Conservation genetics of wild coffee (*Coffea* spp.) in Madagascar

by

Sarada Krishnan

B.Sc., Tamil Nadu Agricultural University, 1985

M.S., Colorado State University, 1989

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Ecology and Evolutionary Biology

2011

This thesis entitled:
Conservation Genetics of Wild Coffee (*Coffea* spp.) in Madagascar
written by Sarada Krishnan
has been approved for the Department of Ecology and Evolutionary Biology

____
Tom A. Ranker

____
Michael Grant

____
Robert Guralnick

____
Andrew P. Martin

____
Michelle L. Sauther

____
Aaron P. Davis

Date _______________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Krishnan, Sarada (PhD., Ecology and Evolutionary Biology)
Conservation Genetics of Wild Coffee (*Coffea* spp.) in Madagascar
Thesis directed by Tom A. Ranker

ABSTRACT

The genus *Coffea* L. (Rubiaceae) consists of two economically important species for the production of the beverage coffee: *Coffea arabica* and *C. canephora*. Madagascar has 59 described species of which 42 are listed as Critically Endangered, Endangered or Vulnerable by criteria of the Red List Category system of the World Conservation Union (IUCN). The National Center of Applied Research and Rural Development (FOFIFA), the main agricultural research agency in Madagascar, manages and operates the Kianjavato Coffee Research Station which has a vast *ex situ* collection of various Madagascan coffee species. In an attempt to understand the genetic diversity of Madagascan coffee species, this study was undertaken using the collections maintained at the Kianjavato Coffee Research Station's *ex situ* field genebank and extant, natural *in situ* populations. As part of my dissertation, I studied four species: *C. kianjavatensis, C. montis-sacri, C. vatovavyensis,* and *C. commersoniana.* Parentage analysis of *ex situ* propagated offspring of *C. kianjavatensis* and *C. montis*-sacri was performed to understand if outcrossing with other *Coffea* species maintained in the field genebank is compromising the genetic integrity of the collection.

I found the overall genetic diversity of wild Madagascan coffee species to be similar to or even higher than other cultivated and wild *Coffea* species. For the three species endemic to the Kianjavato region, *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis*, higher genetic diversity was observed in the *ex situ* populations than in *in situ* populations. For *C.*

commersoniana, an endemic species of the littoral forests of southeastern Madagascar and soon to be impacted by mining activities in that region, the *in situ* populations showed higher genetic diversity than the *ex situ* population. Parentage analysis of seed-propagated offspring of *C. kianjavatensis* and *C. montis-sacri* revealed that cross contamination with pollen from other *Coffea* species in the *ex situ* field genebank is occurring.

These results have significant implications for the conservation management of wild *Coffea* species. The higher genetic diversity of the *ex situ* collections which were originally made in the early 1960's could be indicative of a sampling of what was present at that time and as a result of collection from multiple origins. It could also be a result of cross contamination from pollen transfer from another species resulting in hybridization when seedlings are used in replanting lost plant collections. The genetic partitioning among the two *in situ* populations of *C. commersoniana* was high enough to warrant that these two populations be kept separate for restoration purposes. Based on these findings, recommendations for conservation management are made.

This dissertation research is the first study to characterize the genetic diversity of Madagascan *Coffea* held at the *ex situ* field genebank and comparing this with extant wild populations. The parentage study is also the first to quantify the extent of cross-species contamination of collections held in this or any other *Coffea* genebank. This study has fundamental implications for the future of *ex situ* and *in situ* conservation of *Coffea* and provides a framework for future conservation research for Madagascan and other *Coffea* species.

ACKNOWLEDGEMENTS

My doctoral program and this dissertation would not have been possible without the help and support of many individuals. Tom Ranker, my advisor, has been a great support throughout the process, without which this dissertation would not have been possible. Aaron Davis has been invaluable in providing knowledge and advice about Madagascar, Madagascan coffee, field work and local contacts, making my dream a reality. All my other committee members, Mike Grant, Andy Martin, Rob Guralnick, and Michelle Sauther provided assistance throughout my program. I am extremely grateful to each and everyone for all your support and guidance.

In Madagascar, I would not have been able to achieve all that I did without the assistance of many people. At FOFIFA, the welcome and cooperation provided by Jean Jacques (JJ) Rakotomala and Yvonne Rabenantoandro were indispensable. JJ went above and beyond to provide all the assistance I needed, including helping with all the permitting processes. Jeannot Ramelison of FOFIFA/Bioversity International CWR Program was my first contact in Madagascar and helped me connect with all the right people. My stay at RBG, Kew, Madagascar was made possible with assistance from Aaron Davis and Helene Ralimanana. Assistance in the field was provided by Franck Rakotonasolo, an amazing botanist, and Tiana Ranarivelo. At the Kianjavato Coffee Research Station, assistance was provided by M. Alfred Rabemiafara and JoePrince Rajaomanitra. At Tolagnaro, the QMM staff was very helpful and hospitable. Notable among them are Johny Rabenantoandro, Faly Randriatafika, and Rivo Rajoharison. In addition to FOFIFA, access to the germplasm at Kianjavato Coffee Research was made possible by Ueshima Coffee Company (UCC). My sincere thanks to all of you.

Funding for my research came from many sources. My sincere thanks to: Association of Zoological Horticulture - Conservation Grant; University of Colorado Museum – Walker Van

Riper Fund and Museum Awards Program grants; Beverly Sears Graduate Student Grant; and University of Colorado, Department of Ecology and Evolutionary Biology – Graduate Student Research Grant. Thanks to Tom Ranker for providing funding assistance for lab supplies and genetics work.

Many others provided assistance during various phases of my project. Notable among them are Dr. Rajinder Ranu who patiently trained me on the DNA extraction process; my fellow graduate students, Melissa Islam, who showed me all that I needed to know in the lab, and Loren Sackett, who taught me how to use GeneMapper®; Jenny Neale for sharing her knowledge of conservation genetics; Brian Vogt and Denver Botanic Gardens for allowing me to balance a demanding job with the demands of a doctoral program; Hari Krishnan for his support with the children; and Andrew Sky and Sharon Schwartzkopf for their word processing assistance that helped me complete this document . Big thanks to all of you and countless others who have cheered me on along the way.

Finally, I could not have done this without support from my family. I am greatly indebted to my late parents, M. S. Ramdas and Subalakshmi Ramdas, who constantly encouraged and supported my deep interest in nature and my adventurous spirit. Their memory and all that they taught me will always guide me throughout my life. I feel privileged and honored that my uncle Dr. M. S. Swaminathan, despite his busy schedule, took time to review my work and give me constructive comments. My children Vinay Krishnan and Vilok Krishnan have been amazing throughout my whole doctoral program. They never complained when I was busy with school work and was away in the evenings and weekends. It is to them that I dedicate this dissertation. May this serve as an inspiration to both of you to achieve your own dreams.

TABLE OF CONTENTS

CHAPTER ONE

AN ASSESSMENT OF THE GENETIC INTEGRITY OF *EX SITU* GERMPLASM COLLECTIONS OF THREE RARE SPECIES OF *COFFEA* FROM MADAGASCAR............ 1

INTRODUCTION.......... 1

BACKGROUND.......... 1

MADAGASCAN COFFEE 3

EX SITU FIELD GENEBANK 5

CONSERVATION GENETICS AND MOLECULAR MARKERS 7

PARENTAGE ANALYSIS 9

OBJECTIVES 11

MATERIALS AND METHODS 12

PLANT MATERIAL.......... 12

DNA EXTRACTION AND MOLECULAR MARKERS 16

ANALYSIS OF GENETIC DIVERSITY 17

ANALYSIS OF POPULATION GENETIC STRUCTURE.......... 19

PARENTAGE ANALYSIS.......... 19

RESULTS.......... 20

GENETIC DIVERSITY 20

Coffea kianjavatensis.......... 20

Coffea montis-sacri.......... 25

Coffea vatovavyensis 29

POPULATION GENETIC STRUCTURE.......... 32

PARENTAGE ANALYSIS.......... 32

Coffea kianjavatensis.......... 35

Coffea montis-sacri.......... 38

DISCUSSION 38

GENETIC DIVERSITY PATTERNS.......... 38

PARENTAGE ANALYSIS.......... 44

CONSERVATION AND MANAGEMENT IMPLICATIONS 46

CHAPTER TWO

GENETIC DIVERSITY PATTERNS OF *COFFEA COMMERSONIANA*, A RARE AND ENDAGERED MALAGASY ENDEMIC 50

INTRODUCTION 50

BACKGROUND 50

OBJECTIVES 54

MATERIALS AND METHODS 55

PLANT MATERIAL 55

DNA EXTRACTION AND MOLECULAR MARKERS 57

ANALYSIS OF GENETIC DIVERSITY 58

ANALYSIS OF POPULATION GENETIC STRUCTURE 58

RESULTS 59

GENETIC DIVERSITY 59

POPULATION GENETIC STRUCTURE 63

DISCUSSION 65

GENETIC DIVERSITY PATTERNS 65

CONSERVATION AND MANAGEMENT IMPLICATIONS 68

CHAPTER THREE

SUMMARY AND CONCLUSIONS 71

GENETIC DIVERSITY 73

PARENTAGE ANALYSIS 76

CONSERVATION IMPLICATIONS 77

LITERATURE CITED 81

APPENDICIES

Appendix 1: List of herbarium collections. 94

Appendix 2: *Coffea* DNA Extraction Procedure 100

Appendix 3: *Coffea* PCR Procedure 101

Appendix 4: Timetable proposed by QIT Madagascar Minerals (QMM) for conducting their mining activities in the Tolagnaro region of southeastern Madagascar 102

Appendix 5: Parental allocation to offspring – identification of each designated parent with the actual collection number for *C. kianjavatensis* and *C. montis-sacri* 103

LIST OF TABLES

Table 1.1: Geographic Locations and IUCN Red List Categories of Madagascan *Coffea* spp 4

Table 1.2: Populations sampled and their locations 13

Table 1.3: GenBank EMBL Accession number, locus code, primer sequences, repeat motif structure, product size, reference, and amplification detection of the 12 primer pairs tested ... 18

Table 1.4: Genetic variability of *C. kianjavatensis* populations at six microsatellite loci 21

Table 1.5: Allele frequencies at six polymorphic loci for *C. kianjavatensis* 23

Table 1.6: Hardy-Weinberg probability test for *C. kianjavatensis*. 24

Table 1.7: Genetic variability of *C. montis-sacri* populations at six microsatellite loci. 26

Table 1.8: Allele frequencies at six polymorphic loci for *C. montis-sacri* 27

Table 1.9: Hardy-Weinberg probability test for *C. montis-sacri*. 28

Table 1.10: Genetic variability of *C. vatovavyensis* populations at six microsatellite loci. 30

Table 1.11: Allele frequencies at six polymorphic loci for *C. vatovavyensis* 31

Table 1.12: Hardy-Weinberg probability test for *C. vatovavyensis* 33

Table 1.13: AMOVA results examining genetic partitioning between *ex situ* and wild populations for *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis.* 34

Table 1.14: Parentage allocation percentage classified on a parental pair basis for *C. kianjavatensis* and *C. montis-sacri* 36

Table 2.1: Populations sampled and their locations 56

Table 2.2: Genetic variability of *C. commersoniana* populations at six microsatellite loci. 60

Table 2.3: Allele frequencies at six polymorphic loci for *C. commersoniana* 61

Table 2.4: Hardy Weinberg probability test for *C. commersoniana* 64

Table 2.5: AMOVA results examining genetic partitioning between *ex situ* and wild populations for *C. commersoniana* 66

LIST OF FIGURES

Figure 1.1: Allocation of offspring per potential parent for *C. kianjavatensis* 37

Figure 1.2: Allocation of offspring per potential parent for *C. montis-sacri* 39

Figure 2.1: Map showing the littoral forest areas of southeastern Madagascan 53

CHAPTER ONE

AN ASSESSMENT OF THE GENETIC INTEGRITY OF *EX SITU* GERMPLASM COLLECTIONS OF THREE RARE SPECIES OF *COFFEA* FROM MADAGASCAR

INTRODUCTION

BACKGROUND

Coffee is one of the most economically important crops, and is produced in about 80 tropical countries, with an annual production of nearly seven million tons of green beans (Musoli et al. 2009). It is the second most valuable commodity exported by developing countries after oil with over 75 million people depending on it for their livelihood (Vega et al. 2003; Pendergrast 2009). *Coffea* L. (Rubiaceae, Ixoroideae, Coffeeae) consists of 125 species distributed in Africa, Madagascar, the Comoros, and the Mascarene Islands (La Réunion and Mauritius), tropical Asia, and Australia (Davis et al. 2006; Davis and Rakotonasolo 2008; Davis 2010; Davis in press). The *Coffea* genus was very recently expanded from 103 species to 125 with the inclusion of the genus *Psilanthus* Hook.f. within *Coffea*, (Davis 2010; Davis in press). Of these, two species are economically important for the production of the beverage coffee, *Coffea arabica* (Arabica coffee) and *C. canephora* (robusta coffee), and to a lesser extent, *C. liberica* (Liberian or Liberica coffee, or excelsa coffee) (Davis et al. 2006). Higher beverage quality is associated with *C. arabica*, which accounts for about 70% of world coffee production (Lashermes et al. 1999). *Coffea arabica* is a tetraploid ($2n = 4x = 44$) and self-fertile, whereas all other *Coffea* species are diploid ($2n = 2x = 22$) and mostly self-sterile (Pearl et al. 2004).

Of the 125 species of *Coffea,* 45 are endemic to Africa, 59 endemic to Madagascar, one endemic to the Comoros, four endemic to the Mascarenes (Reunion and Mauritius), 15 in

southern and southeastern Asia, and one species endemic to Australasia (Davis et al. 2006; Davis 2010; Davis et al. 2010). Differing greatly in morphology, size, and ecological adaptation, all species are perennial trees or woody shrubs (Davis et al. 2005; Anthony et al. 2010). Generally occurring in humid, evergreen forests, *Coffea* species in Africa and Madagascar inhabit diverse forest types, which may correlate with the high species diversity of this genus (Davis et al. 2006).

Based on plastid and ITS sequence analysis, six geographic groupings have been identified for *Coffea* (before the inclusion of *Psilanthus*): Upper Guinea (UG) clade, Lower Guinea/Congolian (LG/C) clade, East African-Indian Ocean (EA-IO) clade, East-Central Africa (E-CA) clade, East Africa (EA) clade, and Mascarenes (MAS) clade with the Madagascan species included within the EA-IO clade (Maurin et al. 2007). Due to limited ranges of Madagascan species, Davis et al. (2006) suggested radial and rapid speciation of these species as well as a recent origin for the genus. Colonization of volcanic islands of the Indian Ocean such as the Comoros Islands, La Réunion, and Mauritius most likely occurred through long distance dispersals (Cros et al. 1998; Maurin et al. 2007). A single dispersal event from Africa is suggested for the origin of Madagascan *Coffea* (Maurin et al. 2007).

Most of the scientific research undertaken for *Coffea* has focused on the economically important species with very little research on non-commercial species (Davis et al. 2006). Native populations are being negatively impacted by land use conversion, overexploitation, or the introduction of exotics, which is leading to the erosion of genetic diversity from wild races and species (Hein & Gatzweiler 2005). In an earlier study before the inclusion of *Psilanthus* into *Coffea,* it was found that 72 of 103 species of *Coffea* (70%) were threatened with extinction, as a result of decline in the quantity and quality of habitat (Davis et al. 2006).

MADAGASCAN COFFEE

Of the 59 species endemic to Madagascar, seven are listed as Critically Endangered, 27 as Endangered, eight as Vulnerable, seven as Near Threatened, six as Least Concern, three as Data Deficient, and one as Not Evaluated by the criteria of the Red List Category system of the World Conservation Union (IUCN) (Davis et al. 2006; Davis et al. 2010). Table 1.1 lists the described *Coffea* species of Madagascar with their geographic locations and IUCN Red List Categories.

Most *Coffea* species in Madagascar have narrow distributions with *C. perrieri* being the most widely distributed species (Davis et al. 2006). This narrow distribution of species is a major concern since the quantity and quality of habitat are in decline with high estimates of extinction threat (Davis et al. 2006). The *ex situ* field genebank maintained by the National Center of Applied Research and Rural Development (FOFIFA) in Kianjavato is a resource preserving the Madagascan *Coffea* germplasm. Other than this, no attempt for conservation of *Coffea* genetic resources has been made *in situ* (Dulloo et al. 1998). Given the high threat status of Madagascan *Coffea* species, there is an urgent need to assess the genetic diversity preserved in the Kianjavato *ex situ* collections and *in situ* populations, and based on these findings initiate new collecting programs to enhance the field collections to fill the gaps (Dulloo et al. 1998).

Table 1.1: Geographic Locations and IUCN Red List Categories of Madagascan *Coffea* spp

Species	**Location**	**Status**
C. abbayesii	South-east Madagascar	Endangered
C. alleizettii	Central Madagascar	Endangered
C. ambanjensis	North-west Madagascar	Endangered
C. ambongensis	West Madagascar	Endangered
C. andrambovatensis	East Madagascar	Critically Endangered
C. ankaranensis	North Madagascar	Endangered
C. arenesiana	East Madagascar	Near Threatened
C. augagneurii	North Madagascar	Endangered
C. bertrandii	South Madagascar	Vulnerable
C. betamponensis	East Madagascar	Endangered
C. bissetiae	West Madagascar	Data Deficient
C. boinensis	West Madagascar	Critically Endangered
C. boiviniana	North Madagascar	Near Threatened
C. bonnieri	North Madagascar	Endangered
C. buxifolia	Central Madagascar	Near Threatened
C. commersoniana	South-east Madagascar	Endangered
C. coursiana	East Madagascar	Vulnerable
C. decaryana	West Madagascar	Endangered
C. dubardii	North Madagascar	Least Concern
C. farafanganensis	South-east Madagascar	Vulnerable
C. fragilis	Madagascar	Not Evaluated
C. gallienii	North Madagascar	Critically Endangered
C. grevei	West Madagascar	Least Concern
C. heimii	North Madagascar	Vulnerable
C. homollei	East Madagascar	Least Concern
C. humbertii	South-west Madagascar	Endangered
C. jumellei	North Madagascar	Endangered
C. kianjavatensis	East Madagascar	Endangered
C. labatii	West Madagascar	Endangered
C. lancifolia	East Madagascar	Near Threatened
C. leroyi	East Madagascar	Near Threatened
C. liaudii	East Madagascar	Endangered
C. littoralis	North-east Madagascar	Critically Endangered
C. mangoroensis	East Madagascar	Vulnerable
C. manombensis	South-east Madagascar	Endangered
C. mcphersonii	North-east Madagascar	Endangered
C. millotii	East Madagascar	Least Concern
C. minutiflora	South-east Madagascar	Data Deficient
C. mogenetii	North Madagascar	Endangered
C. montis-sacri	East Madagascar	Critically Endangered
C. moratii	West Madagascar	Endangered
C. perrieri	West Madagascar	Least Concern
C. pervilleana	North-east Madagascar	Vulnerable
C. pterocarpa	West Madagascar	Critically Endangered
C. rakotonasoloi	East Madagascar	Critically Endangered
C. ratsimamangae	North Madagascar	Endangered
C. resinosa	East Madagascar	Near Threatened
C. richardii	East Madagascar	Near Threatened
C. sahafaryensis	North-east Madagascar	Endangered
C. sakarahae	South Madagascar	Vulnerable
C. sambavensis	North-east Madagascar	Endangered
C. tetragona	North-west Madagascar	Vulnerable
C. toshii	East Madagascar	Data Deficient
C. tricalysioides	North Madagascar	Least Concern
C. tsirananae	North Madagascar	Endangered
C. vatovavyensis	East Madagascar	Endangered
C. vavateninensis	East Madagascar	Endangered
C. vianneyi	South-east Madagascar	Endangered
C. vohemarensis	North-east Madagascar	Endangered

To combat this loss of genetic diversity, Bioversity International (formerly International Plant Genetic Resources Institute - IPGRI) launched a project in 2005 for the conservation of genetic resources of crop wild relatives (CWRs). The aims of this project were to improve those aspects of knowledge of agricultural biodiversity that are important for the livelihoods of poor people, making conservation and management of agricultural biodiversity more effective at the gene pool and ecosystem levels. In Madagascar, wild species of *Coffea* are among Bioversity International's target genera. The present study was undertaken in collaboration with FOFIFA, the body that administers Bioversity International's CWR conservation project on *Coffea* in Madagascar.

EX SITU FIELD GENEBANK

For the long-term conservation of genetic resources of crop plants, collections are maintained in genebanks around the world for ease of access by plant breeders, researchers, and other users (Van Hintum et al. 2000). Even though considerable progress has been made in assembling and conserving these genetic resources over the past three decades, many of the germplasm collections are now facing major problems of size and organization (Van Hintum et al. 2000). Collecting missions over the past few decades have helped establish *Coffea* genebanks in various countries with at least 11,700 accessions representing 70 *Coffea* species represented in these various filed genebanks (Anthony et al. 2007a). Many of these *Coffea* genebanks are experiencing genetic erosion due to loss of trees resulting from aging, cultivation in non-conducive climates, and inappropriate cultivation methods (Anthony et al. 2007b; Vega et al. 2008). In addition, hybridization in *ex situ* collections may compromise of the genetic makeup, integrity and value of the collection (Maunder et al. 2003).

FOFIFA is the main agricultural research agency in Madagascar. The Kianjavato Coffee Research Station (KCRS) managed by FOFIFA has a vast *ex situ* collection of various Madagascan coffee species with their total collections encompassing 2,649 specimens in 132 accessions (J. J. Rakotomalala, pers. comm.). The KRCS was established in 1954 with the main aim of improvement of *Coffea canephora* (robusta coffee) through selection and making these improved genotypes available to coffee growers in southeastern Madagascar, and serving as a resource imparting advice on improved cultivation practices. In 1960, the United Nations' Food and Agriculture Organization (FAO) initiated wild collecting of wild Madagascan species for *ex situ* germplasm preservation, which was continued until 1974 by French teams such as ORSTOM (Office de la Recherche Scientifique et Technique d'Outre-Mer) and IRCC (Institut de Recherches du Cafe, du Cacao). From 1974 to 1982, FOFIFA was funded by the Malagasy government for coffee research and germplasm preservation. From 1982 to 2002, government funding stopped and the collections were maintained as best as was possible with limited funding, though during this time many individual plants per accession were lost. Since 2002, the Ueshima Coffee Corporation (UCC) of Japan has funded approximately 90% of the maintenance of the field genebank, allowing the preservation of this valuable germplasm resource. During the past few years, many missing plants have been replaced. Replacement of individuals has been predominantly through seeds collected from each accession (J. J. Rakotomalala, pers. comm.). The replenishment of the lost germplasm has been conducted without knowledge of the genetic diversity of the collection with seed selection for propagation from parents at random without knowledge of out-crossing with other species maintained in the collections. This practice can be problematic for outcrossing species, which would lead to loss of genetic integrity. Knowledge of

the existing genetic diversity is critical in management of the existing collection as well as providing direction for future improvement.

CONSERVATION GENETICS AND MOLECULAR MARKERS

Wild populations are exposed to a wide range of threats from deterministic and stochastic factors that drive them to extinction (Frankham et al. 2002). Deterministic factors either directly or indirectly associated with human activities include loss of habitat, over-exploitation, pollution, introduced species, etc. and stochastic factors are influenced by demography, environment, catastrophe, and genetics (Frankham et al. 2002). The two primary threats due to genetic stochasticity faced by declining populations are slow erosion of genetic variability by drift and the short-term lowering of fitness caused by inbreeding depression (Amos and Balmford 2001). Genetic variation within individuals (heterozygosity), genetic differences among individuals within a population, and genetic differences among populations are important to fitness and adaptive change, the loss of which are of serious concern to conservationists (Meffe and Carroll 1994). Loss of genetic diversity could lead to reduced evolutionary flexibility, decline in fitness, loss of local adaptations, and increased probability of population or species extinction (Meffe and Carroll 1994).

