Mitochondrial DNA

in Ancient Human Populations

of Europe

Clio Der Sarkissian

Australian Centre for Ancient DNA
Ecology and Evolutionary Biology
School of Earth and Environmental Sciences
The University of Adelaide
South Australia

A thesis submitted for the degree of Doctor of Philosophy at The University of Adelaide

July 2011

TABLE OF CONTENTS

Abstract **10**

Thesis declaration **11**

Acknowledgments **12**

General Introduction **14**

RECONSTRUCTING PAST HUMAN POPULATION HISTORY USING MODERN MITOCHONDRIAL DNA 15

Mitochondrial DNA: presentation 15

Studying mitochondrial variation 16

Genetic variation 16

Phylogenetics and phylogeography 16

Dating using molecular data, and its limits 17

Population genetics 19

The coalescent theory and coalescent simulations 21

Reconstructing past genetic history of Europeans using mitochondrial data 23

African origins of humans 23

Out-of-Africa 25

Genetic origins of Europeans 26

Limits of modern DNA studies 29

RECONSTRUCTING PAST HUMAN POPULATION HISTORY USING ANCIENT DNA 29

First ancient DNA studies 30

The nature of ancient DNA 30

Post-mortem degradation of DNA 31

Contamination 33

Authenticating ancient DNA data 34

Survival of DNA 38

Applications of ancient DNA 39

Phylogeny and phylogeography of extant and extinct species 40

Domestication processes 41
Evolution of biological functions 41
Temporal population dynamics 41
Other applications 42
Future directions of ancient DNA studies 42
Ancient human DNA studies 43
The problem of authenticity in ancient DNA studies 43
Application of ancient human DNA studies 43
Investigation of ancient human mitochondrial DNA diversity in Europe 46
Limitations of ancient and modern-day human DNA 48
PRESENTATION OF THE PROJECT 50
General aims of the study 50
Samples and organisation of the thesis 50
Methodology 53
HVR-I sequencing 53
Typing of 22 coding region SNPs (GenoCore22 reaction) 54
Comparative genetic database 57
Principal Component Analysis 58
Classical multidimensional scaling 58
Haplotype sharing 58
Coalescence simulation: Bayesian Serial SimCoal and Approximate Bayesian Computation 59
Expected outcomes 59
REFERENCES 60

Chapter One - Ancient Mitochondrial DNA Unravels Complex Population History in North East Europe 80

ABSTRACT 81
INTRODUCTION 82
Prehistoric, historical and cultural background of north east Europe 82

Reconstructing the population history in north east Europe using modern-day mitochondrial data 83

Reconstructing the population history in north east Europe using ancient mitochondrial data 84

MATERIALS AND METHODS 86

Sample description and archaeological context 86

Sample preparation and DNA extraction 87

Hypervariable-Region I sequencing 88

Typing of mtDNA coding region single nucleotide polymorphisms (SNPs) by GenoCoRe22 multiplex PCR 89

Cloning 89

Quantitative Real-Time PCR 90

Authentication of the mtDNA data 91

Populations used in comparative analyses 93

Principal Component Analysis 94

Genetic distance mapping 94

Haplotype sharing analysis 95

Coalescent simulations 95

RESULTS 97

Amplification success and authentication of the ancient DNA data 97

Haplogroup distribution in modern-day populations of north Eurasia 100

Mesolithic Uznyi Oleni Ostrov/Popovo compared to modern-day Eurasian populations 100

Bronze Age Bolshoy Oleni Ostrov compared to modern-day Eurasian populations 105

18th century A.D. Chalmny-Varre compared to modern-day Eurasian populations 106

Comparison among ancient Eurasian populations 107

Testing population history hypothesis using Bayesian Serial SimCoal 108

DISCUSSION 111

Genetic discontinuity between prehistoric populations and Saami 111

Siberian influence in the mtDNA gene pool of North East Europeans 112

Genetic link with extant populations of the Volga-Ural 112

Singularity of Uznyi Oleni Ostrov within the Mesolithic diversity 114

Importance of haplogroup U in the Palaeolithic/Mesolithic substratum 115

Limited genetic impact of ancient north east European foragers on modern-day populations 116

