Origins and Dispersal of the Sweet Potato and Bottle Gourd in Oceania: Implications for Prehistoric Human Mobility

Andrew Christopher Clarke

2009

Origins and Dispersal of the Sweet Potato and Bottle Gourd in Oceania: Implications for Prehistoric Human Mobility

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Plant Biology

at Massey University, Palmerston North,

New Zealand.

Andrew Christopher Clarke

2009

What manner of men were they who by surpassing the achievements of the Phœnicians in the Mediterranean and the Vikings of the north Atlantic are worthy of being called the supreme navigators of history?

Sir Peter Buck (Te Rangi Hiroa)
***Vikings of the Sunrise* (1938, p. 13)**

ABSTRACT

The origins of the sweet potato (*Ipomoea batatas*) and bottle gourd (*Lagenaria siceraria*), two important commensal species in prehistoric Polynesia, have remained elusive. Most recently, a South American origin has been favoured, which prompts a number of interesting questions surrounding how, when, from where and by whom these species dispersed into the Pacific. For this project, hypotheses were formulated based on existing archaeological, linguistic and maritime evidence, and tested using a molecular approach. For both species, extensive marker development was necessary.

For the bottle gourd, a set of seven molecular markers (two chloroplast and five nuclear) was developed to test the hypothesis of a South American origin for the Polynesian bottle gourd. These were sequenced in 36 accessions of bottle gourd from Asia, the Americas and New Zealand. Analyses of these markers support a dual origin for the Polynesian bottle gourd: the chloroplast markers identify an Asian origin, but the nuclear markers reveal alleles that originate in both the Americas and Asia. By combining information from a number of sources, a model for the domestication(s) and global dispersal of the bottle gourd is proposed.

For the sweet potato, the amplified fragment length polymorphism (AFLP) technique was used. First, using a new procedure that will be applicable to other studies, AFLP scoring parameters were optimised to improve phylogenetic resolution. Second, to elucidate sweet potato dispersal in Oceania, AFLP profiles were generated for 270 unique accessions of sweet potato from Asia, Island Melanesia, Polynesia and the Americas. The putative *kumara* lineage, which represents a prehistoric, Polynesian-mediated introduction from South America, was identified. Sweet potato accessions from Asia to Western Polynesia were found to be genetically diverse, and the relationships between them are more complex than previously recognised. The phylogenetic positions of the Māori varieties ‘Hutihuti’, ‘Rekamaroa’ and ‘Taputini’ are inconsistent with these accessions representing pre-European cultivars; instead it is more likely that they are early European introductions.

To answer questions about the prehistoric dispersal patterns of the bottle gourd, future work could make use of high resolution markers and ancient DNA (aDNA) from archaeological and early historic-period samples. Future work on the sweet potato is needed to narrow down the point of Polynesian contact on the South American coast, and to answer this question more intensive sampling is required. Integration of genetic, linguistic, historical and morphological data will also be important.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Professor David Penny. The opportunity to work with you, and in your lab, has been an exciting, challenging and rewarding experience. Your unyielding enthusiasm for biology, relentless pursuit of interesting questions, and blatant disregard for interdisciplinary boundaries have been particularly appreciated. Thank you also for providing funding to complete research here at Massey and to travel and meet the large number of people who have contributed to this project.

I would also like to thank my team of co-supervisors: Dr John McCallum, Professor Peter Lockhart and Professor Lisa Matisoo-Smith. John, thank you for allowing me to spend time in your lab, and for lots of interesting ideas. Pete, thank you for advice on data analysis and encouragement along the way. Lisa, thank you for your advice in interpreting data, and helping me see my research in an anthropological context.

Trish McLenachan helped with everything from PCR-fixing, cloning-fixing and sequencing-fixing, to experimental design, emotional support ("It'll be fine" when I needed it), editing and proof-reading and even the occasional meal and alcoholic beverage. All have been appreciated.

For this project I have relied a great deal on those who have generously provided samples. For bottle gourd samples I especially thank Charles Heiser and Mike Burtenshaw, as well as Richard Cross, John Palmer and Peter Matthews. For allowing me to sample the Yen Sweet Potato Collection in Japan, I thank Makoto Nakatani and staff at the National Institute of Crop Science, Tsukuba. Steve Lewthwaite also provided many sweet potato samples of New Zealand material, Pam Fletcher provided tissue culture material from the Yen Collection, Mary Taylor provided modern samples from the Pacific, and Richard Walter provided samples from Mangaia, Cook Is.

At the National Museum of National History, Smithsonian Institution, I would like to thank Bruce Smith, Noreen Tuross (now of Harvard University), Dave Erickson, Sarah Zabriskie and members of the Laboratories of Analytical Biology. Your generosity in hosting me for an internship was greatly appreciated, as was the opportunity to write a paper with you afterwards.

Roger Green generously provided me with draft copies of his 2005 *Sweet Potato Transfers in Polynesian Prehistory* paper, which greatly helped develop the background material for the project. Roger was also very generous with his time — lending his vast knowledge of Pacific prehistory to the interpretation of the bottle gourd and sweet potato results in archaeological, historical and linguistic contexts.

Heidi Meudt helped with many aspects of the project, including high-throughput DNA extractions, AFLP fingerprinting, scoring and analysis of AFLP data, and proof-reading of various chapters. Heidi's patience even extended to writing an AFLP review with me, an opportunity for which I am very grateful — even if I sometimes wondered whether writing a review paper concurrently with a thesis was really a good idea!

Barbara Holland assisted with the data analysis for both the bottle gourd and sweet potato, lending her extensive phylogenetic expertise to a daunting task. I have particularly appreciated Barbara's ability to see research potential in even the scariest data matrix. If David's "tame mathematician" exists then she is certainly Barbara.