In addition to being crucial for the adaptation of species to a dynamic environment (Wise et al. 2002), genetic diversity also has economic value related to the potential benefits offered through breeding of new varieties and crop improvement of economically important crops (Hein & Gatzweiler 2005). In 1970, a catastrophic outbreak of coffee rust in Brazil caused great losses leading to higher coffee world market prices (Scarascia-Mugnozza and Perrino 2002). In India in the early part of the 20th century, coffee rust resulted in serious losses of Arabica coffee, which

resulted in extensive plantings of robusta coffee (M. S. Swaminathan, pers. comm.). In instances like these, conserving genetic diversity of wild races and wild species may play a critical role in developing improved varieties of coffee due to the potential of these genetic resources for increasing resistance to plant diseases.

Conservation and sustainable use of plant genetic resources have become focal points of many national and international agendas with tremendous success in developing methods for the conservation of genetic resources *ex situ* while *in situ* conservation is still inadequate (Gole et al. 2002). This inadequacy is mainly due to socio-economic factors, lack of policy and political will, and a lack of scientific understanding of the natural environment and biological characteristics of species (Gole et al. 2002). The greatest threats to biodiversity arise from habitat destruction, alien species invasion, and genetic homogeneity (Swaminathan 2000). In order to achieve results in conservation and enhancement of natural resources, an integrated gene management practice encompassing biopartnerships, participatory forest management, community gene management, biosphere management, and genetic resources enhancement and sustainable use should be implemented (Swaminathan 2000). The Convention on Biological Diversity (CBD) at the tenth meeting of the Conference of the Parties held October 18 to 24, 2010 in Nagoya, Aichi Prefecture, Japan, adopted a revised and updated Strategic Plan for Biodiversity (http://www.cbd.int/sp/). The Aichi Biodiversity Targets outlines five strategic goals to be implemented through 20 major targets such as addressing the underlying cause of biodiversity loss, reducing direct pressures on biodiversity, safeguarding biodiversity at the ecosystem level, enhancing the benefits provided by biodiversity, and providing for capacity building. To conserve the whole genetic diversity of a taxon, knowing the genetic structure of populations is

essential and hence this should become one of the principal strategies in the conservation efforts of species to ensure success (Gole et al. 2002; Shapcott et al. 2007).

In my genetic diversity studies, I used microsatellite markers, which are a powerful method for identifying highly polymorphic Mendelian markers (Avise 2004). Microsatellites are codominant, easily reproducible, and are used widely for genetic mapping, genetic diversity assessment, and population genetics (Poncet et al. 2004). Microsatellite markers have been developed for *C. arabica* and *C. canephora*. Combes et al. (2000) surveyed 13 *Coffea* taxa to examine cross amplification of eleven primer pairs designed for *C. arabica* in detecting microsatellite loci. Eight of these primer pairs were amplified across all four Madagascan species tested and can be used as markers for genetic variation studies. Poncet et al. (2004 and 2007) also report good transferability of microsatellite primers within the *Coffea* genus. In addition to these studies, the transferability of microsatellite markers across different *Coffea* species and their high levels of polymorphisms have been demonstrated by various other authors (Baruah et al. 2003; Coulibaly et al. 2003; Hendre et al. 2008). Microsatellites have become valuable in analysis of genetic diversity and structure in various *Coffea* species and identification of cultivars (Cubry et al. 2008; Maluf et al. 2005; Moncada and McCouch 2004; Prakash et al. 2005; Silvestrini et al. 2007).

PARENTAGE ANALYSIS

Gene flow shapes the diversity of species and is an important feature of population genetics (Gerber et al. 2000). Genetic markers are used in the study of actual gene flow by reconstructing relationships between parental and offspring generations (Gerber et al. 2000). In

ex situ collections, maintaining the genetic integrity of the germplasm require knowledge of gene flow to ensure that genetic erosion due to contamination by foreign pollen is decreased (Suso *et al.* 2006). Based on genetic information, parentage analysis allows determination of parental genotypes of each set of offspring genotypes (Deuchesne et al. 2008). The study and use of molecular markers for parentage analysis has exploded over the past decade with the introduction of microsatellite markers and more refined statistical techniques (Jones et al. 2010). In a review of plant parentage, pollination, and dispersal using microsatellites, Ashley (2010) found 41 papers measuring pollen dispersal and paternity assignment in a total of 36 different species. Microsatellites, which are highly variable, have made direct estimation of gene flow more feasible (Ouborg et al. 1999) and give rise to highly accurate parentage assignments due to their characteristic of high levels of codominant polymorphism (Gerber et al. 2000). By using the exclusion method and assuming Mendelian inheritance, any putative parent that fails to share an allele with the offspring of interest can be eliminated as a true parent (Jones et al. 2010). By determining the genotypes of all reproductive adults in a population and comparing those with seedling genotypes using maximum likelihood methods or paternity exclusion analysis, direct estimation of gene flow can be achieved (Ouborg et al. 1999).

Cross-species hybridization has been reported for the genus *Coffea*. *Coffea arabica* has been demonstrated to be an amphidiploid that arose from natural hybridization between *C. canephora* and *C. eugenioides*, or ecotypes related to these diploid species (Lashermes et al. 1999). Using ITS and plastid sequence data, Maurin et al. (2007) report the hybrid origin of the *C. heterocalyx* accession from IRD-Montpellier (JC66) resulting from either introgression in the wild or chance crossing in cultivation between *C. eugenioides* and *C. liberica.* The cultivated

Timor Hybrid is a spontaneous interspecific cross between *C. arabica* and *C. canephora*, with 50% of the hybrid genome represented by the *C. canephora* genome (Lashermes et al. 2000).

At FOFIFA's KCRS, the replenishment of the lost germplasm has been conducted without knowledge of the genetic diversity of the collection and selection of seed parents at random without knowledge of outcrossing with other species maintained in the collections, which would lead to genetic erosion. By sampling all possible parents of a given *Coffea* species in the *ex situ* population, any offspring with an unmatched parent can be considered as being contaminated by foreign pollen, i.e., pollen from another *Coffea* species maintained in the collection. Using microsatellite allelic diversity information, parentage analysis of offspring of *C. kianjavatensis* and *C. montis-sacri* in the *ex situ* collections is presented.

OBJECTIVES

For assessment of genetic diversity of wild populations, evaluating the entire *Coffea* genus would not be feasible within the scope of my study and hence I am concentrating on locally endemic endangered species of the Kianjavato region. The three endemic species of this region include: *C. kianjavatensis* (Endangered), *C. montis-sacri* (Critically Endangered), and *C. vatovavyensis* (Endangered).

Genetic diversity studies of the existing *ex situ* germplasm at the KCRS or of *in situ* populations of wild coffee have not been performed on Madagascan *Coffea* species. In collaboration with the FOFIFA managers of the KCRS, I explored levels and patterns of genetic diversity of the *ex situ* and *in situ* populations and partitioning of genetic diversity within- and among-populations for all three study species. In addition, I also conducted parentage analysis of seedlings of *C. kianjavatensis* and *C. montis-sacri* in the *ex situ* collections to assess the degree to which the offspring from individuals from the *ex situ* genebank accessions might be the result

of interspecific crosses and, thus, potentially decreasing the value of the collections. From the results of this study, recommendations are made for the long-term conservation of the wild *Coffea* gene pool. All study species are represented in the *ex situ* collections held at the KCRS.

Specific questions addressed were: 1) What levels of genetic diversity are present within and across each *ex situ* and wild (= *in situ*) population? 2) How is genetic diversity structured among these populations? 3) Are progeny always purebred or are some the result of interspecific mating?

This is the first study examining the genetic diversity of Madagascan *Coffea* species in *ex situ* field genebanks and natural populations in order to make recommendations for the conservation of this valuable resource of agrobiodiversity.

MATERIALS AND METHODS

PLANT MATERIAL

I conducted fieldwork in Madagascar to collect plant specimens during December 2007 and November 2008. During the first visit, I visited the Kianjavato region to collect leaf and herbarium samples of *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis.* The entire *ex situ* germplasm at the KRCS was sampled for these three species. Aaron Davis (Royal Botanic Gardens, Kew) provided geographic location information of wild populations of *Coffea* species from herbarium collections data (A. Davis, unpubl. data). Using this information, I collected samples from wild populations of *C. kianjavatensis* and *C. montis-sacri* in 2007. I could not locate wild populations of *C. vatovavyensis* at that time, although this species was collected during my 2008 visit. Table 1.2 lists the *ex situ* and *in situ* populations sampled. The location

Table 1.2: Populations sampled and their locations.

Species and Population	Location	Number of Individuals	Latitude/Longitude
C. kianjavatensis A. 213	FOFIFA Kianjavato Coffee Research Station	84	21°22'25"S 47°51'57"E
C. kianjavatensis A. 602	FOFIFA Kianjavato Coffee Research Station	35	21°22'25"S 47°51'57"E
C. kianjavatensis In situ	Mt. Vatovavy	63	21°24'36"S 47°56'32"E
C. montis-sacri A. 321	FOFIFA Kianjavato Coffee Research Station	16	21°22'29"S 47°51'57"E
C. montis-sacri In situ	Mt. Vatovavy	6	21°24'37"S 47°56'35"E
C. vatovavyensis A. 308	FOFIFA Kianjavato Coffee Research Station	24	21°22'28"S 47°52'00"E
C. vatovavyensis A. 830	FOFIFA Kianjavato Coffee Research Station	25	21°22'30"S 47°52'04"E
C. vatovavyensis A. 1009	FOFIFA Kianjavato Coffee Research Station	6	21°22'31"S 47°52'02"E
C. vatovavyensis In situ	Sangasanga Forest	36	21°22'27"S 47°52'07"E

coordinates (collected using WGS 84 map datum using a Magellan Meridian Color Handheld GPS) given are representative of the location of the first few samples collected for each population. Several leaves of each individual plant were collected and placed in a plastic bag with silica gel. Voucher specimens of selected samples were collected in duplicates of four wherever possible, one each for herbaria at Royal Botanic Gardens, Kew (K), University of Colorado Museum (COLO), Parc Botanique et Zoologique de Tsimbazaza (TAN), and FOFIFA: National Center of Applied Research and Rural Development (TEF). The list of herbarium specimens is listed in Appendix 1.

The germplasm collections at FOFIFA's KCRS are identified with an 'A' number, assigned to a group of plants of each species collected from a particular geographic location. About 70% of the herbarium vouchers of Kianjavato 'A' numbered accessions are housed at the Muséum National d'Histoire Naturelle, Paris (P). Planted in rows, each individual plant in an accession group is assigned a line number and a plant number. From discussions with local FOFIFA officials, it was my understanding that when individual plants within an accession died, they were predominantly replaced with seedlings of seeds collected from plants within that same accession group and rarely with plants propagated by cuttings.

Collections of *C. kianjavatensis* consisted of 84 individuals belonging to the *ex situ* accession A. 213; 35 individuals belonging to the *ex situ* accession A. 602; and 63 individuals collected from wild populations at Mt. Vatovavy in Mananjary District in Fianarantsoa Province. The natural habitat of wild populations in Mt. Vatovavy is humid evergreen forest. Associated plant species in this habitat include *Chassalia* (Rubiaceae), *Garcinia verucosa* (Clusiaceae), *Oncostemum* (Myrsinaceae), *Dracaena* (Ruscaceae), *Pandanus* (Pandanaceae), *Dypsis* (Arecaceae), *Diporidium* (Ochnaceae), and *Polysphaeria* (Rubiaceae). The specimens were

collected from altitudes ranging from 418 to 455 meters and the slopes were steep and rocky. The KCRS is located at altitudes ranging from 56 to151 meters. Accession A. 213 in the KCRS was established originally from seeds collected from two wild populations in Mananjary District (Charrier 1978) in 1962/1963. The collections made from the wild populations for this study came from one of these locations. A. 602 is documented to have been collected in 1967/1968 from Isaka-Ivandro forest in Tolagnaro District in Southeastern Madagascar (Charrier 1978). During my field visit in 2008, I attempted to visit the forest from where this collection was thought to have originated, but I was not able to locate any *C. kianjavatensis* populations at that location.

Collections of *C. montis-sacri* consisted of accession A. 321 with 16 individuals and a single wild population of six individuals collected from Mt. Vatovavy in the same humid evergreen forest habitat as *C. kianjavatensis*. The wild samples were collected from altitudes ranging from 450 – 477 meters and the associated plant species were bamboo and other Gramineae with *Clidemia hirta* (Melastomataceae), *Dianella* (Liliaceae), *Pandanus* (Pandanaceae), and *Dypsis* (Arecaceae). The plants were located on rocky slopes. *Ex situ* accession of *C. montis-sacri* (A. 321) was originally collected in 1964 from Vatovavy/Vatolahy in Mananjary District.

Collections of *C. vatovavyensis ex situ* germplasm consisted of three accessions, A. 308, A. 830, and A. 1009 with 24, 25, and 6 individuals respectively. Thirty-six samples from a wild population were collected from the Sangasanga Forest, a humid evergreen forest, adjacent to the KCRS in Mananjary District, in Fianarantsoa Province. Plants were located in rocky habitat on steep slopes with altitudes ranging from 168 – 182 meters. The associated plant species in this location include *Dracaena* sp. (Ruscaceae), *Canarium madagascariensis* (Burseraceae),

Ravenala madagascariensis (Strelitziaceae), and *Maranta* sp. (Marantaceae). The plants in the *ex situ* A. 308 accession group were originally collected from Vondrozo district in Fianarantsoa Province in 1964, A. 830 from Fananehana/Mananara areas in Toamansia Province in 1969, and the geographic location of the origin of A. 1009 is unknown.

For parentage analysis, during the visit of December 2007, I requested that FOFIFA officials germinate seeds of *C. kianjavatensis* and *C. montis-sacri* so that I could collect leaf samples from the seedlings during my 2008 visit. Seeds of *C. kianjavatensis* were collected on 1 February 2008 and germinated in the nursery in a seedling tray, from which individual seedlings were transferred to individual pots on 28 August 2008. From this seedling population, I sampled 50 seedlings for parentage analysis. *Coffea montis-sacri* seeds were collected on 17 November 2007 and after germination the seedlings were grafted onto seedling rootstocks of *C. perrieri* (A. 12) on 20 February 2008. The grafted seedlings were transferred to individual pots on 25 August 2008. From this grafted seedling population, I sampled 34 seedlings for parentage analysis. Several leaves of each individual seedling were placed into a plastic bag with silica gel to preserve the DNA required for analysis.

DNA EXTRACTION AND MOLECULAR MARKERS

Genomic DNA was extracted from 10 mg of silica-dried leaf material using GenCatch™ Plant Genomic DNA Purification kit by Epoch Biolabs. Slight modifications were made to the extraction protocols. A detailed account of the extraction procedure is described in Appendix 2. Extracted DNA was sent to Nevada Genomics, Reno, Nevada for quantification, optimization, and fragment analysis using SSR markers. Twelve microsatellite markers were originally selected based on Combes et al. (2000) and Poncet et al. (2004) and tested, of which six markers

(M253, M254, M256, M780, M784, and M883) had low signal and did not amplify well. These six markers were discarded and the remaining six (M255, M257, M258, M259, M260, and M746) were used in this study (Table 1.3).

The DNA was quantified and normalized to 5.0 ng/µl. PCR amplifications were carried out using an MJ thermocycler. Each 10.0 µl PCR amplification reaction contained 4.0 µl of 5.0 ng/µl genomic DNA, 1.0 µl Primer Panel mix, and 5.0 µl Qiagen Multiplex PCR Mix. The amplifications were performed using a "touchdown" PCR profile as described in Coulibaly et al. (2003), which is listed in Appendix 3. The only modification was the time for the initial denaturation at 94°C was increased to 15 minutes due to the use of Qiagen Multiplex PCR Mix, which is a hot-start *Taq* DNA Polymerase. The samples were run on an Applied Biosystems Prism 3730 DNA Analyzer. The filter set used was G5, which detects the fluorescent dyes 6-FAM, VIC, NED, and PET. The samples were run with the 500 MW size standards labeled with LIZ. The six microsatellite loci were amplified in a single 6-primer panel.

The fragment analysis results were scored using GeneMapper® Software Version 4.0 by Applied Biosystems.

ANALYSIS OF GENETIC DIVERSITY

Data analysis to assess genetic diversity was performed using GENEPOP v.4.0 software (Rousset 2008). Parameters used to estimate genetic diversity included: number of alleles per locus (**A**), the mean observed ($\mathbf{H_o}$) and the mean expected ($\mathbf{H_e}$) heterozygosities based on Hardy-Weinberg assumptions, the allelic fixation index ($\mathbf{F_{is}}$), and the number of observed genotypes per

Table 1.3: GenBank EMBL Accession number, locus code, primer sequences, repeat motif structure, product size, reference, and amplification detection of the 12 primer pairs tested (Y-yes; N-no).

EMBL Accession #	Locus Code	Primer Sequence	Repeat Motif	Product Size (bp)	Reference	Amplification (Y/N)
AJ250253	M253	F: CTTGGTTCTTTCTTTCGGGT R: TTTCCCTCCCAATGTCTGTA	$(GA)_5(GT)_8TT(GT)_4TT(GT)_7$ $(GA)_{11}(TC)_2(CT)_3GT$	240-270	Combes *et al.* 2000	N
AJ250254	M254	F: GGCTCGAGATATCTGTTTAG R: TTTAATGGGCATAGGGTCC	$(CA)_{15}(CG)_4CA$	132-166	Combes *et al.* 2000	N
AJ250255	M255	F: CCCTCCCTGCCAGAAGAAGC R: AACCACCGTCCTTTTCCTCG	$(GT)_5CT(GT)_2/(GT)_{12}$	160-170	Combes *et al.* 2000	Y
AJ250256	M256	F: AGGAGGGAGGTGTGGGTGAAG R: AGGGGAGTGGATAAGAAGG	$(GT)_{11}$	118-134	Combes *et al.* 2000	N
AJ250257	M257	F: GACCATTACATTTCACACAC R: GCATTTTGTTGCACACTGTA	$(CTCACA)_4/(CA)_9$	103-122	Combes *et al.* 2000	Y
AJ250258	M258	F: AACTCTCCATTCCCGCATTC R: CTGGGTTTTCTGTGTTCTCG	$(CA)_3/(CA)_3/(CA)_{18}$	89-135	Combes *et al.* 2000	Y
AJ250259	M259	F: ATCCGTCATAATCCAGCGTC R: AGGCCAGGAAGCATGAAAGG	$(GT)_3/(GT)_7$	72-103	Combes *et al.* 2000	Y
AJ250260	M260	F: TGATGGACAGGAGTTGATGG R: TGCCAATCTACCTACCCCTT	$(CT)_9(CA)_8/(CT)_4/(CA)_5$	100-132	Combes *et al.* 2000	Y
AJ308746	M746	F: GGCCTTCATCTCAAAAACCT R: TCTTCCAAACACACGGAGACT	$(CT)_{12}/(CA)_{11}$	378	Rovelli *et al.* 2000	Y
AJ308780	M780	F: ATTCTCTCCCCCTCTCTG R: GTTAGTATGTGATTTGGTGTGG	$(CA)_6$	95	Rovelli *et al.* 2000	N
AJ308784	M784	F: TTGCTTGCTTGTTCTGTTAT R: TGACACGAGAGTTAGAAATGA	$(GT)_7/(GC)_7/(GT)_7$	126	Rovelli *et al.* 2000	N
AJ308883	M883	F: CGTCTCGTTTCACGCTCTCT R: GATCTGCATGTACTGGTGCTTC	$(GT)_{15}$	237	Rovelli *et al.* 2000	N

population per locus. GENEPOP v.4.0 was also used to calculate the allele frequencies at each locus with the private alleles for each population identified. Conformance to Hardy-Weinberg equilibrium by population was performed by assessing the significance of the $\mathbf{F_{is}}$ values by means of Fisher's exact tests implemented in GENEPOP v.4.0 by the Markov Chain (MC) method of 10,000 dememorization steps, followed by 20 batches of 5,000 iterations per batch. Where the number of alleles is less than five, the default in the batch mode is complete enumeration rather than MC method, where no standard error is computed. The $\mathbf{F_{is}}$ reported is based on Weir and Cockerham's (1984) estimate.

ANALYSIS OF POPULATION GENETIC STRUCTURE

Hierarchical genetic structure was examined through an analysis of molecular variance (AMOVA) (Excoffier et al. 1992) as implemented in Arlequin v.3.5.1 (Schneider et al. 2000). AMOVA was applied to estimate the components of variance among and within populations based on $\mathbf{\Phi_{st}}$, a statistic analogous to $\mathbf{F_{st}}$ for each of the species to test the significance against the null hypothesis of no structure. For each locus, 20% missing data was allowed.

PARENTAGE ANALYSIS

Parentage analysis was performed using the computer program Parental Allocation of Singles in Open Systems (PASOS) 1.0 (Duchesne et al. 2005). PASOS is a parental allocation program that detects missing parents when a proportion of them have not been collected by identifying collected parents based on individual multi-locus genotypes (Duchesne et al. 2005). The approach used by the program is a combination of parental pair likelihoods with a

subsequent filtering procedure with the allocation of an offspring starting with the search for the most likely pair among all the potential pairs of collected parents (Duchesne et al. 2005). Of the multitude of parentage analysis programs available, PASOS seems to be the best program for my application since I was testing to see if the seedlings propagated (in the FOFIFA Kianjavato Coffee Research Station Nursery) for the use of repopulating dead plants within an accession are contaminated by pollen from another *Coffea* species maintained in the field genebank. The identification of cross-species contamination within an individual plant or an accession would, of course, mean that the genetic integrity of the collection had been compromised. Since the entire collections of *C. kianjavatensis* and *C. montis-sacri* adult plants were sampled, detection of any missing plants can be presumed to be that of a contributing parent belonging to another species, thereby confirming genetic contamination. The possibility of contamination from wild *Coffea* species occurring nearby (within 500 km) in wild populations (e.g. *C. vatovavyensis*) cannot be overruled.