LIST OF SUPPLEMENTARY MATERIALS 117

ACKNOWLEDGEMENTS 118

REFERENCES 118

SUPPLEMENTARY MATERIALS 124

Chapter Two - Mitochondrial Genome Sequencing in Mesolithic North East Europe Unearths a New Sub-clade Within the Broadly Distributed Human Haplogroup C1 147

ABSTRACT 148

INTRODUCTION 149

Current phylogeography of the human mitochondrial haplogroup C 149

Origins for the sub-clade C1 in Europe 152

Complete mitochondrial genome sequencing in ancient and present-day populations 153

MATERIAL and METHODS 154

DNA extraction 154

Enrichment of ancient human mitochondrial DNA 155

Resequencing using the MitoChip v.2.0 (Affymetrix®) 156

SNP confirmation by direct sequencing and minisequencing 157

Authentication of the ancient mtDNA sequence 158

Phylogeny of the C1 clade 159

RESULTS/DISCUSSION 159

Use of MitoChip v.2.0 for resequencing of ancient mitochondrial genomes 159

Resequencing base call rate 162

Phylogeny of the mitochondrial C1 lineage in Mesolithic Uznyi Oleni Ostrov 163

Under-sampling of whole mitochondrial genomes in Eurasia 166

Effect of post-Mesolithic population dynamics 167

A proposed shared genetic history for the Icelandic-specific C1e and the Mesolithic C1f European sub-clades 169

CONCLUSION 170

LIST OF SUPPLEMENTARY MATERIALS 170

ACKNOWLEDGMENTS 170

REFERENCES 171

SUPPLEMENTARY MATERIALS 176

Chapter Three - The Mitochondrial Gene Pool of Scythians of the Rostov Area, Russia: A Melting Pot of Eurasian Influences 181

ABSTRACT 182

INTRODUCTION 183

The Scythians 183

The Bronze Age in the central Eurasian Steppe 183

The Iron Age in the central Eurasian Steppe 184

The origins of Scythians 185

Cultural and genetic homogeneity among ancient nomads of the Eurasian Steppe 185

Genetic diversity of present-day Eurasian populations 186

Ancient DNA from central Eurasia 188

MATERIAL AND METHODS 189

Sample description and archaeological context 189

Sample preparation and DNA extraction 190

Hypervariable-Region I sequencing and coding region GenoCore22 typing 190

Cloning 190

Quantitative Real-Time PCR 190

Authentication of the mtDNA data 191

Ancient populations used in comparative analyses 191

Present-day populations used in comparative analyses 193

Map of haplogroup frequencies 194

Principal Component Analysis (PCA) 195
Haplotype-based analyses of the mtDNA data 195
Fixation index (F_{ST}) calculations and Analysis of the Molecular Variance (AMOVA) 195
Classical Multi-Dimensional Scaling (MDS) 196
Haplotype-sharing analysis 196
RESULTS 197
Success rate for the amplification of authenticated ancient mtDNA 197
Problems associated with the independent replication of one ancient mtDNA haplotype 197
Scythian sample set used in this study 199
Mitochondrial haplogroup structure of the Scythians and comparison with modern-day populations of Eurasia 200
Mitochondrial haplogroup structure of Iron Age populations of Eurasia 205
Haplotype-based analyses 205
Informative haplotypes of Scythians shared with Central Asians 207
Informative haplotypes of Scythians absent from modern-day Central Asians 209
Mitochondrial homogeneity among ancient populations of central Eurasia 212
Mitochondrial continuity in eastern Siberian populations 213
DISCUSSION 215
Mitochondrial makeup of Scythians 215
Western mtDNA substratum in Bronze Age nomads of central Eurasia 215
Genetic input from the East into ancient nomadic populations of central Eurasia 216
Homogeneity among Iron Age populations of the central Eurasian Steppe 218
'Western' genetic influence in the Bronze Age Tarim Basin 220
CONCLUSION 221
LIST OF SUPPLEMENTARY MATERIALS 222
ACKNOWLEDGMENTS 223
REFERENCES 223
SUPPLEMENTARY MATERIALS 230

