# Cultural Integration: Genetics, Archaeology and the impact of the Viking Diaspora on the Isle of Man

**Thesis submitted for the degree of**

**Doctor of Philosophy**

**at the University of Leicester**

**By**

**Hayley Dunn MSc (UCL)**

**School of Archaeology and Ancient History**

**& Department of Genetics**

**University of Leicester**

**January 2015**

# Cultural Integration: Genetics, Archaeology and the impact of the Viking Diaspora on the Isle of Man

Hayley Dunn

## Abstract

Over the last two decades many in the archaeological community have developed a degree of scepticism and suspicion towards using genetics to study the past. My project aims to dispel some of these concerns, highlighting the inferences about demographic and social history which can be made from genetic data gathered from modern populations. Centred around the British-Irish Isles and in particular, the Isle of Man, this study investigates for the first time, differing demographic histories in the male and female gene-pools.

The focus of this research is the 400 years of Norwegian rule in the Irish Sea region following the Viking visitations of 900CE, when the Isle of Man was an important centre in the politics of the Irish Sea kingdom. DNA samples were collected from male volunteers, followed by the extraction of high resolution mtDNA and Y chromosome data. Volunteers were enlisted on the basis of surnames taken from 16$^{th}$-century documentation to ensure deep historical roots on the Island. The high quality genetic data generated for the Isle of Man not only gave greater distinguishing power for looking at closely related populations, but also allowed separate male and female population histories to be explored.

Approximate Bayesian Computation (ABC) modelling was used to explore the

demographic history of the Isle of Man and how the impact of the Norwegian diaspora differed for men and women, providing a powerful statistical and probabilistic approach to admixture analysis. Whilst drawing evidence from history, archaeology and incorporating genetic data to provide indications of where the Island draws its genetic influences from, this project provided a case study for how the fields of archaeology and genetics can be better integrated for the exploration of the past.

# ACKNOWLEDGEMENTS

It takes a team to complete a PhD and I'm indebted to a great number of people who have helped and supported me through this process. Firstly I would like to thank my supervisors, Prof. Mark Jobling and Prof. Simon James for taking an active interest, not only in my work, but in my well being. Their comments and suggestions have been invaluable in shaping this thesis. I feel lucky to have had such kind and considerate mentors to guide me through this process.

Special thanks are extended to the Department of Genetics and the School of Archaeology and Ancient History for providing the necessary funding for this project. I would also like to thank my examiners Dr Ed Hollox and Dr Stephan Shennan for their thoughtful questions and insights, which made my viva an enjoyable experience.

Thanks also goes to Dr. Turi King and Dr. Rita Rasteiro, without whom I would have been unable to carry out the technical aspects of this work. Thank you for generously sharing your knowledge. I would also like to extend my gratitude to Prof. Jo Story and the rest of the Diasporas team. Their insights and alternative perspectives were essential for achieving clarity in my work.

I am deeply grateful to all the students and academic staff of the SAAH over the years, who have helped me consider my work from an archaeological view-point. In particular, Julia Nikolaus, Dr. Dave Edwards, Dr. Sophie Adams, Dr. Chantal Beilmen, and Sergio Gonzalez Sanchez for their friendship and the long discussions in the Marquis.

I would also like to thank the members of Mark Jobling's and Ed Hollox's labs, especially Gurdeep Lall, Dr. Chiara Batini, Dr. Pille Hallast, Dr Daniel Zadik and Dr.

Barbara Ottolini for their advice and generosity. I'm am so grateful that they always made me feel welcome, despite my somewhat unusual circumstances which often kept me away from the lab.

Completing this PhD would have been a lonely business if it was not for my adopted Manchester family, who helped me to juggle the roles of new mother and research student, without loosing my marbles. A million thank-yous to Tasha, Leona and Sarah for ensuring my precious boy was always well cared for, enabling me to concentrate on higher things. Amongst others and in no particular order, to Emma, Viv, Rich, Lianne, Daisy, Kat, Niki and Cathy, who helped me to managed my stress levels in many fun and varied ways; thanks for believing in me, I love you guys. For all those others who listened when all I could talk about was PhDs or babies, I thank you.

Matt, my partner, co-parent and rock, whose steadfast support, patience and understanding provided the secure base from which I could carry out my work, I am forever grateful for his love during this journey. Finally, Freddie, my unexpected gift. Despite the apparent inconvenient timing of his arrival, Freddie's spirit and affection not only kept me grounded but his presence in my life gave me the fresh perspective needed to maintain the passion and drive I required to complete this work. I feel very blessed.

# TABLE OF CONTENTS

***Cultural Integration: Genetics, Archaeology and the impact of the Viking Diaspora on the Isle of Man*** ***1***
***Thesis submitted for the degree of*** ***1***
***Doctor of Philosophy*** ***1***
***at the University of Leicester*** ***1***
***By*** ***1***
***Hayley Dunn MSc (UCL)*** ***1***
***School of Archaeology and Ancient History*** ***1***
***& Department of Genetics*** ***1***
***University of Leicester*** ***1***
***January 2015*** ***1***
**Cultural Integration: Genetics, Archaeology and the impact of the Viking Diaspora on the Isle of Man 2**
Hayley Dunn 2
Abstract 2
**Acknowledgements 4**
**Chapter One: Introduction and Research aims 8**
1.1 Introduction 8
1.2 Research Background 9
1.3 The Case Study 11
1.3.1 Materials and Methods 13
1.4 Chapter Outline 15
1.5 Terminology 17
1.6 Scope and assumptions 19
**Chapter Two: Population genetics and population history 21**
2.1 Introduction 21
2.2 Essential Genetics 22
2.2.1 Genetic diversity 23
2.2.2 Mitochondrial DNA (mtDNA) 28
2.2.3 Y Chromosome (NRY) 32
2.3 Description and inference 39
2.3.1 Classical Population Genetics 40
2.3.2 Phylogeography 42
2.3.3 Modelling and Simulations 44
2.4 Historical relationship between Genetics and Archaeology 48
2.4.1 In the beginning 50

2.4.2 Creating Historical Narratives 54
2.4.3 New approaches 57
**Chapter Three: The Isle of Man in context 64**
3.1 Introduction 64
3.1.1 Geography and demography 65
3.1.2 Language and Place-names 67
3.2 Archaeological context 69
3.2.1 Initial colonisations 70
3.2.2 The transition to a farming economy 73
3.2.3 Metal Ages and exchange networks 77
3.2.4 Christians and Pagans 83
3.2.5 The Post-Norse era 93
3.3 Genetic context 95
3.3.1 British and Irish Isles 96
3.3.2 Impact of the Neolithic Transition 102
3.4 Summary and conclusions 106
**Chapter Four: Data Sources and Methodologies 108**
4.1 Introduction 108
4.2 Data Collection 109
4.2.1 Background to sampling strategy 110
4.2.2 Criteria for Surname Selection 111
4.2.3 Volunteer Recruitment 114
4.2.4 Comparative NRY Data Sets 117
4.4.2 Comparative mtDNA Data Sets 117
4.3 NRY DNA Purification and Typing 121
4.3.1 Introduction 121
4.3.2 DNA Extraction 123
4.3.3 STR Typing 124
4.3.3 SNP Typing 127
4.4 Data Analysis 131
4.4.1 Introduction 131
4.4.2 Classical Population Genetics 133
4.4.3 Phylogeographic Analysis 135
4.4.4 Computer Modelling 136
**Chapter Five: Results 146**
5.1 Introduction 146
5.2 The Genetic landscape 146
5.2.1 NRY 146
5.2.2 mtDNA 159
5.3 Hypothesis Building 167
5.3.1 NRY 167
5.3.2 mtDNA scenarios 173
**Chapter Six: Discussion and conclusions 179**
6.1 Introduction 179

6.2 Discussion 179
6.2.1 The genetic landscape of Atlantic Europe 180
6.2.2 Demographic History of Irish Sea populations 183
6.2.3 The Scandinavian Diaspora and the Isle of Man 187
6.2.4 Genetic data for exploring the past 191
6.3 Conclusions 193
6.4 Future prospects 195
**Appendix A 198**
**University Road 201**
***Department of Genetics 202***
**University Road 202**
**CONSENT FORM 202**
***Department of Genetics 203***
**University Road 203**
***WHITE 205***
***ASIAN, including 205***
**Asian English, Asian Scottish, Asian Welsh, Asian British 205**
***BLACK, including 205***
**Black English, Black Scottish, Black Welsh, Black British 205**
***MIXED 205***
**Caribbean 205**
**D1 205**
**Chinese English, Chinese Scottish, Chinese Welsh, Chinese British 205**
***(Please specify) 205***
***OTHER ETHNIC GROUP 205***
***(Please specify) 205***
Appendix A4 Replace this sheet with PDF of Project information sheet (1 page) 206
replace this sheet with results certificate 213
Replace this sheet with info sheet 1 214
Replace this sheet with info sheet 2 214
**Appendix B 221**
**Glossary 245**
**Bibliography 247**

# CHAPTER ONE: INTRODUCTION AND RESEARCH AIMS

## 1.1 Introduction

Genetics offers archaeology an exciting new tool to explore the past, providing a fresh perspective and a unique source of data to test current archaeological theories against. It is notoriously difficult to draw conclusions about ancient migrations and population demographics from archaeological data but genetics has the potential to directly investigate these types of questions and more (Shennan 2008 173-175; Steel 2008, 177-178; Zvelebil and Zvelebil 1988). Genetics can also offer insight into the different demographic histories of men and women, providing clues about cultural mating practices such as polygamy versus monogamy (Rasteiro and Chikhi 2013), hypergamy (Thomas et al. 2006) and patrilocality (Marks et al. 2012; Seielstad et al. 1998).

However, the integration of genetic data into archaeological theories about the past has not been straightforward. The reasons for this are complicated, but cultural differences between the disciplines have played a major role in preventing an easy union between the fields. Genetics, firmly rooted in the sciences with its emphasis on data, generalised models and patterns, has often been at odds with humanities-based archaeological theories which argued for the uniqueness and complexity of each context (Pluciennik 2006; Renfrew 2010; Thomas 2006).

Furthermore, the exchange of ideas has been limited by radically different publication preferences. Archaeologists tend to write books or publish in monographs, whereas geneticists prefer to publish in peer-reviewed journals (Jobling 2012). A lack of systematic collaboration and communication has

resulted in little sympathy or understanding of the developments in each other's fields.

This thesis aims to bring the fields of archaeology and genetics closer together, through the examination of the Norse impact on the Isle of Man. So by using genetics to gain a better understanding of the integration of Norse and Irish Sea cultures, it will help integrate genetics into archaeological analysis. Incorporating archaeological theories into genetics models creates a more holistic approach to inferring the past, by playing to each discipline's strengths. Furthermore the problems of circularity and over-interpretation which have often plagued previous studies, can be avoided by testing correlations and hypotheses within a statistical framework designed to incorporate data and theories from both fields.

## 1.2 Research Background

Until recently genetic data have mostly been used to look at large-scale population movement and general patterns of behaviour change over time, ignoring regional contexts. The earliest works attempting to find correlations between genetics and ancient human cultural changes were necessarily on a grand scale due to technological constraints (e.g. Cavalli-Sforza et al. 1996). However, contemporary archaeologists were growing more interested in exploring fine-scale regional variations in human culture (Thomas 2006).

The infancy of methodologies designed to extrapolate population history from population genetic data has meant that much of the published work to date, has focused on the Palaeolithic and Neolithic Ages, when populations were smaller, and the impact of demographic changes were most dramatic. Less work has been done on the later events of the Holocene (the last 10,000 years), when

population dynamics were much more complex (Renfrew 2010). The exploration of later periods require larger, more localised, higher-resolution data sets, which have not been available until recently.

These difficulties have resulted in some archaeologists feeling a certain amount of scepticism about the value of genetic data for answering the questions that concern them, questions relating to the way that past peoples lived, interacted and moved within their own communities and beyond, particularly during post-Neolithic periods (Brown and Pluciennik 2001). On the other hand, most geneticists have continued to work on their projects, apparently little concerned with current archaeological debates or theories on relevant time periods (Pluciennik 2006).

However, genetics and archaeology have complementary strengths and weaknesses. Genetic data can provide direct evidence of past migrations and social structures, which are otherwise invisible to archaeologists. The problem is that patterns found within gene-pools could have multiple causes, and are difficult to date. Archaeological material culture, on the other hand, can indicate when and where contacts between cultures could have occurred but cannot reveal the exact nature of these interactions (Shennan 2008, 173). By working together, a clearer picture of past demographic events and social interactions is possible.

Furthermore, advances in molecular technology have increased the resolution and volume of genetic data, enabling closely related populations to be examined in more detail. Here, I use the Isle of Man as a case study, to demonstrates how regional, high-resolution genetic data can be used in tandem with archaeology to inform theories about the past, for a relatively small geographical area, such as the Irish Sea. In particular, the Viking Age Norwegian diaspora is an

archaeologically and historically attested migration period in this region, which can be tested against genetic data. It provides evidence that genetic methodologies in combination with archaeology has the power to identify movements between genetically distinct populations, as well as elucidating sex-specific behaviours.