RESULTS

GENETIC DIVERSITY

Coffea kianjavatensis

All six loci were polymorphic across all populations. The genetic diversity parameters of $\mathbf{A}$, $\mathbf{H_o}$, $\mathbf{H_e}$, and number of observed genotypes were higher for the *ex situ* populations A. 213 and A. 602 than in the wild population (Table 1.4). Accession A. 213 showed higher diversity in mean number of alleles across loci at 4.50 (ranging from 3 – 6 for individual locus) and a mean number of genotypes across loci at 8.00 (ranging from 4 genotypes at locus M260 to 12 at locus

Table 1.4: Genetic variability of *C. kianjavatensis* populations at six microsatellite loci. N = sample size per locus, A = allele numbers per locus, H_o = the observed heterozygosity, H_e = the expected heterozygosity, and F_{is} = the allelic fixation index for polymorphic loci

Population	Microsatellite Locus	Genetic Diversity Parameters					
		N	A	H_o	H_e	F_{is}	No. obs. Genotypes
A. 213	M255	84	4	0.4762	0.5943	0.1987	9
	M257	84	5	0.4286	0.706	0.3930	9
	M258	84	4	0.2381	0.2389	0.0033	5
	M259	84	6	0.6905	0.7198	0.0407	12
	M260	83	3	0.3133	0.421	0.2558	4
	M746	83	5	0.3373	0.4722	0.2855	9
	Mean	83.67	4.50	0.4143	0.5257	0.2118	8.00
A. 602	M255	35	4	0.6857	0.5181	-0.3236	4
	M257	35	4	0.1143	0.7176	0.8407	4
	M258	35	4	0.6286	0.5500	-0.1429	5
	M259	35	5	0.9143	0.7155	-0.2777	8
	M260	35	3	0.6000	0.5613	-0.0689	4
	M746	35	4	0.6857	0.4727	-0.4507	4
	Mean	35.00	4.00	0.6048	0.5892	-0.0264	4.83
In situ	M255	63	3	0.6508	0.5189	-0.2541	5
	M257	63	3	0.3175	0.5781	0.4508	6
	M258	63	2	0.0952	0.0914	-0.0420	2
	M259	63	3	0.5714	0.5682	-0.0056	6
	M260	63	2	0.3175	0.2903	-0.0935	3
	M746	63	5	0.2698	0.286	0.0564	6
	Mean	63.00	3.00	0.3704	0.3888	0.0475	4.67

M259), which were significantly higher than *in situ* populations (two-tailed *t*-test, $t = 2.57$, $P<0.05$). Estimates of $\mathbf{H_o}$ and $\mathbf{H_e}$ were highest for accession A602 with a mean $\mathbf{H_o}$ of 0.60 and mean $\mathbf{H_e}$ of 0.59 across loci, resulting in a low $\mathbf{F_{is}}$ of -0.03 suggesting heterozygosity excess, with the $\mathbf{H_e}$ being significantly higher than the *in-situ* population (Table 1.4). The value of $\mathbf{F_{is}}$ was highest for population A. 213 with a mean $\mathbf{F_{is}}$ of 0.21, suggesting heterozygote deficiency, which significantly deviated from Hardy-Weinberg equilibrium (Table 1.6). The *in situ* population had a mean number of alleles of 3.0 and the mean number of genotypes observed was 4.7. The mean $\mathbf{H_o}$ (0.37) was slightly lower than mean $\mathbf{H_e}$ (0.39) resulting in a mean inbreeding coefficient ($\mathbf{F_{is}}$) of 0.05 which significantly deviated from Hardy-Weinberg equilibrium at the 0.05 level of significance (Table 1.4 and 1.6).

For the six polymorphic loci surveyed, there were a total of 29 alleles, ranging from 2 to 6 per locus (Table 1.5). Of these, five alleles were unique to individual populations with three private alleles observed in *ex situ* population A. 213 at locus M257, M259, and M260, one private allele in *ex situ* population A. 602 at locus M260, and one private allele in the *in situ* population at locus M746 (Table 1.5). In addition, both *ex situ* populations had seven alleles that were unique to them that were not present in the wild population. These were present in all loci except locus M260 (Table 1.5).

The Hardy-Weinberg probability tests by populations showed that the $\mathbf{H_o}$ values of all three *C. kianjavatensis* populations significantly deviated from Hardy-Weinberg equilibrium with the *ex situ* A. 213 and *in situ* populations showing a significant deficiency of heterozygotes and the *ex situ* population A. 602 showing a significant excess of heterozygotes (Table 1.6).

Table 1.5: Allele frequencies at six polymorphic loci for *C. kianjavatensis* (*Private alleles

Locus	Allele	Allele Frequency			
		A. 213	A. 602	*In situ*	Mean
M255	177	0.1190	0.6571	-	0.2587
	181	0.1607	0.0143	0.1190	0.0980
	183	0.5952	0.1429	0.6429	0.4603
	187	0.1260	0.1857	0.2381	0.1833
M257	123	0.1726	0.0571	0.1429	0.1242
	125	0.4048	0.3429	0.5714	0.4397
	129	0.1012	0.3143	-	0.1385
	135	0.0119*	-	-	0.0040
	139	0.3095	0.2857	0.2857	0.2936
M258	95	0.0536	0.0143	0.0476	0.0385
	101	0.8690	0.6143	0.9524	0.8119
	103	0.0655	0.2571	-	0.1075
	105	0.0119	0.1143	-	0.0421
M259	107	0.0238*	-	-	0.0079
	113	0.4226	0.2143	0.5952	0.4107
	115	0.0595	0.4143	-	0.1579
	117	0.0595	0.2571	-	0.1055
	119	0.2619	0.0857	0.2063	0.1846
	121	0.1726	0.0286	0.1984	0.1332
M260	107	0.0120*	-	-	0.0040
	117	0.7108	0.3000	0.8254	0.6121
	119	0.2771	0.5857	0.1746	0.3458
	125	-	0.1143*	-	0.0381
M746	355	0.0241	-	0.0238	0.0160
	357	0.1245	0.6571	-	0.2605
	359	-	-	0.0159*	0.0053
	361	0.0542	0.0143	0.0397	0.0361
	363	0.7108	0.3143	0.8413	0.6221
	365	0.0904	0.0143	0.0794	0.0614

Table 1.6: Hardy-Weinberg probability test for *C. kianjavatensis*. F_{is} = allelic fixation index for polymorphic loci, P = probability value, SE = standard error, Chi^2 = Chi square value using Fisher's method, and df = degrees of freedom

Population		Locus						Chi^2	df	P
		M255	M257	M258	M259	M260	M746			
A. 213	F_{is}	0.1987	0.393	0.0033	0.0407	0.2558	0.2855	Infinity	12	HS*
	P	0.0000	0.0000	0.6331	0.0000	0.0360	0.0000			
	S.E.	-	0.0000	-	0.0000	-	0.0000			
A. 602	F_{is}	-0.3236	0.8407	-0.1429	-0.2777	-0.0689	-0.4507	Infinity	12	HS*
	P	0.0852	0.0000	0.0035	0.0835	0.0002	0.0127			
	S.E.	-	-	-	0.0051	-	-			
In situ	F_{is}	-0.2541	0.4508	-0.0420	-0.0056	-0.0935	0.0564	32.503	12	0.0012*
	P	0.0749	0.0000	1.0000	0.6794	0.6722	0.0615			
	S.E.	-	-	-	-	-	0.0069			

* Significantly deviates from Hardy Weinberg equilibrium at 0.05 level of significance.
HS - Highly significant (P<0.001)

Coffea montis-sacri

All six loci were polymorphic across all populations. The genetic diversity parameters of **A**, $\mathbf{H_o}$, $\mathbf{H_e}$, and the number of observed genotypes were higher for the *ex situ* population A. 321 than in the wild population (Table 1.7). Population A. 321 showed a mean of **A** across loci of 5.2 (ranging from 4 to 8 for individual locus) and a mean number of genotypes across loci of 6.7 (ranging from 4 genotypes at locus M259 to 10 at locus M260), which were significantly higher at the 0.05 level of significance when a two-tailed *t*-test was performed. Estimates of $\mathbf{H_o}$ and $\mathbf{H_e}$ were also higher for the *ex situ* accession A. 321 with a mean $\mathbf{H_o}$ of 0.59 and mean $\mathbf{H_e}$ of 0.70 compared to the *in situ* population which had a mean $\mathbf{H_o}$ of 0.44 and a mean $\mathbf{H_e}$ of 0.48, though only the $\mathbf{H_e}$ was significant at the 0.05 level of significance. The fixation index was higher for population A. 321 with a mean $\mathbf{F_{is}}$ of 0.16 compared to *in situ* population with a mean $\mathbf{F_{is}}$ of 0.08. The *in situ* population had a mean **A** of 2.7 and the mean number of genotypes observed was 3.0 (Table 1.7).

For the six polymorphic loci surveyed, there were a total of 32 alleles, ranging from 2 to 8 per locus (Table 1.8). Of these, 16 alleles were unique to the *ex situ* population A. 321 across all loci and one private allele was unique to the *in situ* population at locus M746 (Table 1.8).

The Hardy-Weinberg probability tests by populations showed that the $\mathbf{H_o}$ values of *ex situ C. montis-sacri* population A. 321 deviated from Hardy-Weinberg equilibrium with a significant deficiency in heterozygotes at the 0.05 level of significance while those from the *in situ* population did not deviate significantly from Hardy-Weinberg equilibrium (Table 1.9).

Table 1.7: Genetic variability of *C. montis-sacri* populations at six microsatellite loci. N = sample size per locus, A = allele numbers per locus, H_o = the observed heterozygosity, H_e = the expected heterozygosity, and F_{is} = the allelic fixation index for polymorphic loci

Population	Microsatellite Locus	Genetic Diversity Parameters					
		N	A	H_o	H_e	F_{is}	No. obs. Genotypes
A. 321	M255	16	4	0.7500	0.6375	-0.1765	6
	M257	16	5	0.5000	0.7167	0.3023	6
	M258	16	5	0.3125	0.6979	0.5522	7
	M259	16	4	0.8125	0.5917	-0.3732	4
	M260	16	8	0.6250	0.8146	0.2327	10
	M746	16	5	0.5625	0.7625	0.2623	7
	Mean	16.00	5.17	0.5938	0.7035	0.1560	6.67
In situ	M255	6	3	0.6667	0.5833	-0.1429	3
	M257	6	2	0.5000	0.4000	-0.2500	2
	M258	6	4	0.5000	0.6500	0.2308	5
	M259	6	2	0.3333	0.3000	-0.1111	2
	M260	6	2	0.3333	0.3000	-0.1111	2
	M746	6	3	0.3333	0.6500	0.4872	4
	Mean	6.00	2.67	0.4444	0.4806	0.0751	3.00

Table 1.8: Allele frequencies at six polymorphic loci for *C. montis-sacri* (*Private alleles

Locus	Allele	Allele Frequency		
		A. 321	*In situ*	Mean
M255	177	0.2812	0.3333	0.3073
	179	0.5312	0.5833	0.5573
	181	0.0938*	-	0.0469
	183	0.0938	0.0833	0.0886
M257	101	0.4375	0.7500	0.5938
	107	0.0625*	-	0.0313
	123	0.0625*	-	0.0313
	125	0.3125	0.2500	0.2813
	139	0.1250*	-	0.0625
M258	101	0.1250*	-	0.0625
	107	0.5000	0.5833	0.5417
	109	0.0625	0.0833	0.0729
	115	0.2500	0.0833	0.1667
	117	0.0625	0.2500	0.1563
M259	107	0.0625*	-	0.03125
	111	0.1250*	-	0.0625
	113	0.2188	0.1667	0.1928
	115	0.5938	0.8333	0.7136
M260	111	0.0625*	-	0.0313
	113	0.0938*	-	0.0469
	115	0.2188	0.1667	0.1928
	123	0.0625*	-	0.0313
	129	0.0625*	-	0.0313
	131	0.0625*	-	0.0313
	137	0.3750	0.8333	0.6042
	139	0.0625*	-	0.0313
M746	359	0.0625*	-	0.0313
	361	0.2188*	-	0.1094
	363	0.4062	0.1667	0.2865
	369	-	0.2500*	0.1250
	371	0.1250	0.5833	0.3542
	373	0.1875*	-	0.0938

Table 1.9: Hardy-Weinberg probability test for *C. montis-sacri*. F_{is} = allelic fixation index for polymorphic loci, P = probability value, SE = standard error, Chi^2 = Chi square value using Fisher's method, and df = degrees of freedom

Population		Locus						Chi^2	df	P
		M255	M257	M258	M259	M260	M746			
A. 321	F_{is}	-0.1765	0.3023	0.5522	-0.3732	0.2327	0.2623	49.2971	12	0.0000*
	P	0.6714	0.0107	0.0010	0.3149	0.0014	0.0061			
	S.E.	-	0.0017	0.0004	-	0.0009	0.0015			
In situ	F_{is}	-0.1429	-0.2500	0.2308	0.1111	0.1111	0.4872	4.5602	12	0.9711
	P	0.6364	1.0000	0.7576	1.0000	1.0000	0.2121			
	S.E.	-	-	-	-	-	-			

* Significantly deviates from Hardy-Weinberg equilibrium at 0.05 level of significance.

Coffea vatovavyensis

Three loci (M255, M259, and M746) were polymorphic across all populations. Two loci, M257 and M258 were monomorphic for the *ex situ* population A. 1009 and locus M260 was monomorphic for *ex situ* population A. 830. The *ex situ* populations A. 308 and A. 830 had higher values for the means of **A**, $\mathbf{H_o}$, $\mathbf{H_e}$, and mean number of observed genotypes than the *ex situ* population A. 1009 and the *in situ* population (Table 1.10). These two populations (A. 308 and A. 830) did not show significant differences from each other for these four parameters when tested with a two-tailed *t*-test. All four populations showed mean heterozygote deficiencies and had a positive mean $\mathbf{F_{is}}$. Accession A. 308 showed the highest diversity with mean **A** across loci at 5.0 (ranging from 4 – 7 for individual locus), mean $\mathbf{H_o}$ and $\mathbf{H_e}$ of 0.48 and 0.64, respectively, and a mean number of genotypes across loci of 5.5 (ranging from 3 genotypes at locus M259 to 7 at locus M746). The parameters of mean **A** and mean number of observed genotypes were significantly higher for A. 308 than population A. 1009 and *in situ* population. A. 1009 showed the lowest mean **A** (1.8) and mean number of observed genotypes (2.0). This was due to the monomorphic nature of two loci, M257 and M258. The mean $\mathbf{F_{is}}$ was lowest for the *in situ* population with a mean value of 0.07. The *in situ* population had a mean value of **A** of 3.0 and the mean number of genotypes observed were 3.5 (Table 1.10).

For the six loci surveyed, there were a total of 26 alleles, ranging from 1 to 7 per locus (Table 1.11). Of these, there were a total of 21 private alleles with eight private alleles observed in *ex situ* population A308 at locus M255, M257, M260, and M746, six private alleles in *ex situ* population A. 830 at locus M255, M257, and M746, one private allele in *ex situ* population A. 1009 at locus M746, and six private alleles in the *in situ* population at locus M255, M257, M259, M260, and M746 (Table 1.11). Locus M258 did not have any private alleles in any of the

Table 1.10: Genetic variability of *C. vatovavyensis* populations at six microsatellite loci. N = sample size per locus, A = allele numbers per locus, H_o = the observed heterozygosity, H_e = the expected heterozygosity, and F_{is} = the allelic fixation index for polymorphic loci

Population	Microsatellite Locus	Genetic Diversity Parameters					
		N	A	H_o	H_e	F_{is}	No. obs. Genotypes
A. 308	M255	24	5	0.5833	0.7047	0.1722	5
	M257	24	4	0.625	0.6793	0.0800	6
	M258	24	5	0.9167	0.7455	-0.2296	6
	M259	24	4	0.0417	0.5181	0.9196	3
	M260	24	5	0.5417	0.5217	-0.0382	6
	M746	24	7	0.1667	0.6821	0.7556	7
	Mean	24.00	5.00	0.4792	0.6419	0.2535	5.50
A. 830	M255	25	5	0.9200	0.6933	-0.3269	7
	M257	25	5	0.2400	0.6083	0.6055	7
	M258	25	4	0.6000	0.6250	0.0400	7
	M259	25	5	0.3600	0.7967	0.5481	8
	M260	25	1	0.0000	0.0000	-	1
	M746	25	5	0.3600	0.7050	0.4894	9
	Mean	25.00	4.17	0.4133	0.5714	0.2766	6.50
A. 1009	M255	6	2	0.5000	0.4000	-0.2500	2
	M257	6	1	0.0000	0.0000	-	1
	M258	6	1	0.0000	0.0000	-	1
	M259	6	2	0.0000	0.5333	1.0000	2
	M260	5	2	0.6000	0.4500	-0.3333	2
	M746	5	3	0.4000	0.6250	0.3600	4
	Mean	5.67	1.83	0.2500	0.3347	0.2711	2.00
In situ	M255	36	4	0.6111	0.5782	-0.0570	6
	M257	36	3	0.1944	0.2254	0.1373	4
	M258	36	3	0.0000	0.1095	1.0000	3
	M259	36	3	0.0278	0.0556	0.5000	2
	M260	32	2	0.3438	0.3286	-0.0460	3
	M746	35	3	0.2286	0.2118	-0.0794	3
	Mean	35.17	3.00	0.2343	0.2515	0.0720	3.50

Table 1.11: Allele frequencies at six polymorphic loci for *C. vatovavyensis* (*Private alleles)

Locus	Allele	Allele Frequency				
		A. 308	A. 830	A. 1009	*In situ*	Mean
M255	171	0.2083	0.3600	0.7500	-	0.3296
	173	0.0625*	-	-	-	0.0156
	175	0.0208*	-	-	-	0.0052
	181	-	0.2000	0.2500	0.3611	0.2028
	187	0.4375*	-	-	-	0.1094
	189	0.2708	0.3800	-	0.5417	0.2981
	191	-	0.0400	-	0.0694	0.0274
	193	-	0.0200*	-	-	0.0050
	221	-	-	-	0.0278*	0.0070
M257	119	0.0417	-	-	0.8750	0.2292
	121	-	-	-	0.0139*	0.0035
	125	-	-	-	0.1111*	0.0278
	129	-	0.4200*	-	-	0.1050
	131	0.4167	0.0200	1.0000	-	0.3592
	133	0.1875*	-	-	-	0.0469
	135	-	0.0200*	-	-	0.0050
	137	0.3542	0.0600	-	-	0.1036
	141	-	0.4800*	-	-	0.1200
M258	99	0.1875	-	-	0.0278	0.0538
	101	0.3750	0.2200	1.0000	0.0278	0.4057
	105	0.1250	0.5600	-	-	0.1713
	107	0.0417	0.1400	-	0.9444	0.2815
	111	0.2708	0.0800	-	-	0.0877
M259	107	-	-	-	0.0139*	0.0035
	109	0.3333	0.2400	-	0.9722	0.3864
	111	0.0208	0.1200	-	-	0.0352
	115	0.6250	0.2800	-	-	0.2263
	117	0.0208	0.1000	0.6667	0.0139	0.2004
	119	-	0.2600	0.3333	-	0.1483
M260	105	-	-	-	0.7969*	0.1992
	107	0.0833*	-	-	-	0.0208
	109	0.0208*	-	-	-	0.0052
	111	0.6667	1.0000	0.7000	-	0.5917
	113	0.0417	-	0.3000	0.2031	0.1362
	117	0.1875*	-	-	-	0.0469
M746	339	-	0.0400*	-	-	0.0100
	353	0.0208	0.2200	-	-	0.0602
	361	0.3958*	-	-	-	0.0990
	363	0.1042	0.4800	0.3000	-	0.2211
	365	0.0208	0.1000	-	-	0.0302
	367	-	0.1600*	-	-	0.0400
	377	0.0208	-	-	0.0571	0.0195
	379	-	-	-	0.0571*	0.0143
	385	0.0208	-	-	0.8857	0.2266
	399	0.4167	-	0.6000	-	0.2542
	401	-	-	0.1000*	-	0.0250

populations. In addition, the *ex situ* populations had 13 alleles that were unique to them that were not present in the wild population. These were present in all loci (Table 1.11).

The Hardy-Weinberg probability tests by population showed that the $\mathbf{H_o}$ of two *ex situ* populations (A. 308 and A. 830) and of the *in situ* population of *C. vatovavyensis* deviated from Hardy- Weinberg equilibrium with significant deficiencies of heterozygotes at the 0.05 level of significance. The *ex situ* population A. 1009 was in Hardy-Weinberg equilibrium since the mean $\mathbf{F_{is}}$ did not deviate significantly from zero (Table 1.12).

POPULATION GENETIC STRUCTURE

For *C. kianjavatensis*, the majority of the variation was present within populations (84.52%) with the remaining variation distributed among populations (15.48%) (Table 1.13). *Coffea montis-sacri* had a lower among population variation (6.35%) with the vast majority of the variation distributed within populations (93.65%), though the variation was not significant at the 0.05 level of significance suggesting that there is no significant difference between the *ex situ* A. 321 and *in situ* populations. For *C. vatovavyensis*, the partitioning of genetic variation was almost equally distributed among- (47.03%) and within-populations (52.97%) (Table 1.13).

PARENTAGE ANALYSIS

FOFIFA records indicate that the seeds of *C. kianjavatensis* were collected from A. 213-line 8-plant 2, which corresponds to my collection number SK 70 and the *C. montis-sacri* seeds were collected from A. 321-line 68-plant 2, which corresponds to my collection number SK 109,

Table 1.12: Hardy-Weinberg probability test for *C. vatovavyensis*. F_{is} = allelic fixation index for polymorphic loci, P = probability value, SE = standard error, Chi^2 = Chi square value using Fisher's method, and df = degrees of freedom

Population		Locus**						Chi^2	df	P
		M255	M257	M258	M259	M260	M746			
A. 308	F_{is}	0.1722	0.0800	-0.2296	0.9196	-0.0382	0.7556	Infinity	12	HS*
	P	0.0000	0.0000	0.0000	0.0000	0.1408	0.0000			
	S.E.	0.0000	-	0.0000	-	0.0104	0.0000			
A. 830	F_{is}	-0.3269	0.6055	0.04	0.5481	N/A	0.4894	Infinity	10	HS*
	P	0.1094	0.0000	0.1115	0.0000	N/A	0.0002			
	S.E.	0.0076	0.0000	-	0.0000	N/A	0.0002			
A. 1009	F_{is}	-0.2500	N/A	N/A	1.0000	-0.3333	0.3600	7.7522	8	0.4382
	P	1.0000	N/A	N/A	0.0303	1.0000	0.6190			
	S.E.	-	N/A	N/A	-	-	-			
In situ	F_{is}	-0.0570	0.1373	1.0000	0.5000	-0.0460	-0.0794	32.9950	12	0.001*
	P	0.0549	0.4337	0.0002	0.0141	1.0000	1.0000			
	S.E.	-	-	-	-	-	-			

* Significantly deviates from Hardy-Weinberg equilibrium at 0.05 level of significance.

HS - Highly significant (P<0.001)

**Monomorphic loci are designated as N/A

Table 1.13: AMOVA results examining genetic partitioning between *ex situ* and wild populations for *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis.*

Species	Source of variation	df	Sum of Squares	Variance components	Percentage of variation
Coffea kianjavatensis	Among populations	2	64.466	0.2683	15.48***
	Within populations	361	528.869	1.46501	84.52
Coffea montis-sacri	Among populations	1	4.202	0.1304	6.35^{ns}
	Within populations	42	80.844	1.9249	93.65
Coffea vatovavyensis	Among populations	3	150.799	1.1626	47.03***
	Within populations	178	233.096	1.3095	52.97

*** $P < 0.001$

^{ns} – not significant

both of which would, therefore, be the maternal parents. After I performed the genetic diversity analysis, comparison of alleles of the supposed maternal parent for both species with the offspring revealed that some of the offspring did not have any alleles of the putative mother. Hence it was concluded that there were discrepancies in record keeping at the nursery where seeds had been mixed up and so I decided to perform the parentage analysis without assigning the maternal parent a priori.

