, km6·al O	250	1	1	1	1	3
13750	LCT-EEN	375	s	23/02/84	0 29 S	77 18 W	Ecuador	Napo	Rio Payamino	Comuna San Jose de Payamino	350	1	1	1	1	3
13760	LCT-EEN	376	c	24/02/84	0 29 S	77 18 W	Ecuador	Napo	Rio Payamino	Comuna San Jose de Payamino	350	1	1	1	1	3
13770	LCT-EEN	377	c	24/02/84	0 29 S	77 18 W	Ecuador	Napo	Rio Payamino	Comuna San Jose de Payamino	350	1	1	0	1	3
13780	LCT-EEN	378	c	25/02/84	0 29 S	77 18 W	Ecuador	Napo	Rio Payamino	Comuna San Jose de Payamino	350	1	1	1	1	3
13790	LCT-EEN	379	c	25/02/84	0 29 S	77 18 W	Ecuador	Napo	Rio Payamino	Comuna San Jose de Payamino	350	1	1	1	1	3
13800	LCT-EEN	380	s	09/05/84	0 29 S	77 18 W	Ecuador	Napo	Rio Payamino	Comuna San Jose de Payamino	350	1	1	1	1	3
13810	LCT-EEN	381	b	29/05/84	1 13 S	71 80 W	Colombia	Amazonas	Rio Caqueta	Las Palmas	100	1	2	1	4	8
13820	LCT-EEN	382	c	04/06/84	1 4 S	70 14 W	Colombia	Amazonas	Rio Apaporis	Centro Providencia	100	1	1	1	4	9
13830	LCT-EEN	383	o	07/06/84	1 15 S	69 43 W	Colombia	Amazonas	Rio Caqueta	Puerto Cordoba	100	1	1	1	4	9
13840	LCT-EEN	384	c	06/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	camino a Tutinentza	500	1	1	1	3	5
13850	LCT-EEN	385	c	06/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	camino a Tutinentza	500	1	1	1	3	5
13860	LCT-EEN	386	c	07/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	camino hacia SO	500	1	1	1	3	5
13870	LCT-EEN	387	c	08/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	camino hacia Rio Wawaimi	500	1	1	1	3	5
13880	LCT-EEN	388	c	08/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	camino hacia Rio Wawaimi	500	1	1	1	3	5
13890	LCT-EEN	389	c	08/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	camino hacia Rio Wawaimi	500	1	1	1	3	5
13900	LCT-EEN	390	c	08/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	camino Rio Wawaimi-Mision	500	1	1	1	3	5

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
13910	LCT-EEN	391	c	08/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	camino al S de la Mision	500	1	2	1	3	5
13920	LCT-EEN	392	s	09/12/84	2 23 S	77 32 W	Ecuador	Morona-Santiago	Taisha	carretera al NO, N de San Luis	500	1	2	1	3	5
13930	LCT-EEN	393	c	19/04/85	0 26 S	76 47 W	Ecuador	Napo	Rio Napo	frente Isla Anaconda	350	1	2	1	1	3
13940	LCT-EEN	394	c	11/07/85	0 26 S	76 47 W	Ecuador	Napo	Rio Napo	Hacienda Primavera: 5km al N	230	1	1	0	1	3
13950	LCT-EEN	395	c	11/07/85	0 26 S	76 47 W	Ecuador	Napo	Rio Napo	Hacienda Primavera: 5km al N	230	1	1	1	1	3
13960	LCT-EEN	396	c	11/07/85	0 26 S	76 47 W	Ecuador	Napo	Rio Napo	Hacienda Primavera: 5km al N	230	1	1	1	1	3
13970	LCT-EEN	397	c	11/07/85	0 26 S	76 47 W	Ecuador	Napo	Rio Napo	Hacienda Primavera: 5km al N	230	1	1	0	1	3
13980	LCT-EEN	398	c	11/07/85	0 26 S	76 47 W	Ecuador	Napo	Rio Napo	Hacienda Primavera: 5km al N	230	1	1	0	1	3
13990	LCT-EEN	399	c	11/07/85	0 26 S	76 47 W	Ecuador	Napo	Rio Napo	Hacienda Primavera: 5km al N	230	1	1	1	1	3
14000	LCT-EEN	400	c	11/07/85	0 26 S	76 47 W	Ecuador	Napo	Rio Napo	Hacienda Primavera: 5km al N	230	1	1	1	1	3
14010	LCT-EEN	401	c	27/08/85	0 50 S	75 32 W	Ecuador	Napo	Rio Napo	1km rio abajo de Tiputini	200	1	1	1	1	3
14020	LCT-EEN	402	g	27/08/85	0 57 S	75 28 W	Ecuador	Napo	Rio Yasuni	1km de la desembocadura	200	1	2	0	1	3
14030	LCT-EEN	403	c	27/08/85	0 57 S	75 28 W	Ecuador	Napo	Rio Yasuni	1km de la desembocadura	200	1	3	1	1	3
14040	LCT-EEN	404	c	27/08/85	0 57 S	75 28 W	Ecuador	Napo	Rio Yasuni	1km de la desembocadura	200	1	2	1	1	3
14050	LCT-EEN	405	c	27/08/85	0 57 S	75 28 W	Ecuador	Napo	Rio Yasuni	1km de la desembocadura	200	1	2	1	1	3
14060	LCT-EEN	406	c	27/08/85	0 57 S	75 28 W	Ecuador	Napo	Rio Yasuni	media hora de la desembocadura	200	1	1	1	1	3
14070	LCT-EEN	407	c	29/08/85	0 55 S	75 23 W	Ecuador	Napo	Rio Napo	abajo de Nuevo Rocafuerte	200	1	1	0	1	3
14080	LCT-EEN	408	c	29/08/85	0 55 S	75 23 W	Ecuador	Napo	Rio Napo	abajo de Nuevo Rocafuerte	200	1	3	1	1	3
14090	LCT-EEN	409	c	29/08/85	0 58 S	75 13 W	Ecuador	Napo	Rio Napo	abajo de Nuevo Rocafuerte	200	1	1	1	1	3
14100	LCT-EEN	410	b	29/08/85	0 58 S	75 13 W	Ecuador	Napo	Rio Napo	abajo de Nuevo Rocafuerte	200	3	1	1	1	3
14110	LCT-EEN	411	c	29/08/85	0 55 S	75 13 W	Ecuador	Napo	Rio Aguarico	media hora de la desembocadura	200	1	1	1	1	3
14120	LCT-EEN	412	c	29/08/85	0 55 S	75 13 W	Ecuador	Napo	Rio Aguarico	una hora de la desembocadura	200	1	1	1	1	3
14130	LCT-EEN	413	c	29/08/85	0 55 S	75 13 W	Ecuador	Napo	Rio Aguarico	una hora de la desembocadura	200	1	1	1	1	3
14140	LCT-EEN	414	c	29/08/85	0 55 S	75 13 W	Ecuador	Napo	Rio Aguarico	una hora de la desembocadura	200	1	1	1	1	3
14150	LCT-EEN	415	c	29/08/85	0 55 S	75 13 W	Ecuador	Napo	Rio Aguarico	una hora de la desembocadura	200	1	1	1	1	3
14160	LCT-EEN	416	c	19/09/85	3 1 S	77 56 W	Ecuador	Morona-Santiago	Rio Santiago	Kusumaza: abajo de Santiago	210	1	1	0	3	5
14170	LCT-EEN	417	c	19/09/85	3 1 S	77 56 W	Ecuador	Morona-Santiago	Rio Santiago	Kusumaza: abajo de Santiago	210	1	1	0	3	5
14180	LCT-EEN	418	c	19/09/85	3 1 S	77 56 W	Ecuador	Morona-Santiago	Rio Santiago	Kusumaza: abajo de Santiago	210	1	2	1	3	5
14190	LCT-EEN	419	c	19/09/85	3 1 S	77 56 W	Ecuador	Morona-Santiago	Rio Santiago	Kusumaza: abajo de Santiago	210	1	1	1	3	5
14200	LCT-EEN	420	c	19/09/85	3 1 S	77 56 W	Ecuador	Morona-Santiago	Rio Santiago	Kusumaza: abajo de Santiago	260	1	1	1	3	5

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
14210	LCT-EEN	421	c	19/09/85	3 1 S	77 56 W	Ecuador	Morona-Santiago	Rio Santiago	Kusumaza: abajo de Santiago	260	1	1	1	3	5
14220	LCT-EEN	422	c	20/09/85	3 3 S	78 6 W	Ecuador	Morona-Santiago	Rio Santiago	Tsuis: arriba de Santiago	260	1	1	1	3	5
14230	LCT-EEN	423	c	20/09/85	3 3 S	78 6 W	Ecuador	Morona-Santiago	Rio Santiago	Tsuis: arriba de Santiago	260	1	1	0	3	5
14240	LCT-EEN	424	c	20/09/85	3 3 S	78 6 W	Ecuador	Morona-Santiago	Rio Santiago	Tsuis: arriba de Santiago	260	1	2	1	3	5
14250	LCT-EEN	425	c	20/09/85	3 3 S	78 6 W	Ecuador	Morona-Santiago	Rio Santiago	Tsuis: arriba de Santiago	260	1	1	0	3	5
14260	LCT-EEN	426	c	20/09/85	3 3 S	78 6 W	Ecuador	Morona-Santiago	Rio Santiago	Tsuis: arriba de Santiago	260	1	1	1	3	5
14270	LCT-EEN	427	c	21/09/85	2 57 S	77 53 W	Ecuador	Morona-Santiago	Rio Yaupi	cerca desembocadura	210	1	1	1	3	5
14280	LCT-EEN	428	c	21/09/85	2 57 S	77 53 W	Ecuador	Morona-Santiago	Rio Yaupi	cerca desembocadura	210	1	1	1	3	5
14290	LCT-EEN	429	c	21/09/85	2 57 S	77 53 W	Ecuador	Morona-Santiago	Rio Yaupi	media hora de la desembocadura	200	1	1	1	3	5
14300	LCT-EEN	430	c	22/09/85	3 2 S	78 5 W	Ecuador	Morona-Santiago	Rio Santiago	4km al oeste de Santiago	260	1	2	1	3	5
14310	LCT-EEN	431	c	22/09/85	3 2 S	78 5 W	Ecuador	Morona-Santiago	Rio Santiago	4km al oeste de Santiago	260	1	1	0	3	5
14320	LCT-EEN	432	c	22/09/85	3 2 S	78 5 W	Ecuador	Morona-Santiago	Rio Santiago	4km al oeste de Santiago	260	1	1	1	3	5
14330	LCT-EEN	433	c	22/09/85	3 2 S	78 5 W	Ecuador	Morona-Santiago	Rio Santiago	4km al oeste de Santiago	260	1	1	0	3	5
14340	LCT-EEN	434	c	22/09/85	3 2 S	78 5 W	Ecuador	Morona-Santiago	Rio Santiago	4km al oeste de Santiago	260	1	1	1	3	5
14350	LCT-EEN	435	c	22/09/85	3 3 S	78 0 W	Ecuador	Morona-Santiago	Rio Santiago	Santiago: cerca de la Mision	210	1	1	0	3	5
20010	EBC	1	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	1	0	4	7
20020	EBC	2	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	1	0	4	7
20030	EBC	3	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	1	0	4	7
20040	EBC	4	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	1	0	4	7
20050	EBC	5	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	2	1	4	7
20060	EBC	6	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	2	1	4	7
20070	EBC	7	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	1	0	4	7
20080	EBC	8	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	2	0	4	7
20090	EBC	9	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	2	1	4	7
20100	EBC	10	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	2	1	4	7
20110	EBC	11	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	2	0	4	7
20120	EBC	12	c	23/05/84	0 33 S	72 10 W	Colombia	Amazonas	Rio Caqueta	Isla Clemencia	100	1	1	1	4	7
20130	EBC	13	b	24/05/84	0 47 S	72 2 W	Colombia	Amazonas	Rio Caqueta	Casa de Sebastian	100	3	2	0	4	7
20140	EBC	14	b	24/05/84	0 47 S	72 2 W	Colombia	Amazonas	Rio Caqueta	Casa de Sebastian	100	3	2	0	4	7
20150	EBC	15	c	24/05/84	0 45 S	71 59 W	Colombia	Amazonas	Rio Caqueta	Punta Sebastian	100	1	1	0	4	7