## 1.3 The Case Study

The focus of this research is the Isle of Man, located in the Irish Sea between Great Britain and Ireland. The island's location has periodically made it of importance as part of ancient and more recent exchange networks, which moved valuable items around the Irish Sea and beyond (reviewed in Cunliffe 2001, 34, 59, 211, 505-506). However, it has been largely free of the most recent European influences which have been important in Britain's history, such as the Roman, Anglo-Saxon and Norman occupations.

Previous genetic studies of the British-Irish Isles noted that western regions of this archipelago are genetically similar to each other but distinct from eastern parts, especially to England (Capelli et al. 2003; Hellenthal et al. 2014; Weale et al. 2002). It is not clear whether this pattern is the result of a long-standing shared population history among western regions which is distinct from England, or whether the difference between the regions is due to the more recent European influences which affected eastern Britain but not the western parts (Pattison 2011 2008; Thomas et al. 2008; Weale et al. 2002). A detailed examination of the Isle of Man provides another reference point from which to explore these questions.

As well as considering the genetic data of the Isle of Man within the wider

context of the British-Irish Isles and western Europe, this research includes detailed examination of the impact of the Norwegian diaspora on the island during the Viking Age. Records of Norwegian contact with Irish Sea communities began in the 8th century and lasted until the middle of the 13th century (reviewed in McDonald 2008).

Initially taking the form of raiding parties, these excursions later became migrations. There is also evidence of significant blurring of ethnic boundaries throughout the Irish Sea region, through inter-marriage, creating a mixed Norse and Gaelic population known locally as *Gall-Gaedhil* meaning 'foreign Gael' (Crawford 1987; Smyth 1977). A contemporary commentator complained that British locals were copying the Scandinavian hair-styles and dress (Æfric 1000, trans. Mary Clayton 2002), while archaeological excavations of putatively Scandinavian domestic sites and graves around the British and Irish Isles revealed objects of both Norwegian and local designed and manufacture (Bersu and Wilson 1966; Crawford 1987).

On the Isle of Man, a wealth of Norwegian influence is revealed through archaeological evidence (reviewed in Wilson 2008) and numerous stone carvings and crosses bearing Scandinavian art and inscriptions (Kermode 1994) There is a legacy of Norse place-names (Moore 1890), as well as Norse activities in the area being well documented in contemporary historical texts (Munch 1874).

There is some existing genetic evidence of Norwegian ancestry in the male lineages (Bowden et al. 2008; Capelli et al. 2003) but the female lineages have never been examined. Estimates for Norwegian admixture vary depending on the methods used and the single data-set which has been used to make these calculations (Capelli et al. 2003) is considered of low resolution by today's standards. Also this Norwegian genetic influence is always assumed to be from

the Viking Age, rather than earlier or later undocumented migrations, with little way of testing which is the case. This research partially addresses this, by using modelling to estimate when the Norwegian lineages arrived.

Archaeological and historical evidence points to a male Viking presence on the Isle of Man from at least the 10th century (Wilson 2008, 25-55). However, the extent and longevity of the Viking colonisation is unclear, as is whether the male Viking migrants were of Norse descent rather than Irish, Scots, or Manx who had adopted the Viking lifestyle and culture. Nor do we know if there was a significant influx of Norwegian women, or perhaps women from elsewhere around the Irish Sea. This research used a model-based approach to explore these questions, examining both male and female lineages at higher resolution than previously available. In doing so, it describes a more detailed population history for the Isle of Man than has been previously attempted using genetic data.

### 1.3.1 Materials and Methods

This research focused on the uni-parentally inherited markers, mitochondrial DNA (mtDNA) and the Y chromosome (NRY), to enable the comparison of female and male genetic patterns respectively and to investigate the effect that local and more distant populations may have had on the island's gene-pool. I collected DNA samples from men living on the Isle of Man. To avoid the clouding effect of more recent migrations to the island, I carefully selected individuals with ancestry going back to at least the 16th century, to achieve a deeply rooted population sample. Their male-defining NRY was typed in Prof. Mark Jobling's laboratory by Dr Turi King, although I extracted DNA and typed STRs for about a third of the samples.

MtDNA haplotypes were analysed for a subset of these individuals who had two generations of maternal ancestry on the island; this work was carried out by a student as part of her Master's thesis (Boquete Vilariño 2012). I gathered comparative samples with a similar degrees of compatible resolution, from the literature, for populations within Great Britain and Ireland, as well as Spain, Portugal, France, Netherlands, Denmark, Sweden and Norway. Most of these comparative populations had also been sampled on the basis of two generations of ancestors born within the targeted area.

The genetic markers were analysed to a much higher resolution than previously used for this region, which enabled a detailed examination of the Isle of Man's population history. The NRY samples were screened for 17 Y-filer STRs and for 20 haplotype-defining SNPs, which included sub-division of the high frequency R1b haplogroup. Likewise, the entire control region (>1000 base pairs) of mitochondrial DNA was sequenced. The higher resolution of the data allowed a greater degree of differentiation between the closely related populations of the Irish Sea region than has been previously possible.

I used a number of tools to describe the genetics of the Isle of Man and compare it to other populations within the Atlantic region of western Europe. Genetic distance measures were calculated and converted into spatial distances using multidimensional scaling (MDS) which allows the genetic landscape to be visualised. Network analysis was used to examine the relationships between closely related lineages shared between the populations. These tools highlight patterns and reveal affinities between populations to provide clues about influences on the Isle of Man gene-pool.

Computer modelling and simulations were used to shed light on the extent of male and female Norwegian genetic input into the island during the Viking Age.

Using Approximate Bayesian Computation (ABC), I explored a range of potential admixture scenarios drawn from archaeology and history to be tested in light of the genetic data. The sex-specificity of the inheritance of NRY and mtDNA enabled different population histories for men and women to be inferred, providing insights into sex-mediated migration patterns and other sex-specific behaviours which could have impacted on the gene-pool of the area.

## 1.4 Chapter Outline

The first half of Chapter 2 explains the basic principles underlying population genetics theory. The regions of the genome used to carry out this research are introduced along with the methodologies used to infer population history from these genome regions. The second half provides an overview of the history of population genetics for population history inference. This section describes attempts to integrate genetics into archaeology, discussing cultural differences and technological issues which have historically hindered an easy synthesis, and how these are being overcome.

Chapter 3 introduces the case study, describing the history of the Isle of Man as it is understood from archaeology, history and linguistics. It also describes previous investigations which have explored the genetic landscape of the Irish Sea and surrounding regions, and the current understanding of the population history of the area including the Isle of Man, based on this research.

The methods and methodologies which were used to obtain, extract and analyse the genetic material and used to infer the history of the Isle of Man are described in Chapter 4 in three sections. The first section details the Isle of Man sampling strategy, including how DNA donors were selected and recruited for

their Manx ancestry on the basis of their surnames. Details of comparative data sets used are also listed in this first section. DNA extraction, sequencing and typing methods are described in the second section. The third section details the descriptive, analytical and inferential methods used for comparing and explaining the genetic data.

Chapter 5 contains the results of the various analytical methods used to describe and explain the patterns found in the genetic data of the Isle of Man and surrounding populations. This chapter is divided into two parts, the first covering the results of descriptive analyses, revealing the genetic landscape of Western Europe and the context of the Isle of Man. The second part details the findings of computer modelling through simulations used to test various demographic scenarios for both the NRY and mtDNA, for Irish Sea and Scandinavian populations.

Chapter 6 contains a discussion of the results in light of the archaeological evidence discussed in Chapter 3 and paying close attention to the issues discussed in the second half of Chapter 2. The discussion is split into four parts. The first discusses the genetic landscape of Atlantic Europe and the Isle of Man place within this picture. The second looks at the demographic history of the Irish Sea region, while the third looks at the impact of the Scandinavian diaspora upon the region. The fourth is concerned with the insights gained through this research and its implications for the integration of genetic data into archaeological theories about the human past and population history.

Chapter six contains a summary of the main findings of this research, detailing the original contribution that this has to made to knowledge of the population history of the Isle of Man and the Irish sea region. This chapter also describes the potential future prospects for expanding and improving this research and how

its methods may be adapted and applied to other regions and temporal periods.

## 1.5 Terminology

This thesis is unique at the present time, combining knowledge and data from both archaeology and population genetics, and requires an understanding of the technical terms specific to each field. To this end I have taken care to explain the meaning of terms in the text, but for convenience have also included a glossary in the appendix section. However, there is some ambiguous terminology which needs defining here to avoid misunderstandings later on.

The labelling of the geographical region of interest to this thesis is not as simple as one might first assume. There are historical and political divides which mean that the commonly used names actually exclude important areas under investigation. 'The UK', the common abbreviation for the United Kingdom of Great Britain and Northern Ireland (OED Online 1924), refers to the political unit which does not include the Republic of Ireland or the Isle of Man (Davies 1999, xxvii- xlii; Oxford Dictionaries Online 2014), for these reasons the term is not used in this thesis.

According to the Oxford English Dictionary, 'The British Isles' is the geographic term used to describe the island of Great Britain or simply, Britain, which is comprised of the countries England, Scotland and Wales, and associated smaller islands, including Ireland and the Isle of Man (OED Online 2008). However, the Government of Eire disagrees, stating:

> The British Isles is not an officially recognised term in any legal or inter-governmental sense. It is without any official status. The Government, including the Department of Foreign Affairs, does not use this term. Our officials in the Embassy of Ireland, London, continue to monitor the media in Britain for any abuse of the official terms as set out in the [406]

> Constitution of Ireland and in legislation. These include the name of the State, the President, Taoiseach and others. (Ahern 2005).

So, with this in mind, I used the term 'British and Irish Isles' to refer to Britain, Ireland and all the associated islands of this archipelago. 'Britain' is used to indicate the landmass incorporating England, Scotland and Wales, and 'Ireland' refers to the entire island incorporating both Northern Ireland and the Republic of Ireland. Further, the people of the Republic of Ireland or Eire, and Northern Ireland are refered to as 'Irish', and peoples of the Britain and associated islands, (excluding Ireland) are refered to as British (Davies 1999, xxvii- xlii; OED 2014).

Other terminology which needs defining is 'Norse' and 'Vikings'. 'Viking' is the name given to the Scandinavian raiders who terrorised the coastlines of the British and Irish Isles from the late 8th century onwards. The following centuries of Scandinavian rule have become known as the Viking Age (Richards, 2007). In this thesis, the term 'Viking' to refers to a particular set of cultural traits of Scandinavian origin carried by people of uncertain, possibly mixed ancestry. 'Norse' which derives from 'Northmen' refers specifically to those of known Viking Age Scandinavian heritage, whereas lineages which are traced to Norway without any temporal certainty are simply 'Norwegian'.

When referring to dates, the secular BCE/CE is used in place of the Christian BC/AD convention. BCE stands for 'Before Common Era' and equates to BC (Before Christ) whereas CE, 'Common Era', replaces AD, 'Anno Domini', which means 'in the year of our lord'. The dates, however, remain the same, so the year of writing this thesis would be 2014CE, rather than AD2014. This system is widely used by non-Christian scholars and I considered it more appropriate for use in our secular society.

## 1.6 Scope and assumptions

As a PhD project it was not possible to collect genetic data from all the locations that may have been relevant to this research. Instead I had to rely on published data that were available. Unfortunately the published data for the British Isles are of very low resolution or use genetic markers which are incompatible with my data and so can only be used for very basic comparisons. The Irish data included high resolution STR data, but lacked geographical and phylogenetic resolution. This meant that some of the more detailed comparisons between the Isle of Man's nearest neighbours were beyond the scope of this project. However, this is reflected in the type of inferences that it was possible to make with the available data.

It is important to note that since modern genetic data are used as a proxy for ancient data, there is an assumption that populations have not changed dramatically in the intervening years. Smaller populations, and in particular island populations, tend to be more affected by random processes which change the genetics of a population over time. A modelling approach overcomes these issues to a degree by taking into account these stochastic processes.

However, models tend to be overly simplistic because even simple models involve large numbers of parameters. Also, given time constraints, I was only able to test a limited number of likely scenarios. The models examined here, were unable to account for background migration between the regions under examination, which would undoubtedly have occurred. This is not to say that these models cannot tell us something about the population's history. This work provides the foundations for later work which can use more complex models to build upon these findings to develop more realistic scenarios, which fit the data better.

The exceptionally fast pace of technological development in genetics means that at the time of writing, some of the methods used in this research were already beginning to look a little out-dated. The price and speed of sequencing means that it is increasingly common to look at the entire human genome to estimate admixture and infer population history. This field, called population genomics, compares the genomes of individuals across populations. Large blocks of similarity carried by individuals between populations can indicate shared population history and admixture. Much more information about a person's ancestry, and therefore a population's history, can be gained from much smaller data sets. However, it was beyond the scope of this project to begin this type of investigation.

It was also beyond the scope of this project to investigate DNA from ancient sources. Although technology has improved the reliability of results obtained from ancient DNA (aDNA), it is still a relatively rare and expensive data source. The British and Irish Isles seems to be particularly poor in aDNA data, probably due to the damp climate and acidic soil types which can inhibit DNA preservation. However, aDNA data are likely to become more abundant in the future, and will provide an alternative source of information which can be used to better inform model parameters to build upon this research.