I would like to thank the many others who helped during the course of this project. Mike Burtenshaw was an excellent collaborator on the bottle gourd project. Peter Matthews assisted with the sweet potato sampling in Japan, and the bottle gourd manuscript. Peter's extensive knowledge of Pacific ethnobotany and tendency to ask really hard questions were both appreciated. Ashley Gould provided excellent guidance on the historical aspects of the sweet potato in New Zealand. Meeghan Pither-Joyce and the other members of the Genome Research Lab at Crop & Food Research (now Plant & Food Research), Lincoln taught me many molecular techniques. Ruth Morrison assisted with germinating bottle gourd seeds to obtain fresh tissue. Sue Gardiner's lab at HortResearch (now Plant & Food Research), particularly Mike Cook, helped grind the sweet potato tissue for DNA extraction. Paul Hocquard, Karen Sinclair and Joy Wood

made sure I had all the consumables I required, often at short notice. Warwick Allen wrote several time-saving Perl scripts. Tim White assisted with the network analysis of the bottle gourd. Lorraine Berry did all the capillary electrophoresis of the DNA sequencing and AFLP products, and was always helpful with troubleshooting when things didn't go quite according to plan.

Māori consultation was an important part of the project, and I would like to acknowledge the guidance of Nick Roskruge, Jonathan Procter and Rangitāne O Manawatu, and the late Dell Wihongi.

For stipend funding I would like to acknowledge a Massey University Doctoral Scholarship and the generosity of Pete Lockhart and Mike Hendy. For travel funding to attend conferences, conduct field work and complete lab work away from Massey, I would like to thank the Institute of Molecular BioSciences, the Allan Wilson Centre for Molecular Ecology and Evolution, the Royal Society of New Zealand, the Smithsonian Institution, John McCallum and the New Zealand Institute for Crop and Food Research.

I would also like to thank all the staff, students and visitors — both past and present — of the Allan Wilson Centre and Institute of Molecular BioSciences who have contributed towards this project, especially the occupants of the Farside and PLEB Labs, the Boffin Lounge and the Fish Bowl. I have benefited enormously from everything from help with lab work and literature searches, to interesting discussions about molecular biology, data analysis and anthropology (amongst *many* other things). You have provided a stimulating, supportive and entertaining environment in which to work. I would especially like to acknowledge the good friends I have made during my time here: Richard Carter, Fred Delsuc, Chrissie Down, Barbara Holland, Anu Idicula, Kim McBreen, Heidi Meudt, Matt Phillips, Melanie Pierson, Renae Pratt, Scott Roy and Alice Storey. Of course there are many others I could name, including outside the Allan Wilson Centre.

I would like to acknowledge and thank Andrea Bryant, whose love, support and tolerance were often far in excess of what I deserved. I would also like to thank Andrea's family for being so supportive and welcoming, especially Ralph and Raewyn, and fellow PhD-ers Jeremy and Michelle.

Finally, I would like to acknowledge the support of my family. Mum and Dad, your support — in all areas — has been unwavering for much longer than the course of this degree, but has been particularly appreciated during the last few years. Thank you also to my sister Sarah, who is always encouraging despite still having no idea what it is I actually *do*.

PREFACE

The sweet potato (*Ipomoea batatas*) and bottle gourd (*Lagenaria siceraria*) were two crop species fundamental to many agricultural systems in Polynesia. A long-standing scientific interest in these species centres on their origins; both are thought to have arrived in the Pacific from South America. If the prehistoric Oceanic lineages of these species are of South American origin, then this prompts a number of interesting questions surrounding how, when, from where and by whom these species were dispersed into the Pacific. Based on existing archaeological, linguistic, botanical and sailing technology evidence we are able to formulate testable hypotheses around the dispersal of these species. These hypotheses are tested in this thesis using a molecular (DNA) approach.

The thesis is divided into five chapters. The first chapter is a general introduction that frames the remaining research by providing an overview of prehistoric human settlement and mobility in the Pacific, with the focus on the contribution of molecular studies of humans and their commensals (animals and plants) to understanding human mobility in the Pacific. The discussion then narrows to the question of contact between Eastern Polynesia and the Americas, specifically the evidence and likelihood of such contact. Chapter One concludes with the aims and hypotheses of the research.

Chapter Two describes the molecular strategies employed for the bottle gourd and sweet potato. The chapter begins by outlining the considerations and approaches for developing appropriate molecular markers for a given taxonomic group and scientific question, with a focus on closely-related taxa where genetic variation is relatively low. In the context of these considerations, the molecular methods of choice are justified for both the bottle gourd and sweet potato. For the bottle gourd, PCR and sequencing markers derived from inter-SSR multilocus genetic fingerprints (so-called ISSR-derived SCAR markers) were chosen as an appropriate marker system. The development of these, which employed a novel combination of existing molecular methods, is described in detail. For the sweet potato, AFLP fingerprinting was chosen as an appropriate

marker system. Although the AFLP technique is firmly established and widely used, there are several aspects of the method which require further development — primarily scoring, i.e., methods for converting raw AFLP profiles into a binary data matrix. During the course of the AFLP work it was found, in collaboration with Drs Barbara Holland and Heidi Meudt, that the widely-used parameters for generating binary data matrices from unedited AFLP profiles are non-optimal, and that by adjusting these parameters the number of characters can be significantly increased and the amount of homoplasy significantly decreased.