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
20160	EBC	16	c	24/05/84	0 45 S	71 59 W	Colombia	Amazonas	Rio Caqueta	Punta Sebastian	100	1	1	0	4	7
20170	EBC	17	c	24/05/84	0 45 S	71 59 W	Colombia	Amazonas	Rio Caqueta	Punta Sebastian	100	1	1	0	4	7
20180	EBC	18	c	24/05/84	0 45 S	71 59 W	Colombia	Amazonas	Rio Caqueta	Punta Sebastian	100	1	1	0	4	7
20190	EBC	19	c	24/05/84	0 45 S	71 59 W	Colombia	Amazonas	Rio Caqueta	Punta Sebastian	100	1	1	0	4	7
20200	EBC	20	c	24/05/84	0 45 S	71 59 W	Colombia	Amazonas	Rio Caqueta	Punta Sebastian	100	1	1	0	4	7
20210	EBC	21	c	24/05/84	0 45 S	71 59 W	Colombia	Amazonas	Rio Caqueta	Punta Sebastian	100	1	1	0	4	7
20220	EBC	22	c	24/05/84	0 45 S	72 4 W	Colombia	Amazonas	Rio Caqueta	Isla Mariname	100	1	1	0	4	7
20230	EBC	23	c	24/05/84	0 45 S	72 4 W	Colombia	Amazonas	Rio Caqueta	Isla Mariname	100	1	1	0	4	7
20240	EBC	24	c	24/05/84	0 41 S	72 0 W	Colombia	Amazonas	Rio Caqueta	Quebrada del Arroz	100	1	1	0	4	7
20250	EBC	25	c	24/05/84	0 41 S	72 0 W	Colombia	Amazonas	Rio Caqueta	Quebrada del Arroz	100	1	1	0	4	7
20260	EBC	26	c	24/05/84	0 41 S	72 0 W	Colombia	Amazonas	Rio Caqueta	Quebrada del Arroz	100	1	1	0	4	7
20270	EBC	27	c	24/05/84	0 41 S	72 0 W	Colombia	Amazonas	Rio Caqueta	Quebrada del Arroz	100	1	1	0	4	7
20280	EBC	28	c	24/05/84	0 41 S	72 0 W	Colombia	Amazonas	Rio Caqueta	Quebrada del Arroz	100	1	1	0	4	7
20290	EBC	29	c	24/05/84	0 41 S	72 0 W	Colombia	Amazonas	Rio Caqueta	Quebrada del Arroz	100	1	2	1	4	7
20300	EBC	30	c	24/05/84	0 41 S	72 0 W	Colombia	Amazonas	Rio Caqueta	Quebrada del Arroz	100	1	2	1	4	7
20310	EBC	31	b	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	3	2	1	4	7
20320	EBC	32	c	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	1	1	1	4	7
20330	EBC	33	c	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	1	1	0	4	7
20340	EBC	34	c	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	1	1	0	4	7
20350	EBC	35	c	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	1	1	0	4	7
20360	EBC	36	c	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	1	1	0	4	7
20370	EBC	37	c	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	1	1	0	4	7
20380	EBC	38	c	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	1	1	0	4	7
20390	EBC	39	c	25/05/84	0 45 S	71 57 W	Colombia	Amazonas	Rio Caqueta	fr Isla Ma Cristina	100	1	1	0	4	7
20400	EBC	40	c	25/05/84	0 46 S	71 53 W	Colombia	Amazonas	Rio Caqueta	Puerto Engano	100	1	1	0	4	7
20410	EBC	41	c	25/05/84	0 46 S	71 53 W	Colombia	Amazonas	Rio Caqueta	Puerto Engano	100	1	1	0	4	7
20420	EBC	42	c	25/05/84	0 46 S	71 53 W	Colombia	Amazonas	Rio Caqueta	Puerto Engano	100	1	1	0	4	7
20430	EBC	43	c	25/05/84	0 49 S	71 48 W	Colombia	Amazonas	Rio Caqueta	El Bufeo	100	1	1	0	4	7
20440	EBC	44	c	25/05/84	0 49 S	71 48 W	Colombia	Amazonas	Rio Caqueta	El Bufeo	100	1	1	0	4	7
20450	EBC	45	c	25/05/84	0 49 S	71 48 W	Colombia	Amazonas	Rio Caqueta	El Bufeo	100	1	1	0	4	7

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
20460	EBC	46	c	25/05/84	0 49 S	71 48 W	Colombia	Amazonas	Rio Caqueta	El Bufeo	100	1	1	0	4	7
20470	EBC	47	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	1	0	4	7
20480	EBC	48	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	2	1	4	7
20490	EBC	49	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	2	0	4	7
20500	EBC	50	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	1	0	4	7
20510	EBC	51	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	1	0	4	7
20520	EBC	52	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	2	0	4	7
20530	EBC	53	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	1	0	4	7
20540	EBC	54	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	1	0	4	7
20550	EBC	55	c	26/05/84	0 55 S	71 48 W	Colombia	Amazonas	Rio Caqueta	Raudal Quinche	100	1	2	0	4	7
20560	EBC	56	c	26/05/84	0 58 S	71 40 W	Colombia	Amazonas	Rio Caqueta	Dos Islas	100	1	2	0	4	7
20570	EBC	57	c	27/05/84	1 2 S	71 18 W	Colombia	Amazonas	Rio Caqueta	Cano Coca	100	1	1	0	4	8
20580	EBC	58	c	27/05/84	1 2 S	71 18 W	Colombia	Amazonas	Rio Caqueta	Cano Coca	100	1	2	1	4	8
20590	EBC	59	c	27/05/84	1 2 S	71 18 W	Colombia	Amazonas	Rio Caqueta	Cano Coca	100	1	1	0	4	8
20600	EBC	60	c	27/05/84	1 2 S	71 18 W	Colombia	Amazonas	Rio Caqueta	Cano Coca	100	1	1	0	4	8
20610	EBC	61	c	27/05/84	1 2 S	71 18 W	Colombia	Amazonas	Rio Caqueta	Cano Coca	100	1	1	0	4	8
20620	EBC	62	c	27/05/84	1 2 S	71 18 W	Colombia	Amazonas	Rio Caqueta	Cano Coca	100	1	1	0	4	8
20630	EBC	63	c	27/05/84	1 2 S	71 18 W	Colombia	Amazonas	Rio Caqueta	Cano Coca	100	1	1	0	4	8
20640	EBC	64	c	27/05/84	1 3 S	71 17 W	Colombia	Amazonas	Rio Caqueta	Remanso de Tijereta	100	1	1	0	4	8
20650	EBC	65	c	27/05/84	1 3 S	71 17 W	Colombia	Amazonas	Rio Caqueta	Remanso de Tijereta	100	1	1	0	4	8
20660	EBC	66	c	27/05/84	1 3 S	71 13 W	Colombia	Amazonas	Rio Caqueta	Raudal de Tijereta	100	1	1	0	4	8
20670	EBC	67	c	28/05/84	1 6 S	71 8 W	Colombia	Amazonas	Rio Caqueta	El Venado	100	1	1	0	4	8
20680	EBC	68	c	28/05/84	1 6 S	71 8 W	Colombia	Amazonas	Rio Caqueta	El Venado	100	1	1	0	4	8
20690	EBC	69	c	28/05/84	1 6 S	71 8 W	Colombia	Amazonas	Rio Caqueta	El Venado	100	1	1	0	4	8
20700	EBC	70	c	28/05/84	1 6 S	71 8 W	Colombia	Amazonas	Rio Caqueta	El Venado	100	1	1	0	4	8
20710	EBC	71	c	28/05/84	1 5 S	71 8 W	Colombia	Amazonas	Rio Caqueta	El Venado	100	1	1	0	4	8
20720	EBC	72	c	28/05/84	1 6 S	71 6 W	Colombia	Amazonas	Rio Caqueta	El Brazuelo	100	1	1	0	4	8
20730	EBC	73	c	28/05/84	1 6 S	71 6 W	Colombia	Amazonas	Rio Caqueta	El Brazuelo	100	1	1	0	4	8
20740	EBC	74	c	28/05/84	1 6 S	71 6 W	Colombia	Amazonas	Rio Caqueta	El Brazuelo	100	1	1	0	4	8
20750	EBC	75	h	29/05/84	1 8 S	71 6 W	Colombia	Amazonas	Rio Caqueta	Bellavista	100	1	1	0	4	8

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
20760	EBC	76	c	29/05/84	1 11 S	71 2 W	Colombia	Amazonas	Rio Caqueta	Solarte: E Mirana	100	1	2	1	4	8
20770	EBC	77	c	29/05/84	1 11 S	71 2 W	Colombia	Amazonas	Rio Caqueta	Solarte: E Mirana	100	1	2	1	4	8
20780	EBC	78	c	29/05/84	1 11 S	71 2 W	Colombia	Amazonas	Rio Caqueta	Solarte: E Mirana	100	1	2	1	4	8
20790	EBC	79	c	29/05/84	1 11 S	71 2 W	Colombia	Amazonas	Rio Caqueta	Solarte: E Mirana	100	1	2	1	4	8
20800	EBC	80	c	29/05/84	1 13 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas	100	1	1	0	4	8
20810	EBC	81	c	29/05/84	1 13 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas	100	1	1	0	4	8
20820	EBC	82	c	29/05/84	1 13 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas	100	1	1	0	4	8
20830	EBC	83	c	29/05/84	1 13 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas	100	1	1	0	4	8
20840	EBC	84	c	29/05/84	1 16 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas: J Bora	100	1	2	1	4	8
20850	EBC	85	c	29/05/84	1 16 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas: J Bora	100	1	2	1	4	8
20860	EBC	86	c	29/05/84	1 16 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas: J Bora	100	1	2	1	4	8
20870	EBC	87	c	29/05/84	1 16 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas: J Bora	100	1	2	1	4	8
20880	EBC	88	c	29/05/84	1 16 S	71 0 W	Colombia	Amazonas	Rio Caqueta	Las Palmas: J Bora	100	1	2	0	4	8
20890	EBC	89	c	29/05/84	1 15 S	70 59 W	Colombia	Amazonas	Rio Caqueta	casa A Mirana	100	1	1	0	4	8
20900	EBC	90	c	29/05/84	1 15 S	70 59 W	Colombia	Amazonas	Rio Caqueta	casa A Mirana	100	1	1	0	4	8
20910	EBC	91	c	29/05/84	1 15 S	70 59 W	Colombia	Amazonas	Rio Caqueta	casa A Mirana	100	1	2	1	4	8
21000	EBC	100	c	30/05/84	1 17 S	70 55 W	Colombia	Amazonas	Rio Caqueta	fr Islas Miranas	100	1	2	1	4	8
21010	EBC	101	c	31/05/84	1 24 S	70 35 W	Colombia	Amazonas	Rio Caqueta	Puerto Maria Manteca	100	1	2	1	4	8
21020	EBC	102	g	31/05/84	1 24 S	70 35 W	Colombia	Amazonas	Rio Caqueta	Puerto Maria Manteca	100	3	2	0	4	8
21030	EBC	103	c	01/06/84	1 33 S	70 8 W	Colombia	Amazonas	Rio Caqueta	Isla Manacaro	100	1	1	0	4	9
21040	EBC	104	c	01/06/84	1 33 S	70 8 W	Colombia	Amazonas	Rio Caqueta	Isla Manacaro	100	1	1	0	4	9
21050	EBC	105	c	02/06/84	1 29 S	70 0 W	Colombia	Amazonas	Rio Caqueta	Isla Zumaeta	100	1	1	0	4	9
21060	EBC	106	c	02/06/84	1 29 S	70 0 W	Colombia	Amazonas	Rio Caqueta	Isla Zumaeta	100	1	1	0	4	9
21070	EBC	107	c	02/06/84	1 29 S	70 0 W	Colombia	Amazonas	Rio Caqueta	Isla Zumaeta	100	1	1	0	4	9
21080	EBC	108	c	02/06/84	1 29 S	70 0 W	Colombia	Amazonas	Rio Caqueta	Isla Zumaeta	100	1	1	0	4	9
21090	EBC	109	c	02/06/84	1 12 S	69 53 W	Colombia	Amazonas	Rio Caqueta	Puerto Miriti	100	1	1	0	4	9
21100	EBC	110	c	02/06/84	1 12 S	69 53 W	Colombia	Amazonas	Rio Caqueta	Puerto Miriti	100	1	1	0	4	9
21110	EBC	111	h	02/06/84	1 12 S	69 49 W	Colombia	Amazonas	Rio Caqueta	Isla Miriti	100	1	1	0	4	9
21120	EBC	112	c	04/06/84	1 4 S	70 14 W	Colombia	Amazonas	Rio Apaporis	Centro Providencia	100	1	1	0	4	9
21130	EBC	113	c	04/06/84	0 55 S	70 18 W	Colombia	Amazonas	Rio Apaporis	cerca Providencia	100	1	1	0	4	9