# CHAPTER TWO: POPULATION GENETICS AND POPULATION HISTORY

## 2.1 Introduction

Just as an individual's genome contains information about their ancestry, the collective DNA of a population can reveal something of its history, if one knows where and how to look. Scientists in the early 20th century began to recognise this, as they attempted to reconcile Mendel's work on inheritance with Darwin's theories of evolution. In the 1930s and 1940s, the forefathers of population genetics designed mathematical models to predict the forces shaping evolution through the changing patterns of heritable traits in populations over generations (Fisher 1930; Wright 1943 1943). These models still form the basis of modern population genetics today (Charlesworth and Charlesworth 2009).

It is only with technological advances of the last 30 years that the forces predicted by these early models could be measured through detailed observations at the molecular level. The molecular age has allowed particular regions of the human genome to be targeted and compared between individuals, identifying differences that have no physical effects on the carrier (Crow 2000). Advances in computing enabled researchers to compare molecular data against modified versions of the original population genetics models, to begin to make inferences about the events which may have shaped modern gene-pools (Cox et al. 2008; Matsmura and Forster 2008 1-2).

It was in light of these new possibilities for exploring human prehistory that the term 'archaeogenetics' was coined (Amorim 1999) and Colin Renfrew optimistically wrote about the burgeoning field as a ‘new synthesis’ of

archaeology and genetics (Renfrew 2000a, 2000b). However, the fusing of the fields has not been simple. Relationships forged between geneticists and archaeologists, until recently, have been largely episodic in nature. The tools for reading the past from genetic data have been slow to develop, and sometimes researchers may have been too quick to jump to conclusions based on perceived correlations between genetic and archaeological data. Archaeologists too, have been impatient with the 'broad-brush' approach to explaining genetic patterns over large areas (Pluciennik 2006a; Thomas 2006a) and sceptical about the benefits of genetic data for providing any useful insight into the intricacies of human interactions at the regional level.

The third part of this Chapter focuses on efforts to pool resources and knowledge between geneticists and archaeologists, in the development of this new field. I will describe how difficulties initially arose due to cultural differences between the fields, relating to their approach to understanding the past. Within the disciplines too, debates about how best to interpret their own data have hindered cross-disciplinary interaction. However, I believe we are entering a new phase, where philosophical and methodological developments are finally making for a viable synthesis between archaeology and genetics. Paramount to the successful integration of genetics and archaeology is a deeper understanding of each other's subject matter. To this end, I will begin by explaining some of the basic concepts and models underpinning population genetics theory.

## 2.2 Essential Genetics

The study of population genetics is essentially the study of genetic diversity. Most tools and methodologies developed within the field are designed to identify, measure and describe relative population diversity. Genetic diversity

changes over time in response to particular forces, and these forces in turn can be influenced by demographic processes and social behaviour. Through understanding the interplay of forces and processes, and their effect on patterns in genetic diversity, inferences can be made about population history.

I will begin this section by describing what genetic diversity is, why it is useful and how it is identified and measured. I will introduce the five different forces, describing the way they effect diversity through their interactions with each other and how this is impacted by human behaviour and demographic changes. The following sections introduce the regions of the genome which, due to their special properties, have become popular tools for exploring these processes through examining diversity patterns in modern gene-pools.

### 2.2.1 Genetic diversity

Genetic diversity measures the degree to which individuals differ to one another, within and between populations. The patterns of genetic variation seen in a population's gene-pool can reveal something about its history, if the processes affecting diversity levels can be disentangled and understood. However, despite occupying a wide range of environments, ranging across the entire planet, human genomes are remarkably similar to each other, differing at only about 1/1000 sites (Li and Sadler 1991) Even when individuals are from very geographically distant locations the average difference between them is only slightly higher than for individuals from the same population (Li et al. 2008). So it is important to identify regions of the genome where this limited diversity is greatest, for population studies.

Typically, regions of the genome that code for important proteins are protected by natural selection so vary little within species and sometimes are even

conserved between species (Bejerano et al. 2004). Random changes to regions coding for important proteins are more likely to render them non-functioning or pathological than to be an improvement. Through natural selection these deleterious mutations are removed or kept at low frequency in the gene-pool because bearers tend to fail to pass them on to the next generation due to reduced ability to reproduce or premature death (Crow 2000). Fortunately, genomes also contain vast swathes of DNA which is non-functional, and these regions tend to accumulate the highest levels of diversity (Chakravarti 1999 1998; Nickerson et al. 1998).

A new mutation may arise in a single person through faulty DNA replication, but over generations, under certain circumstances, the numbers carrying this particular mutation may increase. If people with the mutation migrate, they carry it to other populations. Changing patterns in frequency and distribution of new mutations over time and space are ultimately dependent on the number of descendants the original bearer has and how they behave. The relative frequencies between different variants is understood through identifying and measuring five processes: mutation, recombination, natural selection, migration and genetic drift (Tishkoff and Verrelli 2003).

Of the five processes, drift and migration are the most affected by demography and social behaviour, so are of particular interest to those interested in population history. It is important to note that migration need not equate with mass movements of people. In genetics, migration can simply mean gene-flow, which is actually the result of migrants passing their genes on to the next generation in their new population. Over longer time frames it is possible to see how genes spread without individuals needing to move very far at all. Over the course of several generations genetic variants can move, disperse and increase

in frequency through the action of drift, without the need for large-scale movements of people (Wakeley 1999; Wright 1943).

Genetic drift is the name given to the stochastic way that a variant behaves over time, due to reproductive chance. It is highly dependent on population size, for example new variants tend to disappear faster in smaller populations, decreasing genetic diversity. It is important to note that it is the effective population size ($N_e$) that is important to this process, rather than the census size, which may be far larger. $N_e$ includes only the number of people within a population that are reproducing (Wright 1931), so excludes the very old and the very young.

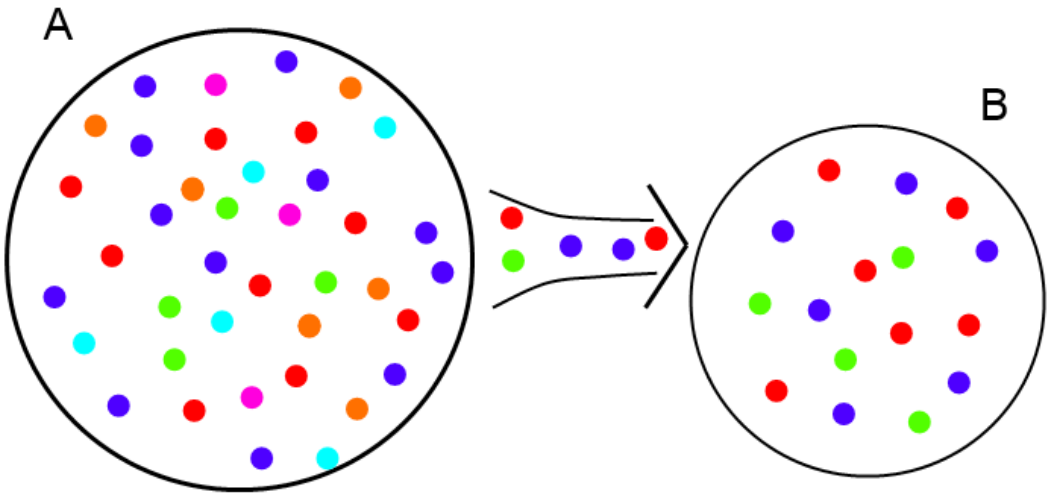



*Figure 2.1: Simplified diagram of founder effect. The large circles represent populations and the coloured dots represent individuals. A few individuals leave population A and form population B, so B has only a subset of A's diversity.*

For example, chimpanzees have a census size of 200,000 (Jobling et al. 2013, 245) and an effective population size of 35,000 (Tishkoff and Verrelli 2003). However, the worldwide effective population size for humans has been calculated as approximately 10,000 despite the census size being nearer to 7 billion (Takahata and Satta 1997). This is thought to be due to a relatively recent

worldwide expansion out of Africa from a small number of founders (founder effect, see Fig. 2.1; Cox et al. 2008). Similarly, population bottlenecks (i.e. a rapid decline followed by renewed expansion), for example due to disease, can also have long-lasting effects on the effective population size (Hunley et al. 2009).

Drift is also sensitive to social factors, if they affect the number of children produced and by whom. In this way, changes in cultural behaviour affecting reproduction can leave lasting patterns in a population's genetic diversity (François et al. 2010). For example, in polygynous populations, relative reproductive success is dimorphic between the sexes. Males have a smaller effective population than females because a few males reproduce with many females. By contrast in monogamous societies, one would expect males and females to have approximately equal effective population sizes (Dupanloup et al. 2003). In this way social structure also impacts on genetic diversity.

If demographic and social impacts on a population are to be understood, confounding factors that impact on a population's diversity but are not related to demographic or social factors should be avoided or accounted for in some way. Two of these processes are natural selection and recombination. Natural selection can work in several ways: decreasing diversity by removing new deleterious mutations (purifying selection) or increasing beneficial mutations to fixation (carried by 100% of the population), or maintaining diversity through balancing selection. It is important to remember that natural selection also impacts on areas of DNA close to the region under selection, removing or increasing their frequency in a gene-pool through a 'selective sweep'. In this way the effect of natural selection can mimic the effect of, for example a range expansion from a small number of individuals (Currat et al. 2006).

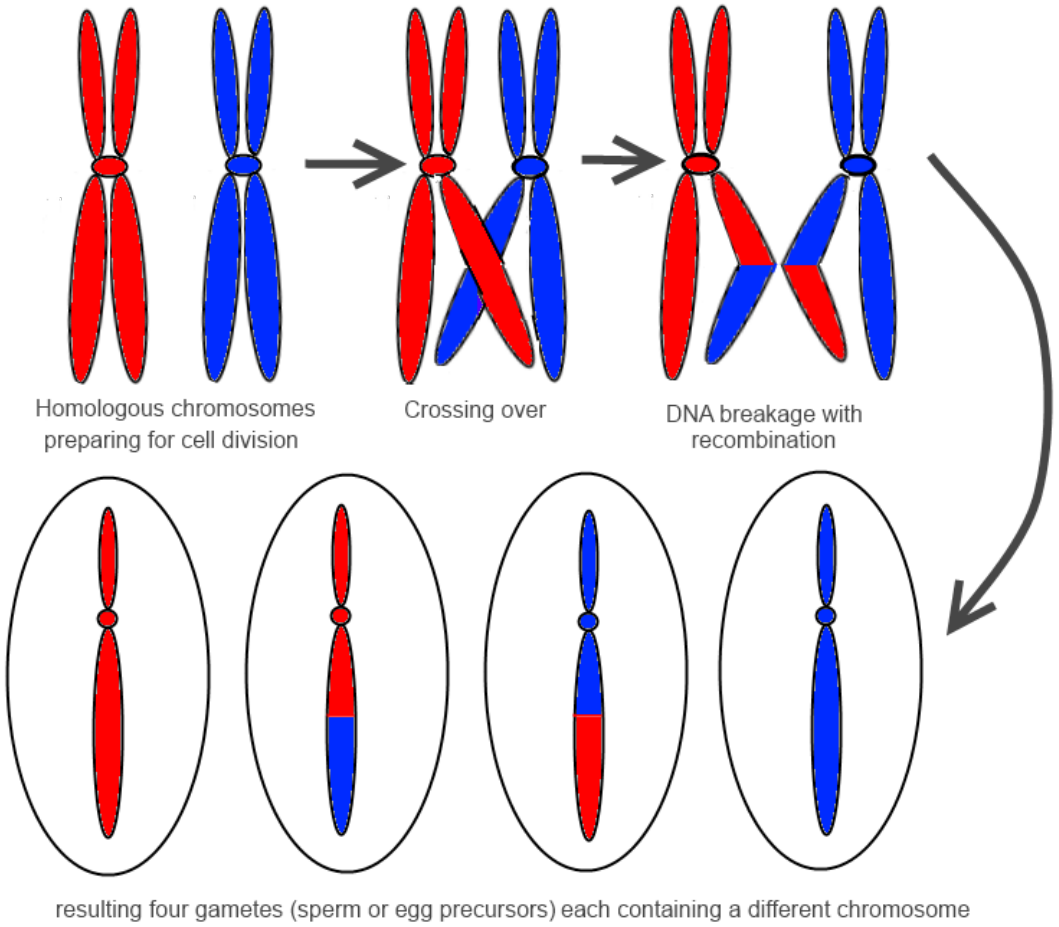



*Figure 2.2: Recombination. At meiosis the homologous chromosomes replicate and form matching pairs. These matching chromosomes then undergo recombination. Following recombination the cell divides producing haploid sperm or ovum ready for reproduction.*

Recombination is the process of chromosome shuffling which helps to maintain diversity within a species (see Fig. 2.2). Recombination occurs between the matching pairs of chromosomes in the germ-line cells before they divide to become spermatozoa (sperm) or ova (eggs). It involves the reshuffling of the genetic material between the pairs of homologous chromosomes, so that each sperm or egg receives one chromosome which is a unique mixture of genetic material from both chromosomes. This reshuffling of material can make it difficult to directly trace particular lineages back in time. However, the impact of recombination is not uniform across the genome; there are hotspots as well as

zones where it is rare (Lercher and Hurst 2002).

Fortunately, by carefully choosing specific regions of the genome, recombination and natural selection become less of a problem, because some regions do not undergo recombination, and also appear to be essentially selectively neutral. To be useful in population studies the regions also need to contain sufficient inter-individual diversity to allow useful population comparisons. The Y chromosome and mitochondrial DNA are two such regions and I will be utilising them for this research. They also have a number of other interesting qualities which have made them popular tools over the last few decades for studying population history.