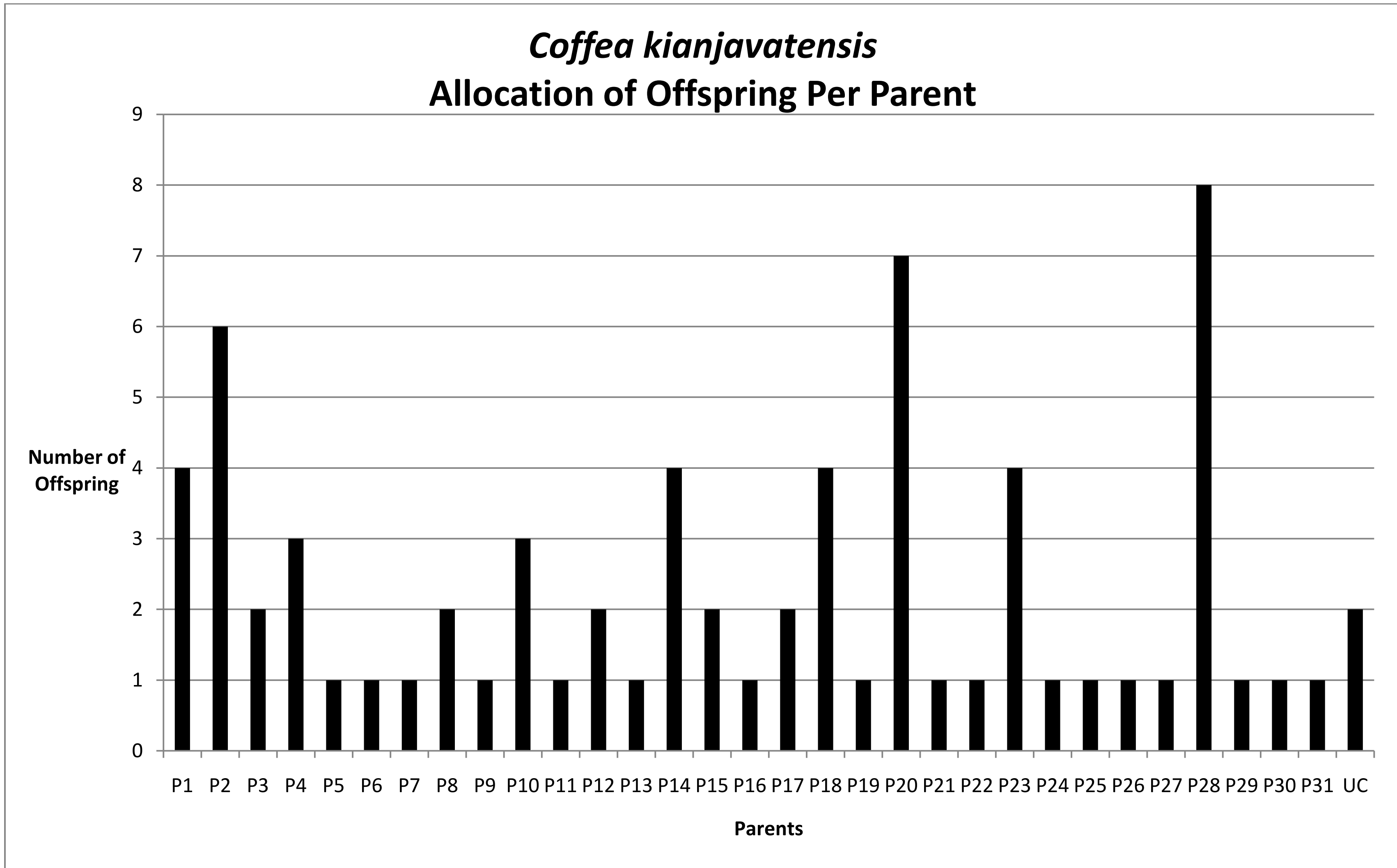
Coffea kianjavatensis

Of the 50 seedlings of *C. kianjavatensis,* 13 seedlings did not amplify across all six loci and so data for these seedlings were discarded from the study. When an initial parentage analysis was performed, one seedling showed both parents as uncollected and I decided to discard data from that seedling as well, under the assumption that this must have been caused by mixing of seeds from other species in the nursery since the entire possible *C. kianjavatensis* parental population had been sampled. So in the final parentage analysis only 36 seedlings were used. When the alleles were assigned on a parental pair basis, 34 (94%) had both parents identified in the parental population and two (6%) had one parent identified and one parent unidentified (Table 1.14). Parentage assignment showed 31 parents (26.05%) contributing to the genotypes of the 36 offspring from a total of 119 individuals in the parental population belonging to accessions A. 213 and A. 602 (Figure 1.1). The identification of each parent with my actual collection number is listed in Appendix 5. Of these 31 parents, accession A. 213 contributed to 42 offspring haplotypes and A. 602 contributed to 28 offspring haplotypes. Two offspring haplotypes were from uncollected parents, i.e. parents belonging to other species.

Table 1.14: Parentage allocation percentage classified on a parental pair basis for *C. kianjavatensis* and *C. montis-sacri*

	C. kianjavatensis	*C. montis-sacri*
Both parents collected	94.44%	66.67%
One parent collected + one parent uncollected	5.56%	33.33%

Figure 1.1: Allocation of offspring per potential parent for *C. kianjavatensis* UC = uncollected parent



Coffea montis-sacri

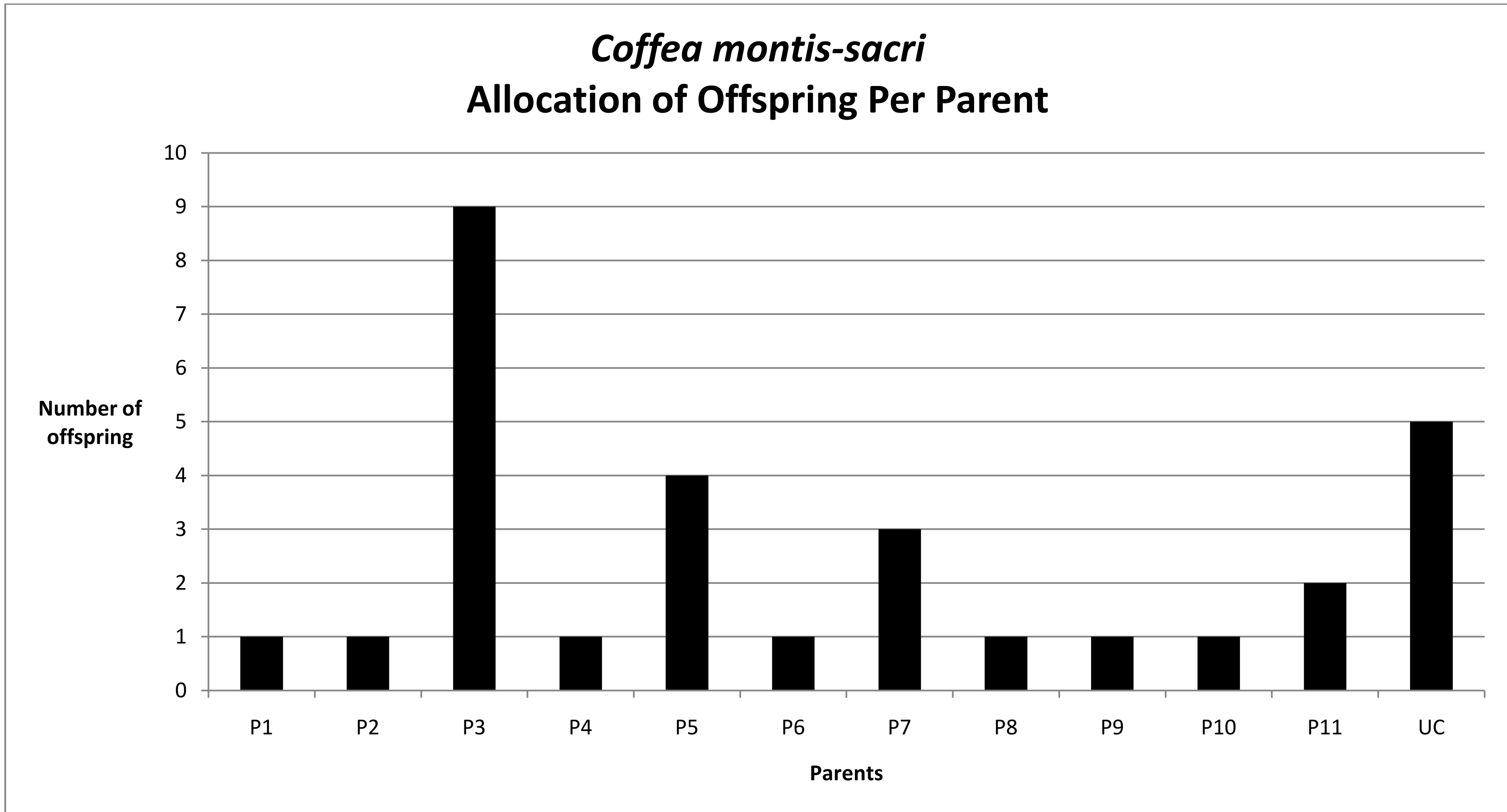
Of the 34 seedlings of *C. montis-sacri,* nine seedlings did not amplify across all six loci and so these seedlings were discarded from the study. When an initial parentage analysis was performed, 10 seedlings showed both parents as uncollected and I decided to discard these seedlings as well under the assumption that this must have been caused by the mixing of seeds from other species in the nursery since the entire possible *C. montis-sacri* parental population had been sampled. So in the final parentage analysis only 15 seedlings were used. When the alleles were assigned on a parental pair basis, 10 (67%) had both parents identified in the parental population and five (33%) had one parent identified and one parent unidentified (Table 1.14). Parentage assignment showed 11 parents (69%) contributing to the genotypes of the 15 offspring from a total of 16 individuals in the parental population belonging to accession A. 321 (Figure 1.2). The identification of each parent with my actual collection number is listed in Appendix 5. Five offspring haplotypes were from uncollected parents, i.e. a parent belonging to another species.

DISCUSSION

GENETIC DIVERSITY PATTERNS

The KCRS operated by FOFIFA in Madagascar is a unique and valuable *ex situ* field genebank housing a vast collection of various Madagascan coffee species with their total collections encompassing 2,649 specimens in 132 accessions (J. J. Rakotomalala, pers. comm.).

Figure 1.2: Allocation of offspring per potential parent for *C. montis-sacri*. UC = uncollected parent



With natural habitats being lost at a fast pace, conserving the genetic resource of these wild crop relatives in this *ex situ* field genebank will become a valuable alternative. Understanding the genetic diversity represented in these genebanks is key in developing strategies for optimum management of these genetic resources. There have been no previous attempts to quantify the genetic diversity of this germplasm. My present study examining three wild species of coffee, *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis*, all narrowly endemic to the Kianjavato region is the first attempt at quantifying the genetic diversity and gene flow of these species in the *ex situ* genebank and comparing with extant wild populations from nearby forests in order to make recommendations for improving and managing the *ex situ* germplasm.

The high mutation rate of microsatellite markers make them valuable tools for assessment of genetic structure and diversity within species (Cubry et al. 2008). Microsatellite markers suitable for genetic studies have been developed for *Coffea* species by various researchers and have shown good transferability across diploid species (Combes et al. 2000; Coulibaly et al. 2003; Cubry et al. 2008; Poncet et al. 2004; Poncet et al. 2007). Combes et al. (2000) developed 11 primer pairs for 11 microsatellite loci for *C. arabica*, which were tested for cross-species amplification across 11 diploid *Coffea* species and two related *Psilanthus* species (now placed in *Coffea*; Davis 2010). Results showed good cross amplification across species including four Madagascan species. Good transferability of microsatellite markers developed for *C. arabica* was demonstrated across genetically distant species such as *C. canephora, C. heterocalyx, C. pseudozanguebariae* (Coulibaly et al. 2003; Poncet et al. 2004; Poncet et al. 2007), *C. liberica, C. eugenioides* (Poncet et al. 2004), and *C. canephora* hybrids with *C. heterocalyx* and *C. pseudozanguebariae* (Poncet et al. 2007). Curby et al. (2008) showed good transferability of microsatellite markers developed in *C. arabica* and *C. canephora* to 15 other *Coffea* species.

Based on these studies, 12 microsatellite markers were selected to test for cross amplification across the three Madagascan study species, *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis,* of which six markers showed good amplification and so these six markers were used for genetic diversity assessment (Table 1.3).

The higher levels of genetic diversity found in the two *ex situ* populations of *C. kianjavatensis* compared to the *in situ* population are most likely remnants of the diversity from the original collections that were collected and established in 1962-1963. The lower genetic diversity of the extant *in situ* population may indicate that genetic diversity has been lost in the wild possibly due to habitat loss and highlights the importance of preserving the plants currently in the *ex situ* collections. Because the wild population harbors no novel allelic diversity compared to the *ex situ* populations, making additional collections from that population would not be of any value in terms of supplementing the genetic diversity of the *ex situ* populations.

For *C. montis-sacri* also, the *ex situ* population (A. 321) showed higher genetic diversity than the *in situ* population (Table 1.7) and harbored a much higher number of private alleles (Table 1.8). Similar to *C.* kianjavatensis, most of the genetic variation was within populations (Table 1.13) again reflecting the primarily outcrossing nature of this species. As a Critically Endangered species (IUCN, 2001), I was able to locate only six plants in the wild. There are only 16 plants in the *ex situ* collection and so it is important that the existing germplasm is preserved and additional collections made from the extant wild population. In addition, attempts to scout for locations of new populations should be made.

The partitioning of most of the genetic diversity within single populations of both *C. kianjavatensis* and *C. montis-sacri* is typical of primarily outcrossing species and is comparable

to the genetic structure observed in an international cacao (*Theobroma cacao*) collection where the within-group variation accounted for 84.6%, and the variation between accession groups accounted for 15.4% of the total molecular variation (Zhang et al. 2009). This pattern of genetic partitioning suggests that conspecific populations are not significantly differentiated from each other and crossing of individuals between populations will not cause any problems, such as outbreeding depression.

For *C. vatovavyensis,* two *ex situ* populations (A. 308 & A. 830) exhibited higher genetic diversity than the *in situ* population. The low diversity of A. 1009 was mainly due to having only six individuals in that population as well as the monomorphic nature of two loci (M257 and M258). *Coffea vatovavyensis* differed from the other two species surveyed in having lower within population (53%) genetic variation and a higher among population (47%) differentiation (Table 1.13). This is comparable to the genetic structure observed in the threatened Chilean vine *Berberidopsis coralliana*, which exhibit an among-population variation of 54.83% when *ex situ* and *ex situ* populations are compared (Etisham-Ul-Haq et al. 2001). The presence of 21 private alleles across all populations, an indication of a unique genetic diversity assemblage of each population, necessitates that each population be preserved so as not to lose this diversity. The *in situ* population had six private alleles and so wild collecting missions should be undertaken to enhance the *ex situ* germplasm collection and preserve this genetic diversity before this population is lost.

The overall genetic diversity of wild Madagascan *Coffea* species seems to be similar or even higher than other cultivated and wild coffee species. The mean ranges of $\mathbf{H_o}$ for *C. kianjavatensis* was 0.37 – 0.60, for *C. montis-sacri* was 0.44 – 0.59, and for *C. vatovavyensis* was 0.23 – 0.47 with mean **A** ranging from 3.0 – 4.5, 2.7 – 5.2, and 1.8 – 5.0, respectively. In

comparison, the mean $\mathbf{H_o}$ and mean **A** for *C. arabica* (both cultivated and wild accessions) was 0.49 and 2, for *C. canephora* (genotypes from different genetic groups such as Congolese, Guinean and Ugandan) was 0.29 and 5, for *C. congensis* (accessions from different Central African regions) was 0.27 and 3, and for *C. liberica* (genotypes from different varieties such as *C. liberica* var. *liberica* and *C. liberica* var. *dewevrei*) was 0.34 and 4, respectively using 60 microsatellite loci (Cubry et al. 2008). Silvestrini et al. (2007) using 16 SSR primers reported an average **A** of 2.5 for *C. eugenioides*, 2.8 for *C. canephora*, and 2.1 for *C. racemosa* from plants maintained at the *Coffea* Germplasm Collection of IAC (Instituto Agronomico de Campinas, Brazil). The mean $\mathbf{H_o}$ ranged from 0.30 – 0.45 and the mean **A** ranged from 6.42 – 7.83 for *C. canephora* populations from six different regions of Uganda (Musoli et al. 2009), where $\mathbf{H_o}$ was comparable to the present study. Musoli et al. (2009) also report higher genetic diversity of the cultivated genotypes compared to wild populations and attribute this to the multiple origins of the cultivated plants and successive hybridization of wild material with introduced genotypes. The higher diversity of the Madagascan *Coffea* species in the *ex situ* genebank could also be due to the multiple origins of the material during collection as well as representation of sampling from populations which were larger during the 1960s than the current extant wild populations, representing higher genetic diversity. Another cause of higher diversity could also be due to cross-contamination of germplasm with pollen from another species arising from the practice of replacing lost plants in the collection with seedlings germinated from open-pollinated seeds as revealed in the parentage analysis.

PARENTAGE ANALYSIS

Understanding the mechanisms and extent of gene flow within and among plant populations and species has practical implications for the conservation and utilization of plant genetic resources (Dawson et al. 1997). Maintaining the genetic integrity of germplasm collections of open-pollinated commercial crops is a challenge if wild populations of compatible related species are within pollination distance (Suso et al. 2006). If *ex situ* collections are used for reintroduction and restoration, interspecific hybridization could jeopardize the genetic integrity of endangered species, irrevocably contaminating the gene pool (Zhang et al. 2010). Hence knowledge of genetic diversity and gene flow within the germplasm becomes essential in maintaining and managing *ex situ* genebanks. A main objective of any field genebank is to maintain the germplasm without any genetic erosion due to contamination by foreign pollen (Suso et al. 2006). This becomes a challenge when managing outcrossing species like those of *Coffea* where several species are held within the field genebank in close proximity to each other. By understanding gene flow patterns, management strategies can be developed so that the genetic integrity of the germplasm is maintained.

For *C. kianjavatensis*, parentage analysis indicates that 5.56% of the offspring were contaminated with pollen from another *Coffea* species, whereas the contamination percentage for *C. montis-sacri* was 33.33%. The higher contamination rate for *C. montis-sacri* could be due to the small number of individuals (16) in the *ex situ* population. In addition, only 15 seedlings were used in the parentage analysis study. Future study should be undertaken using a larger number of seedlings for parentage analysis.

Numerous studies have looked at assigning paternity in populations where the maternal plants are known. In a paternity analysis study using microsatellites in bur oak (*Quercus*

macrocarpa, Fagaceae), Dow and Ashley (1998) estimated at least 57% of acorns resulted from fertilization by trees outside the stand. Using six microsatellites in a gene flow study of a Malagasy *Eucalyptus grandis* (Myrtaceae) seed orchard, pollen flow from outside the stand was determined to be nearly 40% (Chaix et al. 2003). Paternity analyses have also been performed in coniferous trees such as limber pine (*Pinus flexilis*, Pinaceae) using allozymic loci (Schuster and Mitton 2000) and in Japanese red pine (*Pinus densiflora*, Pinaceae) using microsatellite loci (Lian et al. 2001), with both studies estimating pollen dispersal distances. In *Pinus flexilis*, a wind pollinated species, paternity analysis indicated pollen immigration of 6.5% from populations 2 km to 100+ km away (Schuster and Mitton 2000). In the *Pinus densiflora* study, paternity analysis indicated that at least 31% of the offspring were fertilized by pollen from trees outside the stand (Lian et al. 2001). For *Sinojackia xylocarpa*, an extinct tree species in the wild, 32.7% of the seeds collected from maternal trees maintained in an *ex situ* collection were reported to be hybrids as a result of pollen contamination from another related species, *S. rehderiana* (Zhang et al. 2010). Paternity analysis for this insect-pollinated species in the *ex situ* collection revealed long-distance pollination, with average pollen dispersal distance of 184.6 m and a maximum distance of 600 m, which was comparable to other insect-pollinated tree species (Zhang et al. 2010). Though the present study is not looking at paternity assignments or pollen dispersal distances, the main objective was to perform parentage analysis regardless of identification of maternal and paternal parents. This will help determine if the practice of replacement planting in the *ex situ* field genebank using seed propagated material is compromising the integrity of the gene pool by contamination from pollen of other *Coffea* species maintained in the genebank and if so, to what degree.

This is the first study examining pollination and gene flow patterns in an *ex situ* coffee field genebank. The results indicate that open pollinated seed propagation in the coffee field genebank is unambiguously contaminated by pollen from other species of *Coffea* and that the level of extra species cross-fertilization was variable depending on the species sampled. The results showed that some offspring of both species were a result of interspecific hybridization. These results have serious implications concerning the value and future management of ex situ *Coffea* germplasm collections. These collections should be managed in such as way as to preserves alleles present in populations that are unique to a particular species and not lost due to outcrossing with other species.

CONSERVATION AND MANAGEMENT IMPLICATIONS

To adequately conserve the full range of a target species' genetic diversity, no single conservation technique applied alone is adequate (Dulloo et al. 1998). There are essentially two basic conservation strategies: *in situ* and *ex situ* (Dulloo et al. 1998). For commercial plant species, in addition to maintaining plants *in situ* in their natural habitats, maintenance of plants in an *ex situ* field genebank offers feasible medium and long-term storage, conservation of genetic diversity of target taxa which could be lost in the wild due to vulnerability to natural and anthropogenic disasters, and easy access for characterization, evaluation and breeding purposes (Dulloo et al. 1998). However, conservation in field genebanks pose threats of their own such as exposure to pests and diseases, natural hazards such as drought, weather damage, human error, vandalism, and genetic erosion (Dulloo et al. 1998; Engelmann and Dulloo 2007). Another important consideration for conservation in *ex situ* collections is the possibility of inter-specific hybridization that could occur when common barriers to interspecific crossing such as

geographic isolation is removed when plants are cultivated together in an artificially sympatric living collection (Ye et al. 2006). In addition to conservation in field genebanks, other *ex situ* conservation measures such as *in vitro* storage, pollen storage, DNA storage, and seed storage should be more thoroughly evaluated and optimized (Engelmann and Dulloo 2007).

Traditionally long term germplasm conservation of *Coffea* species has been done through *ex situ* field collections since coffee seeds are recalcitrant and not amenable to conventional seed storage methods (Vega et al. 2008). Exploration and collecting missions of wild *C. arabica, C. canephora* and other wild *Coffea* species was undertaken in the 1960s, 1970s, and 1980s by various organizations such as FAO (Food and Agriculture Organization), ORSTOM (Office de la Recherche Scientifique et Technique Outre-Mer; renamed Institute de Recherche pour le Developpement (IRD) in 1998), CIRAD (Centre de Cooperation Internationale en Recherche Agronomique pour le Developpement) and IPGRI (International Plant Genetic Resources Institute), leading to the establishment of *ex situ* field genebanks in Africa (Cameroon, Ethiopia, Ivory Coast, Kenya, Tanzania), Madagascar, India, and the Americans (Brazil, Colombia, and Costa Rica) (Vega et al. 2008). These *ex situ* germplasm collections are vital as the future of coffee crop improvement depends on the use of untapped genes found in this wild gene pool (Vega et al. 2008). Many of these field genebanks have been experiencing substantial losses of plant accessions due to age of the trees, unsuitable cultivation practices, climatic conditions (Vega et al. 2008), and lack of funding. Before accessions and the genetic diversity contained in them are lost, characterizing the genetic diversity held within these collections will be a first step in prioritizing *ex situ* conservation.

To increase efficiency in space utilization and to ensure optimal representation of genetic diversity, genetic studies should assess genetic redundancy contained within the collection so

that redundant genotypes can be removed, making room for new collections representing genetic diversity not currently present in the collection similar to that performed for an international *Cacao* germplasm collection in Costa Rica (Zhang et al. 2009). Maintenance of collections in field genebanks become prohibitively expensive requiring considerable inputs such as land, labor, and materials (Engelmann and Dulloo 2007). In order to overcome this, developing a core collection should become a priority (van Hintum et al. 2000). A core collection is defined as a limited set of accessions representing, with a minimum of repetitiveness, the genetic diversity of a crop species and its wild relatives (Frankel 1984). Van Hintum et al. (2000) describe the process of establishing a core collection. Passport data on the accessions in the genebank should also be maintained which should include information on the genetic origin of the accessions and information about provenance of the plant material (Anthony et al. 2007b). During establishment of a genebank collection, attempts should be made at collecting herbarium voucher specimens as well as DNA samples of each accession, which will provide baseline information for future assessment of genetic integrity.