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
21140	EBC	114	c	04/06/84	0 55 S	70 18 W	Colombia	Amazonas	Rio Apaporis	cerca Providencia	100	1	1	1	4	9
21150	EBC	115	c	04/06/84	0 55 S	70 18 W	Colombia	Amazonas	Rio Apaporis	cerca Providencia	100	1	1	0	4	9
21160	EBC	116	c	04/06/84	0 55 S	70 18 W	Colombia	Amazonas	Rio Apaporis	cerca Providencia	100	1	1	0	4	9
21170	EBC	117	c	04/06/84	0 55 S	70 18 W	Colombia	Amazonas	Rio Apaporis	cerca Providencia	100	1	1	0	4	9
21180	EBC	118	c	04/06/84	0 55 S	70 18 W	Colombia	Amazonas	Rio Apaporis	cerca Providencia	100	1	1	1	4	9
21190	EBC	119	c	04/06/84	1 4 S	70 14 W	Colombia	Amazonas	Rio Apaporis	Centro Providencia	100	1	1	0	4	9
21200	EBC	120	c	06/06/84	1 17 S	69 33 W	Colombia	Amazonas	Rio Caqueta	fr Isla Baranoa	100	1	1	0	4	9
21210	EBC	121	c	06/06/84	1 17 S	69 32 W	Colombia	Amazonas	Rio Caqueta	Isla Baranoa	100	1	2	1	4	9
21220	EBC	122	c	06/06/84	1 17 S	69 32 W	Colombia	Amazonas	Rio Caqueta	Isla Baranoa	100	1	1	1	4	9
21230	EBC	123	c	03/06/84	1 17 S	69 41 W	Colombia	Amazonas	Rio Caqueta	fr Isla Cordoba	100	1	2	1	4	9
21240	EBC	124	c	03/06/84	1 17 S	69 41 W	Colombia	Amazonas	Rio Caqueta	fr Isla Cordoba	100	1	1	0	4	9
21250	EBC	125	c	03/06/84	1 17 S	69 41 W	Colombia	Amazonas	Rio Caqueta	fr Isla Cordoba	100	1	2	1	4	9
21260	EBC	126	c	03/06/84	1 17 S	69 41 W	Colombia	Amazonas	Rio Caqueta	fr Isla Cordoba	100	1	2	1	4	9
21270	EBC	127	c	03/06/84	1 17 S	69 41 W	Colombia	Amazonas	Rio Caqueta	fr Isla Cordoba	100	1	2	1	4	9
21280	EBC	128	c	03/06/84	1 17 S	69 41 W	Colombia	Amazonas	Rio Caqueta	fr Isla Cordoba	100	1	1	0	4	9
21290	EBC	129	c	04/06/84	1 17 S	69 44 W	Colombia	Amazonas	Rio Caqueta	fr Puerto Cordoba	100	1	1	0	4	9
21300	EBC	130	c	04/06/84	1 17 S	69 44 W	Colombia	Amazonas	Rio Caqueta	fr Puerto Cordoba	100	1	1	0	4	9
21310	EBC	131	c	04/06/84	1 13 S	69 49 W	Colombia	Amazonas	Rio Caqueta	fr Isla Miriti	100	1	1	0	4	9
21320	EBC	132	c	04/06/84	1 13 S	69 49 W	Colombia	Amazonas	Rio Caqueta	fr Isla Miriti	100	1	1	0	4	9
21330	EBC	133	c	04/06/84	1 18 S	69 41 W	Colombia	Amazonas	Rio Caqueta	Remanso de Cordoba	100	1	1	0	4	9
21340	EBC	134	c	04/06/84	1 18 S	69 41 W	Colombia	Amazonas	Rio Caqueta	Remanso de Cordoba	100	1	1	0	4	9
21350	EBC	135	c	04/06/84	1 18 S	69 41 W	Colombia	Amazonas	Rio Caqueta	Remanso de Cordoba	100	1	2	1	4	9
21360	EBC	136	c	04/06/84	1 15 S	69 43 W	Colombia	Amazonas	Rio Caqueta	Puerto Cordoba	100	1	1	0	4	9
21370	EBC	137	c	04/06/84	1 15 S	69 43 W	Colombia	Amazonas	Rio Caqueta	Puerto Cordoba	100	1	1	0	4	9
21380	EBC	138	c	05/06/84	1 21 S	69 26 W	Colombia	Amazonas	Rio Caqueta	Isla Guillermina	100	1	1	0	4	9
21390	EBC	139	c	05/06/84	1 27 S	69 27 W	Colombia	Amazonas	Rio Caqueta	Isla Colombia	100	1	2	1	4	9
21400	EBC	140	c	05/06/84	1 28 S	69 28 W	Colombia	Amazonas	Rio Caqueta	Isla Colombia	100	1	1	0	4	9
21410	EBC	141	c	05/06/84	1 27 S	69 27 W	Colombia	Amazonas	Rio Caqueta	fr Isla Colombia	100	1	1	0	4	9
21420	EBC	142	c	05/06/84	1 27 S	69 27 W	Colombia	Amazonas	Rio Caqueta	Isla Colombia	100	1	1	1	4	9
21430	EBC	143	c	05/06/84	1 27 S	69 27 W	Colombia	Amazonas	Rio Caqueta	Isla Colombia	100	1	2	1	4	9

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
21440	EBC	144	c	05/06/84	1 28 S	69 28 W	Colombia	Amazonas	Rio Caqueta	fr Isla Colombia	100	1	1	0	4	9
21450	EBC	145	c	06/06/84	1 19 S	69 27 W	Colombia	Amazonas	Rio Caqueta	fr Isla Barahona	100	1	1	1	4	9
21460	EBC	146	c	06/06/84	1 16 S	69 25 W	Colombia	Amazonas	Rio Caqueta	Isla Barahona	100	1	1	0	4	9
21470	EBC	147	c	06/06/84	1 14 S	69 28 W	Colombia	Amazonas	Rio Caqueta	fr Isla Barahona	100	1	2	1	4	9
21480	EBC	148	c	05/06/84	1 28 S	69 28 W	Colombia	Amazonas	Rio Caqueta	fr Isla Colombia	100	1	2	1	4	9
31250	MORONA	125	c	01/01/38	5 0 S	77 0 W	Peru				100	-1	1	1	6	11
40480	NANAY	48	c	01/01/38	4 0 S	74 0 W	Peru				100	-1	1	1	6	12
40690	NANAY	69	c	01/01/38	4 0 S	74 0 W	Peru				100	-1	1	1	6	12
41730	NANAY	173	c	01/01/38	4 0 S	74 0 W	Peru				100	-1	1	1	6	12
42550	NANAY	255	c	01/01/38	4 0 S	74 0 W	Peru				100	-1	1	1	6	12
50130	PARINARI	13	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
50160	PARINARI	16	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
50290	PARINARI	29	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
50300	PARINARI	30	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
50350	PARINARI	35	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
50390	PARINARI	39	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
50460	PARINARI	46	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
50810	PARINARI	81	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
51210	PARINARI	121	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
51211	PARINARI	121	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
51500	PARINARI	150	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
51670	PARINARI	167	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
51690	PARINARI	169	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
51890	PARINARI	189	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
52850	PARINARI	285	c	01/01/38	5 0 S	74 0 W	Peru				100	-1	1	1	6	13
60060	SCAVINA	6	c	01/01/38	3 0 S	74 0 W	Peru				100	-1	1	1	6	14
60120	SCAVINA	12	c	01/01/38	3 0 S	74 0 W	Peru				100	-1	1	1	6	14
70050	IMC	5	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70060	IMC	6	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70100	IMC	10	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15

Descriptor Reference (see Appendix 1):

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15	1.17	1.18
Access	Collpref	Num	Spe	Date	Latitud	Longitu	Country	Province	Locality 1	Locality 2	Alti	Stat	Mate	Surv	Pop1	Pop2
70110	IMC	11	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70140	IMC	14	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70200	IMC	20	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70230	IMC	23	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70270	IMC	27	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70310	IMC	31	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70420	IMC	42	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70440	IMC	44	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70530	IMC	53	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70570	IMC	57	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70600	IMC	60	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70670	IMC	67	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
70760	IMC	76	c	01/01/38	4 0 S	73 0 W	Peru				100	-1	1	1	6	15
80010	POUND	1	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	15
80040	POUND	4	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	16
80050	POUND	5	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	16
80070	POUND	7	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	16
80071	POUND	7	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	16
80080	POUND	8	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	16
80120	POUND	12	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	16
80121	POUND	12	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	16
80122	POUND	12	c	01/01/43	3 0 S	73 0 W	Peru				100	-1	1	1	6	16
90190	EET	19	c	01/01/60	1 30 S	80 0 W	Ecuador				100	-1	1	1	5	10
90960	EET	96	c	01/01/60	1 30 S	80 0 W	Ecuador				100	-1	1	1	5	10
91030	EET	103	c	01/01/60	1 30 S	80 0 W	Ecuador				100	-1	1	1	5	10
100010	NACIONAL	1	c	01/01/60	2 0 S	80 0 W	Ecuador				100	-1	1	1	5	10

APPENDIX 3 DATA SETS 2 AND 6: POD, SEED AND PRODUCTION DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
LCT-EEN	1	3	5	7	5	1	0	181	85	0.47	21	543	25	2.08	24.3	12.0	10.5	0.49	1	1	0	25.77	1.55
LCT-EEN	4	5	1	5	5	1	0	163	81	0.50	16	482	36	1.53	22.6	13.2	8.2	0.58	0	0	1	22.94	1.21
LCT-EEN	5	3	5	7	7	1	0	168	99	0.59	18	625	33	1.79	26.2	13.3	8.4	0.51	1	1	1	15.43	1.96
LCT-EEN	6	0	5	7	5	1	0	163	105	0.64	22	710	19	2.16	25.4	12.5	9.8	0.49	1	1	1	26.88	1.96
LCT-EEN	8	0	5	5	5	1	0	170	103	0.61	25	771	12	2.00	23.3	12.3	10.0	0.53	0	0	1	25.77	3.23
LCT-EEN	10	3	5	7	5	1	0	149	87	0.58	21	497	16	2.50	24.2	14.1	9.3	0.58	0	0	1	31.65	1.98
LCT-EEN	11	0	5	7	5	1	0	116	79	0.68	16	298	13	1.83	23.3	12.8	10.4	0.55	1	0	0	49.02	1.57
LCT-EEN	14	0	1	7	5	1	0	108	100	0.93	18	412	14	1.71	22.8	12.7	9.3	0.56	1	0	0	41.67	1.71
LCT-EEN	15	3	5	7	5	1	0	136	87	0.64	16	420	23	1.55	23.0	12.1	8.4	0.53	1	1	1	26.04	1.67
LCT-EEN	18	0	5	7	5	1	0	155	80	0.52	16	431	20	1.79	23.5	13.4	9.8	0.57	0	1	0	23.81	2.10
LCT-EEN	19	0	1	7	5	1	0	205	110	0.54	20	981	33	2.29	24.8	13.6	10.9	0.55	0	1	0	15.06	2.01
LCT-EEN	20	3	5	7	5	1	0	158	84	0.53	17	459	22	2.06	22.9	13.8	9.7	0.60	1	1	0	23.81	1.91
LCT-EEN	22	0	5	7	7	1	0	178	95	0.53	21	661	36	1.59	23.0	11.9	8.7	0.52	0	0	1	17.73	1.57
LCT-EEN	23	3	1	5	7	1	0	190	96	0.51	18	788	38	2.38	26.0	13.1	9.7	0.50	0	1	1	13.02	2.02
LCT-EEN	24	0	1	5	7	1	0	170	95	0.56	22	550	29	1.50	22.6	13.6	9.0	0.60	1	1	0	22.94	1.50
LCT-EEN	25	0	5	7	5	1	0	161	101	0.63	25	705	25	2.28	26.2	14.8	8.9	0.56	0	1	0	24.27	1.65
LCT-EEN	26	0	5	5	5	1	0	180	103	0.57	17	713	40	1.84	27.6	12.9	7.7	0.47	0	0	1	11.85	2.11
LCT-EEN	27	3	5	0	3	1	5	151	74	0.49	14	364	30	1.00	20.6	11.2	7.7	0.54	0	0	1	32.05	1.04
LCT-EEN	28	3	5	5	5	1	0	160	88	0.55	22	525	20	1.93	22.7	13.4	9.1	0.59	0	1	0	32.89	1.52
LCT-EEN	29	0	1	7	5	1	0	120	96	0.80	22	542	21	1.93	25.2	15.0	8.3	0.60	0	1	1	27.17	1.75
LCT-EEN	31	3	5	5	5	1	0	215	107	0.50	21	988	37	2.82	26.8	14.9	11.3	0.56	1	1	0	11.47	2.36
LCT-EEN	36	0	5	7	7	1	0	160	91	0.57	18	571	14	2.20	25.1	13.6	10.1	0.54	1	1	1	24.75	2.89
LCT-EEN	37	0	1	7	7	1	0	220	121	0.55	26	1325	37	2.28	27.3	13.3	10.2	0.49	0	0	1	10.37	2.61
LCT-EEN	38	5	5	5	5	1	0	260	120	0.46	27	1450	31	2.88	30.3	13.9	11.3	0.46	0	1	0	13.30	2.43
LCT-EEN	39	3	1	5	5	1	0	147	78	0.53	15	435	29	-1.00	23.4	13.4	8.2	0.57	0	1	1	25.51	1.35
LCT-EEN	44	0	5	5	5	1	0	150	89	0.59	22	453	10	3.20	27.9	14.9	11.8	0.53	1	1	0	39.68	2.52
LCT-EEN	46	3	9	7	7	1	0	176	84	0.48	17	348	21	1.35	21.1	11.4	9.0	0.54	0	0	1	44.64	1.07
LCT-EEN	48	0	5	5	5	1	0	163	96	0.59	22	632	30	-1.00	-1.0	-1.0	-1.0	-1.00	1	1	0	22.12	1.51
LCT-EEN	58	0	1	0	1	1	0	135	91	0.67	20	451	14	1.86	23.8	13.3	9.4	0.56	1	0	0	50.00	1.43
LCT-EEN	59	0	9	7	7	1	0	210	112	0.53	34	1315	28	2.53	28.6	13.9	9.8	0.49	0	1	0	12.95	2.76