### 2.2.2 Mitochondrial DNA (mtDNA)

MtDNA is unique in the genome, because it is not a part of the nuclear DNA, but instead is contained within the cell's energy-producing organelles, mitochondria (Fig. 2.3). Each mitochondrion contains many identical copies of its DNA and each cell contains many mitochondria. This high copy number makes mtDNA a particularly suitable target for ancient DNA analysis, as it increases the probability of recovering information from ancient samples (Jobling et al. 2013), 125).

Mitochondria and their DNA are inherited exclusively down the female line, from mother to daughter. Males also inherit mitochondria from their mothers, but do not contribute any to the next generation (Birky 1995; Birky et al. 1989). The maternal inheritance pattern has made mtDNA invaluable for studying female-specific population history.

This uniparental mode of inheritance makes mtDNA a haploid locus, meaning only one version is effectively carried per individual. The majority of human

chromosomes (autosomes) are diploid, as each person carries two versions, one from each parent. The exception to this is the sex-determining chromosomes, in which the X chromosome is diploid in females, and haploid in males along with the Y chromosome (described in Section 2.2.3).

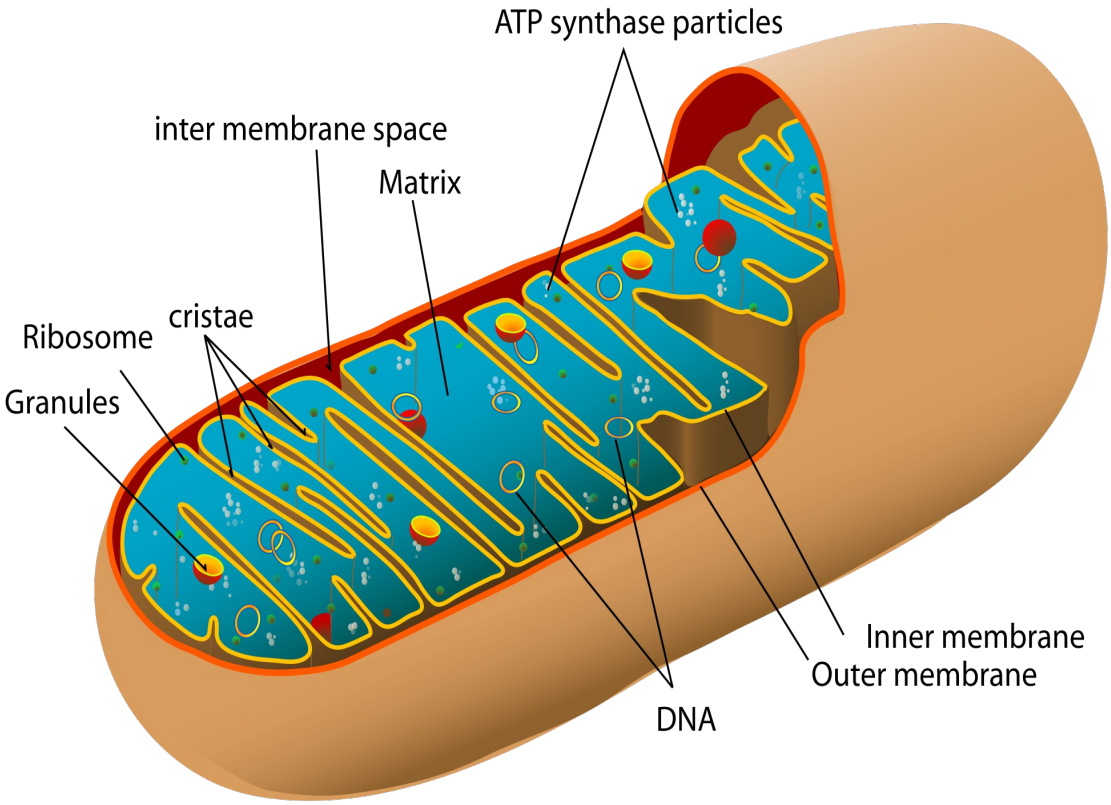



*Figure 2.3: The mitochondrion, the energy-producing cellular organelle, contains many copies of its own circular DNA. Illustration by Shanel, WikiCommons 2014.*

The haploid status of mtDNA has implications for its effective population size, and therefore its susceptibility to the effects of drift. For every male and female pair, there are four of each autosome (non-sex chromosomes) which can be inherited by the next generation but only one mtDNA. This makes the effective population size of mtDNA a quarter that of the majority of the genome. Hence, mtDNA tends to lose diversity at a faster rate than the rest of the genome through drift. For example, a woman who has only sons will not pass on her

mtDNA to the next generation (Fig. 2.7; Birky et al. 1989). However, this characteristic can also result in geographical differentiation, which is beneficial in population studies.

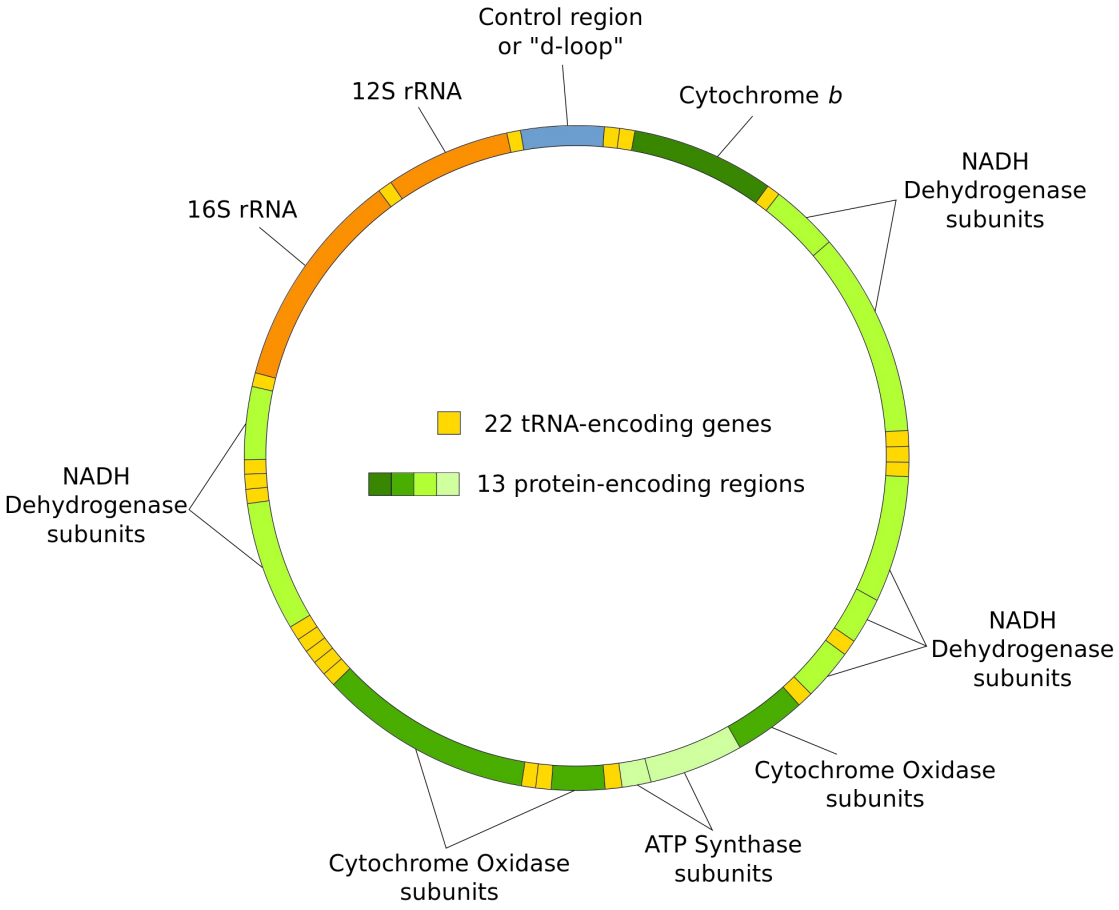



*Figure 2.4. Mitochondrial DNA (mtDNA). Schematic diagram showing the major regions of circular mtDNA molecule. From WikiCommons 2014.*

Researchers wanting to analyse past populations are generally more interested in the mutations in the non-coding regions because as 'neutral' polymorphisms they are considered to be unaffected by natural selection and therefore to reflect population history. However, although markers under inspection for population studies are probably themselves evolutionary neutral, that is not directly under the control of selection, the genes that mtDNA does contain, are likely to have been under natural selection at some point in the past (Zhang and Singh 2014).

For example, mtDNA is involved in thermoregulation, which means it may have had an adaptive role to the colder climates of the north, therefore affecting the distribution of different types (Mishmar et al. 2003). As a non-recombining entity this means that natural selection would have acted upon the mtDNA molecule as a whole, with an associated increase in frequency of particular variants through selective sweeps (Hurst and Jiggins 2005).

When describing genetic diversity, early research focused on the control region (Fig. 2.4) of the mtDNA chromosome, which contains the most variation, in particular hyper-variable region 1 (HVR 1; eg. Richards et al. 1996). Different patterns of mutational differences were identified between individuals. Variant lineages, as defined by different groups of mutations, are called haplotypes. Haplotypes are clustered into broader haplogroups which shared a recent common ancestor (Soares et al. 2010), although all haplogroups are ultimately related at some point in the past (Cann et al. 1987).

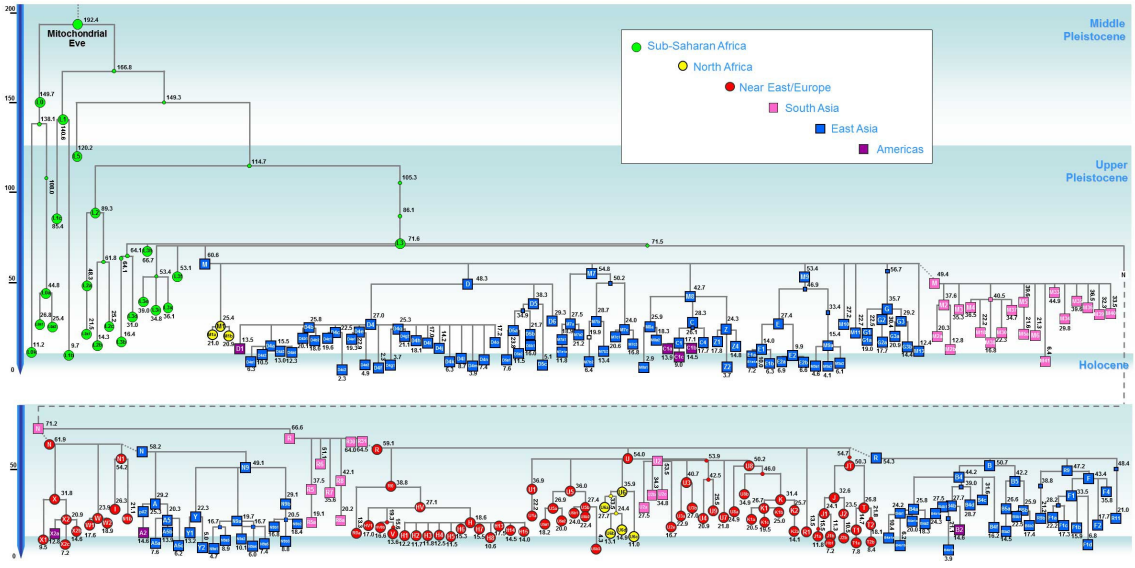



*Figure 2.5: mtDNA Phylogenetic Tree, with estimates for the ages of lineages and regions of origins (Soares et al 2009).*

Mitochondrial DNA mutates about ten times faster than nuclear DNA due to its

energy-producing function which puts it into close contact with mutagenic free-radicals. Its high replication rate increases the chance of errors, and it is vulnerable to damage which is also compounded by a lack of the protective protective structure found in nuclear chromosomes. The HVRs are non-coding so have ten-fold higher mutation rate again with an average ~$10^{-6}$ per base per generation, compared to ~$10^{-8}$ per base per generation for the rest of the nuclear genome (Jobling et al. 2013).

This high mutation rate means that back mutations in the HVRs are relatively common, making it difficult to define evolutionary trees (phylogenies) from this region alone. However, as the price of sequencing has reduced and its speed has increased, it has become more common to sequence larger regions, encompassing HVR 2, the whole control region and in some cases the entire mtDNA genome. This increased resolution has enabled greater understanding of the evolutionary relationships between extant lineages (see Fig. 2.5). It has also increased the reliability of mutation rate estimates. Both of these factors are important for predicting when and how demographic or social factors occurred from patterns in modern mtDNA diversity (Soares et al. 2010).

### 2.2.3 Y Chromosome (NRY)

The Y chromosome (NRY) is contained in the cell's nucleus and is one of two sex chromosomes, the other being the X chromosome. The NRY is male-determining, so is only carried by males and as such is it is passed exclusively from father to son down the male line, making it complementary to mtDNA (Fig. 2.6), by allowing male-specific populations histories to be infered. Like mtDNA, the NRY is haploid, as each male cell contains only one copy. So the effective population size of the NRY is a third that of the X chromosome (the other sex-

chromosome), a quarter that of the rest of the genome, but theoretically equal to mtDNA (Caballero 1995).