Chapter Three describes research on the origins and dispersal of the bottle gourd in Oceania. The introduction brings together disparate and rarely synthesised bottle gourd research, allowing the Oceanic origins of this species to be placed in context: as with humans themselves, the introduction of the bottle gourd into the Pacific represents the most recent event in the global dispersal of this species. Because the bottle gourd occurs in Africa, Asia and the Americas, gourds from all areas are germane to the origins in Oceania. A range of bottle gourd accessions from Asia, the Americas and Oceania were obtained and these are described. This is followed by the methods, which describe how the five SCAR markers developed in Chapter Two were used, along with two additional chloroplast markers, to amplify and sequence polymorphic loci from 36 accessions of bottle gourd. The analyses of these data are presented, followed by a discussion of Oceanic dispersal scenarios compatible with the results. The implications of bottle gourd research using ancient DNA, which was undertaken at the Smithsonian Institution, Washington, D.C. at the same time as my own research (and with which I was involved), are also discussed. The discussion is extended to the global dispersal and domestication(s) of the bottle gourd, as the use of the chloroplast markers in two outgroup species and the reinterpretation of the archaeological data allow some tentative conclusions to be made in these areas.

Chapter Four deals with the origins and dispersal of the sweet potato in Oceania. The introduction reviews the literature in this area. Unlike the bottle gourd, a reasonably clear and consistent picture is emerging for the sweet potato; the field was brought together in 1974 with the publication of Douglas Yen's landmark *The Sweet Potato and Oceania: An Essay in Ethnobotany* (Yen, 1974), and was resynthesised in a multidisciplinary, multi-author volume in 2005 (Ballard *et al.*, 2005). Over 400 sweet

potato accessions from Oceania were sampled for this study, and these are described. This is followed by the methods, which describe the application of the amplified fragment length polymorphism (AFLP) multilocus fingerprinting technique to the analysis of approximately 300 of the sweet potato accessions. The optimised scoring parameters determined in Chapter Two were used to convert the raw AFLP profiles into a binary data matrix AFLP. The matrix was used to construct phylogenetic trees that have been interpreted, with assistance from Emeritus Professor Roger Green, in the context of available linguistic and historical data. The large number of taxa, wide geographic coverage in Oceania, narrow geographic coverage in the Americas, and large amount of historical, linguistic and morphological information throw open a multitude of avenues for further research and interpretation. While some of these avenues were pursued here, more detailed work is necessary in certain areas, e.g., the origins of prehistoric and historic sweet potato in New Zealand.

Chapters Two, Three and Four, which form the core of the thesis, are published, or intended for publication, as scientific papers (full bibliographic information below). Chapter Two includes papers published in *Trends in Plant Science* (Meudt & Clarke, 2007) and *Systematic Biology* (Holland *et al.*, 2008), and a manuscript intended for publication in *Plant Systematics and Evolution* or similar (Clarke & McLenachan, *in prep*). All three of these publications have been, or are being, written in collaboration with other researchers, and therefore contain additional material not presented in the thesis. Chapter Three is already published as two papers — one in *Molecular Biology and Evolution* (Clarke *et al.*, 2006), and one in the *Proceedings of the National Academy of Sciences of the United States of America* (Erickson *et al.*, 2005). The Erickson *et al.* (2005) paper, which was written in collaboration with researchers at the Smithsonian Institution, contains material not presented in the thesis. A third bottle gourd publication (on the origin, domestication and global dispersal of the species) is being prepared for *Horticultural Reviews* (Clarke & Penny, *in prep*). Chapter Four is intended for publication in a multi-disciplinary science journal.

Chapter Five includes a general summary, with the goal of bringing together research described in the other chapters and placing it in the context of human mobility in the Pacific and the extent to which it addresses the question of contact between prehistoric Polynesia and the Americas. Chapter Five also suggests a number of avenues

of future research for both the sweet potato and bottle gourd — to further elucidate the origins and dispersal of these species in Oceania and in other regions in which they are found. The potential of other crop species as markers of prehistoric human contact between Polynesian and the New World is briefly outlined.

Appendices 1–10 are provided as hard copies at the end of the thesis, and appendices 11–21 are on the enclosed CD. Appendices include accession lists for the sweet potato, oligonucleotide sequences, DNA sequence data matrices for bottle gourd, AFLP matrices for sweet potato, and reprints of published papers.

Relevant PhD Publications

	Appendix
Meudt, H. M., and **Clarke, A. C.** 2007. Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science* **12:** 106–117.	7
Holland, B. R.,† **Clarke, A. C.**,† and Meudt, H. M.† 2008. Optimizing automated AFLP scoring parameters to improve phylogenetic resolution. *Systematic Biology* **57:** 347–366. † *Equal contributors*	8
Clarke, A. C., Burtenshaw, M. K., McLenachan, P. A., Erickson, D. L., and Penny, D. 2006. Reconstructing the origins and dispersal of the Polynesian bottle gourd (*Lagenaria siceraria*). *Molecular Biology and Evolution* **23:** 893–900.	9
Erickson, D. L., Smith, B. D., **Clarke, A. C.**, Sandweiss, D. H., and Tuross, N. 2005. An Asian origin for a 10,000-year-old domesticated plant in the Americas. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 18315–18320.	10

TABLE OF CONTENTS

Abstract v
Acknowledgements vii
Preface xi
Abbreviations xxiii
List of Figures xxv
List of Tables xxvii