APPENDIX 3 DATA SETS 2 AND 6: POD, SEED AND PRODUCTION DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
LCT-EEN	60	0	9	7	5	1	0	190	110	0.58	21	937	38	2.26	24.6	13.8	10.3	0.56	0	1	0	11.96	2.20
LCT-EEN	61	0	5	5	5	1	0	183	90	0.49	21	683	22	1.74	24.1	13.2	9.6	0.55	0	1	1	23.36	1.95
LCT-EEN	62	5	5	5	3	1	5	118	63	0.53	4	140	40	1.11	19.2	11.1	6.2	0.58	0	0	1	41.67	0.60
LCT-EEN	64	0	1	7	5	1	0	130	83	0.64	20	501	11	1.63	25.6	11.5	9.4	0.45	1	0	0	64.10	1.42
LCT-EEN	65	0	5	7	7	1	0	145	100	0.69	20	640	16	2.75	26.3	15.1	10.9	0.57	1	0	0	30.12	2.08
LCT-EEN	66	0	1	5	5	1	0	151	91	0.60	18	532	32	1.40	24.5	12.9	6.8	0.53	1	0	0	20.83	1.50
LCT-EEN	67	0	5	5	5	1	5	138	83	0.60	15	450	36	1.35	23.3	13.5	7.1	0.58	1	0	1	21.37	1.30
LCT-EEN	68	0	5	5	5	1	0	147	87	0.59	18	435	23	1.85	23.7	12.6	10.2	0.53	0	0	1	32.05	1.36
LCT-EEN	69	0	5	5	5	1	0	158	102	0.65	28	590	15	2.31	24.7	12.9	10.8	0.52	1	0	0	29.41	2.27
LCT-EEN	70	3	5	5	3	1	0	153	87	0.57	20	530	31	1.50	-1.0	-1.0	-1.0	-1.00	1	0	1	31.25	1.03
LCT-EEN	71	0	5	5	5	1	0	143	82	0.57	20	412	14	1.75	25.1	13.5	10.1	0.54	1	0	1	48.08	1.49
LCT-EEN	72	0	1	5	5	1	0	188	102	0.54	19	725	36	1.89	25.6	12.5	11.7	0.49	1	1	1	17.86	1.56
LCT-EEN	73	0	5	5	5	1	0	165	99	0.60	19	704	46	1.92	25.8	14.0	8.3	0.54	0	1	1	13.23	1.64
LCT-EEN	74	0	1	5	5	1	0	130	102	0.78	15	354	17	1.80	27.5	12.0	7.3	0.44	1	0	1	46.30	1.27
LCT-EEN	76	0	1	5	5	1	0	160	88	0.55	17	492	41	1.00	23.4	12.7	7.9	0.54	0	0	1	27.17	0.90
LCT-EEN	77	0	5	5	5	1	0	165	97	0.59	19	670	34	2.14	25.1	13.3	9.0	0.53	1	0	1	17.01	1.73
LCT-EEN	78	0	5	0	5	1	3	160	100	0.62	28	745	10	2.22	24.8	12.4	9.6	0.50	1	0	1	38.46	2.60
LCT-EEN	79	3	5	5	5	1	0	167	86	0.51	20	543	26	1.74	23.0	12.5	9.0	0.54	1	0	1	25.00	1.54
LCT-EEN	80	0	5	7	7	1	0	240	95	0.40	25	928	34	2.00	23.1	13.2	11.4	0.57	0	1	0	19.69	1.49
LCT-EEN	81	0	5	7	7	1	0	195	87	0.45	21	762	73	2.00	25.2	13.4	8.9	0.53	1	1	0	10.82	1.27
LCT-EEN	82	3	1	5	7	1	0	167	98	0.59	27	675	14	1.75	24.6	13.5	10.1	0.55	1	1	1	44.64	1.60
LCT-EEN	83	0	5	5	5	1	0	161	100	0.62	20	770	32	2.25	25.8	13.8	9.2	0.53	0	0	1	13.89	2.25
LCT-EEN	84	0	5	0	5	1	0	129	88	0.68	19	483	24	1.67	23.4	13.1	8.8	0.56	0	1	1	30.12	1.38
LCT-EEN	85	0	1	5	5	1	0	156	96	0.62	19	633	30	1.75	24.5	14.2	9.6	0.58	1	1	1	23.81	1.40
LCT-EEN	86	0	5	5	5	1	0	153	91	0.59	16	638	45	1.75	22.6	12.8	8.3	0.57	0	0	1	17.24	1.29
LCT-EEN	87	0	5	5	5	1	0	125	77	0.62	11	345	23	2.14	24.4	14.1	9.1	0.58	1	0	1	26.32	1.65
LCT-EEN	88	0	1	5	5	1	0	159	95	0.60	16	575	36	2.25	26.1	14.4	8.6	0.55	1	0	0	16.34	1.70
LCT-EEN	89	0	5	5	5	1	0	151	99	0.66	16	628	33	2.50	28.7	14.7	9.0	0.51	1	1	0	16.34	1.85
LCT-EEN	90	3	1	5	5	1	0	135	86	0.64	16	353	24	1.50	23.1	12.6	8.3	0.55	0	0	1	43.10	0.97
LCT-EEN	92	0	1	5	5	1	0	200	98	0.49	18	850	43	1.50	22.6	12.6	8.4	0.56	0	0	1	18.52	1.26

APPENDIX 3 DATA SETS 2 AND 6: POD, SEED AND PRODUCTION DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
LCT-EEN	93	0	5	5	5	1	0	150	102	0.68	24	783	28	2.25	26.3	14.5	8.5	0.55	1	0	1	20.33	1.76
LCT-EEN	94	3	7	7	7	1	0	193	86	0.45	15	522	38	1.80	20.0	12.7	8.1	0.64	0	1	1	14.45	1.82
LCT-EEN	97	5	9	5	5	1	0	175	90	0.51	18	558	35	1.98	23.2	13.9	9.6	0.60	1	1	0	17.48	1.63
LCT-EEN	101	0	5	7	5	1	0	140	85	0.61	17	409	27	1.50	21.5	12.2	9.4	0.57	1	0	0	28.09	1.32
LCT-EEN	107	5	7	7	7	1	0	218	85	0.39	16	582	41	1.80	24.3	11.7	8.6	0.48	0	1	1	17.01	1.43
LCT-EEN	108	5	5	7	7	1	0	197	91	0.46	20	700	34	2.30	24.5	14.9	9.8	0.61	0	1	1	11.85	2.48
LCT-EEN	109	0	5	7	5	1	0	120	86	0.72	13	282	15	1.67	24.3	12.3	9.3	0.51	1	0	0	48.08	1.39
LCT-EEN	113	0	1	5	5	1	0	150	118	0.79	35	968	4	3.00	24.5	15.0	13.0	0.61	1	0	0	83.33	3.00
LCT-EEN	115	0	5	5	3	1	0	185	99	0.54	18	-1	-1	2.57	26.7	16.7	8.8	0.63	1	1	0	-1.00	-1.00
LCT-EEN	123	3	5	5	5	1	0	180	84	0.47	18	510	10	3.12	27.3	15.3	12.2	0.56	1	1	0	52.08	1.92
LCT-EEN	125	0	5	5	5	1	0	220	127	0.58	29	1215	27	2.50	27.3	15.2	10.5	0.56	0	1	0	19.84	1.87
LCT-EEN	127	0	5	5	5	1	0	190	108	0.57	22	828	42	1.40	25.5	11.8	8.5	0.46	0	0	1	24.27	0.98
LCT-EEN	128	0	5	5	3	1	0	110	90	0.82	17	310	4	3.00	29.3	15.3	12.0	0.52	1	0	0	89.29	2.80
LCT-EEN	129	0	5	7	5	1	0	140	98	0.70	20	424	9	2.88	29.6	15.2	11.8	0.51	1	0	0	52.08	2.13
LCT-EEN	130	3	5	5	5	1	0	180	108	0.60	20	668	29	2.10	27.0	13.3	9.8	0.49	1	0	0	21.74	1.59
LCT-EEN	131	0	5	7	5	1	0	150	98	0.65	22	500	17	2.36	24.3	14.6	10.6	0.60	1	0	0	43.10	1.36
LCT-EEN	132	5	9	7	5	1	0	180	107	0.59	19	686	25	2.66	25.4	15.2	9.7	0.60	1	0	1	19.23	2.08
LCT-EEN	133	0	0	7	7	1	0	192	106	0.55	18	867	34	2.40	25.2	15.0	9.3	0.60	1	1	1	11.26	2.61
LCT-EEN	134	0	9	7	5	1	0	155	91	0.59	14	356	22	2.00	24.6	13.3	9.4	0.54	1	1	0	31.25	1.45
LCT-EEN	136	0	1	7	5	1	0	168	106	0.63	24	895	33	2.50	25.8	14.6	8.8	0.57	1	0	0	14.71	2.06
LCT-EEN	142	0	1	5	5	1	0	143	100	0.70	18	431	23	1.23	23.0	12.1	8.8	0.53	1	0	0	37.31	1.17
LCT-EEN	145	0	9	7	5	1	0	185	92	0.50	15	549	24	2.20	24.9	13.9	9.6	0.56	1	0	0	25.77	1.62
LCT-EEN	148	0	5	5	5	1	0	125	76	0.61	12	269	21	1.09	20.6	11.5	7.9	0.56	1	0	0	54.35	0.88
LCT-EEN	151	3	5	7	5	1	0	177	97	0.55	13	479	34	2.53	26.6	14.5	9.7	0.55	0	1	1	15.43	1.91
LCT-EEN	152	0	1	0	3	1	0	139	92	0.66	17	547	34	1.65	25.9	12.7	7.6	0.49	1	0	0	20.66	1.42
LCT-EEN	154	3	5	5	5	1	0	160	89	0.56	14	563	40	1.55	25.2	11.9	7.6	0.47	1	0	1	14.79	1.69
LCT-EEN	156	3	9	7	7	1	0	195	91	0.47	19	714	14	2.15	22.8	13.5	11.5	0.59	1	0	0	24.51	2.91
LCT-EEN	157	0	5	5	3	1	0	165	105	0.64	15	648	32	2.72	29.3	15.6	9.5	0.53	1	0	1	12.25	2.55
LCT-EEN	158	0	5	7	7	1	0	216	113	0.52	24	1041	32	2.16	25.4	14.0	9.9	0.55	1	0	0	13.44	2.33
LCT-EEN	159	0	5	5	5	1	0	105	88	0.84	19	413	10	3.25	27.0	15.1	11.4	0.56	1	0	0	48.08	2.08

APPENDIX 3 DATA SETS 2 AND 6: POD, SEED AND PRODUCTION DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
LCT-EEN	160	0	5	7	7	1	0	210	122	0.58	25	1369	38	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	1	10.50	2.51
LCT-EEN	162	3	5	7	5	1	0	150	94	0.63	17	552	26	2.40	24.7	14.6	9.7	0.59	1	0	0	20.00	1.92
LCT-EEN	163	0	5	5	5	1	0	155	90	0.58	18	557	22	2.09	23.6	13.9	10.5	0.59	1	0	0	20.16	2.25
LCT-EEN	166	0	5	5	5	1	0	189	118	0.62	22	1008	35	2.92	29.0	16.1	10.1	0.56	1	1	0	10.20	2.80
LCT-EEN	169	0	5	7	7	1	0	208	128	0.62	20	1437	41	2.90	29.8	15.2	9.0	0.51	1	0	1	7.23	3.38
LCT-EEN	175	5	9	7	7	1	0	170	93	0.55	17	521	34	1.92	23.3	12.4	8.4	0.53	0	0	1	17.36	1.69
LCT-EEN	178	0	5	5	5	1	0	149	87	0.58	14	431	40	1.76	25.0	13.0	8.8	0.52	1	1	1	19.38	1.29
LCT-EEN	179	0	5	7	5	1	0	195	99	0.51	18	726	37	1.67	25.2	12.9	8.2	0.51	1	0	0	16.34	1.65
LCT-EEN	181	0	1	0	5	1	0	190	100	0.53	16	270	40	2.00	26.3	12.4	9.2	0.47	0	1	0	15.24	1.64
LCT-EEN	182	3	9	7	5	1	0	156	81	0.52	15	360	38	1.05	20.7	10.3	7.4	0.50	1	1	1	30.49	0.86
LCT-EEN	185	5	5	5	5	1	0	207	83	0.40	12	547	36	1.75	25.6	11.7	8.3	0.46	1	0	1	18.94	1.47
LCT-EEN	188	3	5	5	5	1	0	145	74	0.51	11	354	41	1.20	21.9	10.9	7.1	0.50	0	1	1	18.66	1.31
LCT-EEN	189	5	9	5	5	1	0	177	68	0.38	13	318	32	0.97	19.3	10.2	8.3	0.53	1	1	1	37.88	0.83
LCT-EEN	194	0	9	5	3	1	0	126	72	0.57	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
LCT-EEN	196	3	5	3	5	0	0	142	68	0.48	14	316	23	0.82	16.8	8.7	7.4	0.52	0	0	1	75.76	0.57
LCT-EEN	197	0	9	0	3	0	5	76	46	0.61	6	57	12	0.53	14.9	8.3	6.7	0.56	0	1	1	227.27	0.37
LCT-EEN	199	3	9	0	3	0	3	150	67	0.45	8	240	49	0.86	19.3	9.3	6.4	0.48	0	0	1	32.47	0.63
LCT-EEN	201	3	5	7	5	1	0	180	97	0.54	19	688	38	1.78	25.9	12.2	8.3	0.47	0	1	0	14.12	1.86
LCT-EEN	202	3	5	7	5	1	0	185	115	0.62	18	812	40	2.29	28.1	14.3	8.5	0.51	0	0	1	14.12	1.77
LCT-EEN	203	0	1	7	5	1	0	110	92	0.84	17	432	27	1.52	24.8	13.5	8.0	0.54	1	1	1	31.25	1.19
LCT-EEN	205	0	5	7	5	1	0	145	96	0.66	20	497	12	2.25	25.8	12.6	9.9	0.49	0	1	1	45.45	1.83
LCT-EEN	212	3	5	0	5	-1	3	-1	-1	-1.00	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	0	0	1	-1.00	-1.00
LCT-EEN	213	0	5	5	7	-1	0	-1	-1	-1.00	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	0	1	0	-1.00	-1.00
LCT-EEN	214	3	5	7	7	-1	0	-1	-1	-1.00	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	0	1	1	-1.00	-1.00
LCT-EEN	215	3	5	0	3	-1	3	-1	-1	-1.00	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	0	0	1	-1.00	-1.00
LCT-EEN	218	0	7	5	5	1	0	184	105	0.57	16	752	39	2.80	26.8	16.2	9.1	0.60	1	1	0	9.36	2.74
LCT-EEN	219	3	5	7	7	1	0	174	109	0.63	22	730	29	2.60	25.1	15.5	9.8	0.62	1	0	0	15.06	2.29
LCT-EEN	220	0	9	7	5	1	0	178	85	0.48	18	587	27	1.25	23.2	11.5	9.1	0.50	1	1	0	17.99	2.06
LCT-EEN	223	3	5	5	5	-1	0	170	92	0.54	15	548	36	2.09	25.9	14.5	8.8	0.56	1	0	0	14.97	1.86
LCT-EEN	237	3	5	7	5	1	0	215	112	0.52	25	974	32	2.90	27.9	16.0	9.1	0.57	0	1	1	12.63	2.48