In Madagascar, the present study characterizing the genetic diversity of the three locally endemic species of the Kianjavato region, *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis*, can serve as a model for characterizing and evaluating the genetic diversity of all other species in the collection. The present study shows that high genetic diversity is represented in this *ex situ* collection, though how much of this is due to introgression in the genebank is unknown. Results confirm that contamination of the existing gene pool in the *ex situ* collection is prevalent, compromising the genetic integrity of the collection. A thorough examination of collection records should be undertaken to determine what percentage of the original collection still remains as well as additional wild collection missions should be considered to augment this

collection to capture some of the genetic diversity that is present in wild populations, but not represented in the field genebank.

Except for *C. arabica*, almost all other coffee species are self-incompatible (Anthony et al. 2007a). *Coffea* species share a common genome making interspecific hybridization possible, which is valuable in the transfer of new characters from diploid coffee species into the genome of *C. arabica* cultivars (Anthony et al. 2007a). In the present study, in an open pollinated system, cross pollination between species was identified using parentage analysis. In order to maintain the genetic integrity of the collection, replacement plantings should be performed with plants propagated either clonally (through cuttings or tissue culture) or through seeds generated by controlled pollination. Knowledge of the extent of outcrossing among species will be critical in designing strategies for *ex situ* germplasm management.

CHAPTER TWO

GENETIC DIVERSITY PATTERNS OF *COFFEA COMMERSONIANA*, A RARE AND ENDAGERED MALAGASY ENDEMIC

INTRODUCTION

BACKGROUND

Since the separation of the tectonic plate from Africa 165 million years ago and from India at the end of the Cretaceous period about 70 million years ago, Madagascar has been evolving in isolation leading to the emergence of numerous and distinct forms of plants, animals, and geological features (Rakotosamimanana 2003). Based on a comprehensive review of phylogenetic studies of the Malagasy biota, Yoder and Nowak (2006) attribute the origin of this distinct diversity to Cenozoic dispersal, predominantly of African origins. Madagascar is home to over 10,000 vascular plant species with about 90% endemism (Moat and Smith 2007). The family Rubiaceae to which the genus *Coffea* belongs to is the largest family of woody plants in Madagascar with 569 species exhibiting 91% endemicity (Davis & Bridson 2003; Davis et al. 2009).

The Vegetation Atlas of Madagascar (Moat and Smith 2007) identifies ten major physiognomic types with 15 mapped vegetation units, namely: humid forest, degraded humid forest, littoral forest, wooded grassland-bushland mosaic, plateau grassland-wooded grassland mosaic, tapia forest, western humid forest, western dry forest, western sub-humid forest, south western dry spiny forest-thicket, degraded south western dry spiny forest, south western coastal bushland, wetlands, mangroves, and cultivation (Moat and Smith 2007). The Madagascan *Coffea* species have narrow distribution ranges occurring in diverse forest types including littoral,

evergreen, gallery (riverine), mixed deciduous, dry, xerophytic, and elfin forests (Davis et al. 2006).

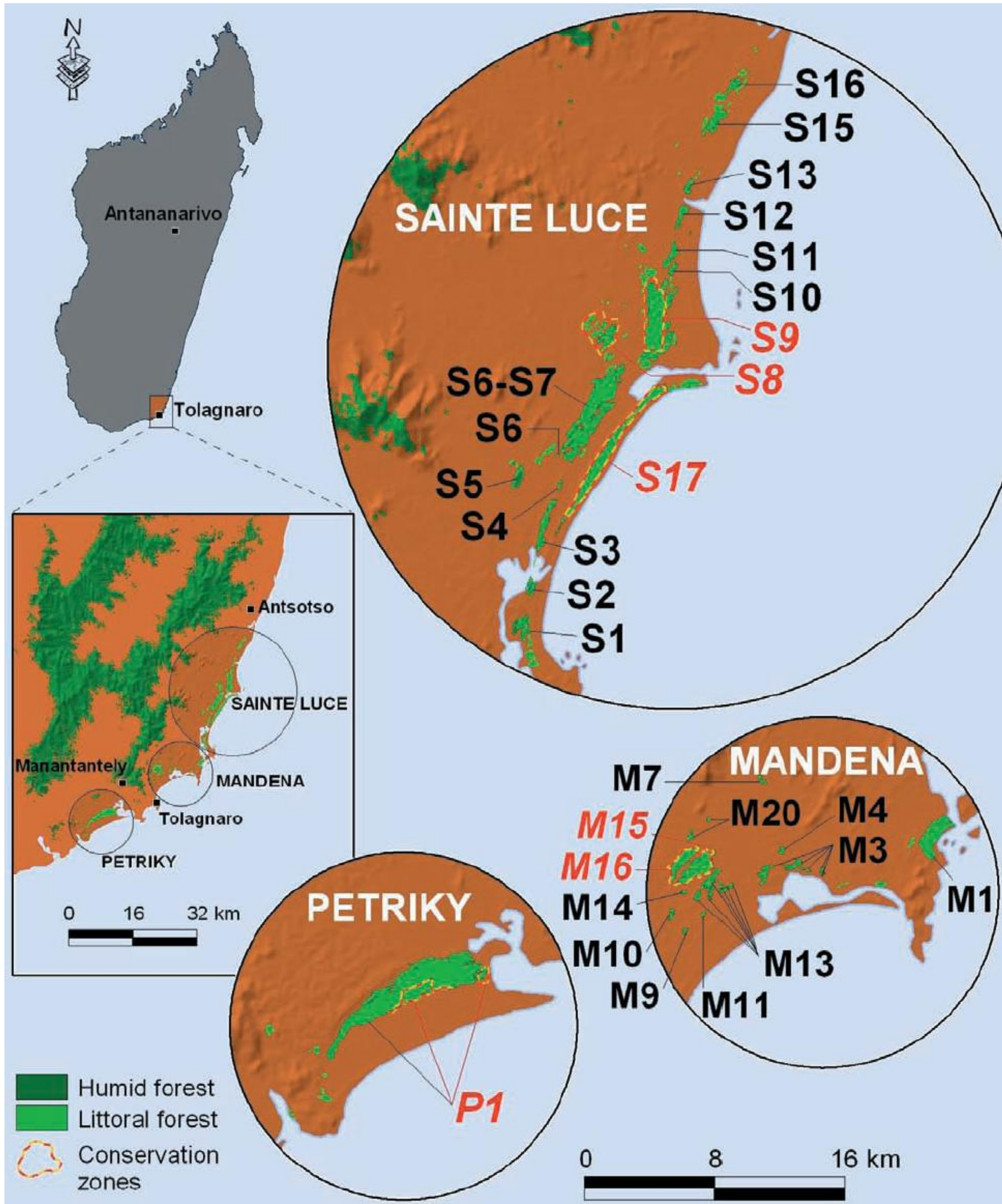
The littoral forest of Madagascar is a distinctive type of humid evergreen forest restricted to unconsolidated sand located within a few kilometers from the Indian Ocean (Lowry et al. 2008). The littoral forests, which once occupied much of the coastal fringe of eastern Madagascar and were contiguous with the dense humid lowland evergreen forests, now persist only in small fragments (de Gouvenain and Silander 2003). The original size of this habitat was less than 1% of Madagascar's total surface area, and today exists in only about 10% of its original range (Moat and Smith 2007) with only about 1.5% of the remaining fragments included within the existing protected areas network (Consiglio et al. 2005). Even though the habitat range is very small, the littoral forests harbor about 13% of Madagascar's total native flora, of which 25% are endemic to this habitat (Moat and Smith 2007). The littoral region of southeastern Madagascar in Tolagnaro (Fort Dauphin) is dominated by the Vohimena Mountains and a rolling coastal plain extending several kilometers to the Indian Ocean (Vincelette et al. 2007b).

One of the most threatened ecosystems in Madagascar with less than 2,835 ha remaining, the littoral forests of the Tolagnaro region are expected to lose numerous plant and animal species in the near future as a result of deforestation and consequent habitat changes (Bollen and Donati 2006). The remaining littoral forests of southeastern Madagascar are under severe pressure from various threats from the local human population such as *tavy* (shifting slash and burn agriculture), bushfires as a result of the practice of *tavy*, harvest of timber and non-timber forest products (e.g. charcoal for cooking, wood for construction) for both subsistence and commercial activities (Bollen and Donati 2006; Vincelette et al. 2003). The three main remaining groups of littoral forest fragments are located in Mandena, Petriky, and Sainte Luce

with fragment sizes ranging from 1 to 377 ha (Bollen and Donati 2006) (Figure 2.1). The most imminent threat to these forests is the plan to extract ilmenite by QIT Madagascar Minerals (QMM) (Bollen and Donati 2006).

QMM, a company jointly owned by Rio Tinto, UK, and the Malagasy State represented by the Office des Mines Nationales et des Industries Strategiques de Madagascar (OMNIS) started an extensive exploration program in 1986 for heavy mineral sands containing titanium dioxide in the form of ilmenite and rutile along the eastern coast of Madagascar (Vincelette et al. 2007a). Major sediments were located underneath the littoral forests in Mandena, Sainte Luce, and Petriky (Lowry et al. 2008). Over the following 20 years, before the start of mining activities in 2009, QMM performed an extensive biodiversity assessment project addressing the potential impact of mining on economic, technical, and cultural issues with ramifications for environmental conservation (Vincelette et al. 2007a). Mining activities were to start in Mandena in 2009 and in Petriky and Sainte Luce 20-45 years later, lasting up to 60 years (Bollen and Donati 2006). The impact of these activities would result in the loss of littoral forests in Mandena, Sainte Luce, and Petriky at 62.8 ha, 661.8 ha, and 705.8 ha, respectively (Bollen and Donati 2006). To mitigate this loss, the environmental impact assessment conducted by QMM has led to the establishment of tree nurseries and plantations, seed banks, and extensive research into reforestation (Bollen and Donati 2006). The detailed timetable of the QMM mining project is given in Appendix 4.

Figure 2.1: Map showing the littoral forest areas of southeastern Madagascar in Sainte Luce, Mandena and Petriky (light green) and newly established conservation zones. (Figure reproduced from Lowry et al. 2008).



Coffea commersoniana (Baill.) A.Chev., known by the vernacular name 'kotofotsy', is the only coffee species endemic to the littoral forests of southeastern Madagascar (Rabenantoandro et al. 2007). Its distribution is restricted to the humid, evergreen littoral forest of the Tolagnaro region and is considered Endangered (Davis et al. 2006) by the IUCN Categories and Criteria system (IUCN 2001). *Coffea commersoniana* has been recorded at Petriky, Mandena, and Sainte Luce littoral forests (Rabenantoandro et al. 2007) and various sites nearby (A. Davis, pers. comm.). The mining operation will lead to a great reduction in population extent and size ('area of occurrence' and 'area of extent' (IUCN 2001), respectively) of this species and hence understanding the genetic diversity is critical for taking necessary steps to mitigate the impacts of mining through *ex situ* and *in situ* (restoration ecology) conservation.

OBJECTIVES

My primary objective was to evaluate the genetic diversity of extant wild populations of *C. commersoniana* from the littoral forests of Mandena and Petriky and to compare those with a population maintained at the *ex situ* field genebank at the FOFIFA Kianjavato Coffee Research Station (KCRS) using microsatellite markers. Specific questions addressed were:. 1) What levels of genetic diversity are present within and across each *ex situ* and *in situ* population? 2) How is genetic diversity structured among these populations? Based on these results, implications and recommendations for conservation management are discussed.

MATERIALS AND METHODS

PLANT MATERIAL

I conducted fieldwork in Madagascar to collect plant specimens during November 2008. I visited the KCRS in Kianjavato and sampled the entire *ex situ* population of *C. commersoniana.* The *ex situ* collection only had one accession (A. 302) with 28 plants. Wild populations were sampled from Mandena and Petriky in the littoral forests in the Tolagnaro region in southeastern Madagascar. In Mandena, I sampled from the two conservation zones, M15 and M16, and in Petriky from the conservation zone P1 (Figure 2.1). These conservation zones are being managed by QMM officials and QMM provided transportation and other logistics to accomplish my fieldwork. While at QMM's facility at Mandena, I also sampled *C. commersoniana* seedlings propagated for restoration purposes at the company's nursery. Table 2.1 lists the *ex situ* and *in situ* populations sampled. The location coordinates (collected using WGS 84 map datum using a Magellan Meridian Color Handheld GPS) given are representative of the location of the first few samples collected for each population. Several leaves of each individual plant were collected and placed in a plastic bag with silica gel. Voucher specimens of selected samples were collected in replicates of four, one each for the herbaria at Royal Botanic Gardens, Kew (K), University of Colorado Museum (COLO), Parc Botanique et Zoologique de Tsimbazaza (TAN), and FOFIFA: National Center of Applied Research and Rural Development (TEF). The list of herbarium specimens is listed in Appendix 1.

The accession of *C. commersoniana* (A. 302) from the KRCS was collected in 1964 from the Tolagnaro region, although the precise locality of collection was not recorded. In 2008, the associated dominant plant species recorded with *C. commersoniana* in the littoral forests of Mandena include: *Malleastrum mandenense* (Meliaceae), *Psorospermum* sp. (Clusiaceae),

Table 2.1: Populations sampled and their locations.

Species and Population	Location	Number of Individuals	Latitude/Longitude
C. commersoniana A. 302	FOFIFA Kianjavato Coffee Research Station	28	21°22’40”S / 47°52’13”E
C. commersoniana Mandena Population	QMM Mandena Conservation Zones M15 & M16	91	24°57’15”S / 47°00’15”E (M15) 24°56’56”S / 46°59’54”E (M16)
C. commersoniana Petriky Population	QMM Petriky Conservation Zone P1	70	21°24’36”S 47°56’32”E
C. commersoniana QMM Seedlings	QMM Nursery at Mandena	15	Not recorded

Phymatosorus scolopendria (Polypodiaceae), and *Campylospermum obtusifolium* (Ochnaceae). In the Petriky littoral forest, the associated dominant species recorded were *Euphorbia* sp. (Euphorbiaceae), *Polycardia phyllanthoides* (Celastraceae), *Senecio antandroi* (Asteraceae), *Campylospermum obtusifolium* (Ochnaceae), and *Diospyros* sp. (Ebenaceae).

DNA EXTRACTION AND MOLECULAR MARKERS

Genomic DNA was extracted from 10 mg of silica-dried leaf material using GenCatch™ Plant Genomic DNA Purification kit by Epoch Biolabs. Slight modifications were made to the extraction protocols. A detailed account of the extraction procedure is described in Appendix 2. Extracted DNA was sent to Nevada Genomics, Reno, Nevada for quantification, optimization and fragment analysis using SSR markers. Six microsatellite markers (M255, M257, M258, M259, M260, and M746) were used in this study (Table 1.3).

The DNA was quantified and normalized to 5.0 ng/µl. PCR amplifications were carried out using an MJ thermocycler. Each 10.0 µl PCR amplification reaction contained 4.0 µl of 5.0 ng/µl genomic DNA, 1.0 µl Primer Panel mix, and 5.0 µl Qiagen Multiplex PCR Mix. The amplifications were performed using a "touchdown" PCR profile as described in Coulibaly et al. (2003), which is listed in Appendix 3. The only modification was the time for the initial denaturation at 94°C was increased to 15 minutes due to the use of Qiagen Multiplex PCR Mix, which is a hot-start *Taq* DNA Polymerase. The samples were run on an Applied Biosystems Prism 3730 DNA Analyzer. The filter set used was G5, which detects the fluorescent dyes 6-FAM, VIC, NED, and PET. The samples were run with the 500 MW size standards labeled with LIZ. The six microsatellite loci were amplified in a single 6-primer panel.

The fragment analysis results were scored using GeneMapper® Software Version 4.0 by Applied Biosystems.

ANALYSIS OF GENETIC DIVERSITY

Data analysis to assess genetic diversity was performed using GENEPOP v.4.0 software (Rousset 2008). Parameters used to estimate genetic diversity included number of alleles per locus (**A**), the mean observed ($\mathbf{H_o}$) and the mean expected ($\mathbf{H_e}$) heterozygosities based on Hardy-Weinberg assumptions, the allelic fixation index ($\mathbf{F_{is}}$), and the number of observed genotypes per population per locus. GENEPOP v.4.0 was also used to calculate the allele frequencies at each locus with the private alleles for each population identified. Conformance to Hardy-Weinberg equilibrium by population was performed by assessing the significance of the $\mathbf{F_{is}}$ values by means of Fisher's exact tests implemented in GENEPOP v.4.0 by the Markov Chain (MC) method employing 10,000 dememorization steps, followed by 20 batches of 5,000 iterations per batch. Where the number of alleles is less than five, the default in the batch mode is complete enumeration rather than MC method, where no standard error is computed. The $\mathbf{F_{is}}$ reported is based on Weir and Cockerham's (1984) estimate.

ANALYSIS OF POPULATION GENETIC STRUCTURE

Hierarchical genetic structure was examined through an analysis of molecular variance (AMOVA) (Excoffier et al. 1992) as implemented in Arlequin v.3.5.1 (Schneider et al. 2000). AMOVA was applied to estimate the components of variance among and within populations based on $\mathbf{\Phi_{st}}$, a statistic analogous to $\mathbf{F_{st}}$ to test the significance against the null hypothesis of no

structure. AMOVA was performed to estimate genetic partitioning among and within populations with all populations included as well as some excluded. Locus M746 had about 21.6% missing data across populations when the *ex situ* and seedling populations were excluded and so for this analysis, 22% missing data was allowed.

RESULTS

GENETIC DIVERSITY

The *ex situ* population showed very low diversity with one locus (M257) being monomorphic for a single allele and the remaining loci having two alleles each, with each plant in the population carrying the same two alleles (Tables 2.2 and 2.3). This inflated the allelic fixation index ($\mathbf{F_{is}}$) value (-1.00) suggesting an excess of heterozygosity. The mean value of **A** and the mean number of genotypes observed were 1.83 and 1.0, respectively, which were significantly lower compared to the other three populations. In comparison, the *in situ* and seedling populations had higher diversity with all six loci exhibiting polymorphism. The Mandena population showed the highest genetic diversity with a mean **A** of 7.83, ranging from 4 alleles at locus M260 to 11 alleles at M258 and 16.67 mean number of observed genotypes with 29 observed genotypes at locus M259 (Table 2.2). These were significantly higher than the *ex situ* and QMM seedling populations when tested using a two-tailed *t*-test at 0.05 level of significance. The mean value of **A** and the mean number of observed genotypes for the Petriky population were 6.67 (ranging from 3 at locus M259 to 9 at locus M257) and 13.0 (ranging from 4 genotypes at M259 to 23 genotypes at M257), respectively, and for the QMM seedlings were

Table 2.2: Genetic variability of *C. commersoniana* populations at six microsatellite loci. N = sample size per locus, A = allele numbers per locus, H_o = the observed heterozygosity, H_e = the expected heterozygosity, and F_{is} = the mean allelic fixation index for polymorphic loci.

Population	Microsat. Locus	Genetic Diversity Parameters					
		N	A	H_o	H_e	F_{is}	No. obs. Genotypes
A. 302	M255	28	2	1.0000	0.5000	-1.0000	1
	M257	28	1	0.0000	0.0000	-	1
	M258	28	2	1.0000	0.5000	-1.0000	1
	M259	28	2	1.0000	0.5000	-1.0000	1
	M260	28	2	1.0000	0.5000	-1.0000	1
	M746	28	2	1.0000	0.5000	-1.0000	1
	Mean	28.00	1.83	0.8333	0.4167	-1.0000	1.00
Mandena	M255	88	6	0.5000	0.7006	0.2864	14
	M257	79	8	0.5823	0.7695	0.2433	16
	M258	84	11	0.6429	0.7925	0.1889	23
	M259	81	10	0.6914	0.8535	0.1900	29
	M260	81	4	0.1605	0.3557	0.5488	6
	M746	74	8	0.1757	0.7322	0.7601	12
	Mean	81.17	7.83	0.4588	0.7007	0.3375	16.67
Petriky	M255	67	8	0.5373	0.6263	0.1421	19
	M257	61	9	0.8525	0.829	-0.0283	23
	M258	65	6	0.5231	0.5333	0.0192	9
	M259	61	3	0.4262	0.5031	0.1529	4
	M260	55	6	0.4364	0.4704	0.0723	9
	M746	50	8	0.4800	0.6706	0.2842	14
	Mean	59.83	6.67	0.5460	0.6055	0.0980	13.00
QMM Seedlings	M255	15	4	0.4667	0.6619	0.2950	5
	M257	9	5	0.1111	0.7500	0.8519	5
	M258	12	8	0.8333	0.7538	-0.1055	10
	M259	11	7	0.8182	0.8227	0.0055	9
	M260	12	4	0.7500	0.5341	-0.4043	4
	M746	10	3	0.0000	0.5111	1.0000	3
	Mean	11.50	5.17	0.4966	0.6723	0.2224	6.00

Table 2.3: Allele frequencies at six polymorphic loci for *C. commersoniana* (*Private alleles)

Locus	Allele	Allele Frequency				
		A. 302	Mandena	Petriky	QMM Seedlings	Mean
M255	173	-	0.0114*	-	-	0.0029
	175	-	-	-	0.1333*	0.0333
	177	-	0.4261	0.0672	0.4333	0.2317
	179	-	0.1818	0.5896	0.0333	0.2012
	181	0.5000	0.2898	0.1045	0.4000	0.3236
	183	0.5000	0.0739	0.0821	-	0.1640
	185	-	0.0170	0.0448	-	0.0155
	187	-	-	0.0821*	-	0.0205
	189	-	-	0.0224*	-	0.0056
	191	-	-	0.0075*	-	0.0019
M257	107	-	-	0.0410*	-	0.0103
	111	-	-	0.0492	0.0556	0.0262
	113	-	0.2025	0.2623	-	0.1162
	115	-	-	0.2459*	-	0.0615
	117	-	0.0063	0.1148	-	0.0303
	119	-	0.0886	0.1557	-	0.0611
	121	-	0.1076	0.0574	0.2222	0.0968
	123	-	0.0127*	-	-	0.0032
	125	-	0.3671	-	0.5000	0.2168
	127	-	0.2025	0.0574	0.1111	0.0928
	129	-	-	0.0164*	-	0.0041
	131	1.0000*	-	-	-	0.2500
	135	-	0.0127	-	0.1111	0.0310
M258	97	-	0.0714	-	0.0417	0.0283
	101	0.5000	-	0.0231	-	0.1308
	103	-	0.0060	0.0077	0.0417	0.0139
	105	-	0.1786	0.6077	0.0417	0.2070
	107	0.5000	0.1786	0.3154	0.4167	0.3527
	109	-	0.0119	0.0077	-	0.0049
	111	-	0.3631	0.0385	0.2917	0.1733
	113	-	0.0298	-	0.0417	0.0179
	115	-	0.0060*	-	-	0.0015
	119	-	0.0833*	-	-	0.0208
	123	-	0.0060*	-	-	0.0015
	125	-	0.0655	-	0.0417	0.0268
	131	-	-	-	0.0833*	0.0208
M259	105	-	0.0185	0.5656	-	0.14603
	107	-	0.0864*	-	-	0.0216
	109	0.5000*	-	-	-	0.1250
	111	-	0.0309	0.0082	0.0455	0.0212
	113	-	0.1296	0.4262	-	0.1390
	117	0.5000	0.1852	-	0.3182	0.2509
	119	-	0.0247	-	0.1364	0.0403
	121	-	0.0617	-	0.0909	0.0382
	123	-	0.2593	-	0.0909	0.0876

	125	-	0.1173	-	0.2727	0.0975
	127	-	0.0864	-	0.0455	0.0330
M260	119	-	-	-	0.0417*	0.0104
	123	-	0.7901	0.0273	0.6250	0.3606
	125	0.5000	0.1420	0.7091	0.2917	0.4107
	127	-	0.0123	0.1364	-	0.0372
	129	-	-	0.1091*	-	0.0273
	131	-	-	0.0091*	-	0.0023
	133	0.5000	-	0.0091	-	0.1273
	135	-	0.0556	-	0.0417	0.0243
M746	357	-	0.0135*	-	-	0.0034
	359	0.5000	0.0270	-	-	0.1318
	361	0.5000*	-	-	-	0.1250
	367	-	0.0270*	-	-	0.0068
	369	-	0.2635	-	0.2000	0.1159
	371	-	0.0270	0.1200	0.1000	0.0618
	373	-	0.3851	0.0200	0.7000	0.2763
	375	-	0.2365	0.0300	-	0.0666
	377	-	0.0203	0.0400	-	0.0151
	379	-	-	0.5300*	-	0.1325
	381	-	-	0.1900*	-	0.0475
	383	-	-	0.0600*	-	0.0150
	389	-	-	0.0100*	-	0.0025

5.17 (ranging from 3 alleles at locus M746 to 8 alleles at locus M258) and 6.0 (ranging from 3 genotypes at locus M746 to 10 genotypes at locus M258), respectively, with both populations showing significantly higher diversity for these parameters than the *ex situ* population. The two *in situ* populations and the seedling population had lower mean $\mathbf{H_o}$ value compared to mean $\mathbf{H_e}$ resulting in a positive fixation index ($\mathbf{F_{is}}$) with values of 0.338 for Mandena, 0.098 for Petriky and 0.222 for the QMM seedlings populations (Table 2.2).