CHAPTER 1: GENERAL INTRODUCTION 1
1.1 Chapter Overview — Research Context 2
1.1.1 A Note on Dates 3
1.2 Human Origins, Migration and Mobility in Oceania 4
1.2.1 Introduction 4
1.2.2 The Settlement of Oceania 4
1.2.2.1 The First Wave (from 60–40,000 yr BP) 4
1.2.2.2 The Austronesian Expansion and the Settlement of Remote Oceania 6
1.2.2.3 Human Genetic Variation in Oceania 10
1.2.2.4 The Lapita Cultural Complex and Commensal Species 12
1.3 Commensal Animals and Plants in Oceania 14
1.3.1 The Commensal Model 14
1.3.2 Application of Commensal Models 15
1.4 Sailing from Polynesia to the New World in Prehistory 17
1.4.1 Evidence of Polynesia–New World Contact 17
1.4.1.1 Crop Species 17
1.4.1.2 The Chicken (*Gallus gallus*) 18
1.4.1.3 Cultural Evidence 19
1.4.1.4 Human Genetic Evidence 20
1.4.2 Sailing to the New World Coast (and Back) 21
1.5 Aims and Hypotheses 23
CHAPTER 2: MOLECULAR MARKERS IN BOTTLE GOURD AND SWEET POTATO 25
2.1 Chapter Overview 26
2.1.1 A Note on Attribution 27
2.2 Introduction 28
2.2.1 Molecular Markers and the Detection of Genetic Variation 28
2.2.1.1 Multilocus DNA Fingerprinting 29
2.2.1.2 Low-Copy Nuclear Markers versus Multilocus Markers 30

2.2.2 Considerations in Choosing Molecular Marker Systems 31
2.2.3 Molecular Marker Systems: Strategies in Crop Plants 32
2.2.4 Molecular Markers in Bottle Gourd: Considerations, Available Markers and Strategies 34
2.2.5 Molecular Markers In Sweet Potato: Considerations, Available Markers and Strategies 36
2.3 SCAR Markers: Introduction and Methodology 38
2.3.1 An Introduction to SCAR Markers 38
2.3.1.1 Development of SCAR Markers 38
2.3.1.2 SCAR Markers Versus Multilocus Fingerprinting 41
2.3.1.3 Sources of SCAR Markers 41
2.3.2 SCAR Markers Derived from Inter-SSR (ISSR) Fingerprinting 42
2.3.2.1 ISSR Fingerprinting: Introduction 42
2.3.2.2 ISSR Fingerprinting: Methodology 43
2.3.3 Problems and Strategies in SCAR Marker Development 44
2.3.3.1 Obtaining Flanking Regions to Improve SCAR Marker Development 46
2.3.4 Using TAIL PCR to Improve SCAR Marker Development 48
2.3.4.1 Thermal Asymmetric Interlaced (TAIL) PCR: Introduction 48
2.3.4.2 TAIL PCR: Methodology 48
2.3.5 Overview of ISSR-Derived SCAR Marker Development 52
2.3.6 Aims 53
2.4 Development of Bottle Gourd SCAR Markers: Methods 54
2.4.1 Preliminary Technical Remarks 54
2.4.2 ISSR Fingerprinting and Isolation of Polymorphic ISSRs 54
2.4.2.1 ISSR Fingerprinting 54
2.4.2.1.1 Materials and DNA Extraction 54
2.4.2.1.2 ISSR PCR 55
2.4.2.1.3 Betaine as an Enhancing Agent for PCR 56
2.4.2.1.4 Agarose Gel Electrophoresis of ISSR PCR Products 57
2.4.2.2 Polyacrylamide Gel Electrophoresis (PAGE) of ISSR PCR Products 58
2.4.2.2.1 Preparation of Polyacrylamide Electrophoretic Glass Plates 58
2.4.2.2.2 Preparation of Acrylamide Gel Solution 59
2.4.2.2.3 Setup of the Electrophoresis Apparatus 59
2.4.2.2.4 Preparation of DNA Size Ladder 60
2.4.2.2.5 Preparation of ISSR PCR Products for PAGE 60
2.4.2.2.6 Loading and Running of ISSR PAGE Gels 60
2.4.2.2.7 Detection of DNA Fragments by Silver Staining of the ISSR Gels 61
2.4.2.3 Isolation and Re-Amplification of Polymorphic ISSRs 62
2.4.2.3.1 Excision of Polymorphic Bands From the ISSR PAGE Gels 62
2.4.2.3.2 Re-Amplification from Excised Band Eluate 62
2.4.2.3.3 Agarose Gel Extraction of Re-Amplification PCR Products 63
2.4.2.3.4 NanoDrop® Spectrophotometry 63
2.4.3 Conversion of ISSRs into SCAR Markers 64
2.4.3.1 Cloning and DNA Sequencing of ISSRs 64
2.4.3.1.1 Overview of the TA Cloning and Blue–White Screening Techniques 64
2.4.3.1.2 Ligation of Re-Amplified ISSRs into pGEM®-T Easy 65