DATA SETS 2 AND 6: POD, SEED AND PRODUCTION DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
LCT-EEN	239	0	1	7	3	1	-1	185	110	0.59	25	960	35	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	12.14	2.35
LCT-EEN	241	0	5	7	7	1	0	180	104	0.58	19	782	32	2.58	25.0	16.3	11.1	0.65	1	0	0	12.95	2.41
LCT-EEN	242	0	9	5	5	1	0	170	93	0.55	16	597	23	2.75	27.7	13.8	11.5	0.50	1	0	0	16.13	2.70
LCT-EEN	243	0	5	0	5	1	0	110	84	0.76	20	334	-1	1.50	23.0	11.6	8.6	0.50	1	0	0	58.14	-1.00
LCT-EEN	244	3	9	7	5	1	0	205	110	0.54	21	689	28	2.44	26.2	13.7	10.5	0.52	1	0	0	16.23	2.20
LCT-EEN	245	0	5	7	7	1	0	193	-1	-1.00	-1	932	-1	-1.00	-1.0	-1.0	-1.0	-1.00	1	0	0	-1.00	-1.00
LCT-EEN	246	3	5	7	7	1	0	165	89	0.54	20	461	13	2.90	26.5	14.8	11.4	0.56	1	0	0	30.12	2.55
LCT-EEN	248	0	5	5	7	1	0	180	107	0.59	25	847	18	1.47	20.5	11.3	9.6	0.55	1	1	1	39.06	1.42
LCT-EEN	249	0	7	7	7	1	0	235	88	0.37	14	970	-1	2.49	24.2	14.0	11.4	0.58	0	0	1	9.29	-1.00
LCT-EEN	250	0	5	5	5	1	0	165	108	0.65	19	802	40	1.84	26.3	13.8	8.0	0.52	0	0	1	11.42	2.19
LCT-EEN	251	3	5	5	5	1	0	182	107	0.59	20	904	45	1.85	27.2	13.6	8.3	0.50	1	0	0	12.38	1.80
LCT-EEN	252	0	5	5	7	1	0	215	108	0.50	24	986	34	2.16	26.1	13.8	9.5	0.53	1	0	1	15.62	1.88
LCT-EEN	254	0	5	5	7	1	0	165	92	0.56	14	441	37	2.05	24.6	15.1	8.5	0.61	0	1	1	18.52	1.46
LCT-EEN	257	5	9	5	5	1	0	230	95	0.41	19	873	48	2.35	28.6	14.3	9.3	0.50	1	0	1	13.30	1.57
LCT-EEN	258	0	5	5	7	1	0	222	103	0.46	22	937	46	2.12	24.9	15.1	8.5	0.61	1	1	0	10.59	2.05
LCT-EEN	269	0	5	5	5	1	0	148	92	0.62	16	527	32	2.00	25.2	12.5	8.9	0.50	1	1	0	14.79	2.11
LCT-EEN	271	0	5	5	5	1	0	183	112	0.61	22	863	30	1.76	21.9	13.3	9.4	0.61	1	0	0	21.01	1.59
LCT-EEN	280	0	5	5	5	1	0	178	95	0.53	16	675	36	2.00	23.4	13.0	9.0	0.56	1	1	1	11.96	2.32
LCT-EEN	281	3	5	7	7	1	0	174	89	0.51	16	489	24	1.92	23.1	13.6	9.6	0.59	0	1	1	20.83	2.00
LCT-EEN	288	0	5	7	5	1	0	175	104	0.59	17	477	27	2.29	24.1	14.0	9.8	0.58	1	0	0	32.05	1.16
LCT-EEN	295	3	9	5	5	1	0	173	79	0.46	13	352	40	1.34	21.2	11.6	8.8	0.55	1	0	0	35.21	0.71
LCT-EEN	296	3	5	5	5	1	0	165	88	0.53	11	398	38	1.75	24.3	13.0	8.7	0.53	1	0	0	25.77	1.02
LCT-EEN	300	0	9	0	3	0	0	180	97	0.54	15	497	24	2.67	26.0	14.3	11.1	0.55	1	0	0	26.88	1.55
LCT-EEN	302	0	9	7	5	1	0	175	92	0.53	12	400	38	2.09	23.7	15.0	9.5	0.63	1	0	0	19.53	1.35
LCT-EEN	306	0	1	5	5	1	0	143	88	0.62	13	506	-1	1.70	24.8	13.5	8.0	0.54	1	0	0	-1.00	-1.00
LCT-EEN	313	0	5	5	5	1	0	165	98	0.59	22	585	12	3.25	32.1	14.2	11.2	0.44	1	0	0	27.17	3.07
LCT-EEN	320	0	5	5	5	1	0	155	85	0.55	16	448	15	2.40	27.1	14.1	10.8	0.52	1	0	0	27.47	2.43
LCT-EEN	321	3	9	5	5	1	0	185	96	0.52	15	527	30	1.91	22.9	14.0	9.2	0.61	1	1	1	18.94	1.76
LCT-EEN	322	3	5	5	5	1	0	173	90	0.52	16	505	34	1.73	23.1	12.4	9.0	0.54	0	0	1	20.33	1.45
LCT-EEN	323	5	5	5	5	1	0	218	97	0.44	19	697	38	2.14	27.4	14.1	8.8	0.51	1	1	0	13.97	1.88

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
LCT-EEN	324	3	5	5	5	1	0	159	83	0.52	16	395	29	1.25	21.1	11.4	8.0	0.54	1	0	1	43.86	0.79
LCT-EEN	325	0	9	5	5	1	0	208	87	0.42	18	541	27	2.13	22.2	13.7	11.3	0.62	1	0	0	32.47	1.14
LCT-EEN	326	0	1	5	5	1	3	223	123	0.55	25	1182	40	1.53	21.5	12.7	9.1	0.59	1	0	1	18.52	1.35
LCT-EEN	328	0	5	5	5	1	0	162	97	0.60	14	527	26	2.48	26.1	14.6	10.1	0.56	1	0	0	19.53	1.97
LCT-EEN	329	3	9	5	5	1	0	145	87	0.60	14	346	13	2.36	26.8	13.0	10.3	0.49	1	0	0	37.88	2.03
LCT-EEN	330	0	5	5	5	1	0	135	78	0.58	13	280	9	3.28	27.3	15.5	12.5	0.57	1	0	0	53.19	2.09
LCT-EEN	331	3	5	5	7	1	0	126	75	0.60	15	315	18	1.86	22.2	14.9	9.9	0.67	1	0	0	45.45	1.22
LCT-EEN	332	0	1	5	5	1	0	170	95	0.56	18	645	35	2.80	26.9	15.9	10.2	0.59	1	0	0	14.29	2.00
LCT-EEN	335	0	5	5	3	1	0	165	86	0.52	11	407	39	1.86	23.8	13.7	9.2	0.58	1	0	0	20.49	1.25
LCT-EEN	336	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	-1	-1	-1	1.33	21.3	12.6	8.5	0.59	1	1	0	-1.00	-1.00
LCT-EEN	337	0	5	7	5	1	0	210	106	0.50	19	830	38	2.40	24.2	14.6	9.7	0.60	1	0	0	14.71	1.79
LCT-EEN	338	0	5	5	5	1	0	155	90	0.58	15	505	43	1.85	23.7	14.1	7.5	0.59	1	0	0	15.15	1.53
LCT-EEN	339	3	9	5	5	1	0	190	99	0.52	17	656	40	2.18	24.5	14.0	9.3	0.57	0	1	0	16.67	1.50
LCT-EEN	340	0	9	5	5	1	0	145	86	0.59	11	328	27	1.90	25.5	13.2	8.0	0.52	1	0	0	27.17	1.36
LCT-EEN	341	0	9	5	3	1	0	148	77	0.52	10	296	34	1.65	22.1	13.2	8.1	0.60	1	0	0	21.74	1.35
LCT-EEN	342	0	1	5	5	1	0	160	90	0.56	20	486	21	2.50	25.1	13.8	11.7	0.55	1	0	0	27.17	1.75
LCT-EEN	343	0	9	5	5	1	0	155	76	0.49	10	313	41	1.30	21.2	12.9	7.1	0.61	1	0	0	24.04	1.01
LCT-EEN	344	0	5	5	5	1	0	170	97	0.57	18	535	23	2.32	24.8	14.2	10.7	0.57	1	0	0	25.00	1.74
LCT-EEN	345	3	9	5	5	1	0	165	73	0.44	10	279	16	2.43	25.7	13.6	10.6	0.53	1	0	0	39.68	1.58
LCT-EEN	346	0	9	5	3	1	0	152	88	0.58	10	368	41	1.98	26.8	14.7	8.1	0.55	1	0	0	16.89	1.44
LCT-EEN	347	3	5	5	5	1	0	132	81	0.61	12	253	14	2.31	25.8	14.6	11.8	0.57	1	1	0	49.02	1.46
LCT-EEN	348	0	5	5	5	1	0	135	86	0.64	17	349	18	1.91	27.2	14.2	9.3	0.52	1	0	0	27.47	2.02
LCT-EEN	349	0	5	0	5	1	0	150	89	0.59	16	397	13	2.64	26.6	14.7	11.2	0.55	0	1	0	40.32	1.91
LCT-EEN	350	0	9	7	5	1	0	172	87	0.51	16	353	15	2.65	25.7	14.7	11.5	0.57	0	1	0	36.23	1.84
LCT-EEN	351	3	5	5	5	1	0	135	88	0.65	11	343	25	2.26	26.1	15.3	9.6	0.59	1	1	0	22.32	1.79
LCT-EEN	352	0	9	5	5	1	0	180	88	0.49	10	438	42	2.14	27.3	15.5	8.6	0.57	1	0	0	15.15	1.57
LCT-EEN	353	3	1	7	7	1	0	140	90	0.64	18	383	11	2.49	24.2	14.9	12.6	0.62	1	0	0	49.02	1.85
LCT-EEN	354	0	5	5	3	1	0	145	91	0.63	14	404	26	2.49	26.7	15.8	9.3	0.59	1	0	0	21.93	1.75
LCT-EEN	355	0	9	5	5	1	0	205	106	0.52	19	802	37	2.36	25.2	14.2	9.8	0.56	1	0	0	12.38	2.18
LCT-EEN	356	0	5	5	5	1	0	130	88	0.68	15	355	37	1.54	24.4	13.8	7.2	0.57	1	0	0	28.41	0.95