During meiotic cell division in males, the X and the Y chromosomes pair up, using small matching segments called the pseudo-autosomal regions (PARs), where they undergo crossing-over. However, the majority of the Y chromosome does not cross over and it is this 'non-recombining portion of the Y chromosome' (NRY), which is used in population studies (Fig. 2.5; Jobling and Tyler-Smith 2003).

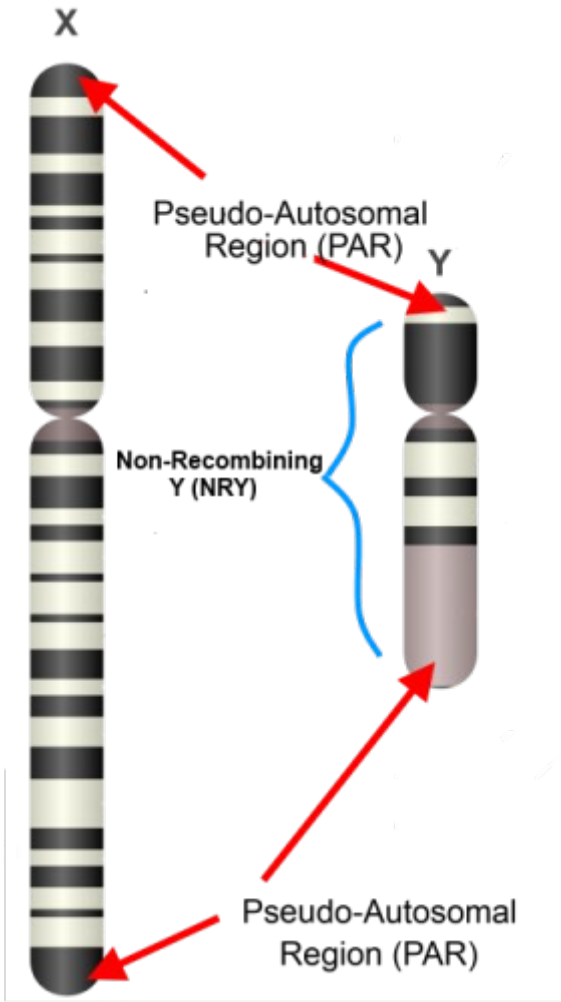



*Figure 2.6: Schematic diagram of sex chromosomes X and Y, showing the pseudo-autosomal regions (PAR) which undergo recombination during gamete formation. The non-recombining region of the Y (NRY) is used for population studies. Illustration adapted from WikiCommons 2014*

The mtDNA and the NRY provide analogous means to study female and male

population histories; however, despite the sex-specificity of their inheritance they cannot be considered equivalent. Male and female behaviour differs sufficiently to result in very different dispersal patterns. The majority of societies (~70%) practice patrilocality, which means that males tend to take a mate from a neighbouring location and bring her back home (Wilkins and Marlowe 2006). Males tend to stay nearer to their birth-places than females, resulting in NRY types being more geographically clustered compared to mtDNA (Jobling and Tyler-Smith 2003). This makes the NRY particularly useful for studying short-range population differences. The larger size of the NRY, ~60 Mega-bases (Mb) compared to mtDNA ~16 kilo-bases (kb) potentially allows for greater resolution and better distinguishing power between closely related populations (Jobling and Tyler-Smith 2003).

The NRY contains about 80 genes which code for proteins, and some of these genes are involved in male fertility, but as yet there has been little or no evidence to suggest natural selection has had any influence on NRY geographical distributions (Jobling and Tyler-Smith 2003), so most models assume neutral evolution. However, there have been incidences where particular NRY haplotypes have risen to unusually high frequencies through social selection. If a man has a social advantage which allows him and his male descendants to pass on their NRY more than other males, their lineage can increase in frequency due to non-genetic advantage (Moore et al. 2006).

This effect was first identified in Asia, where a NRY haplotype showed unusually high frequency and distribution around Mongolia. The proliferation of the NRY haplotype was dated to approximately 1000 years ago, which led researchers to attribute its unusual pattern to the considerable reproductive success of Genghis Khan and his closest male relatives (Zerjal et al. 2003). In Ireland too, a prolific

NRY haplotype, found in uniquely high frequencies in the north-west, was linked to the Uí Néill dynasty. This particular haplotype is only found in other worldwide populations that have historic links with Ireland (Moore et al. 2006).

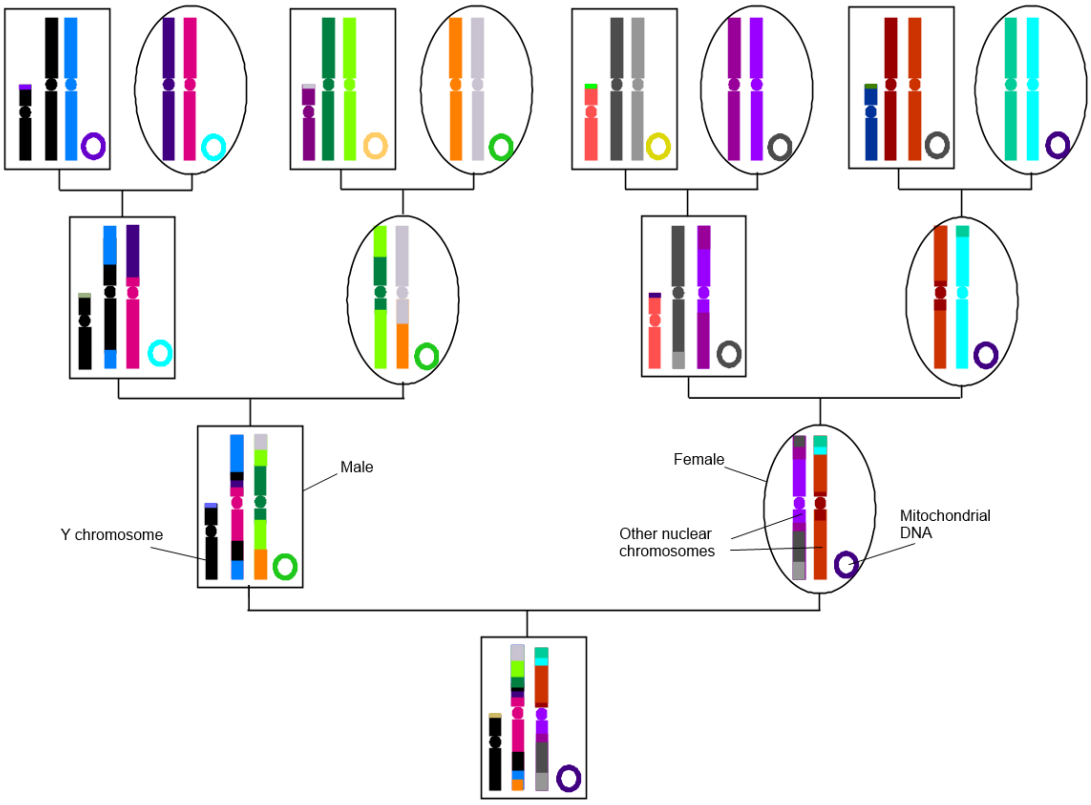



*Figure 2.7: Male- and female-specific inheritance patterns of NRY and mtDNA. Each are inherited as whole units unlike autosomal chromosomes. Image reproduced by permission of M. Jobling.*

For the NRY, single nucleotide polymorphisms (SNPs) define haplogroups. A SNP is a change in a single base in the DNA sequence (Fig. 2.8). SNPs are also sometimes called unique event polymorphisms (UEPs), because, given their low mutation rate (~$10^{-8}$ per base per generation), it is considered highly unlikely that the same base would mutate more than once (Jobling et al. 2004, 509). It is SNPs that form the basis for the phylogeny of the NRY (see Fig. 2.9; Karafet et al.

2008).


*Figure 2.8: Diagram showing (a) how Single Nucleotide Polymorphisms (SNPs) arise as changes in the sequence at a single base. (b) Short Tandem Repeats (STRs) form through changes in repeat-unit copy number. (c) The phylogeny and therefore the evolutionary history of SNPs can be inferred. Haplogroups are defined by their most recent SNP. (d) SNP-based haplogroups and copy numbers of several STRs can be combined to form detailed individual haplotypes.*

When only a few SNPs had been identified a nomenclature developed which indicated the evolutionary relationships between SNPs. However, as more SNPs were discovered, the nomenclature for NRY haplogroups became more convoluted. To make matters worse, as more sub-lineages were identified, their ancestral relationships were better understood, resulting in haplogroups needing to be renamed (Karafet et al. 2008). This made identifying the same haplogroups between research studies very confusing. This was particularly a problem for sub-lineages of the R haplogroup, which is common in western Europe (see Table 2.1).

Most SNPs identified in early studies were chosen for their diversity in well-studied regions such as Europe and Asia. Unfortunately this has implications for studying diversity in other regions, making them appear to be less diverse because the particular set of SNPs under examination are not relevant in these regions. This is known as ascertainment bias and it can have consequences for

inferences made about a population's history. The bias could, in principle, be overcome by sequencing the whole NRY, as has been done for mtDNA, but this has been too expensive to do routinely, due to its much larger size (Jobling and Tyler-Smith 2003). With new sequencing methods, however, this is now changing.

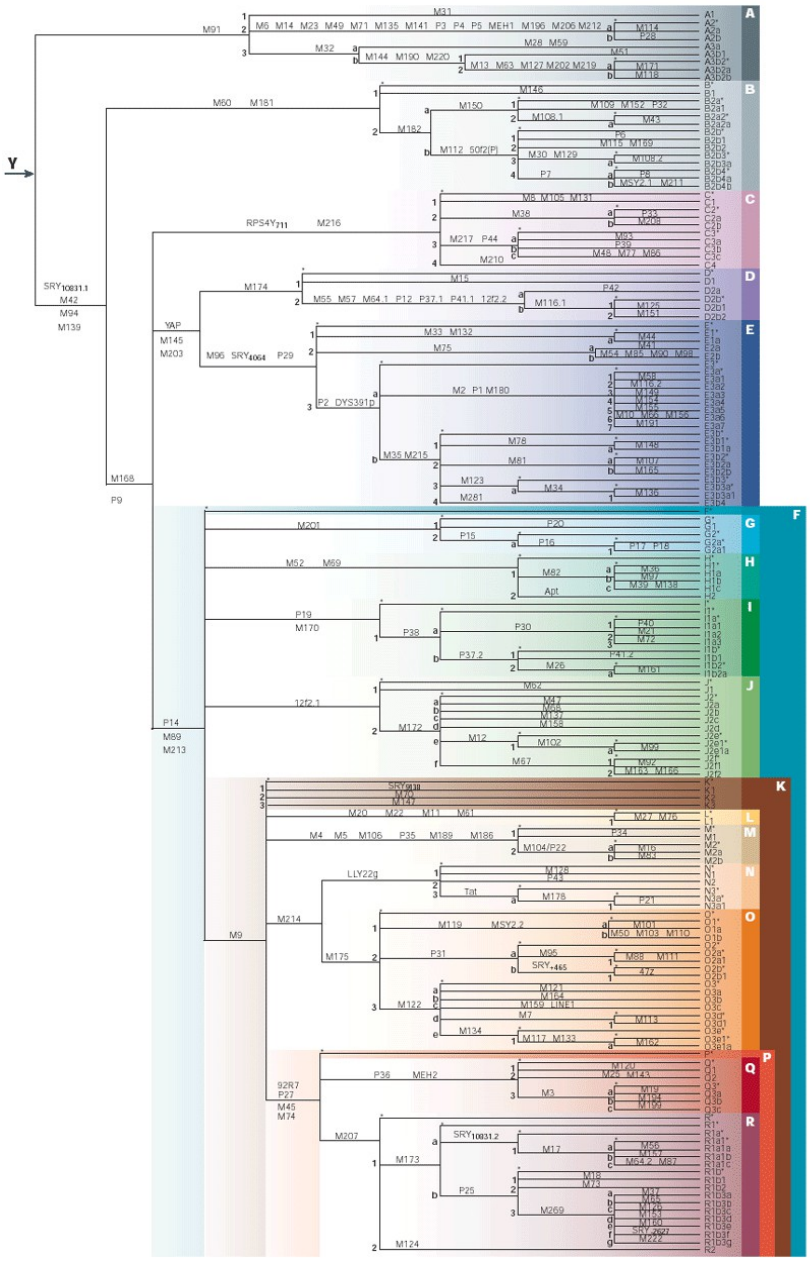


*Figure 2.9: NRY Phylogeny as it was understood in 2003. Although the basic structure of the tree remained the same at the time of writing, the branch tips were becoming better resolved and more numerous (Image taken from Jobling and Tyler-Smith 2003).*

Instead another, highly variable type of mutational variant called STRs were

used, often together with SNPs (see Fig. 2.10). STR stands for short tandem repeat, and they are also known as microsatellites. These regions of DNA have usually identical repeating sequences of 3 – 6 bases whichtend to be variable in all populations. The number of repeats may increase or decrease over time, they can mutate back and forth. Because of this STRs give poor phylogenetic resolution making it difficult to establish their evolutionary relationships. They also change at much faster rate than SNPs, ~$10^{-3}$per STR per generation, accumulating more diversity (Kayser et al. 2000). STR haplotypes are defined by the number of repeats for several STR sites (see Fig. 2.8).