2.4.3.1.3 Transformation of *Escherichia coli* DH5α™ ... 66
2.4.3.1.4 Blue–White Screening, Colony PCR and Inoculation of LB Broth ... 67
2.4.3.1.5 Plasmid DNA Extraction and Digestion ... 68
2.4.3.1.6 DNA Sequencing of Insert DNA ... 69
2.4.3.1.7 Purification of DNA Sequencing Products by Ethanol Precipitation ... 69
2.4.3.1.8 Capillary Electrophoresis of DNA Sequencing Reaction Products ... 70
2.4.3.1.9 DNA Sequence Data Analysis ... 70
2.4.3.2 TAIL PCR of ISSR Flanking Regions ... 71
2.4.3.2.1 Design of TAIL PCR Primers ... 71
2.4.3.2.2 1°, 2° and 3° TAIL PCR ... 71
2.4.3.2.3 Agarose Gel Extraction of TAIL PCR Products ... 72
2.4.3.2.4 Ethanol Precipitation of TAIL PCR Products ... 73
2.4.3.2.5 DNA Sequencing of Gel-Extracted TAIL PCR Products ... 74
2.4.3.2.6 DNA Sequence Data Analysis ... 74
2.4.3.3 Development of ISSR SCAR Markers ... 74
2.4.3.3.1 ISSR SCAR Marker Primer Design ... 74
2.4.3.3.2 Optimisation PCR for ISSR SCAR Markers ... 75
2.4.3.3.3 Trial PCR of ISSR SCAR Markers on Several Accessions ... 76
2.4.3.3.4 SAP/Exo I Cleanup of PCR Products ... 76
2.4.3.3.5 DNA Sequencing of Trial ISSR SCAR Markers ... 76
2.4.3.3.6 DNA Sequence Data Analysis ... 77
2.5 Development of Bottle Gourd SCAR Markers: Results and Discussion ... 78
2.6 Screening of Other Markers for Bottle Gourd ... 84
2.6.1 Screening of Chloroplast Markers ... 84
2.6.2 Screening of Nuclear Markers ... 85
2.6.3 *trn*S–*trn*G Marker from Inter-ccSSR PCR ... 85
2.7 The AFLP Technique: Introduction and Methodology ... 87
2.7.1 Introduction ... 87
2.7.2 Generating an AFLP Fingerprint: Methodology ... 88
2.8 Optimising AFLP Scoring Parameters in Sweet Potato: Introduction ... 93
2.8.1 Aims ... 98
2.9 Optimising AFLP Scoring Parameters in Sweet Potato: Materials and Methods ... 99
2.9.1 Sampling Strategy ... 99
2.9.2 Generation of Raw AFLP Data ... 100
2.9.3 Generation of Datasets using Different Automated Scoring Parameter Settings ... 101
2.9.4 Comparison of Datasets to Determine Optimal Parameter Settings ... 104
2.9.4.1 Measures of Accuracy ... 104
2.9.4.2 Optimal Parameter Settings ... 107
2.9.4.3 Robustness of the Phylogenies to Changes in Parameter Settings ... 107
2.10 Optimising AFLP Scoring Parameters in Sweet Potato: Results and Discussion ... 108
2.10.1 Measures of Accuracy ... 108
2.10.1.1 Phylogenetic Resolution ... 108
2.10.1.2 Correct Assignment of Replicates ... 111
2.10.1.3 Error Rates ... 112

2.10.2 Optimal Parameter Settings....116
2.10.2.1 GeneMapper® 116
2.10.2.2 GeneMarker® 117
2.10.3 Robustness of the Phylogenies to Changes in Parameter Settings 118
2.10.4 Additional Independent Evidence 123
2.11 Optimising AFLP Scoring Parameters in Sweet Potato: Conclusions 124
CHAPTER 3: ORIGINS AND DISPERSAL OF THE BOTTLE GOURD IN OCEANIA 127
3.1 Chapter Overview 128
3.2 Introduction....130
3.2.1 Biology of the Bottle Gourd....130
3.2.1.1 Taxonomy, Systematics and Genetics 131
3.2.1.2 Reproductive Biology....133
3.2.2 Human Uses of the Bottle Gourd, Especially in Polynesia 133
3.2.3 The Bottle Gourd in Africa — Evolution and Domestication....135
3.2.3.1 An African Origin for the Bottle Gourd? 135
3.2.3.2 Discovery of the Wild African Bottle Gourd and its Pollinator?....136
3.2.3.3 Domestication of the Bottle Gourd and the Archaeological Record 137
3.2.4 Global Dispersal of the Bottle Gourd....138
3.2.4.1 The Bottle Gourd in the New World 138
3.2.4.2 The Bottle Gourd in Asia 144
3.2.4.3 A Pre-Austronesian Expansion Pacific Rim Distribution....145
3.2.5 The Bottle Gourd in Island Southeast Asia and Oceania 149
3.2.5.1 The Bottle Gourd in the Western Pacific....149
3.2.5.2 The Wax Gourd in Polynesia....152
3.2.5.3 Distribution of the Polynesian Bottle Gourd 153
3.2.5.4 The Polynesian Bottle Gourd: A New World Origin?....155
3.2.5.5 Natural or Human-Mediated Dispersal from the New World?....158
3.2.5.6 The Bottle Gourd in New Zealand 159
3.2.5.7 A Dual Origin for the Eastern Polynesia Bottle Gourd?....163
3.2.6 Summary 165
3.2.7 Hypothesis and Aims 166
3.2.7.1 Hypothesis....166
3.2.7.2 Aims 167
3.3 Materials and Methods....168
3.3.1 Overview....168
3.3.2 Preliminary Technical Remarks 168
3.3.3 Collection of Bottle Gourd Accessions....168
3.3.4 Genomic DNA Extraction....171
3.3.4.1 Germination of Bottle Gourd Seeds 171
3.3.4.2 CTAB DNA Extraction 172