DATA SETS 2 AND 6: POD, SEED AND PRODUCTION DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
LCT-EEN	357	0	5	5	5	1	0	175	109	0.62	18	725	41	2.35	26.7	15.3	9.9	0.57	1	0	0	13.16	1.85
LCT-EEN	358	0	9	7	5	1	0	152	88	0.58	14	373	19	2.45	27.2	14.8	10.9	0.54	1	1	0	30.86	1.71
LCT-EEN	359	3	5	5	5	1	0	148	99	0.67	18	464	15	2.91	28.6	14.8	11.4	0.52	1	0	0	29.07	2.29
LCT-EEN	360	3	5	5	5	1	0	168	82	0.49	11	357	35	1.85	24.1	13.9	8.7	0.58	1	0	0	21.55	1.33
LCT-EEN	361	0	9	7	5	1	0	150	92	0.61	16	365	18	3.04	30.4	17.2	9.7	0.57	1	0	0	26.88	2.07
LCT-EEN	362	5	5	5	5	1	0	125	66	0.53	9	200	42	0.80	19.7	9.1	5.4	0.46	0	1	1	29.76	0.80
LCT-EEN	364	5	5	5	5	1	0	122	58	0.48	9	165	19	0.80	18.6	8.6	6.9	0.46	0	0	1	58.14	0.91
LCT-EEN	367	5	5	5	5	1	0	139	61	0.44	11	314	32	1.09	22.6	10.7	7.4	0.47	0	0	1	25.77	1.21
LCT-EEN	369	5	0	5	5	1	0	141	68	0.48	9	278	37	0.90	18.9	9.4	7.1	0.50	0	0	1	33.78	0.80
LCT-EEN	371	3	7	5	5	1	0	137	58	0.42	14	325	37	0.75	18.0	8.8	7.1	0.49	0	0	1	22.73	1.19
LCT-EEN	372	3	5	5	5	1	0	180	110	0.61	18	752	39	2.03	28.4	14.5	8.0	0.51	1	0	0	19.23	1.33
LCT-EEN	391	5	1	5	5	1	0	110	85	0.77	19	309	7	2.00	24.8	15.8	12.2	0.64	1	0	0	96.15	1.49
LCT-EEN	393	3	5	5	5	1	0	154	97	0.63	16	516	31	2.31	26.8	14.9	8.7	0.56	0	1	1	19.38	1.66
LCT-EEN	403	3	9	7	7	1	0	160	75	0.47	15	311	29	1.10	18.4	9.8	7.9	0.53	0	1	0	46.30	0.74
LCT-EEN	404	3	5	5	5	1	0	150	72	0.48	11	302	34	1.15	22.3	10.5	6.9	0.47	1	0	1	27.17	1.08
LCT-EEN	405	3	5	5	5	1	0	150	65	0.43	14	264	19	1.06	20.4	9.9	8.0	0.49	0	1	0	54.35	0.97
LCT-EEN	408	3	5	0	3	-1	0	140	65	0.46	11	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	0	1	1	-1.00	-1.00
LCT-EEN	418	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	-1	-1	-1	1.49	21.5	14.0	7.9	0.65	1	0	0	-1.00	-1.00
LCT-EEN	424	0	9	7	7	1	0	255	125	0.49	24	1385	35	2.90	28.7	15.8	9.9	0.55	1	0	1	10.82	2.64
LCT-EEN	430	0	5	5	5	1	0	215	104	0.48	20	883	34	2.68	29.3	14.4	10.1	0.49	1	0	1	15.53	1.89
EBC	5	3	5	0	3	1	0	122	60	0.49	7	165	29	0.71	19.9	9.0	7.6	0.45	0	1	0	41.67	0.83
EBC	6	5	5	0	3	1	0	135	64	0.47	9	170	35	0.77	20.1	8.7	7.2	0.43	0	1	0	55.56	0.51
EBC	8	3	5	0	3	1	0	135	62	0.46	9	-1	34	0.77	19.7	9.0	7.0	0.46	-1	-1	-1	-1.00	-1.00
EBC	9	3	1	0	3	1	0	124	67	0.54	9	185	36	1.00	21.1	10.5	7.7	0.50	0	1	1	47.17	0.59
EBC	10	3	5	5	3	1	0	177	67	0.38	10	280	38	1.03	21.2	9.9	8.1	0.47	0	1	1	39.06	0.67
EBC	11	3	5	0	3	1	0	118	59	0.50	6	-1	28	0.81	19.4	10.0	7.3	0.52	0	1	1	-1.00	-1.00
EBC	29	0	5	0	3	1	0	120	62	0.52	8	150	32	1.05	21.4	10.4	8.0	0.49	0	0	1	45.45	0.69
EBC	30	3	5	5	3	1	0	144	68	0.47	11	215	40	0.64	19.4	9.2	6.6	0.47	0	1	1	50.00	0.50
EBC	48	0	5	0	3	1	0	147	66	0.45	10	230	38	0.91	19.9	9.8	8.0	0.49	0	1	0	43.10	0.61
EBC	49	3	5	5	3	1	0	145	68	0.47	11	-1	35	0.58	17.7	8.7	6.8	0.49	0	1	0	-1.00	-1.00

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
EBC	56	0	5	0	3	1	0	136	58	0.43	10	-1	41	0.60	19.1	8.7	6.5	0.46	0	1	0	-1.00	-1.00
EBC	58	5	5	0	3	1	0	144	67	0.47	8	185	38	0.95	21.1	10.2	7.4	0.48	0	1	0	47.17	0.56
EBC	76	3	5	5	5	1	0	265	100	0.38	20	1050	50	2.28	26.7	13.6	10.4	0.51	0	0	1	13.37	1.50
EBC	77	3	5	5	5	1	0	247	84	0.34	20	825	23	2.62	25.5	10.6	8.9	0.42	0	1	0	28.41	1.53
EBC	78	5	9	5	5	1	0	175	62	0.35	11	250	20	1.40	22.2	11.8	9.8	0.53	0	1	1	44.64	1.12
EBC	79	3	5	5	5	1	0	238	90	0.38	19	875	49	1.57	25.8	12.4	8.9	0.48	0	1	0	17.61	1.16
EBC	84	5	5	5	5	1	0	180	77	0.43	12	400	39	1.45	24.7	10.0	8.1	0.40	0	1	0	24.04	1.07
EBC	85	3	9	0	3	1	0	171	67	0.39	10	300	39	0.96	20.1	10.1	8.1	0.50	0	1	0	37.31	0.69
EBC	86	3	5	5	3	1	0	147	60	0.41	7	185	34	0.96	20.7	9.0	7.7	0.43	0	1	0	47.17	0.62
EBC	91	5	5	0	3	1	0	104	58	0.56	8	120	24	0.94	18.8	10.1	8.5	0.54	0	1	0	65.79	0.63
EBC	100	3	5	0	3	1	0	139	65	0.47	10	200	16	1.63	25.0	10.3	8.2	0.41	0	1	0	64.10	0.98
EBC	101	3	5	0	3	1	0	268	106	0.40	23	1375	54	2.21	28.3	13.4	10.0	0.47	0	1	1	9.92	1.87
EBC	113	5	5	5	5	1	0	125	65	0.52	14	215	9	-1.00	-1.0	-1.0	-1.0	-1.00	0	1	0	75.76	1.47
EBC	121	3	5	0	3	1	0	223	90	0.40	20	900	-1	2.20	24.4	11.3	9.3	0.46	0	0	1	-1.00	-1.00
EBC	123	-1	-1	-1	-1	-1	-1	185	72	0.39	14	-1	-1	1.43	25.3	11.0	8.7	0.43	-1	-1	-1	-1.00	-1.00
EBC	125	3	5	0	3	1	0	200	86	0.43	15	600	35	1.89	24.8	12.3	9.8	0.50	0	1	1	21.93	1.30
EBC	126	0	5	3	3	1	0	207	81	0.39	11	575	48	1.39	24.4	10.7	7.3	0.44	0	1	0	14.29	1.46
EBC	135	0	5	0	5	1	0	224	94	0.42	20	850	47	1.49	24.1	12.4	8.7	0.51	0	1	1	13.97	1.52
EBC	138	3	1	5	5	1	0	172	90	0.52	20	650	20	2.11	26.2	12.6	8.7	0.48	0	0	1	31.25	1.60
EBC	141	5	5	5	3	1	0	154	63	0.41	9	220	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
EBC	142	0	5	0	3	1	0	172	99	0.58	22	825	28	2.15	26.7	14.0	8.7	0.52	0	1	1	21.55	1.66
EBC	145	3	5	0	3	1	0	196	84	0.43	21	625	-1	1.47	22.9	12.3	8.9	0.54	-1	-1	-1	-1.00	-1.00
EBC	147	0	5	5	3	1	0	163	81	0.50	18	500	31	-1.00	-1.0	-1.0	-1.0	-1.00	0	1	1	32.47	0.99
EBC	148	3	5	5	3	1	0	203	95	0.47	16	750	43	2.36	27.1	13.8	9.1	0.51	0	0	1	15.15	1.53
MORONA	125	0	0	0	5	0	0	125	85	0.68	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
NANAY	48	5	5	3	3	1	0	155	80	0.52	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
NANAY	69	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
NANAY	173	3	5	7	7	1	0	185	70	0.38	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
NANAY	255	0	5	5	5	1	0	190	85	0.45	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	13	5	7	7	7	1	0	240	80	0.33	-1	-1	33	1.63	22.0	11.1	9.0	0.50	0	0	1	-1.00	-1.00

APPENDIX 3 DATA SETS 2 AND 6: POD, SEED AND PRODUCTION DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
PARINARI	16	3	5	5	5	1	0	190	85	0.45	-1	-1	34	1.29	20.4	11.2	7.8	0.55	0	1	1	-1.00	-1.00
PARINARI	29	3	7	5	5	1	0	150	70	0.47	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	30	0	7	7	5	1	0	190	75	0.39	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	35	0	7	5	5	1	0	190	95	0.50	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	39	3	7	5	5	1	3	200	85	0.42	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	46	5	7	5	5	1	0	180	75	0.42	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	81	3	7	7	5	1	0	205	100	0.49	-1	-1	32	1.61	22.6	12.0	8.7	0.53	0	1	1	-1.00	-1.00
PARINARI	121	5	7	7	7	1	5	165	80	0.48	-1	-1	29	1.58	21.9	11.3	8.2	0.52	0	0	1	-1.00	-1.00
PARINARI	121	5	7	7	5	1	0	170	90	0.53	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	150	5	7	7	5	1	0	260	95	0.37	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	167	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	169	3	5	5	3	1	0	175	85	0.49	-1	-1	31	1.69	23.4	12.5	9.1	0.53	0	1	1	-1.00	-1.00
PARINARI	189	0	5	5	5	1	0	150	80	0.53	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
PARINARI	285	5	7	5	5	1	3	200	75	0.38	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
SCAVINA	6	3	7	7	7	1	3	160	70	0.44	-1	-1	40	1.19	20.2	10.2	7.9	0.50	0	1	0	-1.00	-1.00
SCAVINA	12	0	7	5	5	1	0	185	70	0.38	-1	-1	38	1.03	21.6	10.6	7.4	0.49	0	0	1	-1.00	-1.00
IMC	5	0	5	3	3	1	0	175	105	0.60	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	6	0	7	5	5	1	0	205	105	0.51	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	10	0	5	5	5	1	0	175	85	0.49	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	11	3	7	3	5	1	0	160	80	0.50	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	14	3	5	5	5	1	0	200	85	0.42	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	20	3	5	5	5	1	0	140	70	0.50	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	23	0	5	7	5	1	0	215	100	0.47	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	27	0	5	5	5	1	0	175	100	0.57	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	31	0	5	3	5	1	0	200	105	0.53	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	42	0	5	5	5	1	0	160	85	0.53	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	44	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	53	0	5	3	5	1	0	215	105	0.49	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	57	0	5	3	5	1	0	160	100	0.62	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
IMC	60	0	5	3	3	1	0	175	100	0.57	-1	-1	44	1.41	22.4	11.8	6.9	0.53	0	0	1	-1.00	-1.00

APPENDIX 3 DATA SETS 2 AND 6: POD, SEED AND PRODUCTION DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.15	2.16	2.17	2.18	2.19	2.20	2.21	2.22	2.23	6.1	6.2
Collpref	Num	Bas	Ape	Sur	Fur	Pai	Pig	Len	Diam	D:LRa	Wall	Podw	Snum	Sweig	Slen	Swid	Sthi	Swlra	Swhi	Sltp	Sdkp	Pindx	Sindx
IMC	67	-1	-1	-1	-1	-1	0	-1	-1	-1.00	-1	-1	42	1.34	22.1	11.4	8.4	0.52	0	0	1	-1.00	-1.00
IMC	76	0	7	5	5	1	0	205	85	0.41	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
POUND	1	0	5	5	3	1	0	205	95	0.46	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
POUND	4	0	7	7	7	1	0	180	75	0.42	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
POUND	5	3	5	5	5	1	0	180	80	0.44	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
POUND	7	0	5	5	5	1	0	195	95	0.49	-1	-1	38	1.50	22.0	12.3	8.5	0.56	0	0	1	-1.00	-1.00
POUND	7	0	5	5	5	1	0	175	100	0.57	-1	-1	38	1.50	22.0	12.3	8.5	0.56	0	0	1	-1.00	-1.00
POUND	8	3	5	5	5	1	5	165	85	0.52	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
POUND	12	3	5	5	5	1	0	175	75	0.43	-1	-1	40	1.32	22.2	10.1	7.5	0.45	0	1	1	-1.00	-1.00
POUND	12	3	5	5	5	1	0	165	80	0.48	-1	-1	40	1.32	22.2	10.1	7.5	0.45	0	1	1	-1.00	-1.00
POUND	12	3	5	3	5	1	0	185	85	0.46	-1	-1	40	1.32	22.2	10.1	7.5	0.45	0	1	1	-1.00	-1.00
EET	19	5	7	7	5	1	3	204	95	0.47	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
EET	96	3	5	7	7	1	3	190	93	0.49	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
EET	103	3	7	7	7	1	0	197	95	0.48	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00
NACIONAL	1	-1	-1	-1	-1	-1	0	-1	-1	-1.00	-1	-1	-1	-1.00	-1.0	-1.0	-1.0	-1.00	-1	-1	-1	-1.00	-1.00