For the six polymorphic loci surveyed, there were a total of 68 alleles, ranging from 1 to 11 per locus (Table 2.3). Of these, 26 alleles were unique across populations with three private alleles in the *ex situ* population, eight private alleles in the *in situ* Mandena population, 12 private alleles in the *in situ* Petriky population, and three private alleles in the seedling population from the QMM nursery (Table 2.3). If just the *in situ* populations of Mandena and Petriky are compared, the allelic diversity was higher in the Mandena population with 20 private alleles compared to 15 private alleles in the Petriky population. The Hardy-Weinberg probability tests by populations showed that all four *C. commersoniana* populations significantly deviated from Hardy-Weinberg equilibrium with a significant deficiency of heterozygotes at the 0.001 level of significance for the Mandena, Petriky, and QMM seedling populations (Table 2.4). The *ex situ* A. 302 population showed a significant excess of heterozygotes.

POPULATION GENETIC STRUCTURE

The hierarchical genetic structure of populations was examined through AMOVA with all four populations and with the *ex situ* population and the QMM seedling population excluded. Populations A. 302 and QMM seedlings were excluded to seek a better understanding of the

Table 2.4: Hardy Weinberg probability test for *C. commersoniana*. F_{is} = allelic fixation index for polymorphic loci, P = probability value, SE = standard error, Chi^2 = Chi square value using Fisher's method, and df = degrees of freedom.

Population		Locus**						Chi^2	df	P
		M255	M257	M258	M259	M260	M746			
A. 302	F_{is}	-1.0000	N/A	-1.0000	-1.0000	-1.0000	-1.0000	Infinity	10	HS*
	P	0.0000	N/A	0.0000	0.0000	0.0000	0.0000			
	S.E.	-	N/A	-	-	-	-			
Mandena	F_{is}	0.2864	0.2433	0.1889	0.1900	0.5488	0.7601	Infinity	12	HS*
	P	0.0021	0.0000	0.0011	0.0000	0.0000	0.0000			
	S.E.	0.0013	0.0000	0.0011	0.0000	-	0.0000			
Petriky	F_{is}	0.1421	-0.0283	0.0192	0.1529	0.0723	0.2842	Infinity	12	HS*
	P	0.3623	0.0710	0.0231	0.2321	0.0773	0.0000			
	S.E.	0.0173	0.0078	0.0039	-	0.0103	0.0000			
QMM Seedlings	F_{is}	0.2950	0.8519	-0.1055	0.0055	-0.4043	1.0000	Infinity	12	HS*
	P	0.0139	0.0000	1.0000	0.7293	0.4346	0.0006			
	S.E.	-	0.0000	0.0000	0.0131	-	-			

*HS - Highly significant ($P<0.001$); significantly deviates from Hardy Weinberg equilibrium

**Monomorphic loci are designated as N/A

impact of habitat fragmentation on among-population genetic partitioning among the wild populations. Genetic variation was partitioned into among-population and within-population variation. When all populations (including A. 302 and QMM seedlings) were tested, the within-population variation was 69.89% and among-population variation was 30.11%, which was significant at the 0.001 level of significance (Table 2.5). When A. 302 and QMM seedlings were excluded, the within-population variation was higher at 74.29% and the among-population variation was lower at 25.71%, which was significant at the 0.001 level of significance.

DISCUSSION

GENETIC DIVERSITY PATTERNS

Knowledge of how genetic diversity is maintained through natural processes is one of many criteria that should be used in developing conservation strategies for maintenance of managed populations of endangered species (Frankham et al. 2002). Other criteria include knowledge of taxonomic uncertainties, total distribution (range), population size and density, population fragmentation, population ecology, and threats (Frankham et al. 2002). Gole et al. (2002) identified knowledge of distribution and population genetic structure as major challenges for coffee research in Ethiopia due to the need for skilled personnel and laboratory equipment. To assess genetic variation, a wide variety of molecular marker technologies are available and many of these are increasingly being applied to complement traditional practices of germplasm and genebank management (Spooner et al. 2005).

Table 2.5: AMOVA results examining genetic partitioning between *ex situ* and wild populations for *C. commersoniana*

Source of variation	df	Sum of Squares	Variance components	Percentage of variation
For all four populations:				
Among populations	3	179.996	0.6654	30.11***
Within populations	394	608.602	1.5447	69.89
With *ex situ* population A. 302 and QMM seedlings excluded:				
Among populations	1	87.994	0.5629	25.71***
Within populations	310	504.102	1.6261	74.29

*** $P < 0.001$

All of the 28 plants from the *ex situ* population were genetically identical to each other suggesting that they are either clones of a single genotype or the result of selfing. In comparison, the *in situ* and seedling populations had higher diversity with all six loci exhibiting polymorphism (Table 2.2). The lower genetic diversity of the QMM seedling population compared to the wild populations could be due to the small sample size of only 15 individuals (Table 2.1). Since all of the wild populations were not sampled, there could be additional alleles that are not represented in my samples from Mandena and Pertriky. This is evident in the seedling population, which had three private alleles not exhibited in the wild populations.

Among-population variation was higher at 30% when all four populations were included compared to among-population variation of 26% when only the two *in situ* populations were used (Table 2.6). The higher among-population genetic partitioning when all four populations were used could have been due to the genetic distinctiveness of the *ex situ* population at the KCRS which was collected in 1964 and which had a high frequency of three private alleles. Even though the Mandena and Petriky populations had eight and 12 private alleles respectively, they were at much lower frequencies. Even though the among-population partitioning (25.7%) of the two wild populations at Mandena and Petriky was lower, such a level of differentiation among populations is generally considered large (Balloux and Lugon-Moulin 2002; Wright 1978), suggesting population divergence since habitat fragmentation and, hence, these populations should be kept separate as much as possible.

CONSERVATION AND MANAGEMENT IMPLICATIONS

An updated deforestation study in 2000 indicated a loss of 60% of the littoral forests over the previous 50 years, of which Mandena has seen 74% of the loss (Vincelette et al. 2003). QMM established a conservation and rehabilitation program in 1990 in an effort to prepare for post-mining rehabilitation (Vincelette et al. 2003). As part of these rehabilitation efforts, a nursery site was developed in Mandena for plant propagation with special care given to finding reliable sources of seeds, to address criteria such as high assured percentage of seed germination, genetically heterogeneous seed trees, high phytosanitary standards, and varied commercial sources for each species (Rarivoson and Mara, 2007). None of the rehabilitation and restoration efforts address the assessment of genetic diversity and the need to ensure that restoration efforts take into consideration the use of diverse genetic material. The present study is the first to examine genetic diversity of one of the impacted littoral forest species endemic to the region, *C. commersoniana*, which can serve as a model for other plant species. By assessing the genetic diversity of *C. commersoniana* seedlings propagated at the QMM Mandena nursery, this study provides insight into the level of genetic diversity that will be represented during restoration.

The plant collection maintained at the *ex situ* gene bank has essentially no genetic diversity because all individuals are identical to each other for the markers employed. Future work should include expansion of this collection to include plant material collected from wild populations at Mandena, Petriky, and Sainte Luce. A major concern about holding *C. commersoniana* germplasm at the FOFIFA KCRS is the lack of compatibility of habitat and growing conditions for this species. *Coffea commersoniana* is adapted to rather seasonal humid forests with unconsolidated sandy soils at sea level, whereas the climatic conditions of the Kianjavato region are humid evergreen forest with primarily basement lavas (91% igneous and

metamorphic) and quartzites with red lateritic soils dominant throughout (Moat and Smith 2007), at elevations ranging from 56 – 151 meters above sea level at the KCRS. Though it is important to continue maintaining this germplasm in the current Kianjavato *ex situ* genebank, it would clearly be beneficial to institute another field genebank in the Tolagnaro region so that diverse genetic material of *C. commersoniana* (and other species) is conserved in conditions closer to those of its natural habitat.

When the wild populations of Mandena and Petriky are compared, they both display considerable diversity comparable to genetic diversity studies of cultivated and wild populations of *C. canephora* from six different regions of Uganda (Musoli et al. 2009). For this study, I was not able to sample *C. commersoniana* from the littoral forests of Sainte Luce. Future research should include quantifying the genetic diversity of this population as well. The seedlings propagated in the QMM Mandena nursery showed moderately high genetic diversity considering that only 15 seedlings were sampled. Since information about the source of seeds used in propagation was not collected, the allele frequency data suggests that seeds were possibly collected from Mandena and Petriky locations and possibly Sainte Luce and grown as an admixed population. The high genetic differentiation of the Mandena and Petriky populations is an important aspect to take into consideration during restoration, which would necessitate keeping these populations separate and propagating the seeds from each of these populations separately without admixture. Thus when post-mining restoration is performed, the Mandena seedlings should be used to rehabilitate the Mandena area while the Petriky seedlings should be used to populate the Petriky area. Admixture of populations in an *ex situ* gene bank should not be a problem as long as proper provenance information for each plant in the collection is recorded.

In addition to conservation of genetic diversity *in situ* and *ex situ*, other complementary conservation strategies should also be implemented. In a study performed on the Critically Endangered (IUCN 2001) *Eligmocarpus cynometroides* (Fabaceae), a tree species restricted to the Petriky region, that evaluated the biology, ecology, and risk of extinction of this species, the following conservation actions were recommended: 1) reduction of threats to existing populations through promoting public awareness and direct control of threats, 2) reinforcement of existing populations, 3) scouting for locations of new populations, 4) creation of new populations (translocations), and 5) *ex situ* conservation in scientific collections and seed banks (Randriatafika et al. 2007).

Significant efforts in biodiversity conservation by QMM through assessment, inventory and development of restoration and conservation strategies over the past two decades have been substantial in understanding the ecology and diversity of the littoral forests of southeastern Madagascar. The present study using one of the endangered species of the littoral forests, *C. commersoniana*, which is also a distant wild relative of a cultivated crop species, has provided much needed information on the genetic variability of wild populations on which conservation management priorities can be more satisfactorily determined.

CHAPTER THREE

SUMMARY AND CONCLUSIONS

Coffee is one of the most economically important crops and is produced in 80 tropical countries and with an annual production of nearly seven million tons of green beans (Musoli et al. 2009). It is the second most heavily traded commodity after oil, with over 75 million people in developing countries depending on coffee for their livelihoods (Pendergrast 2009; Vega et al. 2003). There are two species involved in commercial production, Arabica (*Coffea arabica*), and robusta (*C. canephora*), with higher beverage quality associated with *C. arabica*, which accounts for about 70% of world coffee production (Lashermes et al. 1999). In an attempt to quantify the economic value of *C. arabica* genetic resources in Ethiopian highland forests, Hein and Gatzweiler (2006) conducted a valuation based on an assessment of the potential benefits and costs of the use of *C. arabica* genetic information in breeding programs for developing enhanced coffee cultivars with increased pest and disease resistance, low caffeine contents, and increased yields. The resulting valuation of coffee genetic resources based on comparing costs and benefits for a 30-year discounting period was US $1,458 million and $420 million at discount rates of 5% and 10%, respectively, (taking into account cost reduction and/or productivity enhancement with a delay of 5 and 10 years, respectively, before the benefits are reaped, post-breeding), demonstrating the potential economic importance of Ethiopian coffee genetic resources (Hein and Gatzweiler 2006). Other than a few other coffee species, assessments of the importance of the many other wild species of coffee have not been made. All species, along with old and traditional coffee varieties grown in cultivation, represent the ultimate source of coffee genetic diversity, upon which future crop improvement depends (Engelmann and Dulloo 2007).

Erosion of the *Coffea* genepools has become a significant concern due to the many threats to natural habitats, such as deforestation, encroachment by agricultural activities, population pressures, and economic hardships of the local people that depend on these forests (Engelmann and Dulloo 2007). For *Coffea*, research on the *in situ* conservation of genetic resources has lagged behind compared to developing methods for *ex situ* conservation (Engelmann & Dulloo 2007). Conservation efforts of *Coffea* germplasm in its natural habitats have been very limited with known examples only in Ethiopia and in Mauritius (as cited in Engelmann and Dulloo 2007). Many areas such as the central African region in Gabon and the Central African Republic still remain unexplored and much work still needs to be done in *Coffea* diversity hotspots in Madagascar and mainland Africa, particularly Tanzania (Engelmann and Dulloo 2007). Hence conservation of these valuable genetic resources in *ex situ* collections such as field genebanks becomes imperative as an alternative or back-up strategy for *in situ* conservation measures.

A large amount of coffee genetic diversity has been collected and introduced into field genebanks, though these genetic resources are quickly eroding due to issues such as adaptability problems, vandalism, natural catastrophes, and insufficient funds to maintain the collections, among others (Anthony et al. 2007a). Loss of accessions and consequent genetic erosion is rather universal in *Coffea* field genebanks. The International Coffee Germplasm Center operated by Centro Agronómico Tropical de Investigación y Enseñanza (CATIE) in Costa Rica has experienced substantial losses of coffee plants over the years, resulting in loss of entire accessions (Vega et al. 2008). Many of the wild genotype accessions at CATIE are represented by only one or two individuals (Vega et al. 2008). At CATIE genetic erosion is estimated to range between 2% and 8% in various sections of the genebank (Anthony et al. 2007b). To combat this genetic erosion, CATIE has developed new strategies to ensure conservation of

coffee genetic resources (Anthony et al. 2007b). Genetic diversity information was used to define priorities for conservation, giving a higher priority to genetic groups containing high diversity (Anthony et al. 2007b).

In an attempt to understand the genetic diversity of Madagascan coffee species, this study was undertaken using the collections maintained at the Kianjavato Coffee Research Station's (KCRS) *ex situ* field genebank and extant, natural *in situ* populations. As part of this dissertation, four species were studied: *C. kianjavatensis, C. montis-sacri, C. vatovavyensis,* and *C. commersoniana.* Established in 1954, this *ex situ* field genebank, housing a high percentage of the wild Madagascan coffee species in its collections, has lost many individuals per accession over the last 30 to 40 years. Replenishment of lost germplasm has been conducted without knowledge of the genetic identity of the collection. Quantifying the genetic diversity and understanding gene flow in this collection will play an important role in developing future management strategies, which was the main aim of this dissertation.

GENETIC DIVERSITY

Genetic diversity studies were conducted using microsatellite markers. Because of their high variability, microsatellites have become valuable tools in the assessment of genetic structure and diversity within species and have been used extensively in *Coffea* species with successful transferability across species (Combes et al. 2000; Coulibaly et al. 2003. Cubry et al. 2008, Poncet et al. 2004 and 2007). Six microsatellite loci were used in the assessment of genetic diversity of three species narrowly endemic to the Kianjavato region in eastern Madagascar: *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis* and one species endemic to the littoral

forests of southeastern Madagascar, *C. commersoniana*. *Coffea kianjavatensis* is listed as Endangered, *C. montis-sacri* as Critically Endangered, *C. vatovavyensis* as Endangered, and *C. commersoniana* as Endangered by the criteria of the Red List Category system of the World Conservation Union (IUCN 2001) (Davis et al. 2006). All four species are represented in the FOFIFA KCRS's *ex situ* field genebank collection. Using these collections and wild collections from extant *in situ* populations, genetic diversity was assessed using the parameters of mean number of alleles per locus, mean observed and expected heterozygosities, mean allelic fixation index, mean number of observed genotypes, presence of private alleles, deviations from Hardy Weinberg equilibrium, genetic structure within- and among-populations using AMOVA and test for recent bottlenecks.

My assumption was that both the *ex situ* and *in situ* populations would have low genetic diversity. Contrary to my assumption, I found the overall genetic diversity of wild Madagascan coffee species is similar to or even higher than other cultivated and wild coffee species. For the three species endemic to the Kianjavato region, *C. kianjavatensis, C. montis-sacri,* and *C. vatovavyensis*, higher genetic diversity was observed in the *ex situ* populations than in *in situ* populations. Since these *ex situ* collections were made in the early 1960s, it can be presumed that this diversity is indicative of a sampling of what was present at that time and due to collection from multiple origins. In addition, parentage analysis showed that cross pollination between species is occurring and the higher diversity could be due to cross contamination from pollen transfer from another *Coffea* species resulting in hybridization when seedlings are used in replanting lost plant collections. To understand if this high genetic diversity in the *ex situ* field genebank is an original representation from wild collections or due to hybridization caused by cross contamination, a thorough examination of the genebank records needs to be undertaken.

Hand written accession records have been maintained at the KCRS (J. J. Rakotomalala, pers. comm.), which should be thoroughly examined. The lower genetic diversity of the extant *in situ* populations may indicate that genetic diversity is being lost in the wild possibly due to habitat loss and reiterates the importance of preserving the plants currently in the *ex situ* collections representing genetic diversity that may have already been lost in the wild, unless the *ex situ* collections have been compromised due to genetic mixing. The *in situ* populations of all three species have private alleles not represented in the *ex situ* populations and hence a concerted effort of additional collecting missions should be undertaken to enhance the existing *ex situ* collections, especially since the integrity of the collections is questionable. Efforts to secure *in situ* populations should also become a priority.

Wild populations of *C. commersoniana* face imminent threat from mining operations, which threaten the last natural habitat of this species. To counter loss of species in these littoral forest habitats, QIT Madagascar Minerals (QMM) has undertaken environmental impact assessments (EIAs) that has led to the establishment of tree nurseries and plantations, seed banks, and extensive research into reforestation and natural habitat reserves (Bollen and Donati 2006). *Coffea commersoniana* seedlings are currently being propagated at QMM's Mandena nursery for restoration, post-mining. Genetic diversity assessment of the *ex situ* population held at KCRS field genebank showed very low diversity. The two *in situ* populations from Mandena and Petriky and the QMM seedling population exhibit higher genetic diversity with private alleles represented in all populations. The among population variation of 26% between the Mandena and Pertiky populations suggest that for restoration purposes, these populations should be kept separate by separating the seeds for propagation from these populations without admixture with accurate identification in the nursery. The *ex situ* population in the KCRS field genebank should

be enhanced with collections from new collecting missions. In addition, establishment of another field genebank in southwestern Madagascar, in the natural habitat of *C. commersoniana* will be crucial in maintaining this germplasm in conditions similar to their natural habitat. Another possibility would be re-investment in the Manakara Coffee Research Station, which is situated at sea level in conditions similar to the littoral forests of the Tolagnaro region.

PARENTAGE ANALYSIS

The FOFIFA KCRS has lost several plants within the collections since its establishment in the 1950s. Missing plants have been replaced predominantly through seeds collected from each accession (J. J. Rakotomalala, pers. comm.). The replenishment of the lost germplasm has been conducted without knowledge of the genetic diversity of the collection and selection of seed parents at random without knowledge of out-crossing with other species maintained in the collections, which would lead to loss of genetic integrity.

By performing parentage analysis, my objective was to quantify the rate of genetic contamination in the existing *ex situ* genebank using *C. kianjavatensis* and *C. montis-sacri* seedlings by quantifying the rate of outcrossing with other species. For *C. kianjavatensis*, parentage analysis indicated that 5.56% of the offspring were contaminated with pollen from another coffee species, whereas the contamination percent was higher for *C. montis-sacri* at 33.33%.

The present study is the first to examine pollination and gene flow patterns in an *ex situ* *Coffea* field genebank. The results indicate that open pollinated seed propagation in the *Coffea* field genebank is contaminated by pollen from other species of *Coffea*, either from plants growing in the collection or from wild *Coffea* species occurring nearby in wild populations (e.g.

C. vatovavyensis). The results are variable depending on species. Hence, it is of paramount importance that germplasm collection is managed appropriately, so that alleles unique to a particular species present in populations are not lost due to outcrossing with other species.

CONSERVATION IMPLICATIONS

The present study indicates that considerable genetic diversity is present in Madagascan coffee germplasm held at the *ex situ* field genebank and *in situ* populations, although this variation is variable within species. Conservation strategies should be developed for each individual species as well as for the *ex situ* field genebank as a whole. Based on this study, my conservation recommendations are listed below:

- In order to manage the field genebank for efficient space utilization and to ensure optimal representation of genetic diversity, an assessment of genetic diversity of all species held at the genebank should be performed. Based on these genetic studies, an assessment of genetic redundancy should be made so that redundant genotypes can be removed making room for new collections with genetic diversity not currently present in the collection. An example is the case of *C. commersoniana ex situ* A. 302 population, where all 28 plants sampled had the same combination of alleles at the same frequency at all six loci. In this case, rather than have 28 individuals of the same genotype in the collection, the recommendation would be to retain a few healthy plants and remove the rest to make room for augmenting the collection with new acquisitions from collecting missions.
- New collecting missions from wild populations should be undertaken to enhance the genetic diversity of the existing collections and to capture private alleles from wild populations that are currently not present in the *ex situ* populations.

- The collections accessioning system in the field genebank should be evaluated and remedied so that each individual plant has its own accession number with pertinent details about origin, genetic data, etc., to improve record keeping and assist with making plant selections for breeding programs. Passport data for each individual plant in the collection should be maintained. Attempts should be made to collect DNA samples at the point of origin of accessions when collections are made.
- A concerted effort should be undertaken to examine the current germplasm records at the FOFIFA KCRS to inventory what percentage of the original collections made in 1960s and 1970s still remain. This will give a better understanding about the genetic integrity of the collection.
- In order to maintain the genetic integrity of the collections in the *ex situ* genebank, for replacement planting, plants should be propagated either clonally (through cuttings or tissue culture) or through seeds generated by controlled pollination.
- Since Madagascan *Coffea* species come from diverse habitats, duplicating collections in regions of origin of the different species should be given a priority. For example, in addition to maintaining *C. commersoniana* at the KCRS field genebank, a duplicate collection should be maintained at its region of origin in southeastern Madagascar where the growing conditions are similar to its natural habitat.
- *In situ* conservation should become a priority. Conserving plants in their natural habitats allows plants to continue to evolve in a dynamic environment enabling evolution of traits such as pest and disease resistance, adaptation to climate change, and migration to other areas. A concerted effort should be made by local government officials and plant scientists, in collaboration with local peoples that depend on these forest habitats for their

livelihoods, by developing incentives for conservation similar to what has been undertaken in southeastern Madagascar through creation of alternate employment opportunities. *Coffea* species currently protected within existing reserves and protected areas should be identified and for those species not protected, strategies should be developed for their *in situ* conservation.