3.3.4.3 Qiagen DNeasy® DNA Extraction 175
3.3.4.4 Agarose Gel Electrophoresis of Genomic DNA 176
3.3.5 PCR and Sequencing of Chloroplast and ISSR SCAR Markers 176
3.3.5.1 PCR of Chloroplast and SCAR Markers 177
3.3.5.2 DNA Sequencing of Chloroplast and ISSR SCAR Markers 178
3.3.5.2.1 DNA Sequencing of Chloroplast and ISSR SCAR Markers 179
3.3.5.2.2 Purification of DNA Sequencing Products by CleanSEQ® 179
3.3.5.2.3 Capillary Electrophoresis of Chloroplast and ISSR SCAR Markers 179
3.3.5.3 Editing and Alignment of DNA Sequence Data 180
3.3.6 DNA Sequence Data Analysis 180
3.3.6.1 Genotype Frequency Pie Charts 180
3.3.6.2 Network Analysis 181
3.4 Results 182
3.4.1 DNA Sequence Data Analysis 182
3.4.1.1 Genotype Frequency Pie Charts 182
3.4.1.2 Network Analysis 185
3.4.1.3 Outgroup Analysis 186
3.5 Discussion 187
3.5.1 Origins of the Polynesian Bottle Gourd 187
3.5.1.1 Genetic Evidence 187
3.5.1.2 Reassessing the Evidence Against an Asian Origin 187
3.5.1.3 Summary 188
3.5.2 Reconciling Erickson *et al.* (2005) and Clarke *et al.* (2006) 188
3.5.2.1 Background to Collaboration 189
3.5.2.2 Summary of Erickson *et al.* (2005) 189
3.5.2.3 Implications of a Non-African Origin for the New World Bottle Gourd 192
3.5.2.4 Implications for Clarke *et al.* (2006) 193
3.5.2.4.1 Hypothesis 1: Introgression of Modern Māori Samples 194
3.5.2.4.2 Hypothesis 2: A Non-Asian, Non-American Origin for Polynesian ssp. *siceraria* Alleles? 194
3.5.2.4.3 Hypothesis 3: Archaeological Bottle Gourd Seeds in Peru: An African Origin? 195
3.5.2.5 A Reconciliation? 200
3.5.3 Domestication and Dispersal 201
3.5.4 Assumptions, Limitations and Caveats 201
3.6 Conclusions 204
3.6.1 Molecular Markers in Bottle Gourd 204
3.6.2 The Polynesian Bottle Gourd 204
3.6.3 Domestication and Dispersal 206
CHAPTER 4: ORIGINS AND DISPERSAL OF THE SWEET POTATO IN OCEANIA 209
4.1 Chapter Overview 210
4.2 Introduction 211
4.2.1 The Sweet Potato in Oceania — Hypotheses 211

4.2.2 Sweet Potato Evolution, Domestication and Distribution in the Americas 212
4.2.3 The Sweet Potato in Oceania 214
4.2.3.1 The Linguistic Aspect 214
4.2.3.2 Distribution and Dispersal of the Sweet Potato in Polynesia 216
4.2.4 The Possibility of Natural Transfer? 219
4.2.5 Hypotheses and Aims 220
4.3 Materials and Methods 221
4.3.1 Overview 221
4.3.2 Preliminary Technical Remarks 222
4.3.3 Collection of Sweet Potato Samples 222
4.3.4 Sampling Strategy 223
4.3.5 DNA Extraction 225
4.3.5.1 Leaf Tissue Disruption 225
4.3.5.2 Qiagen DNeasy® DNA Extraction 225
4.3.5.3 Quantification of DNA Yield and Assessment of DNA Quality 227
4.3.5.3.1 Agarose Gel Electrophoresis 227
4.3.5.4 Concentration of DNA for AFLP 227
4.3.6 AFLP Fingerprinting 228
4.3.6.1 Restriction Endonuclease Digestion of DNA 228
4.3.6.2 Ligation of Oligonucleotide Linkers to the Restriction Fragments 229
4.3.6.2.1 Preparation of Oligonucleotide Linkers 229
4.3.6.2.2 Ligation of Linkers to the Restriction Fragments 230
4.3.6.3 Pre-Selective Amplification of Ligation Products 230
4.3.6.4 Selective Amplification of Pre-Selective PCR Products 231
4.3.6.4.1 Primer Screening of Selective PCR Primers 231
4.3.6.4.2 Selective Amplification of Pre-Selective PCR Products 233
4.3.6.5 Capillary Electrophoresis of AFLPs 234
4.3.6.5.1 Poolplexing of Selective Amplification PCR Products 234
4.3.6.5.2 Capillary Electrophoresis of Selective Amplification PCR Products 235
4.3.7 AFLP Scoring 235
4.3.8 Phylogenetic Reconstruction 236
4.3.8.1 Preliminary Phylogenetic Analyses 236
4.3.8.2 Choosing Taxa to Omit for Final Analysis 237
4.3.8.3 Final Phylogenetic Analysis 238
4.3.8.4 Path Lengths Between Replicates 238
4.4 Results and Discussion 240
4.4.1 Proportion of “1” Character States 240
4.4.2 Neighbour-Joining Tree of Complete Dataset 242
4.4.3 Path Lengths Between Replicates 245
4.4.4 Origins of the Oceanic Sweet Potato Lineages 246
4.4.5 Origins of the New Zealand Sweet Potato Lineages 250
4.5 Conclusions 254

CHAPTER 5: SUMMARY AND FUTURE WORK 257
5.1 Summary 258
5.2 Future Work: Bottle Gourd 259
5.2.1 The Oceanic Bottle Gourd 259
5.2.1.1 Ancient DNA 260
5.2.1.2 Additional Sampling and Germplasm Conservation 264
5.2.1.3 High Resolution DNA Sequencing Markers 264
5.2.1.4 Linguistic Analysis 265
5.2.2 Domestication and Dispersal 267
5.2.2.1 Landrace Sampling 267
5.2.2.2 Ancient DNA and Histological Examination 268
5.3 Future Work: Sweet Potato 270
5.3.1 Sweet Potato Phylogenetics 270
5.3.2 Improved Sampling 270
5.3.3 An Integrated Approach 271
5.4 Commensal Plants and New World Contact 272