APPENDIX 4 DATA SETS 3, 4 AND 5: HABIT, FLUSH COLOUR AND FLOWER DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.5
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
LCT-EEN	1	3.4	5.4	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	2	-1.0	-1.0	-1.0	3	5	3	3	0	10.0
LCT-EEN	4	5.8	5.6	1.8	5	5	5	5	3	9.5
LCT-EEN	5	4.6	5.0	1.9	5	3	3	0	5	6.5
LCT-EEN	6	5.6	5.2	1.8	3	0	0	0	3	7.5
LCT-EEN	8	5.4	5.4	1.4	0	-1	-1	-1	-1	-1.0
LCT-EEN	10	5.6	4.8	2.1	0	0	0	0	0	13.0
LCT-EEN	11	4.2	4.6	1.4	0	-1	-1	-1	-1	-1.0
LCT-EEN	14	6.1	6.6	1.6	0	-1	-1	-1	-1	-1.0
LCT-EEN	15	5.0	5.3	1.9	0	-1	-1	-1	-1	-1.0
LCT-EEN	18	6.0	4.8	1.8	0	-1	-1	-1	-1	-1.0
LCT-EEN	19	5.5	5.0	2.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	20	5.4	5.2	1.9	5	5	3	0	0	10.0
LCT-EEN	21	5.8	5.0	2.1	0	0	0	0	0	8.5
LCT-EEN	22	5.0	4.8	2.2	7	-1	-1	-1	-1	-1.0
LCT-EEN	23	6.2	5.8	1.9	3	-1	-1	-1	-1	-1.0
LCT-EEN	24	5.2	5.2	2.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	25	5.0	5.0	1.4	0	-1	-1	-1	-1	-1.0
LCT-EEN	26	5.5	5.0	2.6	0	0	0	0	0	10.0
LCT-EEN	27	6.5	5.5	2.5	5	7	5	0	7	9.5
LCT-EEN	30	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	31	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	32	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	33	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	36	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	37	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	38	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	44	4.1	3.0	1.7	0	-1	-1	-1	-1	-1.0
LCT-EEN	46	-1.0	-1.0	-1.0	5	5	3	5	5	9.0
LCT-EEN	48	5.2	5.0	1.7	0	-1	-1	-1	-1	-1.0

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.6
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
LCT-EEN	49	-1.0	-1.0	-1.0	0	0	0	0	0	9.5
LCT-EEN	52	-1.0	-1.0	-1.0	3	0	0	0	0	9.5
LCT-EEN	57	-1.0	-1.0	-1.0	0	0	0	0	0	9.5
LCT-EEN	58	5.2	4.2	1.4	0	-1	-1	-1	-1	-1.0
LCT-EEN	59	5.4	4.6	1.7	0	-1	-1	-1	-1	-1.0
LCT-EEN	60	-1.0	-1.0	-1.0	3	3	0	0	0	8.5
LCT-EEN	61	5.0	5.0	1.5	5	0	0	5	0	9.0
LCT-EEN	62	6.0	5.0	2.2	5	5	3	5	5	10.0
LCT-EEN	63	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	64	5.0	4.0	1.6	-1	-1	-1	-1	-1	-1.0
LCT-EEN	65	4.8	4.8	1.6	0	0	0	0	0	9.0
LCT-EEN	66	5.3	4.0	1.2	3	-1	-1	-1	-1	-1.0
LCT-EEN	67	5.0	4.4	1.3	5	-1	-1	-1	-1	-1.0
LCT-EEN	68	5.2	5.0	1.3	5	-1	-1	-1	-1	-1.0
LCT-EEN	69	5.0	4.3	1.4	0	-1	-1	-1	-1	-1.0
LCT-EEN	70	6.3	4.3	1.2	5	-1	-1	-1	-1	-1.0
LCT-EEN	71	5.5	4.5	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	72	6.2	5.0	1.5	5	-1	-1	-1	-1	-1.0
LCT-EEN	73	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	74	5.0	5.3	0.9	3	5	3	5	3	7.5
LCT-EEN	76	5.0	5.0	1.1	5	3	0	5	3	8.0
LCT-EEN	77	4.5	3.0	1.3	3	0	0	0	0	8.0
LCT-EEN	78	6.1	5.0	1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	79	5.4	3.8	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	80	-1.0	-1.0	-1.0	5	5	3	3	0	8.5
LCT-EEN	81	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	82	5.4	5.0	1.1	5	-1	-1	-1	-1	-1.0
LCT-EEN	83	5.6	3.8	1.4	3	0	0	3	0	8.5
LCT-EEN	84	5.0	5.2	1.1	3	0	0	0	0	6.5
LCT-EEN	85	4.6	5.2	1.2	0	3	0	3	0	7.5

APPENDIX 4 DATA SETS 3, 4 AND 5: HABIT, FLUSH COLOUR AND FLOWER DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.5
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
LCT-EEN	86	5.0	4.8	1.1	5	5	3	3	5	8.0
LCT-EEN	87	4.8	4.0	1.1	0	0	0	0	0	7.5
LCT-EEN	88	5.6	3.8	1.1	0	-1	-1	-1	-1	-1.0
LCT-EEN	89	5.0	4.3	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	90	5.0	5.4	1.1	3	-1	-1	-1	-1	-1.0
LCT-EEN	91	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	92	5.4	5.4	1.2	3	0	0	3	0	8.0
LCT-EEN	93	5.4	6.0	1.1	3	0	0	0	0	7.5
LCT-EEN	94	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	97	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	101	6.2	5.8	0.8	-1	0	0	0	0	9.0
LCT-EEN	103	-1.0	-1.0	-1.0	0	0	0	0	0	11.0
LCT-EEN	107	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	108	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	109	6.8	6.0	0.9	0	0	0	0	0	11.0
LCT-EEN	113	7.0	5.0	0.8	-1	0	0	0	0	8.0
LCT-EEN	115	5.0	4.8	0.8	-1	-1	-1	-1	-1	-1.0
LCT-EEN	120	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	121	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	122	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	123	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	124	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	125	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	126	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	127	5.2	5.2	0.8	3	3	0	5	0	12.3
LCT-EEN	128	7.0	7.0	0.5	0	0	0	0	0	10.0
LCT-EEN	130	4.6	3.4	0.7	0	-1	-1	-1	-1	-1.0
LCT-EEN	131	6.2	5.4	0.7	-1	-1	-1	-1	-1	-1.0
LCT-EEN	132	6.6	5.0	0.9	5	0	0	0	0	11.8
LCT-EEN	133	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.6
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
LCT-EEN	134	5.0	5.0	1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	136	5.4	5.0	1.0	-1	0	0	0	0	9.0
LCT-EEN	137	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	141	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	142	5.4	4.6	0.9	0	-1	-1	-1	-1	-1.0
LCT-EEN	144	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	145	5.2	5.0	0.8	0	0	0	0	0	7.5
LCT-EEN	146	-1.0	-1.0	-1.0	0	0	0	0	0	6.5
LCT-EEN	148	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	149	-1.0	-1.0	-1.0	0	0	0	0	0	9.5
LCT-EEN	150	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	151	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	152	-1.0	-1.0	-1.0	-1	0	0	0	0	7.5
LCT-EEN	153	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	154	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	155	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	156	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	157	-1.0	-1.0	-1.0	0	0	0	0	0	9.5
LCT-EEN	158	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	159	5.0	4.7	0.8	0	-1	-1	-1	-1	-1.0
LCT-EEN	160	5.4	4.8	0.8	5	-1	-1	-1	-1	-1.0
LCT-EEN	162	4.6	3.8	1.1	0	0	0	0	0	9.0
LCT-EEN	163	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	164	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	165	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	166	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	169	-1.0	-1.0	-1.0	0	0	0	0	0	7.0
LCT-EEN	172	-1.0	-1.0	-1.0	-1	0	0	0	0	9.0
LCT-EEN	175	-1.0	-1.0	-1.0	-1	0	0	0	0	9.0
LCT-EEN	178	-1.0	-1.0	-1.0	3	0	0	0	0	9.0

APPENDIX 4 DATA SETS 3, 4 AND 5: HABIT, FLUSH COLOUR AND FLOWER DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.5
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
LCT-EEN	181	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	182	5.2	4.2	1.2	3	0	0	0	0	8.5
LCT-EEN	185	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	187	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	188	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	189	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	193	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	195	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	196	5.9	5.6	0.7	5	-1	-1	-1	-1	-1.0
LCT-EEN	199	6.6	5.2	0.8	5	5	5	0	3	8.5
LCT-EEN	201	5.0	4.7	1.1	0	0	0	0	0	10.3
LCT-EEN	202	5.0	4.2	0.9	5	0	0	0	0	9.0
LCT-EEN	203	5.0	6.3	1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	205	5.7	5.3	1.0	5	5	5	5	0	7.0
LCT-EEN	212	6.0	5.2	0.9	7	7	7	0	5	9.0
LCT-EEN	214	5.5	5.5	1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	215	5.0	4.2	0.9	7	5	5	0	5	11.0
LCT-EEN	216	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	217	-1.0	-1.0	-1.0	0	0	0	0	0	10.0
LCT-EEN	218	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	219	-1.0	-1.0	-1.0	0	0	0	0	0	8.5
LCT-EEN	220	-1.0	-1.0	-1.0	0	0	0	0	0	6.8
LCT-EEN	221	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	222	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	223	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	227	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	228	-1.0	-1.0	-1.0	0	0	0	0	0	9.5
LCT-EEN	231	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	232	-1.0	-1.0	-1.0	0	0	0	0	0	10.0
LCT-EEN	234	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.6
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
LCT-EEN	236	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	237	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	241	4.3	5.0	1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	242	4.8	4.8	0.9	0	-1	-1	-1	-1	-1.0
LCT-EEN	243	3.0	7.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	244	5.0	5.0	1.1	0	-1	-1	-1	-1	-1.0
LCT-EEN	246	5.2	4.8	0.9	0	-1	-1	-1	-1	-1.0
LCT-EEN	248	5.4	4.8	0.6	3	-1	-1	-1	-1	-1.0
LCT-EEN	249	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	250	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	251	-1.0	-1.0	-1.0	0	0	0	0	0	9.5
LCT-EEN	253	-1.0	-1.0	-1.0	3	0	0	0	0	9.3
LCT-EEN	254	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	255	-1.0	-1.0	-1.0	0	0	0	0	0	7.0
LCT-EEN	257	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	258	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	259	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	264	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	269	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	271	5.4	5.4	0.6	0	-1	-1	-1	-1	-1.0
LCT-EEN	278	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	280	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	281	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	282	-1.0	-1.0	-1.0	0	0	0	0	0	9.0
LCT-EEN	288	5.0	5.0	0.8	0	-1	-1	-1	-1	-1.0
LCT-EEN	295	5.4	5.6	1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	296	5.4	5.7	0.9	3	-1	-1	-1	-1	-1.0
LCT-EEN	297	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	300	6.4	6.8	-1.0	0	0	0	0	0	7.5
LCT-EEN	302	5.0	6.0	0.9	0	-1	-1	-1	-1	-1.0

Descriptor Reference (see Appendix 1):

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.5
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
LCT-EEN	306	5.7	5.9	0.8	0	-1	-1	-1	-1	-1.0
LCT-EEN	312	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	313	5.6	6.0	1.1	0	-1	-1	-1	-1	-1.0
LCT-EEN	320	5.0	5.4	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	321	5.2	5.4	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	323	5.0	4.6	1.3	3	-1	-1	-1	-1	-1.0
LCT-EEN	324	4.6	5.0	1.1	3	-1	-1	-1	-1	-1.0
LCT-EEN	325	5.4	5.7	0.8	0	-1	-1	-1	-1	-1.0
LCT-EEN	326	4.7	6.7	1.3	0	-1	-1	-1	-1	-1.0
LCT-EEN	327	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	328	5.0	6.1	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	329	5.8	5.0	0.7	0	-1	-1	-1	-1	-1.0
LCT-EEN	331	5.7	7.0	0.9	0	-1	-1	-1	-1	-1.0
LCT-EEN	332	6.0	6.0	1.0	0	0	0	0	0	8.0
LCT-EEN	333	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	334	6.0	7.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	335	5.8	5.8	0.8	0	0	0	0	0	8.0
LCT-EEN	336	5.0	5.3	0.9	0	0	0	0	0	9.0
LCT-EEN	337	5.0	5.5	1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	338	5.5	6.3	0.9	0	-1	-1	-1	-1	-1.0
LCT-EEN	339	5.4	5.2	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	340	4.8	5.2	0.8	0	-1	-1	-1	-1	-1.0
LCT-EEN	341	5.0	6.0	1.1	0	-1	-1	-1	-1	-1.0
LCT-EEN	342	5.0	6.5	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	343	5.7	6.8	1.1	0	0	0	0	0	7.5
LCT-EEN	344	4.8	3.5	1.1	0	-1	-1	-1	-1	-1.0
LCT-EEN	345	5.0	4.8	0.9	0	0	0	0	0	8.0
LCT-EEN	346	5.2	5.6	1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	347	5.4	6.1	1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	348	5.0	6.6	0.8	0	-1	-1	-1	-1	-1.0