- Reforestation and restoration of natural habitats should take into account the genetic diversity and partitioning between populations. Decisions regarding the selection of germplasm material to be used for revegetation should be based on this genetic information, in combination with other fundamental conservation data. It is imperative, for example that population integrity is maintained *ex situ* until a restoration program is initiated, although in certain cases an admixture of populations might be beneficial.
- Finally, in addition to conservation *in situ* and in *ex situ* field genebanks, other complementary conservation strategies should be explored and implemented. These include *in vitro* conservation, seed banking for those seeds that are amenable for long-term seed storage, DNA banking, pollen storage, and cryopreservation. Most *Coffea* seeds are recalcitrant, and so far cryopreservation has been the only method for long-term seed banking.

Most of the scientific research undertaken on *Coffea* has focused on the economically important species and cultivars with very limited research on non-commercial wild species (Davis et al. 2006). This dissertation research is the first study to characterize the genetic diversity of Madagascan *Coffea* held at the *ex situ* field genebank and comparing this with extant wild populations. The parentage study is also the first to quantify the extent of cross-species contamination of collections held in this or any other *Coffea* genebank. This study

has fundamental implications for the future of *ex situ* and *in situ* conservation of *Coffea* and provides a framework for future conservation research for Madagascan and other *Coffea* species.

LITERATURE CITED

Amos, W. and A. Balmford. 2001. When does conservation genetics matter? Heredity. 87: 257-265.

Anthony, F., S. Dussert, and E. Dulloo. 2007a. Coffee genetic resources. P-12-22. In: Engelmann, F., M. E. Dulloo, C. Astorga, S. Dussert, and F. Anthony (eds). Conserving coffee genetic resources: complementary strategies for *ex situ* conservation of coffee (*Coffea arabica* L.) genetic resources. A case study in CATIE, Costa Rica. Bioversity International, Rome, Italy.

Anthony, F., C. Astorga, J. Avendaño, and E. Dulloo. 2007b. Conservation of coffee genetic resources in the CATIE field genebank. P-23-34. In: Engelmann, F., M. E. Dulloo, C. Astorga, S. Dussert, and F. Anthony (eds). Conserving coffee genetic resources: complementary strategies for *ex situ* conservation of coffee (*Coffea arabica* L.) genetic resources. A case study in CATIE, Costa Rica. Bioversity International, Rome, Italy.

Anthony F., L. E. C. Diniz, M.-C. Combes, and P. Lashermes. 2010. Adaptive radiation on *Coffea* subgenus *Coffea* L. (Rubiaceae) in Africa and Madagascar. Plant Systematics and Evolution. 285:51-64

Ashley, M. V. 2010. Plant parentage, pollination, and dispersal: how DNA microsatellites have altered the landscape. Critical Reviews in Plant Sciences. 29: 148-161.

Avise, J.C. 2004. Molecular Markers, Natural History and Evolution. Second Edition. Sinauer Associates, Inc. MA, USA.

Balloux, F. and N. Lugon-Moulin. 2002. The estimation of population differentiation with microsatellite markers. Molecular Ecology. 11: 155-165.

Baruah, A., V. Naik, P. S. Hendre, R. Rajkumar, P. Rajendrakumar, and R. K. Aggarwal. 2003. Isolation and characterization of nine microsatellite markers from *Coffea arabica* L., showing wide cross-species amplifications. Molecular Ecology Notes. 3: 647-650.

Bollen, A. and G. Donati. 2006. Conservation status of the littoral forest of south-eastern Madagascar: a review. Oryx. 40(1): 57-66.

Charrier, A. 1978. La structure genetique des cafeiers spontanes de la Region Malagache (*Mascarocoffea*). Memoires ORSTOM No. 87. Paris: ORSTOM.

Chaix, G., S. Gerber, V. Razafimaharo, P. Vigneron, D. Verhaegan, and S. Hamon. 2003. Gene flow estimation with microsatellite in a Malagasy seed orchard of *Eucalyptus grandis*. Theoretical and Applied Genetics. 107: 705-712.

Combes, M. C., S. Andrzejewski, F. Anthony, B. Bertrand, P. Rovelli, G. Graziosi and P. Lashermes. 2000. Characterization of microsatellite loci in *Coffea arabica* and related coffee species. Molecular Ecology. 9: 1178-1180.

Consiglio, T., G. E. Schatz, G. McPherson, P. P. Lowry II, J. Rabenantoandro, Z. S. Rogers, R. Rabevohitra, and D. Rabehevitra. 2005. Deforestation and plant diversity of Madagascar's littoral forests. Conservation Biology. 20(6): 1799-1803.

Coulibaly, I., B. Revol, M. Noirot, V. Poncet, M. Lorieux, C. Carasco-Lacombe, J. Minier, M. Dufour, and P. Hamon. 2003. AFLP and SSR polymorphism in a *Coffea* interspecific backcross progeny [(*C. heterocalyx* x *C. canephora*) x *C. canephora*]. Theoretical and Applied Genetics. 107: 1148-1155.

Cros, J., M.C. Combes, P. Trouslot, F. Anthony, S. Hamon, A. Charrier, and P. Lashermes. 1998. Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. Molecular Phylogenetics and Evolution. 9(1):109-17.

Cubry, P., P. Musoli, H. Legnate, D. Pot, F. de Bellis, V. Poncet, F. Anthony, M. Dufour, and T. Leroy. 2008. Diversity in coffee assessed with SSR markers: structure of the genus *Coffea* and perspectives for breeding. Genome. 51: 50-63.

Davis, A. P. and D. M. Bridson. 2003. Introduction to the Rubiaceae. P-431-434. In: Goodman, S. M. and J. P. Benstead, (eds.). The Natural History of Madagascar. The University of Chicago Press, Chicago.

Davis, A. P., D. Bridson, and F. Rakotonasolo. 2005. A reexamination of *Coffea* subgenus *Baracoffea* and comments on the morphology and classification of *Coffea* and *Psilanthus* (Rubiaceae-Coffeeae). P-398-420. In: Keating, R. C., V. C. Hollowell, and T. Croat (eds). Festschrift for William G. D'Arcy: the legacy of a taxonomist (Monograph in Systematic Botany 104). MBG Press, St. Louis.

Davis, A. P., R. Govaerts, D. M. Bridson and P. Stoffelen. 2006. An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). Botanical Journal of the Linnean Society. 152: 465-512.

Davis, A. P. and F. Rakotonasolo. 2008. A taxonomic revision of the baracoffea alliance: nine remarkable *Coffea* species from western Madagascar. Botanical Journal of the Linnean Society. 158:355-390.

Davis, A. P., R. Govaerts, D. M. Bridson, M. Ruhsam, J. Moat, and N. A. Brummitt. 2009. A global assessment of distribution, diversity, endemism, and taxonomic effort in the Rubiaceae. Annals of the Missouri Botanical Garden. 96(1): 68-78.

Davis, A. P. 2010. Six species of *Psilanthus* transferred to *Coffea* (Coffeeae, Rubiaceae). Phytotaxa. 10: 41-45.

Davis, A. P., F. Rakotonasolo, and P. De Block. 2010. *Coffea toshii* sp. nov. (Rubiaceae) from Madagascar. Nordic Journal of Botany. 28: 134-136.

Davis, A. P. 2011 (In Press). *Psilanthus mannii* Hook.f., the type species of *Psilanthus* Hook.f., transferred to *Coffea* L. Nordic Journal of Botany.

Dawson, I. K., R. Waugh, A. J. Simons, and W. Powell. 1997. Simple sequence repeats provide a direct estimate of pollen-mediated gene dispersal in a tropical tree *Gliricidia sepium.* Molecular Ecology. 6: 179-183.

de Gouvenain, R. C. and J. A. Silander Jr. 2003. Littoral Forest. P-103-111. In: Goodman, S. M. and J. P. Benstead, (eds.). The Natural History of Madagascar. The University of Chicago Press, Chicago.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via the *EM* algorithm (with discussion). Journal of the Royal Statistical Society. 39: 1-38.

Dewoody, J., J. D. Nason, and V. D. Hipkins. 2006. Mitigating scoring errors in microsatellite data from wild populations. Molecular Ecology Notes. 6: 951-957.

Dow, B. D. and M. V. Ashley. 1998. High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. Journal of Heredity. 89: 62-70.

Duchesne, P., T. Castric, and L. Bernatchez. 2005. PASOS (parental allocation of singles in open systems): a computer program for individual parental allocation with missing parents. Molecular Ecology Notes. 5: 701-704.

Duchense, P., T. Meldgaard, and P. Berrebi. 2008. Parentage analysis with few contributing breeders: validation and improvement. Journal of Heredity. 99:323-334.

Dulloo, M. E., L. Guarino, F. Engelmann, N. Maxted, J. H. Newbury, F. Attere, and B. V. Ford-Lloyd. 1998. Complementary conservation strategies for the genus *Coffea*: A case study of Mascarene *Coffea* species. Genetic Resources and Crop Evolution. 45:565-79.

Engelmann, F. and E. Dulloo. Introduction. 2007. P-1-11. In: Engelmann, F., M. E. Dulloo, C. Astorga, S. Dussert, and F. Anthony (eds). Conserving coffee genetic resources: complementary strategies for *ex situ* conservation of coffee (*Coffea arabica* L.) genetic resources. A case study in CATIE, Costa Rica. Bioversity International, Rome, Italy.

Etisham-Ul-Haq, M., T. R. Allnutt, C. Smith-Ramirez, M. F. Gardner, J. J. Armesto, and A. C. Newton. 2001. Patterns of genetic variation in *in situ* and *ex situ* populations of the threatened Chilean vine *Berberidopsis coralline*, detected using RAPD markers. Annals of Botany. 87: 813-821.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 131: 479-491.

Frankel, O. H. 1984. Genetic perspectives of germplasm conservation. P-161-170. In: Arber, W., K. Llimensee, W. J. Peacock, and P. Starlinger (eds). Genetic Manipulation: Impact on Man and Society. Cambridge University Press, Cambridge.

Frankham, R., J. D. Ballou, and D. A. Briscoe. 2002. Introduction to Conservation Genetics. Cambridge Press. Cambridge, UK.

Gerber, S., S. Mariette, R. Streiff, C. Bodénès, and A. Kremer. 2000. Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. Molecular Ecology. 9: 1037-1048.

Gole, T.W., M. Denich, D. Teketay, and P. L. G. Vlek. 2002. Human impacts on *Coffea arabica* genepool in Ethiopia and the need for its *in situ* conservation. P-237-247. In: Engels, J. M. M., V. R. Rao, A. H. D. Brown and M. T. Jackson (eds). Managing Plant Genetic Diversity. CABI Publishing, New York.

Hein, L. and F. Gatzweiler. 2005. The economic value of coffee (*Coffea arabica*) genetic resources. Ecological Economics. 60: 176-185.

Hendre, P.S., R. Phanindranath, V. Annapurna, A. Lalremruata, and R. K. Aggarwal. 2008. Development of new genomic microsatellite markers from robusta coffee (*Coffea canephora* Pierre ex A. Froehner) showing broad cross-species transferability and utility in genetic studies. BMC Plant Biology. 8: 51-69.

IUCN. 2001. IUCN red list categories, Version 3.1. Gland/Cambridge: IUCN Species Survival Commission.

Jones, A. G., C. M. Small, K. A. Paczolt, and N. L. Ratterman. 2010. A practical guide to methods of parentage analysis. Molecular Ecology Resources. 10:6-30.

Lashermes, P., M.-C. Combes, J. Robert, P. Trouslot, A. D'Hont, F. Anthony, and A. Charrier. 1999. Molecular characterization and origin of the *Coffea arabica* L. genome. Molecular and General Genetics. 261:259-66.

Lashermes, P., S. Andrzejewski, B. Bertrand, M. C. Combes, S. Dussert, G. Graziosi, P. Trouslot, and F. Anthony. 2000. Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). Theoretical and Applied Genetics. 100: 130-146.

Lian, C., M. Miwa, and T. Hogetsu. 2001. Outcrossing and paternity analysis of *Pinus densiflora* (Japanese red pine) by microsatellite polymorphism. Heredity. 87: 88-98.

Lowry II, P. P., F. Randriatafika, and J. Rabenantoandro. 2008. Conservation status of vascular plant species from the QMM/Rio Tinto mining area at Mandena, Tolagnaro (Fort Dauphin) region, southeast Madagascar. Madagascar Conservation & Development. 3(1): 55-63.

Luikart, G. and J-M. Cornuet. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology. 12(1): 228-237.

Maluf, M. P., M. Silvestrini, L. M. C. Riggiero, O. G. Filho, and C. A. Colombo. 2005. Genetic diversity of cultivated *Coffea arabica* inbred lines assessed by RAPD, AFLP, and SSR marker systems. Scientia Agricola. 62(4): 366-373.

Maunder, M., C. Hughes, J. A. Hawkins, and A. Culham. 2003. Hybridization in *ex situ* plant collections: conservation concerns, liabilities, and opportunities. P-325-364. In: Guerrant, E. O., K. Havens, and M. Maunder (eds). *Ex Situ* Plant Conservation: Supporting Species Survival in the Wild. Island Press, Washington DC, USA.

Maurin, O., A. P. Davis, M. Chester, E. F. Mvungi, Y. Jaufeerally-Fakim, and M. F. Fay. 2007. Towards a phylogeny for *Coffea* (Rubiaceae): Identifying well-supported lineages based on nuclear and plastid DNA sequences. Annals of Botany. 100: 1565-1583.

Meffe, G.K. and C.R. Carroll. 1994. Principles of Conservation Biology. Sinauer Associates, Inc., Sunderland, MA.

Moat, J. and P. Smith. 2007. Atlas of the Vegetation of Madagascar. Kew Publishing, Royal Botanic Gardens, Kew.

Moncada, P. and S. McCouch. 2004. Simple sequence repeat diversity in diploid and tetraploid *Coffea* species. Genome. 47: 510-509.

Musloi, P., P. Cubry, P. Aluka, C. Billot, M. Dufour, T. De Bellis, D. Pot, D. Bieysse, A. Charrier, and T. Leroy. 2009. Genetic differentiation of wild and cultivated populations: diversity of *Coffea canephora* Pierre in Uganda. Genome. 52: 634-646.

Ouborg N.J., Y. Piquot, and J.M. Van Groenendael. 1999. Population genetics, molecular markers and the study of dispersal in plants. Journal of Ecology. 87: 551-568.

Oosterhout, C.V., W. F. Hutchinson, D. P. M. Willis, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes. 4: 535-538.

Pearl, H. M., C. Nagai, P. H. Moore, D. L. Steiger, R. V. Osgood, and R. Ming. 2004. Construction of a genetic map for arabica coffee. Theoretical and Applied Genetics. 108:829-35.

Pendergrast, M., 2009. Coffee second only to oil? Tea & Coffee Trade Journal. April:38-41.

Piry, S., G. Luikart, and J-M. Cornuet. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. The Journal of Heredity. 90(4): 502-503.

Poncet, V., P. Hamon, J. Minier, C. Carasco, S. Hamon, and M. Noirot. 2004. SSR cross-amplification and variation within coffee trees (*Coffea* spp.). Genome. 47: 1071-1081.

Poncet, V., M. Dufour, P. Hamon, S. Hamon, A. de Kochko, and T. Leroy. 2007. Development of genomic microsatellite markers in *Coffea canephora* and their transferability to other coffee species. Genome. 50: 1156-1161.

Prakash, N. S., D. V. Marques, V. M. P. Varzea, M. C. Silva, M. C. Combes, and P. Lashermes. 2005. Analysis of genetic diversity in Indian robusta genepool (*Coffea canephora*) in

comparison with a representative core collection using SSRs and AFLPs. Genetic Resources and Crop Evolution. 52:333-343.

Rabenantoandro, J., F. Randriatafika, and P. P. Lowry II. 2007. Floristic and structural characteristics of remnant littoral forest sites in the Tolagnaro area. P-65-93. In: Ganzhorn, J. U., S. M. Goodman, and M. Vincelette (eds.). Biodiversity, Ecology and Conservation of Littoral Ecosystems in Southeastern Madagascar, Tolagnoro (Fort Dauphin). Series editor Alfonso Alonso. SI/MAB Series #11. Smithsonian Institution, Washington DC, USA.

Rakotosamimanana, B. 2003. Foreword. In: Goodman, S. M. and J. P. Benstead, (eds.). The Natural History of Madagascar. The University of Chicago Press, Chicago.

Randriatafika, F., J. Rabenantoandro, C. Birkinshaw, and M. Vincelette. 2007. Biology, ecology, risk of extinction, and conservation strategy for *Eligmocarpus cynometroides* (Fabaceae): a priority species at Petriky. P-369-377. In: Ganzhorn, J. U., S. M. Goodman, and M. Vincelette (eds.). Biodiversity, Ecology and Conservation of Littoral Ecosystems in Southeastern Madagascar, Tolagnoro (Fort Dauphin). Series editor Alfonso Alonso. SI/MAB Series #11. Smithsonian Institution, Washington DC, USA.

Rarivoson, C., and R. Mara. 2007. The Mandena tree nursery, and example of plant production adapted for site rehabilitation after mining. P-317-321. In: Ganzhorn, J. U., S. M. Goodman, and M. Vincelette (eds.). Biodiversity, Ecology and Conservation of Littoral Ecosystems in Southeastern Madagascar, Tolagnoro (Fort Dauphin). Series editor Alfonso Alonso. SI/MAB Series #11. Smithsonian Institution, Washington DC, USA.

Rousset, F. 2008. GENEPOP'007: a complete reimplementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources. 8: 103-106

Rovelli, P., R. Mettulio, F. Anthony, F. Anzueto, P. Lashermes, and G. Graziosi. 2000. Microsatellites in *Coffea arabica* L. P-123-133. In: Sera, T., C. R. Soccol, A. Pandey, and S. Roussos (eds). Coffee Biotechnology and Quality. Kluwer Academic Publishers, Netherlands.

Scarascia-Mugnozza, G.T. and P. Perrino. 2002. The history of *ex situ* conservation and use of plant genetic resources. In: Engels, J.M.M., V.R. Rao, A.H.D. Brown and M.T. Jackson (eds). Managing Plant Genetic Diversity. CABI Publishing.

Schlötterer, C. and J. Pemberton. 1998. The use of microsatellites for genetic analysis of natural populations – a critical review. P-71-86. In: DeSalle, R. and B. Schierwater (eds.). Molecular Approaches to Ecology and Evolution. Birkhauser Verlag, Basel, Switzerland.

Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin: A software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Schuster, W. and J. B. Mitton. 2000. Paternity and gene dispersal in limber pine (*Pinus flexilis* James). Heredity. 84: 348-361.

Shapcott, A., M. Rakotoarinivo, R. J. Smith, G. Lysakova, M. F. Fay, and J. Dransfield. 2007. Can we bring Madagascar’s critically endangered palms back from the brink? Genetics, ecology and conservation of the critically endangered palm *Beccariophoenix madagascariensis*. Botanical Journal of the Linnean Society. 154:589-608.

Silvestrini, M., M. G. Junqueira, A. C. Favarin, O. Guerreiro-Filho, M. P. Maluf, M. B. Silvarolla, and C. A. Colombo. 2007. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellite markers. Genetic Resources and Crop Evolution. 54: 1367-1379.

Spooner, D., R. van Treuren, and M. C. de Vicente. 2005. Molecular markers for genebank management. IPGRI Technical Bulletin No. 10.

Suso, M.J., S. Gilsanz, G. Duc, P. Marget and M.T. Moreno. 2006. Germplasm management of faba bean (*Vicia faba* L.): Monitoring intercrossing between accessions with inter-plot barriers. Genetic Resources and Crop Evolution. 53: 1427-1439.

Swaminathan, M. S. 2000. Government-industry-civil society partnerships in integrated gene management. Current Science. 78(5): 555-562.

Van Hintum, T. J. L., A. H. D. Brown, C. Spillane, and T. Hodgkin. 2000. Core collections of plant genetic resources. IPGRI Technical Bulletin. 3: 5-48.

Vega, F., E. Rosenquist, and W. Collins. 2003. Global project needed to tackle coffee crisis. Nature. 425: 343.

Vega, F. E., A. W. Ebert, and R. Ming. 2008. Coffee germplasm resources, genomics, and breeding. P-415-447. In: Plant Breeding Reviews, Volume 30. Janick, J. (ed). John Wiley & Sons, Inc.

Vincelette, M., L. Randrihasipara, J.-B. Ramanamanjato, P. P. Lowry II, and J. U. Ganzhorn. 2003. Mining and environmental conservation: The case of QIT Madagascar Minerals in the southeast. P-1535-1537. In: Goodman, S. M. and J. P. Benstead, (eds.). The Natural History of Madagascar. The University of Chicago Press, Chicago.

Vincelette, M., L. Dean, and J. U. Ganzhorn. 2007a. The QMM/Rio Tinto project history in Tolagnaro and its social and environmental concepts. P-1-8. In: Ganzhorn, J. U., S. M. Goodman, and M. Vincelette (eds.). Biodiversity, Ecology and Conservation of Littoral Ecosystems in Southeastern Madagascar, Tolagnoro (Fort Dauphin). Series editor Alfonso Alonso. SI/MAB Series #11. Smithsonian Institution, Washington DC, USA.

Vincelette, M., J. Dumouchel, J. Giroux, and R. Heriarive. 2007b. The Tolagnaro (Fort Dauphin) region: a brief overview of the geology, hydrology, and climatology. P-9-17. In: Ganzhorn, J. U., S. M. Goodman, and M. Vincelette (eds.). Biodiversity, Ecology and Conservation of Littoral Ecosystems in Southeastern Madagascar, Tolagnoro (Fort Dauphin). Series editor Alfonso Alonso. SI/MAB Series #11. Smithsonian Institution, Washington DC, USA.

Weir, B. S. and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution. 38(6): 1358-1370.

Wise, C. A., T. A. Ranker, and Y. B. Linhart. 2002. Modeling problems in conservation genetics with *Brassica rapa*: genetic variation and fitness in plants under mild, stable conditions. Conservation Biology. 16(6): 1542-1554.

Wright, S. 1978. Evolution and the Genetics of Population, Variability Within and Among Natural Populations. The University of Chicago Press, Chicago.

Ye, Q. G., X-H Yao, S-J Zhang, M Kang, and H-W Huang. 2006. Potential risk of hybridization in *ex situ* collections of two endangered species of *Sinojackia* Hu (Styracaceae). Journal of Integrative Plant Biology. 48(7): 867-872.

Yoder, A. D. & M. D. Nowak. 2006. Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. Annual Review of Ecology, Evolution and Systematics. 37: 405-431.

Zhang, D., S. Mischke, E. S. Johnson, W. Phillips-Mora, and L. Meinhardt. 2009. Molecular characterization of an international cacao collection using microsatellite markers. Tree Genetics and Genomes. 5: 1- 10.

Zhang, J-J., Q-G. Ye, X-H Yao, and H-W Huang. 2010. Spontaneous interspecific hybridization and patterns of pollen dispersal in ex situ populations of a tree species (*Sinojackia xylocarpa*) that is extinct in the wild. Conservation Biology. 24(1): 246-255.

APPENDICIES

Appendix 1: List of herbarium collections submitted to Royal Botanic Gardens, Kew (K), University of Colorado Museum (COLO), Parc Botanique et Zoologique de Tsimbazaza (TAN) and the FOFIFA: National Center of Applied Research and Rural Development (TEF).