*Table 2.1: Showing the changing nomenclacture over time for NRY R1b haplogroup and sub-lineages. Taken from Jobling and Tyler-Smith 2003; Karafet* et al. *2008; (Myres* et al. *2011; Y Chromosome Consortium 2002).*

| SNP | 2002 | 2003 | 2008 | 2010 |
|---|---|---|---|---|
| **R1*-M173** | R1 | R1* | R1* | R1 |
| **R1b-M269** | N/A | R1b3 | R1b1b2* | R1b1b2* |
| **R1b-L11** | N/A | N/A | N/A | R1b1b1a1 |
| **R1b-M65** | R1b3 | R1b3b | R1b1b2b | R1b1b1a1b3 |
| **R1b-S116** | N/A | N/A | N/A | R1b1b1a1b |
| **R1b-M153** | R1b6 | R1b3d | R1b1b2c | R1b1b1a1b4 |
| **R1b-M176** | R1b8 | R1b3f | R1b1b2d | R1b1b1a1b5 |
| **R1b-S145** | N/A | N/A | N/A | R1b1b1a1b2 |
| **R1b-M222** | N/A | R1b3g | R1b1b2e | R1b1b1a1b2a |
| **R1b-U152** | N/A | N/A | R1b1b2h | R1b1b1a1b1 |
| **R1b-U106** | N/A | N/A | R1b1b2g | R1b1b1a1a |
| **R1b-U198** | N/A | N/A | R1b1b2ga | R1b1b1a1a1 |

STR haplotypes and SNP haplogroups are related, since the common ancestor of a SNP haplogroup had a specific STR haplotype. In the time since the SNP arose the STR diversified leading to a subset of STR haplotypes that share similarities.

Indeed, an STR haplotype can usually be used to predict a SNP haplogroup, given a suitable comparison database. However, if sub-lineages of haplogroups are too closely related, STR haplotypes can be too similar to be distinguished using these databases. This is particularly true of sub-lineages of the R1b haplogroup.

So an advantage of using STRs in conjunction with SNPs is that they increase the resolution. This allows for more variation to be identified, in turn, allowing closely related populations to be differentiated more effectively. Forensic research has investigated 187 Y-STRs, identifying 13 highly mutable STRs with sufficient power to differentiate between even closely related males (Ballantyne et al. 2010). However, for population studies, 12 to 17 different STR loci are considered sufficient to differentiate between populations (Moore et al. 2006).

## 2.3 Description and inference

This section aims to provide the non-geneticist with an overview of the methodologies which use mtDNA and NRY diversity to make inferences about the past. I will summarise the methods essential for describing the genetic landscape of a region, drawing hypotheses from patterns found therein and making inferences about the historic events which led to these patterns.

The methods used to explore and elucidate modern gene-pools can be divided into three main categories; classic population genetics, phylogeography and computer simulations, although there is some overlap between the categories. Each has its strengths and weaknesses, but when used under the appropriate conditions they can complement each other, offering alternative ways to describe and interpret genetic diversity.

The Neolithic Age and later periods of human pre-history, often appeared to be associated with changes in human behaviour as they adapted to new technologies. This may have been accompanied by, or even been in response to population growth. Furthermore, any movements between regions would have involved complex interaction between populations, which would not have occurred during the recolonisation of Europe after the glacial periods. This section will introduce the methodologies, highlighting the aspects which are most appropriate for untangling those processes which would have affected diversity during these later periods of the European Holocene.

### 2.3.1 Classical Population Genetics

In classical population genetics simple mathematical models are used to explore the processes shaping genetic diversity. Genetic data are often summarised by statistics encapsulating different aspects of relative genetic diversity, both within and between populations. These methods are largely based on the works of Sewall Wright and Ronald Fisher, developed in the first half of the 20$^{th}$ century.

The Wright-Fisher model (WFM) is a simple model of populations describing the genealogical relationship among genes. It allows predictions of how genetic markers would behave under certain conditions, and how the processes of natural selection, genetic drift, and migration interacted to affect marker frequency over time (Fisher 1930; Wright 1943; 1931). Although the WFM was designed before molecular data was available, it has been revised over time, to accommodate new knowledge about the processes it describes.

Models are necessarily simple, and require idealised, often unrealistic population conditions, such as an infinite constant population size over time, non-overlapping generations, and random mating (Fisher 1930; Wright 1943; 1931).

However, they provide useful ways to describe patterns of diversity and a starting point when looking for explanations for patterns observed. The way that observed diversity diverges from expectations based on these models, can give insight into possible events which shaped the population's gene-pool (Wakeley 1999; Wright 1943; 1931).

However, relationships between populations will always deviate from expectations predicted by models, because humans do not behave in an idealised fashon. We often do not have a constant population size over time and our generations are generally over-lapping. Humans are also often mobile over short distances, and sometimes over greater distances, and in particular we do not practice panmixia (random mating; Hunley et al. 2009). Instead we have cultural practices which may dictate mating practice influencing who reproduces and with whom and these cultural behaviours may change over time (Destro Bisol et al. 2012; Rasteiro and Chikhi 2013; Seielstad et al. 1998). So genetic patterns in modern gene-pools are expected to have been influenced by past demographic and social events, and the ways that they deviate can provide clues about these occurrences.

For example, the 'isolation by distance' model predicts that mating choice between individuals is restricted by distance, so as the geographical distance between populations increases, so does genetic differentiation (Wright 1943). However, if populations are genetically more similar than they should be according to the parameters of the model, then it is likely that they share a history, via long distance migration, for example. Conversely, if they are more different than expected, it could indicate a physical or cultural barrier to genetic mixing (Jobling et al. 2013, 219). As there is often more than one explanation for patterns observed, caution must be exercised when making interpretations

based on observed patterns alone (Thomas et al. 2013).

### 2.3.2 Phylogeography

Phylogeography is concerned with the phylogenetic relationship between haplogroups in time and geographical space. The phylogeny of particular haplogroups are superimposed over a map, in an attempt to trace their dispersal from origin points (Richards et al. 1996). These methods also combine knowledge from mutation models with phylogenies to infer the timing of events that led to patterns observed in the genetic landscape (Soares et al. 2010; Wei et al. 2013a).

Unfortunately some of these methods are not considered useful for populations which have experienced considerable admixture, as is the case in Europe (Nielsen and Beaumont 2009). Also, founder analysis which estimates dates for putative migrations of haplogroups between populations and their subsequent proliferations (Richards et al. 2000) has been heavily criticised in the literature (discussed in Section 2.4.2). However, despite these limitations certain aspects of phylogeographic analysis retains some useful features for highlighting potential demographic events and estimating the age of particular lineages through the use of the molecular clock.

The 'tick' of the molecular clock is the rate at which mutations were acquired so the speed and accuracy of the clock differs depending on the type of loci and its mutation model. The earliest models designed to describe the mode and frequency in which mutations occurs, were often the simpliest, assuming that mutations occur with equal frequency at the same rate throughout the genome (reviewed in Jobling et al. 2013; Kumar 2005). However, these have been adapted to take into account more complex dynamics (Nachman and Crowell 2000;

Rienzo et al. 1994). The archaeological record is being used to provide new calibration points for the molecular clock to improve mutation rate estimates (see Section 2.4.3).

It is important to realise that the age of a lineage does not equate to the timing of its distribution through the gene-pool (Barbujani and Chikhi 2006), and nor does haplogroup divergence equate with population divergence (Barbujani et al. 1998). However, in some cases calculating the age of a lineage may provide a boundary after which the suspected demographic events would have occurred (Wei et al. 2013a).

Traditional phylogenetic tree building uses maximum parsimony or maximum likelihood which means the tree that contains the least number of evolutionary changes is preferred (Fitch 1971). While these methods can be used to infer SNP phylogenies they are not considered appropriate for inferring the phylogenies of HVRs in mtDNA or NRY STR because homoplasy (parallel mutations) adds a major source of structural uncertainty. Even with high-resolution data, extremely large numbers of trees are possible for even a small data-set (Bandelt et al. 1995).

Network analysis was developed to overcome these issues by using median-joining networks which allowed all the most parsimonious trees to be displayed at once, often revealing complex networks (Bandelt et al. 1995). Weighting can be added to particular connections based on the amount of variance at particular loci (Qamar et al. 2002). This reduces the complexity of the network by removing the least likely parallel mutations.

The shape or structure of networks or phylogenies, could signal past demographic events (Bandelt et al. 1995). Star-shaped phylogenies can indicate where haplogroups have undergone rapid proliferation, due to demic

expansions, or selection (Bandelt et al. 1995; Moore et al. 2006; Zerjal et al. 2003). Closely related lineages shared between regions could indicate shared history, such as a period of migration, or it could indicate that this lineage was more common and widespread in the past (Barbujani and Chikhi 2006; Zerjal et al. 2003).

While this method can highlight interesting features or patterns, a more statistical approach is needed to determine the most probable historical scenario which gave rise to them (Nielsen and Beaumont 2009; Thomas et al. 2013). Also while it is theoretically possible to estimate dates for demographic events using the molecular clock given a reliable mutation rate and model and sufficiently large data-set (Soares et al. 2010; Wei et al. 2013a), neither mtDNA HVRs or NRY STRs are considered stable enough to make this type of estimation from their networks (Hallast et al. 2014; Wei et al. 2013b).

### 2.3.3 Modelling and Simulations

Computer simulations have been used to test population genetic theories since the 1960s (Lewontin and Dunn 1960), as well as by anthropologists to investigate the effect of demographic variables on population structure (Kunstader et al. 1963). Simulations are particularly useful for analysing apparently random or stochastic processes because unlike reality, a simulated event can be run hundreds, thousands, even millions of times to give a range of outcomes. Probability distributions can be estimated for effective population size, admixture or migration rates under various demographic scenarios (models). Simulated data provide a framework against which observed data can be tested, giving insight into the processes affecting populations under investigation (Matsmura and Forster 2008, 1-2).

Conversely, computer simulations can also be used to model hypothetical situations to explore their effects on a population. For example, researchers modelling demic expansions showed that rare haplogroups could 'surf' the leading edge, reaching unusually high frequencies at the boundaries, far away from their origin and mimicking the effects of selection. This phenomenon, named allele surfing, is caused populations at the leading edge moving into less densely populated regions, and therefore have a smaller effective population size, this then leads to a loss of diversity through drift (Excoffier and Ray 2008). In this way computer-aided modelling not only provides a way for testing hypotheses within a statistical framework, but they can also be used to identify unexpected patterns caused by demographic or social processes.

As computers have become faster and more powerful, the potential for complexity in models has increased, allowing for more realistic demographic and social scenarios to be tested. Many recent developments in simulation methodologies are rooted in Bayesian statistics. Bayesian statistics are often called conditional statistics because they allow for the probability of something to be established, given what has been observed (Sivia and Skilling 2006).

To illustrate, traditional or frequentist statistics seeks to find the probability, for example, of getting 10 heads, if a coin is tossed 10 times. In Bayesian statistics, the fairness of the coin can also be assessed through the likelihood function. That is, if the coin is tossed 10 times and 10 heads are achieved, what is the likelihood that the coin is fair? Through this the original probability can be calculated more accurately in light of this likelihood function. The 'fairness' of the coin is considered to be a parameter of the coin-toss model and can be influenced by what is known of the situation, for example if the person tossing the coin is disreputable (Sivia and Skilling 2006).

Modelling methods that use this Bayesian methodology are called full-likelihood models and include Markov chain Monte Carlo (MCMC). Although they are powerful for simple models, with complex models the length of time required to calculate the likelihood becomes prohibitive. This is often the case when trying to assess which historical scenarios could have resulted in the patterns observed in a gene-pool (Bertorelle et al. 2010; (Pascual et al. 2007).

Approximate Bayesian Computation (ABC), which is a sub-class of a type of these methods, uses simulations to reconstruct and approximate the likelihood function and therefore the parameters of a scenario (Aeschbacher et al. 2012; Bertorelle et al. 2010; Pascual et al. 2007). ABC provides a statistical way to compare the observed and simulated data. So, in the above example, one would simply toss one thousand fair coins 10 times, and see how often ten heads was achieved. This way the fairness of the original coin could be assessed by comparing it to these results using ABC (Beaumont et al. 2002; Bertorelle et al. 2010).

In population genetics since its not possible to re-run history repeatedly to see how the observed genetic patterns turn out, simulations are used instead (Beaumont et al. 2002). Genealogies are often simulated using coalescent theory which generates phylogenetic trees backwards in time until they coalesce at the most recent common ancestor (Bertorelle et al. 2010). Generating simulations in this way is computationally more efficient, and therefore faster, than simulating forwards in time. This is because only the genealogies of lineages in the sample need to be explored as all lineages must coalesce at some point (see Fig. 2.10; Kingman 1982).

Since the parameters of the scenarios are often unknown, when simulations are generated parameter values are selected at random from a prior distribution.

The prior incorporates knowledge from alternate sources such as archaeology and history and can encompass a wide distribution of values, which are narrowed down by ABC to give a posterior distribution of the most probable values. ABC can also indicate if a prior distribution needs to be changed to produce simulations to match the data, providing a way to test assumptions made about parameters against the genetic data. The width of the posterior distribution indicates the level of certainty associated with these values (Bertorelle et al. 2010).