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.6
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
LCT-EEN	349	6.4	5.6	1.1	3	-1	-1	-1	-1	-1.0
LCT-EEN	350	6.0	5.7	1.1	3	-1	-1	-1	-1	-1.0
LCT-EEN	351	6.0	6.8	0.9	3	0	0	0	0	8.5
LCT-EEN	352	5.9	6.6	1.1	0	0	0	0	0	8.6
LCT-EEN	353	5.5	7.0	1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	354	6.4	6.7	1.2	0	-1	-1	-1	-1	-1.0
LCT-EEN	355	3.0	5.0	1.3	0	-1	-1	-1	-1	-1.0
LCT-EEN	356	6.0	6.5	1.1	0	-1	-1	-1	-1	-1.0
LCT-EEN	357	5.8	6.5	1.1	0	-1	-1	-1	-1	-1.0
LCT-EEN	358	5.7	6.1	0.9	3	-1	-1	-1	-1	-1.0
LCT-EEN	359	5.3	6.4	1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	360	5.9	5.7	1.1	0	-1	-1	-1	-1	-1.0
LCT-EEN	361	5.3	6.7	0.9	0	-1	-1	-1	-1	-1.0
LCT-EEN	362	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	364	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	366	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	367	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	368	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	369	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	370	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	371	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	372	6.3	6.3	0.8	3	-1	-1	-1	-1	-1.0
LCT-EEN	373	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	376	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	382	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	384	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	385	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	386	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	387	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	389	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0

APPENDIX 4 DATA SETS 3, 4 AND 5: HABIT, FLUSH COLOUR AND FLOWER DESCRIPTORS

Descriptor Reference (see Appendix 1):

1.2 CollPref	1.3 Num	3.1 Ban	3.2 Bfo	3.3 Jhe	4.1 Flush	5.1 Ped	5.2 Sep	5.3 Fil	5.4 Ova	5.5 Spl
LCT-EEN	390	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	391	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	393	5.0	5.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	395	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	399	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	406	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	409	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	412	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	414	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
LCT-EEN	415	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	420	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
LCT-EEN	422	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	426	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	427	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	429	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	432	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
LCT-EEN	434	-1.0	-1.0	-1.0	0	-1	-1	-1	-1	-1.0
EBC	5	6.3	7.0	1.0	5	-1	-1	-1	-1	-1.0
EBC	6	6.6	6.6	1.1	-1	-1	-1	-1	-1	-1.0
EBC	9	4.7	5.9	1.2	5	-1	-1	-1	-1	-1.0
EBC	10	7.0	6.6	1.1	5	-1	-1	-1	-1	-1.0
EBC	29	6.1	6.1	1.0	5	-1	-1	-1	-1	-1.0
EBC	30	5.7	5.7	0.8	5	-1	-1	-1	-1	-1.0
EBC	48	5.0	7.0	1.0	5	-1	-1	-1	-1	-1.0
EBC	58	5.8	6.0	1.0	0	-1	-1	-1	-1	-1.0
EBC	76	5.4	6.6	1.2	0	-1	-1	-1	-1	-1.0
EBC	77	5.9	6.4	1.2	5	-1	-1	-1	-1	-1.0
EBC	78	5.4	5.4	1.1	0	-1	-1	-1	-1	-1.0
EBC	79	5.0	5.0	1.1	0	-1	-1	-1	-1	-1.0
EBC	84	5.0	5.0	1.1	0	-1	-1	-1	-1	-1.0

1.2 CollPref	1.3 Num	3.1 Ban	3.2 Bfo	3.3 Jhe	4.1 Flush	5.1 Ped	5.2 Sep	5.3 Fil	5.4 Ova	5.6 Spl
EBC	85	5.5	4.3	1.1	-1	-1	-1	-1	-1	-1.0
EBC	86	5.0	4.7	1.2	-1	-1	-1	-1	-1	-1.0
EBC	91	4.5	4.0	1.1	-1	-1	-1	-1	-1	-1.0
EBC	100	5.5	5.0	1.1	-1	-1	-1	-1	-1	-1.0
EBC	101	6.7	6.4	1.0	-1	-1	-1	-1	-1	-1.0
EBC	114	-1.0	-1.0	-1.0	3	-1	-1	-1	-1	-1.0
EBC	121	6.0	5.7	1.4	-1	-1	-1	-1	-1	-1.0
EBC	125	5.0	6.6	0.8	-1	-1	-1	-1	-1	-1.0
EBC	126	5.0	6.0	1.1	-1	-1	-1	-1	-1	-1.0
EBC	135	5.0	7.0	-1.0	3	-1	-1	-1	-1	-1.0
EBC	138	5.0	5.0	1.2	5	-1	-1	-1	-1	-1.0
EBC	141	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
EBC	142	5.7	5.7	0.6	-1	-1	-1	-1	-1	-1.0
EBC	147	7.0	5.0	-1.0	-1	-1	-1	-1	-1	-1.0
EBC	148	5.0	5.0	1.1	-1	-1	-1	-1	-1	-1.0
MORONA	125	-1.0	-1.0	-1.0	5	5	3	5	0	7.5
NANAY	48	-1.0	-1.0	-1.0	-1	3	3	3	0	7.5
NANAY	69	-1.0	-1.0	-1.0	5	0	3	3	3	7.5
NANAY	173	-1.0	-1.0	-1.0	-1	5	3	5	3	8.0
PARINARI	13	-1.0	-1.0	-1.0	3	0	0	0	0	8.0
PARINARI	16	-1.0	-1.0	-1.0	5	3	3	5	0	7.5
PARINARI	30	-1.0	-1.0	-1.0	5	3	3	0	3	8.5
PARINARI	35	-1.0	-1.0	-1.0	5	5	3	0	3	9.0
PARINARI	39	-1.0	-1.0	-1.0	-1	0	0	0	0	8.5
PARINARI	46	-1.0	-1.0	-1.0	-1	0	0	3	3	8.0
PARINARI	81	-1.0	-1.0	-1.0	5	0	0	3	3	8.5
PARINARI	121	-1.0	-1.0	-1.0	5	3	3	0	3	7.5
PARINARI	121	-1.0	-1.0	-1.0	5	3	0	0	0	7.5
PARINARI	150	-1.0	-1.0	-1.0	3	3	3	0	0	7.0
PARINARI	167	-1.0	-1.0	-1.0	3	3	0	0	3	6.5

APPENDIX 4 DATA SETS 3, 4 AND 5: HABIT, FLUSH COLOUR AND FLOWER DESCRIPTORS

Descriptor Reference (see Appendix 1):

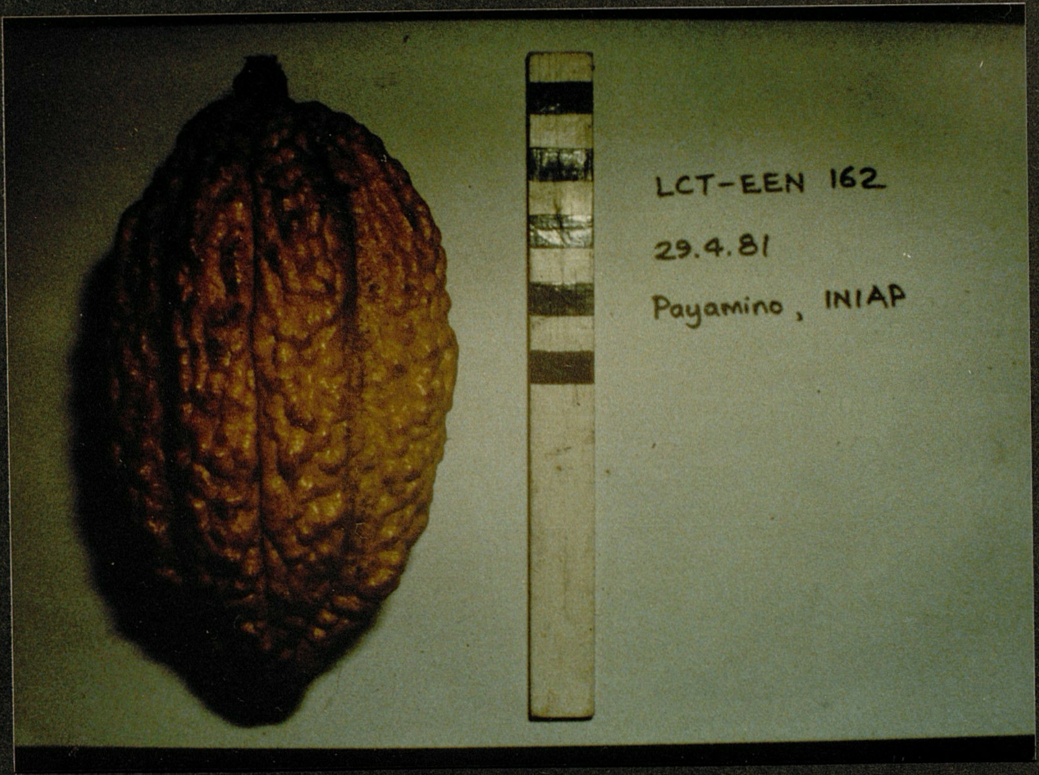
1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.5
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
PARINARI	169	-1.0	-1.0	-1.0	3	0	0	0	0	8.0
PARINARI	189	-1.0	-1.0	-1.0	5	-1	-1	-1	-1	-1.0
PARINARI	285	-1.0	-1.0	-1.0	-1	0	0	0	0	9.0
SCAVINA	6	-1.0	-1.0	-1.0	5	0	0	0	0	9.5
SCAVINA	12	-1.0	-1.0	-1.0	3	0	0	0	0	8.5
IMC	5	-1.0	-1.0	-1.0	3	5	3	3	0	7.5
IMC	6	-1.0	-1.0	-1.0	-1	3	3	3	0	8.5
IMC	10	-1.0	-1.0	-1.0	-1	3	3	3	0	7.5
IMC	11	-1.0	-1.0	-1.0	-1	3	3	3	0	9.5
IMC	14	-1.0	-1.0	-1.0	5	3	0	3	0	10.5
IMC	20	-1.0	-1.0	-1.0	-1	5	3	3	0	8.0
IMC	23	-1.0	-1.0	-1.0	3	0	0	3	0	8.0
IMC	27	-1.0	-1.0	-1.0	3	3	0	5	0	8.5
IMC	31	-1.0	-1.0	-1.0	3	5	3	3	0	7.5
IMC	42	-1.0	-1.0	-1.0	-1	5	3	3	3	8.5
IMC	44	-1.0	-1.0	-1.0	3	5	3	5	0	8.5
IMC	53	-1.0	-1.0	-1.0	-1	5	3	5	0	8.0
IMC	60	-1.0	-1.0	-1.0	5	5	5	5	0	8.0
IMC	67	-1.0	-1.0	-1.0	5	5	3	5	3	9.0
IMC	76	-1.0	-1.0	-1.0	3	5	3	3	3	8.5
POUND	1	-1.0	-1.0	-1.0	5	3	0	5	3	7.0
POUND	4	-1.0	-1.0	-1.0	-1	5	3	0	0	7.5
POUND	5	-1.0	-1.0	-1.0	-1	3	3	5	0	8.0
POUND	7	-1.0	-1.0	-1.0	5	0	0	3	0	7.0
POUND	7	-1.0	-1.0	-1.0	5	0	0	5	3	8.0
POUND	8	-1.0	-1.0	-1.0	5	5	3	0	5	8.0
POUND	12	-1.0	-1.0	-1.0	5	3	0	3	0	8.0
POUND	12	-1.0	-1.0	-1.0	5	3	3	0	0	8.0
POUND	12	-1.0	-1.0	-1.0	5	3	3	0	0	8.0
EET	19	-1.0	-1.0	-1.0	5	5	3	3	3	8.6

1.2	1.3	3.1	3.2	3.3	4.1	5.1	5.2	5.3	5.4	5.6
CollPref	Num	Ban	Bfo	Jhe	Flush	Ped	Sep	Fil	Ova	Spl
EET	96	-1.0	-1.0	-1.0	5	5	5	3	3	8.7
EET	103	-1.0	-1.0	-1.0	5	5	3	5	0	8.4
NACIONAL	1	-1.0	-1.0	-1.0	-1	5	3	5	0	8.0

APPENDIX 5 PHOTOGRAPHS

Photograph 1 *Wild cocoa tree in forest, Río Santiago, August 1985, showing the typical multi-trunked habit of old trees; in this case, the tree is growing on a steep slope.*

Photograph 2 *Young seedling tree in the San Carlos genebank, showing the green flush (lacking anthocyanin pigment) characteristic of wild cocoa from Ecuador.*


LCT-EEN 162
29.4.81
Payamino, INIAP


005
006
008
1/1
2/2
005 1/3
006 5/5
Rio Caquetá
Isla Clemencia
23.5.84
010
500
006
006

Photograph 3 *Pod from wild tree (LCT-EEN 167), showing the heavily-wrinkled surface characteristic of the Ecuadorian population.*

Photograph 4 *Group of pods collected on the Río Caquetá in Colombia (EBC 5,6,8,9,10,11), showing the smoother pod surface characteristic of accessions from the western part of the area explored along the Río Caquetá.*

Photograph 5 A capuchin monkey (Cebus albifrons) eating the contents of a cocoa pod, Isla Anaconda, Río Napo, April 1985 (see Section 6.5).

Photograph 6 Pod fragments found near a wild tree, Sucúa, July 1987; the damage suggested that it had been removed from the tree and opened by a monkey. This would be a common mode of seed dispersal.