REFERENCES 273

APPENDICES 311
Appendix 1: Markers Screened for Bottle Gourd 312
Appendix 2: Bottle Gourd PCR and Sequencing Primers 314
Appendix 3: Bottle Gourd ISSR Sequences 317
Appendix 4: Yen Sweet Potato Accessions 320
Appendix 5: AFLP Oligonucleotides 340
Appendix 6: AFLP Scoring Parameters 341
Appendix 7: Reprint of Meudt & Clarke (2007) 343
Appendix 8: Reprint of Holland *et al.* (2008) 355
Appendix 9: Reprint of Clarke *et al.* (2006) 375
Appendix 10: Reprint of Erickson *et al.* (2005) 383
Appendix 11: Cloning Application ..... CD
Appendix 12: Bottle Gourd NEXUS Files ..... CD
Appendix 13: Table of Bottle Gourd Variable Sites ..... CD
Appendix 14: Bottle Gourd Seed Length and Width Data ..... CD
Appendix 15: Sweet Potato Accessions ..... CD
Appendix 16: Sweet Potato AFLP Profiles ..... CD
Appendix 17: Sweet Potato NEXUS Files ..... CD
Appendix 18: Suppl. Material for Holland *et al.* (2008) ..... CD
Appendix 19: Suppl. Material for Clarke *et al.* (2006) ..... CD
Appendix 20: Support. Info for Erickson *et al.* (2005) ..... CD
Appendix 21: Electronic Copy of Thesis ..... CD

ABBREVIATIONS

A	adenine
AD	*anno Domini*
AD	arbitrary degenerate (PCR primer)
aDNA	ancient DNA
AFLP	amplified fragment length polymorphism
AMOVA	Analysis of Molecular Variance
AMS	accelerator mass spectrometry
APS	ammonium persulfate
a.s.l.	above sea level
AWCGS	Allan Wilson Centre Genome Service
ATP	adenosine 5′-triphosphate
BAC	bacterial artificial chromosome
BC	before Christ (prior to AD 1)
bp	base pair
BW	bin width
°C	degrees Celsius
ca	*circa*
C	cytosine
cal	calorie
CAPS	cleaved amplified polymorphic sequence
ccSSR	consensus chloroplast SSR
cDNA	complementary DNA
CE	capillary electrophoresis
cpDNA	chloroplast DNA
CRoPS	Complexity Reduction of Polymorphic Sequences
CTAB	hexadecyltrimethylammonium bromide
cv.	cultivar
d	day
DArT	Diversity Arrays Technology
DMF	*N,N*′-dimethyl-formamide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide 5′-triphosphate
DTE	dithioerythritol
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid · disodium salt
ERMA	Environmental Risk Management Authority
EST	expressed sequence tag
Exo I	exonuclease I
g	gram
× *g*	times gravity
G	guanine
GMO	genetically modified organism
GTR	GENOTYPER REARRANGER
h	hour
HLA	human leukocyte antigen
Hz	Hertz
indel	insertion–deletion
I-PCR	inverse PCR
IPTG	isopropyl-β-D-thiogalactopyranoside
ISSR	inter-SSR
kb	kilobase pairs (10^3 bp)
L	litre
LB	lysogeny broth/Luria–Bertani
LGDP	local and global detection percentages (GeneMarker®)
LMS	ligation-mediated suppression (PCR)
LSC	long single-copy (cpDNA)
L:W	length:width ratio
m	metre
M	Molar (mol L^{-1})
MAS	marker-assisted selection
Mbp	megabase pairs (10^6 bp)
MDS	multidimensional scaling
MFL	minimum fragment length
min	minute
MLF	multilocus fingerprinting

mtDNA	mitochondrial DNA
nDNA	nuclear DNA
NJ	neighbour-joining
nrITS	nuclear ribosomal internal transcribed spacer
NRY	nonrecombining portion of the Y chromosome
nt	nucleotide
PAGE	polyacrylamide gel electrophoresis
PCA	principal component analysis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PHT	peak height threshold
PNG	Papua New Guinea
POc	proto-Oceanic language
PVP	polyvinyl-pyrrolidone
QTL	quantitative trait locus
RAPD	randomly amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
rfu	relative fluorescent unit
RNA	ribonucleic acid
rpm	revolutions per minute
s	second
s	standard deviation (sample)
SAP	shrimp alkaline phosphatase
SEC	South Equatorial Current
SCAR	sequence-characterised amplified region
SNP	single nucleotide polymorphism
SP	specific (PCR primer)
sp.	species (singular)
SPF	stutter peak filter (GeneMarker®)
spp.	species (plural)
SSC	short single-copy (cpDNA)
SSCP	single-stranded conformation polymorphism
ssp.	subspecies
SSR	simple sequence repeat (microsatellite)
STR	short tandem repeat
syn.	synonym
T	thymine
T_m	melting temperature
TAE	Tris acetate EDTA
TAIL	thermal asymmetric interlaced (PCR)
TBE	Tris borate EDTA
T-DNA	transfer DNA
TEMED	*NNN′N′*-tetramethylethylenediamine
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride
U	unit
UBC	University of British Columbia
UPGMA	Unweighted Pair Group Method with Arithmetic mean
V	Volt
v/v	volume per volume
W	Watt
WGA	whole genome amplification
w/v	weight per volume
w/w	weight per weight
$\bar{x}$	mean (sample)
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
yr BP	calendar years before present (non-radiocarbon date, or calibrated ^{14}C radiocarbon date; by convention before AD 1950)
^{14}C yr BP	years before present (uncalibrated ^{14}C radiocarbon date; by convention before AD 1950)