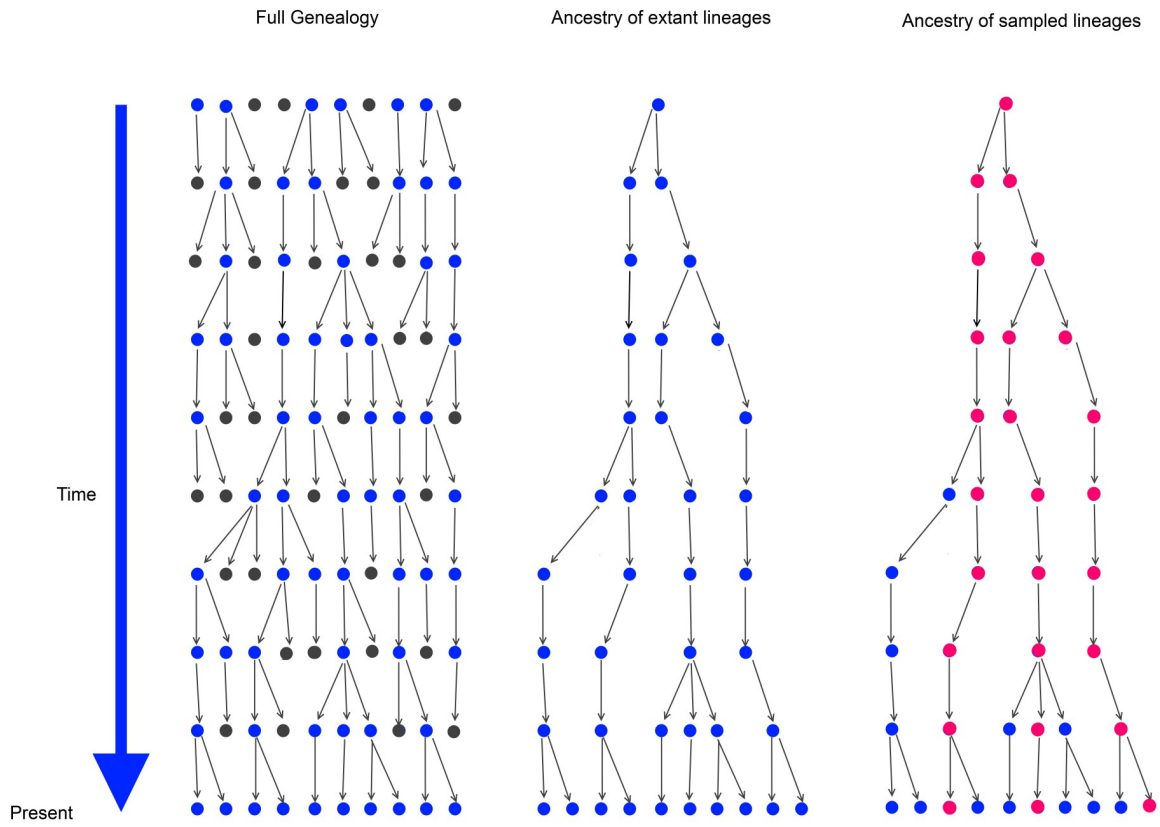



*Figure 2.10: Coalescent Theory allows the genealogy of extant lineages (blue circles) or even just a sampled subset (red circles) to be calculated back in time to the most common ancestor without the need to include all individuals that have ever lived. Mathematical models which use coalescence theory to estimate historic events are far simpler than those drawn from full genealogies as they include genealogical dead ends (black circles; redrawn from Jobling* et al. *2004; 183).*

So, ABC has two useful functions for population genetic analysis; firstly it

provides a way for historic scenarios to be tested against one another by statistically assessing which modelled scenario produces simulated data most similar to the observed data (Beaumont 2008). Secondly the parameters of the models can then be estimated from those simulations closest to the observed data and assumptions about the types of possible demographic events possible and the timings of these possible events are tested in light of the genetic data (Beaumont et al. 2002).

ABC compares summary statistics calculated using classic population genetic methods, for the observed and simulated genetic data. While this helps to increase efficiency and speed of processing and simulating data (Beaumont et al. 2002), it also provides a considerable source of error, as at the time of writing no objective way of choosing which summary statistics to use exists (Beaumont 2008; Burr and Skurikhin 2013). The choice of summary statistics can have an impact on the outcome, and if they are poorly chosen it can decrease the accuracy of ABC. Authors recommend choosing summary statistics on the basis of their relevance to the scenarios under investigation (Csilléry et al. 2010; Sunnaker et al. 2013).

## 2.4 Historical relationship between Genetics and Archaeology

When Colin Renfrew wrote optimistically about the 'new synthesis' between genetics and archaeology (Renfrew 2000b), the study of mtDNA was contributing to our knowledge of the Palaeolithic period of human prehistory; tracing our ancestry back to a single woman in Africa some 150,000 to 200,000 years ago (Cann et al. 1987). Since then genetics has been used to infer the routes that our ancestors took out of Africa, as well as where and when they

entered Europe and how they spread to the rest of the world (Soares et al. 2010). Ancient mtDNA and nuclear DNA extracted from Neanderthal bones have even made it possible to begin to study the relationship between ourselves and our most famous extinct cousins (Hodgson et al. 2010).

However, when examining more recent events in European human prehistory, archaeogenetics has been less impressive, often causing controversy and dividing opinion (Pluciennik 2006; Renfrew 2010). The population events of the European Holocene are complex involving mingling of peoples from different regions. To analyse mixes accurately we need to know something of the origins and histories of migrants and putative host populations. For this information we must turn to archaeological or historical resources.

However, a lack of systematic dialogue between disciplines has meant that geneticists know too little about current evidence and interpretation in archaeology and archaeologists cannot make informed opinions about inferences made from genetic data. Added to this, researchers have often seemed conflicted about what exactly genetic data reveals, offering opposing explanations for the same datasets. It is not surprising that those outside the discipline began to doubt the usefulness of genetic data for shedding light on unanswered archaeological questions (Pluciennik 2006; Renfrew 2010).

Through technological advances and greater collaboration these issues are starting to be resolved. This section summarises the history of the use of genetic data to answer archaeological questions, highlighting cultural differences between fields, as well as methodological issues, debates, new technologies and success stories that are paving the way to a real synthesis of archaeology and genetics.

### 2.4.1 In the beginning

In European prehistory, the Neolithic transition has attracted considerable genetic research activity, and caused a great deal of controversy (Pluciennik 2006b). The earliest attempts to map human diversity in Europe used polymorphic protein products, such as enzymes and antigens found in the blood. Researchers identified a gradient in frequency of protein types, which stretched across Europe from the Near-East to North-Western Europe (Fig. 2.11), mirroring the radiocarbon dates of the Neolithic spread of agriculture in the same direction. This striking similarity was taken as evidence of demic diffusion, whereby farming increased the carry capacity of land allowing farming communities to grow and expand, replacing or absorbing indigenous hunter-gatherers (Ammerman and Cavalli-Sforza 1971).

The mechanism for dramatic cultural changes seen throughout the European Neolithic has been long debated in archaeological circles, and genetic research seemed to offer a potential new source of evidence to shed light on this problem. Competing theories are represented by two models, although few would argue for either extreme (Zvelebil and Zvelebil 1988). In the first model, innovations and materials spread through cultural diffusion as they were adopted by a largely indigenous population. The second, named the wave of advance model, involves population replacement by migrants farmers who carried their own culture and innovations (Pluciennik 2006).

Although the work of Cavalli-Sforza and colleagues seemed to support the wave of advance model, through a demic diffusion of migrant farmers (Ammerman and Cavalli-Sforza 1971), the patterns were ambiguous. There was no way to assign a direction to the gradients observed in blood proteins, nor was there a way to date to the formation of the gradient (François et al. 2010; Jobling et al.

2013, 379-380). There was even evidence that suggested the patterns seen could have just been an artefact of of principal component analysis (PCA; Novembre and Stephens 2008).

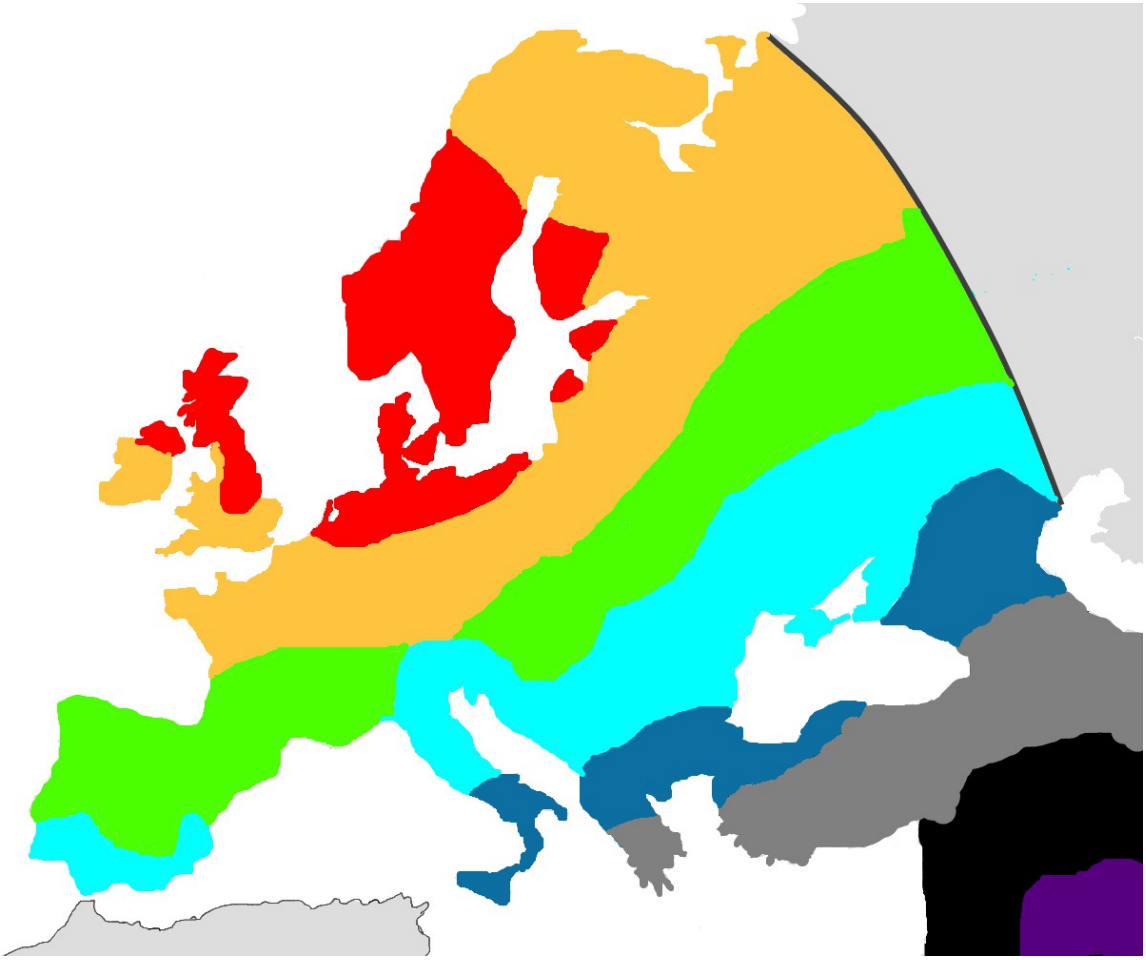


*Figure 2.11: Clines of polymorphic blood proteins through Europe, redrawn from the works of Cavalli-Sforza and colleagues (Cavalli-Sforza* et al. *1996).*

When advances in molecular biology enabled DNA to be examined directly, nuclear DNA and NRY revealed the same large-scale geographical patternings in diversity (Chikhi et al. 2002; Menozzi et al. 1978; Rosser et al. 2000). However, it became clear that these patterns could be the result of a palimpsest of human activity over time, perhaps dating back to the first colonisation events of Europe by modern humans (Barbujani and Goldstein 2004; Chikhi et al. 1998).

Since multiple historical scenarios could explain genetic patterns equally well,

the difficultly lay in determining which scenario was closer the truth. Much of the subsequent work in human population genetics was concerned with devising methods for infering population history from modern genetic diversity and establishing their reliability (Beaumont et al. 2002; Nielsen and Beaumont 2009; Richards et al. 1996). Further, MtDNA appeared to show different patterns from the majority of the genome (Simoni et al. 2000) which had implications that I will return to in the next section.

Of course, not all genetic research was purely concerned with honing methodologies, and geneticists often attempted to ascribe explanations from history and archaeology to patterns observed in gene-pool diversity. However, particularly during the early part of this period of intensive research and development, there was very little collaboration between archaeologists and geneticists. Genetics studies concerning prehistoric events contained few archaeological references and even fewer studies were co-authored by both geneticists and archaeologists. To those outside the field it seemed that geneticists had little interest in collaborating with archaeologists and without insight into current archaeological thought, they were often accused of using overly simplistic interpretations of the archaeological record (Pluciennik 2006).

Parallel to developments in genetics, archaeology was undergoing a paradigm shift in relation to how archaeological evidence could and should be analysed and interpreted (Champion 1991). Many archaeologists of the 1960s and early 1970s had embraced the sciences, forging links with anthropology, mathematics, and geography (Binford 1962). Quantitative methods and modelling became popular as archaeologists sought to understand culture as an adaptive response to environmental processes, such as climate change and geography (Clarke 1968). This philosophy became known as Processual archaeology because of its

view of cultural change as a process affected by other external processes (Champion 1991).