Chapter Four - Local Mitochondrial Continuity In central Sardinia: Ancient DNA Evidence From The Bronze Age 250

ABSTRACT 251

INTRODUCTION 252

Sardinians, European genetic outliers 252

Geographical isolation of Sardinia and initial settlement 252

Archaeology and history of Sardinia 252

Genetic differentiation between modern-day Sardinians and Europeans 253

Mitochondrial differentiation among Sardinians 254

Impact of long-term demographic processes on the Sardinian gene pool 255

Biology of prehistoric Sardinians: cranial morphology and mtDNA 256

MATERIAL AND METHODS 258

Archaeological context 258

Ancient DNA extraction 259

Hypervariable-Region I sequencing and coding region GenoCore22 typing 260

Cloning 260

Comparative mtDNA dataset of modern-day Sardinians 260

Coalescent simulations 260

Phylogenetic network 262

RESULTS/DISCUSSION 262

Amplification success and authentication 262

Comparison with modern-day Sardinian haplogroups 265

Comparison with modern-day Sardinian haplotypes 267

Comparison with other ancient Sardinian populations 269

Test for mtDNA continuity between ancient and modern-day central Sardinians 274

CONCLUSION 277

LIST OF SUPPLEMENTARY MATERIALS 277

ACKNOWLEDGEMENTS 278

REFERENCES 278

SUPPLEMENTARY MATERIALS 283

General discussion - Conclusion 287

METHODOLOGY OF ANCIENT HUMAN DNA STUDY 289

Ancient DNA amplification success rates 289

Authenticity of ancient mtDNA data 290

Replication of ancient genetic data 293

Contextual definition of the archaeological sites sampled for ancient DNA 294

ANALYSIS OF ANCIENT GENETIC DATA 295

Analyses based on haplogroup frequencies 296

Analyses based on haplotypic data 298

CONTRIBUTION OF ANCIENT DNA TO THE RECONSTRUCTION OF HUMAN POPULATION HISTORY 300

Genetic diversity of European Palaeolithic/Mesolithic populations and absence of genetic continuity with Neolithic and present-day populations of Europe 303

Mitochondrial influence of eastern Eurasia in eastern Europe 306

Origins of European genetic outliers: Saami and Sardinians 307

The power of ancient DNA 308

CONCLUSION 310

Significance and contribution to knowledge 310

Problems encountered 311

Future direction 311

REFERENCES 317

Abstract

The distribution of human genetic variability is the result of thousand years of human evolutionary and population history. Geographical variation in the non-recombining maternally inherited mitochondrial DNA has been studied in a wide array of modern populations in order to reconstruct the migrations that have participated in the spread of our ancestors on the planet. However, population genetic processes (e.g., replacement, genetic drift) can significantly bias the reconstruction and timing of past migratory and demographic events inferred from the analysis of modern-day marker distributions. This can lead to erroneous interpretations of ancient human population history, a problem that potentially could be circumvented by the direct assessment of genetic diversity in ancient humans. Despite important methodological problems associated with contamination and post-mortem degradation of ancient DNA, mitochondrial data have been previously obtained for a few spatially and temporally diverse European populations. Mitochondrial data revealed additional levels of complexity in the population history of Europeans that had remained unknown from the study of modern populations. This justifies the relevance of broadening the sampling of ancient mitochondrial DNA in both time and space.

This study aims at filling gaps in the knowledge of the genetic history of eastern Europeans and of European genetic outliers, the Saami and the Sardinians. This study presents a significant extension to the knowledge of past human mitochondrial diversity. Ancient remains temporally-sampled from three groups of European populations have been examined: north east Europeans (200 – 8,000 years before present; N = 76), Iron Age Scythians of the Rostov area, Russia (2,300 – 2,600 years before present; N = 16), Bronze Age individuals of central Sardinia, Italy (3,200 – 3,400 years before present; N = 16). The genetic characterisation of these populations principally relied on sequencing of the mitochondrial control region and typing of single nucleotide polymorphisms in the coding region.