SK 1 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 12, 2007
GPS coordinates: 21˚22'25S / 47˚51'57E
Altitude: 127 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was 2 meters high with a spread of 1 meter, with many spreading branches. Fruits are present. FOFIFA Kianjavato Coffee Research Station Accession Number A.213.

SK 23 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 12, 2007
GPS coordinates: 21˚22'30S / 47˚51'56E
Altitude: 136 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was 3 meters high with a spread of 2 meters. Fruits were observed on a few branches. FOFIFA Kianjavato Coffee Research Station Accession Number A.213.

SK 24 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 12, 2007
GPS coordinates: 21˚22'30S / 47˚51'56E
Altitude: 136 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was 2 meters high with a spread of about 1.5 meters. Fruits were observed on most branches. This is a grafted plant with *C. kianjavatensis* scion grafted on *C. perrieri*. FOFIFA Kianjavato Coffee Research Station Accession Number A.213/A.421.

SK 44 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 12, 2007
GPS coordinates: 21˚22'29S / 47˚51'56E
Altitude: 140 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was 2 meters high with a spread of 2 meters. It had a good spreading habit. Fruits were present. This is a grafted plant with *C. kianjavatensis* scion grafted on *C. perrieri*. FOFIFA Kianjavato Coffee Research Station Accession Number A.213/A305.

SK 101 - *Coffea montis-sacri* A.P.Davis
Date of collection: Dec. 12, 2007
GPS coordinates: 21˚22'29S / 47˚51'57E
Altitude: 152 meters
Habitat: Cultivated at Centre de Recherche FOFIFA

The plant was about 3.5 meters high with a spread of 2 meters. Fruits were present. FOFIFA Kianjavato Coffee Research Station Accession Number A.321.

SK 102 - *Coffea montis-sacri* A.P.Davis
Date of collection: Dec. 12, 2007
GPS coordinates: 21˚22'28S / 47˚51'57E
Altitude: 151 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was about 5.5 - 6.0 meters high with a spread of about 1.0 - 1.5 meters. Fruits were present. This is a grafted plant with *C. montis-sacri* scion grafted on *C. liberica* (=*C. excelsa*) rootstock. FOFIFA Kianjavato Coffee Research Station Accession Number A.321/EX.

SK 109 - *Coffea montis-sacri* A.P.Davis
Date of collection: Dec. 12, 2007
GPS coordinates: 21˚22'28S / 47˚52'01E
Altitude: 133 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was about 3.0 - 3.5 meters high with a spread of about 1.0 meter. Fruits were present. FOFIFA Kianjavato Coffee Research Station Accession Number A.321.

SK 113 - *Coffea vatovavyensis* J.-F.Leroy
Date of collection: Dec. 13, 2007
GPS coordinates: 21˚22'28S / 47˚52'00E
Altitude: 135 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was 3 meters high with a spread of 2 meters. Small-leaved phenotype. Many leaved and branched. Sparse fruiting. FOFIFA Kianjavato Coffee Research Station Accession Number A.308.

SK 131 - *Coffea vatovavyensis* J.-F.Leroy
Date of collection: Dec. 13, 2007
GPS coordinates: 21˚22'28S / 47˚51'59E
Altitude: 142 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was about 4.5 meters high with a spread of 2.5 - 3.0 meters. Large-leaved phenotype. Good fruiting with some fruits ripe. This is a grafted plant with *C. vatovavyensis* scion grafted on *C. perrieri* rootstock. FOFIFA Kianjavato Coffee Research Station Accession Number A.954/A.12.

SK 138 - *Coffea vatovavyensis* J.-F.Leroy
Date of collection: Dec. 13, 2007
GPS coordinates: 21˚22'30S / 47˚52'04E
Altitude: 136 meters
Habitat: Cultivated at Centre de Recherche FOFIFA

The plant had multiple branches with many leaves. Small-leaved phenotype. Fruits were present. This is a grafted plant with *C. vatovavyensis* scion grafted on hybrid of *C. canephora* x *C. congensis* rootstock. FOFIFA Kianjavato Coffee Research Station Accession Number A.830/HB.

SK 173 - *Coffea vatovavyensis* J.-F.Leroy
Date of collection: Dec. 13, 2007
GPS coordinates: 21˚22'31S / 47˚52'02E
Altitude: 127 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was about 2.5 - 3.0 meters high with a spread of 2.0 - 2.5 meters. Small-leaved phenotype. Good branching and leafing. Sooty mold was observed on leaves. This is a grafted plant with *C. vatovavyensis* scion grafted on hybrid of *C. canephora* B x *C. congensis* rootstock. FOFIFA Kianjavato Coffee Research Station Accession Number A.1009/HB.

SK 175 - *Coffea montis-sacri* A.P.Davis
Date of collection: Dec. 13, 2007
GPS coordinates: 21˚22'37S / 47˚52'02E
Altitude: 100 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was about 4.5 - 5.5 meters high with a spread of 1.0 meter. Fruits were present. FOFIFA Kianjavato Coffee Research Station Accession Number A.321.

SK 182 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 13, 2007
GPS coordinates: 21˚22'28S / 47˚51'58E
Altitude: 115 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was 4.5 meters high with a spread of 2.5 - 3.0 meters. Fruits were present. FOFIFA Kianjavato Coffee Research Station Accession Number A.213.

SK 204 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 14, 2007
GPS coordinates: 21˚24'36S / 47˚56'32E
Altitude: 445 meters
Habitat: Humid evergreen forest (Mt. Vatovavy). Flat habitat. Associated spp. - *Chassalia* (Rubiaceae), *Garcinia verucosa* (Clusiaceae), *Oncostemum* (Myrsinaceae), *Dracaena* (Ruscaceae), *Pandanus* (Pandanaceae), *Dypsis* (Arecaceae), *Diporidium* (Ochnaceae), *Polysphaeria* (Rubiaceae)
The plant was about 1.8 meters high with a spread of 1.5 meters. Five branches. Sparse leaves. Apical (new) leaves brownish. No fruits.

SK 210 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 14, 2007
GPS coordinates: 21˚24'36S / 47˚56'32E
Altitude: 445 meters

Habitat: Humid evergreen forest (Mt. Vatovavy). Flat habitat. Associated spp. - *Chassalia* (Rubiaceae), *Garcinia verucosa* (Clusiaceae), *Oncostemum* (Myrsinaceae), *Dracaena* (Ruscaceae), *Pandanus* (Pandanaceae), *Dypsis* (Arecaceae), *Diporidium* (Ochnaceae), *Polysphaeria* (Rubiaceae)
The plant was about 6.0 meters high with a spread of 1.5 meters. No fruits. Parent from which seeds were collected for the A.213 accession at the FOFIFA Kianjavato Coffee Research Station.

SK 241 - *Coffea montis-sacri* A.P.Davis
Date of collection: Dec. 14, 2007
GPS coordinates: 21˚24'37S / 47˚56'35E
Altitude: 450 meters
Habitat: Humid evergreen forest (Mt. Vatovavy). Slopey habitat, drier. Associated spp. - Bamboo & other Gramineae predominent, *Clidemia hirta* (Melastomataceae), *Dianella* (Liliaceae), *Pandanus* (Pandanaceae), *Dypsis* (Arecaceae)
The tree was on a slope and a distant reach and hence was hard to judge the height and spread of the tree.

SK 247 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 14, 2007
GPS coordinates: 21˚24'36S / 47˚56'34E
Altitude: 455 meters
Habitat: Humid evergreen forest (Mt. Vatovavy). Rocky habitat on a steep slope. Associated spp. - *Chassalia* (Rubiaceae), *Garcinia verucosa* (Clusiaceae), *Oncostemum* (Myrsinaceae), *Dracaena* (Ruscaceae), *Pandanus* (Pandanaceae), *Dypsis* (Arecaceae), *Diporidium* (Ochnaceae), *Polysphaeria* (Rubiaceae)
The plant was about 3.5 - 4.5 meters high with a spread of about 2.0 - 2.5 meters. No fruits.

SK 251 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 14, 2007
GPS coordinates: 21˚24'33S / 47˚56'32E
Altitude: 449 meters
Habitat: Humid evergreen forest (Mt. Vatovavy). Rocky habitat on a steep slope. Associated spp. - *Chassalia* (Rubiaceae), *Garcinia verucosa* (Clusiaceae), *Oncostemum* (Myrsinaceae), *Dracaena* (Ruscaceae), *Pandanus* (Pandanaceae), *Dypsis* (Arecaceae), *Diporidium* (Ochnaceae), *Polysphaeria* (Rubiaceae)
The plant was about 3.5 - 4.5 meters high. No fruits.

SK 267 - *Coffea kianjavatensis* J.-F.Leroy
Date of collection: Dec. 14, 2007
GPS coordinates: 21˚24'28S / 47˚56'32E
Altitude: 447 meters
Habitat: Humid evergreen forest (Mt. Vatovavy). Rocky habitat on a steep slope, a little wetter than the other population locations. Associated spp. - *Chassalia* (Rubiaceae), *Garcinia verucosa* (Clusiaceae), *Oncostemum* (Myrsinaceae), *Dracaena* (Ruscaceae), *Pandanus* (Pandanaceae), *Dypsis* (Arecaceae), *Diporidium* (Ochnaceae), *Polysphaeria* (Rubiaceae)

The plant was about 2.5 meters high with a spread of about 2.0 meters. Main branch broken off. No fruits.

SK 270 – *Coffea commersoniana* (Baill.) A. Chev
Date of collection: Nov. 7, 2008
GPS coordinates: 21˚22'40S/ 47˚52'13E
Altitude: 59 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was 1.0 meter high with a spread of about 0.6 meter. FOFIFA Kianjavato Coffee Research Station Accession Number A.302.

SK 271 – *Coffea commersoniana* (Baill.) A. Chev
Date of collection: Nov 7, 2008
GPS coordinates: 21˚22'40S/ 47˚52'13E
Altitude: 59 meters
Habitat: Cultivated at Centre de Recherche FOFIFA
The plant was about 0.6 meter high with a spread of about 0.6 meter. FOFIFA Kianjavato Coffee Research Station Accession Number A.302.

SK 409 – *Coffea vatovavyensis* J.-F. Leroy
Date of collection: Nov. 7, 2008
GPS coordinates: 21˚22'27S/ 47˚52'07E
Altitude: 178 meters
Habitat: Humid evergreen forest (Forêt Sangasanga). Rocky habitat on a steep slope. Associated spp. - *Dracaena* (Ruscaceae), *Canarium madagascariensis* (Burseraceae), *Ravenala madagascariensis* (Strelitziaceae), *Maranta* (Marantaceae). The plant was about 1.5 meters high with a spread of about 1.0 meter. Flowers were present.

SK 416 - *Coffea vatovavyensis* J.-F. Leroy
Date of collection: Nov. 7, 2008
GPS coordinates: 21˚22'27S/ 47˚52'08E
Altitude: 178 meters
Habitat: Humid evergreen forest (Forêt Sangasanga). Rocky habitat on a steep slope. Associated spp. - *Dracaena* (Ruscaceae), *Canarium madagascariensis* (Burseraceae), *Ravenala madagascariensis* (Strelitziaceae), *Maranta* (Marantaceae). The plant was about 1.5 meters high with a spread of about 1.0 meter. Young fruits were present.

SK 445 - *Coffea commersoniana* (Baill.) A. Chev
Date of collection: Nov. 12, 2008
GPS coordinates: 24˚56'56S/ 46˚59'54E
Altitude: 17 meters
Habitat: Littoral Forest. Sandy soil. Associated vegetation: *Malleastrum mandenense* (Meliaceae), *Psorospermum* sp. (Clusiaceae), *Phymatosorus scolopendria* (Polypodiaceae), *Campylospermum obtusifolium* (Ochnaceae).
The plant was about 4 meters high with a spread of about 1.5 meter.

SK 458 - *Coffea commersoniana* (Baill.) A. Chev
Date of collection: Nov. 12, 2008
GPS coordinates: 24˚56'55S/ 46˚59'55E
Altitude: 17 meters
Habitat: Littoral Forest. Sandy soil. Associated vegetation: *Malleastrum mandenense* (Meliaceae), *Psorospermum* sp. (Clusiaceae), *Phymatosorus scolopendria* (Polypodiaceae), *Campylospermum obtusifolium* (Ochnaceae).
The plant was about 4.5 meters high with a spread of about 0.5 meter. Fruits were present

SK 551 - *Coffea commersoniana* (Baill.) A. Chev
Date of collection: Nov. 13, 2008
GPS coordinates: 25˚02'46S/ 46˚51'43E
Altitude: 12 meters
Habitat: Littoral Forest. Sandy soil. Associated vegetation: *Euphorbia* sp. (Euphorbiaceae), *Polycardia phyllanthoides* (Celastraceae), *Senecio antandroi* (Asteraceae), *Campylospermum obtusifolium* (Ochnaceae), *Diospyros* sp. (Ebenaceae).
The plant was about 2.5 meters high with a spread of about 1.2 meter. Multi-branched with many leaves; orchid growing on the trunk. Flowers and fruits were present.

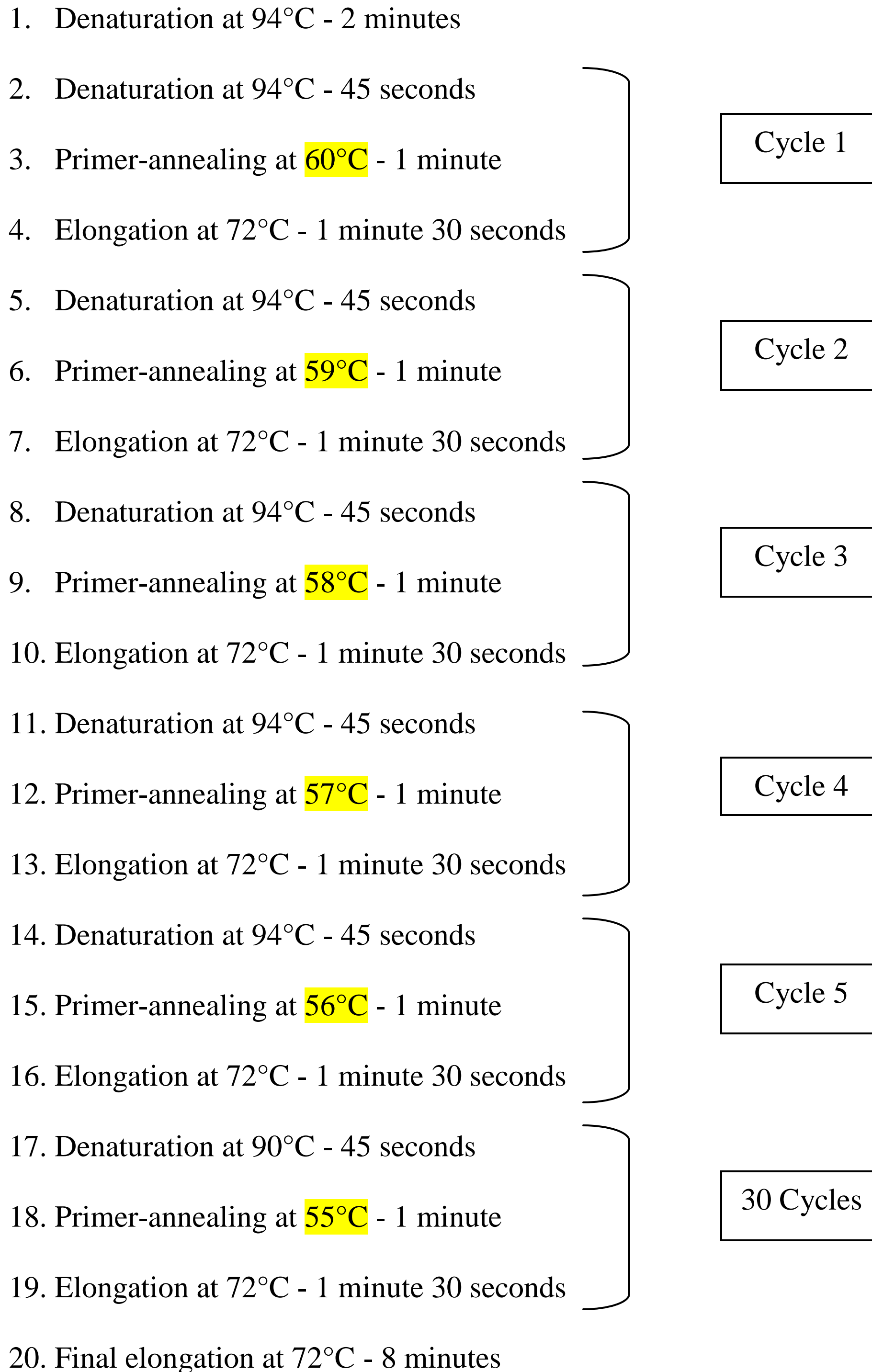
SK 552 - *Coffea commersoniana* (Baill.) A. Chev
Date of collection: Nov, 13, 2008
GPS coordinates: 25˚02'46S/ 46˚51'43E
Altitude: 12 meters
Habitat: Littoral Forest. Sandy soil. Associated vegetation: *Euphorbia* sp. (Euphorbiaceae), *Polycardia phyllanthoides* (Celastraceae), *Senecio antandroi* (Asteraceae), *Campylospermum obtusifolium* (Ochnaceae), *Diospyros* sp. (Ebenaceae).
The plant was about 2.5 meters high with a spread of about 1.2 meter. Multi-branched with many leaves; orchid growing on the trunk. Flowers and fruits were present.

SK 600 - *Coffea commersoniana* (Baill.) A. Chev
Date of collection: Nov. 13, 2008
GPS coordinates: 25˚04'10S/ 46˚51'14E
Altitude: 14 meters
Habitat: Littoral Forest. Sandy soil. Associated vegetation: *Euphorbia* sp. (Euphorbiaceae), *Polycardia phyllanthoides* (Celastraceae), *Senecio antandroi* (Asteraceae), *Campylospermum obtusifolium* (Ochnaceae), *Diospyros* sp. (Ebenaceae).
The plant was about 2.5 meters high with a spread of about 1.0 meter. Fruits were present.

Appendix 2: *Coffea* DNA Extraction Procedure Using *GenCatch*™ Plant Genomic DNA Purification Kit

1. Weigh 10 mg (0.01 g) of silica-dried leaf material and place in FastPrep® tube with Garnet Matrix and 1/4" ceramic bead
2. Place in FastPrep® Cell Disruptor machine and dry run at 4.0 for 20 seconds
3. Add 600µl of **Px1 buffer**
4. Place in FastPrep® machine and run at 4.0 for 20 seconds
5. Incubate at 65°C for 2 hours with occasional finger vortexing
6. Centrifuge at 6,000 for 1 minute
7. Transfer supernatant to 1.5ml tubes
8. Add 200µl of **Px2 buffer**
9. Vortex
10. Incubate in ice for 5 minutes
11. Transfer lysate to shearing tube sitting in a collection tube
12. Centrifuge at 13,000 for 2 minutes
13. Transfer flow-through sample from collection tube to new 1.5ml tube (~650µl)
14. Add 0.5 volume **Px3 buffer** (~325µl) and 1 volume 96 - 100% (use 95%) **Ethanol** (~650µl) to the clear lysate and mix by inverting tube
15. Transfer sample (part of it) from step above to a Plant Genomic Mini Column sitting in a collection tube and close cap
16. Centrifuge at 10,000 for 1 minute
17. Discard filtrate
18. Repeat steps 15 - 17 for the rest of the sample
19. Wash the column twice with 0.7ml (700µl) of **WS Buffer** by centrifuging at 13,000 for 30 seconds
20. Discard filtrate
21. Centrifuge at 13,000 for another 2 minutes to remove traces of WS Buffer
22. Transfer the column to a new collection tube
23. Add 100µl preheated (65°C) ddH2O
24. Incubate at room temperature for 5 minutes
25. Centrifuge at 13,000 for 1 minute to elute DNA
26. Transfer eluted DNA to new 1.5ml tubes
27. Store DNA at -20°C

Appendix 3: *Coffea* PCR Procedure (Reference: Coulibaly et al. 2003)



1. Denaturation at 94°C - 2 minutes
2. Denaturation at 94°C - 45 seconds
3. Primer-annealing at 60°C - 1 minute
4. Elongation at 72°C - 1 minute 30 seconds
5. Denaturation at 94°C - 45 seconds
6. Primer-annealing at 59°C - 1 minute
7. Elongation at 72°C - 1 minute 30 seconds
8. Denaturation at 94°C - 45 seconds
9. Primer-annealing at 58°C - 1 minute
10. Elongation at 72°C - 1 minute 30 seconds
11. Denaturation at 94°C - 45 seconds
12. Primer-annealing at 57°C - 1 minute
13. Elongation at 72°C - 1 minute 30 seconds
14. Denaturation at 94°C - 45 seconds
15. Primer-annealing at 56°C - 1 minute
16. Elongation at 72°C - 1 minute 30 seconds
17. Denaturation at 90°C - 45 seconds
18. Primer-annealing at 55°C - 1 minute
19. Elongation at 72°C - 1 minute 30 seconds
20. Final elongation at 72°C - 8 minutes

Appendix 4: Timetable proposed by QIT Madagascar Minerals (QMM) for conducting their mining activities in the Tolagnaro region of southeastern Madagascar (reproduced from Vincelette et al. 2007a).

Year	Activity
1989-1992	First phase of social and environmental studies: carried out by external consultants on flora, fauna, soil and geology, hydrology, aquatic, and socio-economic issues.
1996-2001	Second phase of studies: establishment of a social, environmental and conservation team within QMM and recruiting consultants and collaborators.
1998	Signature of the "Framework Agreement" with the Malagasy Government.
1998-2001	Elaboration of the Social and Environmental Impact Assessment (SEIA).
2001	Deposition and approval of the SEIA with public consultants organized by the Office National de l'Environnement (ONE).
2002	Elaboration of the Environmental Management Plan (EMP).
2002-2005	Final feasibility studies and semestrial reports to ONE.
August 2005	Investment decision by Rio Tinto.
2006	Start of construction of infrastructure.
2009	Expected start of the mining operation (up to 60 years).
+/- 2070	Mine closure.

Appendix 5: Parental allocation to offspring – identification of each designated parent with the actual collection number for *C. kianjavatensis* and *C. montis-sacri*.

Parent ID	*C. kianjavatensis* collection #	*C. montis-sacri* collection #
P1	SK004 (A. 602)	SK101 (A. 321)
P2	SK006 (A. 602)	SK102 (A. 321)
P3	SK010 (A. 602)	SK104 (A.321)
P4	SK011 (A. 602)	SK106 (A.321)
P5	SK013 (A. 602)	SK107 (A.321)
P6	SK014 (A. 602)	SK108 (A.321)
P7	SK018 (A. 602)	SK109 (A. 321)
P8	SK075 (A. 602)	SK103 (A. 321)
P9	SK083 (A. 602)	SK110 (A. 321)
P10	SK009 (A. 602)	SK112 (A. 321)
P11	SK094 (A. 602)	SK178 (A. 321)
P12	SK008 (A. 602)	
P13	SK016 (A. 602)	
P14	SK060 (A. 213)	
P15	SK035 (A. 213)	
P16	SK049 (A. 213)	
P17	SK031 (A. 213)	
P18	SK080 (A. 213)	
P19	SK064 (A. 213)	
P20	SK093 (A. 213)	
P21	SK003 (A. 213)	
P22	SK040 (A. 213)	
P23	SK001 (A. 213)	
P24	SK034 (A. 213)	
P25	SK022 (A. 213)	
P26	SK026 (A. 213)	
P27	SK048 (A. 213)	
P28	SK186 (A. 213)	
P29	SK071 (A. 213)	
P30	SK062 (A. 213)	
P31	SK070 (A. 213)	