LIST OF FIGURES

Fig. 1.1 Regions of Oceania 5
Fig. 1.2 The Austronesian Language Family 8
Fig. 1.3 The Lapita Expansion into Oceania 9
Fig. 2.1 Traditional Process of SCAR Marker Development 39
Fig. 2.2 Overview of ISSR Fingerprinting 43
Fig. 2.3 Obtaining SCAR Flanking Regions 46
Fig. 2.4 Overview of TAIL PCR Technique 50
Fig. 2.5 Theoretical Results of TAIL PCR Electrophoresis 51
Fig. 2.6 Overview of ISSR-Derived SCAR Marker Development in Bottle Gourd 52
Fig. 2.7 Cucumber Chloroplast Genome Showing Gourd Marker Positions 86
Fig. 2.8 Overview of the AFLP Fingerprinting Technique 89
Fig. 2.9 AFLP Poolplexing and Capillary Electrophoresis 91
Fig. 2.10 Typical Electronic Profiles for AFLP 92
Fig. 2.11 Effects of Changing Bin Width and Peak Height Threshold 97
Fig. 2.12 Flow Chart of Scoring Parameter Optimisation Process 102
Fig. 2.13 Phylogenetic Resolution and Error Rates versus Bin Width 114
Fig. 2.14 Size Difference Between Identical Alleles of Replicate Pairs 115
Fig. 2.15 Consensus Networks of Different AFLP Parameter Settings 119
Fig. 2.16 Consensus Networks of GeneMapper® and GeneMarker® Trees 121
Fig. 2.17 Consensus Networks of Default and Optimised Trees 122
Fig. 3.1 The Bottle Gourd (*Lagenaria siceraria*) 130
Fig. 3.2 Phytoliths from Bottle Gourd and *Cucurbita* 140
Fig. 3.3 Prehistoric Bottle Gourd Floats, Huaca Prieta, Peru 141
Fig. 3.4 Prehistoric Bottle Gourd in Asia, the New World and Oceania 146
Fig. 3.5 Distributions and Dispersal Routes of the Oceanic Bottle Gourd 156
Fig. 3.6 Bottle Gourd Exocarp from Kohika, New Zealand 160
Fig. 3.7 Seeds of Bottle Gourd and Outgroup Accessions 173
Fig. 3.8 Genotype Frequencies for 36 Accessions of Bottle Gourd 183
Fig. 3.9 Spectronet Network of Bottle Gourd Nuclear Data 186
Fig. 3.10 Prehistoric Bottle Gourd Seeds from Five Sites in Peru 197
Fig. 3.11 Hypothesised Global Dispersal and Domestication of Bottle Gourd 207
Fig. 4.1 The Introduction of Sweet Potato into Oceania: The Tripartite Hypothesis 213
Fig. 4.2 The Banks & Solander "Kumala" Sweet Potato (1769) 217
Fig. 4.3 Overview of AFLP Fingerprinting in Sweet Potato 221
Fig. 4.4 Workflow for Generation of AFLP Data 224
Fig. 4.5 Taxa Ranked by Percentage of "1" Character States 241

Fig. 4.6 Sweet Potato Neighbour-Joining Tree.....242
Fig. 4.7 Path Lengths Between AFLP Replicates.....245
Fig. 4.8 Leaf Shapes of Eastern Polynesian Sweet Potato.....247

LIST OF TABLES

Table 2.1 ISSR Primers Used from UBC Primer Set #9 56
Table 2.2 Polymorphic ISSR Bands Isolated for Bottle Gourd 78
Table 2.3 TAIL PCR and SCAR Marker Development 80
Table 2.4 Polymorphisms Present in Developed SCAR Markers 81
Table 2.5 Rates of Polymorphism in TAILed vs. non-TAILed SCAR Markers 82
Table 2.6 30-Sample Dataset for Optimising AFLP Scoring Parameters 99
Table 2.7 Resolution Scores and Numbers of Parsimony Informative Sites 109
Table 2.8 Normalised Resolution Scores 110
Table 2.9 Replicate Error Rates for Representative Datasets 112
Table 2.10 Binning Error Simulation for Identical Fragments 116
Table 3.1 Defining Characters of the *Lagenaria siceraria* subspecies 132
Table 3.2 Bottle Gourd Accession Details 169
Table 3.3 Chloroplast and Nuclear Bottle Gourd PCR Markers 177
Table 3.4 Ancient New World Bottle Gourds from Erickson *et al.* (2005) 190
Table 3.5 Seed L:W Ratios in New Zealand, the New World, Africa and Asia 198
Table 3.6 Differences in Mean Seed Length:Width Ratios Between Regions 198
Table 4.1 Distribution of *Kumara* and Cognate Terms in South America and Polynesia 215
Table 4.2 Results of PCR Primer Combination Screening for AFLP 232
Table 4.3 Eco+3 and Mse+3 Selective Amplification Primer Combinations 233
Table 4.4 Absorbance and Emission Maxima for AFLP Fluorophores 233
Table 5.1 Priority Bottle Gourd Samples for aDNA Analysis 262
Table 5.2 Indigenous Words for Bottle Gourd in the New World and Polynesia 266
Table 5.3 Germplasm Collections of Bottle Gourd Landraces 268
Table A1.1 Chloroplast Markers Screened for Use in Bottle Gourd 312
Table A1.2 Nuclear Markers Screened for Use in Bottle Gourd 313
Table A2.1 pGEM®-T Easy PCR Primers 314
Table A2.2 TAIL PCR Primers 314
Table A2.3 Chloroplast and Nuclear PCR and Sequencing Primers 316
Table A4.1 Yen Sweet Potato Collection (June 2004) 320
Table A5.1 Oligonucleotides Used in AFLP 340
Table A6.1 AFLP Scoring Parameters in GeneMapper® v. 3.7 341