Changes in mitochondrial DNA structure were tracked through time by comparing ancient and modern populations of Eurasia. Analysis of haplogroup data included principal component analysis, multidimensional scaling, fixation index computation and genetic distance mapping. Haplotypic data were compared by haplotype sharing analysis, phylogenetic networks, Analysis of the Molecular Variance and coalescent simulations. The sequencing of a whole mitochondrial genome in a north east European Mesolithic individual lead to defining a new branch within the human mitochondrial tree.

This work presents direct evidence that Mesolithic eastern Europeans belonged to the same Palaeolithic/Mesolithic genetic background as central and northern Europeans. It was also shown that prehistoric eastern Europeans were the recipients of multiple migrations from the East in prehistory that had not been previously detected and/or timed on the basis of modern mtDNA data. Ancient DNA also provided insights in the genetic history of European genetic outliers; the Saami, whose ancestral population still remain unidentified, and the Sardinians, whose genetic differentiation is proposed to be the result of mating isolation since at least the Bronze Age. This study demonstrates the power of aDNA to reveal previously unknown population processes in the genetic history of modern Eurasians.

Thesis declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution to Clio Der Sarkissian and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue, the Australasian Digital Theses Program (ADTP) and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Clio Der Sarkissian

Date:..03./06./11..

Acknowledgments

This work would not have been possible without the financial, technical and intellectual support of The Genographic Project, in collaboration with the National Geographic Society, IBM and the Waitt Family Foundation.

I am very grateful to my principal supervisor, Alan Cooper, for giving me the opportunity to work at the ACAD. I feel very privileged to have been granted access to valuable ancient human remains and to the outstanding work conditions of the ancient DNA laboratory. I would like to thank him for his support and for showing me what it means to be passionate and enthusiastic about science.

I would like to show my gratitude and admiration to my supervisor, Wolfgang Haak, who has taught me so much about ancient human populations and clean ancient DNA laboratory work. I would like to thank him for his wise advice, but also for his patience and understanding in the course of these three years. Danke vielmals, Wolfgang!

I would like to thank my postgraduate coordinators, Robert Reid and John Jennings for their crucial assistance with administrative issues.

I am deeply indebted to all the collaborators who participated in this work. In particular, I would like to thank Oleg Balanovsky, Elena Balanovska, Valery Zaporozhchenko, Guido Brandt, Kurt Alt, Andrew Clarke, David Soria, Carles Lalueza-Fox, Oscar Ramirez, Robin Skeates, Giuseppina Gradoli, Antonio Torroni, Anna Oliveri, Maria Pala, David Caramelli, Alessandra Modi and Martina Lari for their very precious help.

I am particularly grateful to Jeremy Austin for all his help and for the critical review of my work.

I am heartily thankful to Maria Lekis for her administrative help, her friendship and for making me feel at home in Adelaide.

I thank all my past and present fellow co-workers of the Darling Building and the ACAD for their technical and intellectual support. The ACAD has been a pleasant and friendly (and clean!) work environment thanks to all of them. I will also remember fondly of the ACAD events and I feel very fortunate to have developed friendships with so many of my colleagues.

I would like to address very special thanks to my friends in Adelaide: Bastien, Camille, Damien, Doreen, Emma, Gaynor, Grant, Jessica, Julien, Kimiko, Manue, Marjorie, Valentin, and Virginie who have been a real family to me.

I would like to thank Claire C., Claire M., Fanny, Magali, and Thomas for their patience, their support, their advice, and for remaining my friends despite the distance.

Thanks also to my aunt and cousins for their great support during the very last stages of the writing of this thesis.

I owe my deepest gratitude to my family and in particular to my parents Jutta and Yves and my brother Antoine.

À Claire,