However, when Cavalli-Sforza and colleagues were publishing their works, this paradigm, with its emphasis on scientific rigour and data analysis, was being criticised. Archaeologists were recognising the difficulty of describing ancient cultures without being influenced by the view-point of their own cultural ideologies (Champion 1991). In this light, they questioned the implicit assumptions underlying scientific models and methodologies (Brown and Pluciennik 2001a). For many, processualism also missed the richness and uniqueness of individual historical contexts (Champion 1991).

From the mid 1970s, archaeologists moved away from looking for external triggers for cultural change, and instead explored ways in which change could be driven from within societies. The symbolism and ideology behind archaeological artefacts were determined to provide insight into local customs and developments. The regional and historical contexts became more important than looking for large-scale general themes (Champion 1991).

In this light, some archaeologists considered that genetic research ignored the complexities of the Neolithic transition apparent in archaeological data at the regional level (Brown and Pluciennik 2001; Thomas 2006; Zvelebil and Zvelebil 1988). There was also concern that culture, ethnicity and genetics were being equated in a way that was considered not only unrealistic and simplistic, but possibly also dangerous (Pluciennik 1996; Sims-Williams 1998). Archaeologists who rejected genetics as too reductionist (Brown and Pluciennik 2001; Zvelebil 1989), were accused by geneticists of being 'indigenists', thought to be incapable of believing that cultural change could have outside influences (Ammerman 1989). Alternatively, some archaeologists accepted the

interpretation of prehistory offered by geneticists as indisputable, assuming that archaeological view-point must be wrong (Brown and Pluciennik 2001).

A lack of systematic dialogue between fields meant that neither side completely understood the complexities, problems and developments of the other. Unfortunately this issue was not to be remedied quickly. Attempts to bring the fields closer together were episodic rather than a fully integrated interdisciplinary approach. Collaborative collections of works, mostly produced by Colin Renfrew, started to appear around the turn of the century (see Bellwood and Renfrew 2002; Renfrew and Boyle 2000) but even these saw genetics and archaeology separated into different chapters (Pluciennik 2006).

Furthermore, at around the same time a divide within genetics was beginning to appear. The division manifest as those who attributed the widespread European patterns in diversity to Neolithic expansions (eg. Chikhi et al. 2002), versus those who regarded modern Europeans to be largely descended from Palaeolithic peoples (Richards et al. 1996). However, underneath these differing theories was a more fundamental disagreement regarding methodologies and the interpretation of genetic patterns.

### 2.4.2 Creating Historical Narratives

By the mid 1990s the growing consensus amongst geneticists was that the large-scale patterns observed in the modern European gene-pool could be best explained by a Neolithic demic diffusion expansion of farmers originating in the Near-East. Much of the evidence was drawn from computer modelling which attempted to recreate patterns observed in real data by simulating data under a range of known demographic parameters. These simulations suggested that significant clines could only have formed with at least a 66% Neolithic

contribution from the Near-East. Although genetic data did not exclude cultural diffusion occurring in some regions, the implications were that this had little impact on the overall patterns (Barbujani et al. 1995; Cavalli-Sforza et al. 1996; Chikhi 2009).

However, in 1996 researchers exploring European mtDNA diversity found no evidence of the patterns observed in other regions of the genome. Contrary to the prevailing view in genetics, they concluded that modern European populations showed only a small contribution from Neolithic Near-Eastern populations (Richards et al. 1996). Using the phylogeographic method of founder analysis, a relatively new methodological approach, they calculated the age of European lineages and deemed the majority (~85%) too ancient to have arrived during the Neolithic. A 15% Neolithic input was calculated based on frequencies of putatively Neolithic haplotypes in modern samples. They claimed the low frequency supported the cultural diffusion model (Richards et al. 1996)

This study caused concern, firstly because it equated the age of a mutation with the age of population divergence (Barbujani et al. 1998; Cavalli-Sforza and Minch 1997). Critics argued that based on this logic, if European's colonised Mars tomorrow, future researchers would think it had happened in the Palaeolithic (Barbujani et al. 1998). Secondly, there was no reason to think that the frequency of Neolithic lineages in extant populations bore any relation to the amount of admixture occurring several thousands years ago (Goldstein and Chikhi 2002).

Despite these misgivings, over the next decade phylogeography became a popular tool for explaining distributions of mtDNA and NRY haplogroups. The age and distribution of particular haplogroups were matched to local demographic events drawn from history and archaeology. Haplogroups were linked to specific migrations (Richards et al. 2000) and in some cases even

associated with named cultures (Semino et al. 2000). However, assumptions underlying the interpretation of haplogroup distribution patterns were never made explicit and the hypotheses presented were never tested. Critics argued that without suitable testing, conclusions were little more than story-telling (Bertranpetit 2000; Chikhi 2009; Goldstein and Chikhi 2002).

Books were published which claimed to identify genetic markers linked to ancient ethnicities, such as Celt or Viking (Oppenheimer 2007; Sykes 2007). Concurrent with this came the rise of genetic genealogy, where members of the public pay to have their DNA typed to reveal their 'ancient tribal ethnicity', as propounded by associated books (eg www.oxfordancestors.com). These books targetted a lay-audience, so lacked detailed descriptions of methods required for others to repeat and therefore validate their claims. As such, their findings were never subject to the peer-review process.

When phylogeography was tested using simulated data with known parameters, it was found that it performed well in cases where a small population colonised an otherwise empty region. However, the methods were unable to recover the demographic events of populations that had experienced considerable admixture (Nielsen and Beaumont 2009). Perhaps unsurprisingly, when data from ancient sources began to emerge it did not support the phylogeographic view of European prehistory (Pinhasi et al. 2012). However, it was not necessarily true that methodologies were flawed, rather, the problem had been simplistic interpretation of data and a lack of hypothesis testing.

Arguably, the collective indignation at the extent of this type of narrative building through genetics has been instrumental in fostering closer relationships between geneticists and archaeologists. In more recent years genetics has reached a stage in development where collaborations with archaeologists are

imperative for understanding past population history through patterns in genetic data. Changes in archaeological thought, which favour a multi-disciplinary approach to exploring the past (Lucas 1997) mean that this synthesis is also desirable for archaeologists. There is a growing consensus that both computer-modelling and ancient DNA (aDNA) provide a useful way to achieve this integration (Pinhasi et al. 2012; Shennan 2008; Steel 2008; Thomas et al. 2013).

### 2.4.3 New approaches

There are a number of developments and new approaches which have converged within recent times and which finally allow for a merger between genetics and archaeology. In genetics these are ultimately linked to improvements and advances in technology. There are also important philosophical changes within archaeological theory which are more sympathetic to collaborative efforts to understand human prehistory.

One of the most important technological breakthroughs in genetics has been Next Generation Sequencing (NGS). This new method has made sequencing DNA much quicker and cheaper than ever before (see Fig. 2.12; Jobling et al. 2013; Stoneking and Krause 2011). For comparison, the Human Genome Project, took 13 years and approximately £1.7 billion to sequence the first human genome (National Human Genome Research Institute 2010). Through NGS, an entire human genome can now be sequenced in under a week for close to £1000 (Hayden 2014). Whole genome analysis has important implications for the study of human population history through genetics.

The NRY and mtDNA provide relatively cheap and plentiful data for studying populations, but they only provide a limited insight into an individual's total

ancestry. Exploring male or female lineages ignores the multitude of other ancestors which have contributed to a person's genome. Sequencing the entire genome accesses more information about individual ancestry which can then be used to make more accurate and detailed inferences about population history and in particular admixture between populations (Stoneking and Krause 2011). Useful SNPs, identified throughout the genome including the NRY, can be screened directly, reducing costs and time further (Jobling et al. 2013), as well as overcoming problems associated with ascertainment bias.

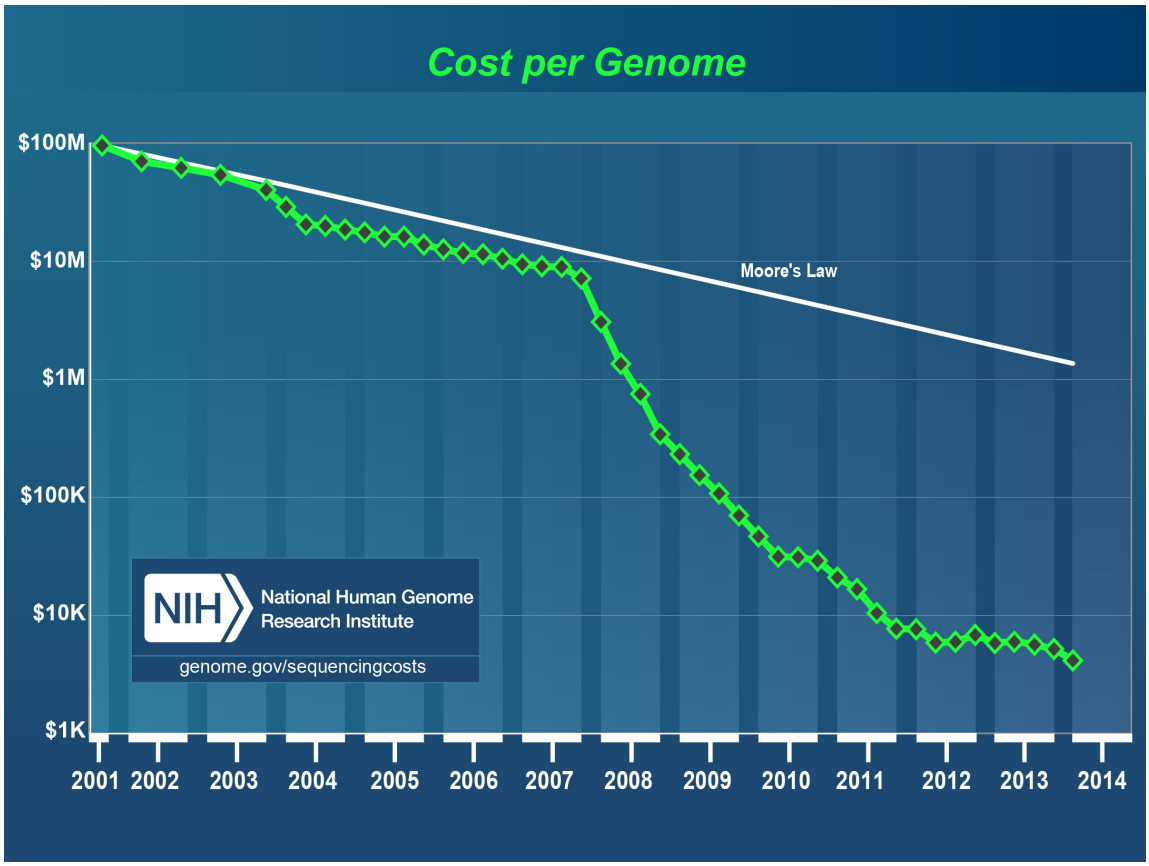



*Figure 2.12: Cost of sequencing showing a dramatic reduction in price since the introduction of Next Generation Sequencing. For comparison, the white line indicates how this trend would look if it were following Moore's law which states that computing power doubled every two years (Taken from National Human Genome Research Institute 2014).*

NGS is also revolutionising ancient DNA research. The methods involved in NGS require DNA to be broken down into small fragments which are then sequenced

in parallel.. This small fragment size makes it particularly useful for sequencing old and degraded DNA like that found in ancient sources. This also removes the problem of contamination because the small fragments of ancient DNA (aDNA) are preferentially sequenced over the much larger contaminant DNA (Stoneking and Krause 2011).

This means that aDNA datasets are increasing rapidly, not only for mtDNA but also NRY, and in some cases entire ancient genomes (Pinhasi et al. 2012). Sufficiently large, well-dated aDNA datasets have the potential to provide temporal and geographical snapshots of ancient gene-pools, creating a genetic stratigraphy which could provide clues about population movements and cultural changes over time at the regional level. However, at the time of writing aDNA is still a rare resource providing only a glimpse of past genetic patterns on a much wider-scale.

Although aDNA data-sets are small and patchy the picture of Europe's past that they are revealing is intriguing. There are distinct regional genetic differences between central and southern Neolithic burials, which have yet to be explained (Deguilloux et al. 2011; Haak et al. 2010; Pinhasi et al. 2012), as well as apparent discontinuations between past and present populations in Central Europe (Bramanti et al. 2009). The use of aDNA alongside stable isotope analysis has revealed complex relationships between members of single assemblage in Scandinavia, providing evidence of that people with distinct gene-pools and lifestyles were being buried together for centuries (Skoglund et al. 2012).

The patterns seen in these examples are described by ancient mtDNA and so reflect female history of Europe. As more NRY aDNA data becomes available it is likely that these patterns will be different again reflecting the different demographic and social conditions affecting men and women (Seielstad et al.

1998). This complex mosaic pattern of varying interaction will come as no surprise to most archaeologists, but at least genetic data now have the resolutions to detect these subtleties and to contribute to archaeological theories.

One of the biggest sources of uncertainty in genetics is associated with the molecular clock. Phylogenetic resolution through more detailed sequencing, has improved matters particularly for mtDNA, which is now regularly sequenced in full (Soares et al. 2010) has been much more difficult for the 3000-fold larger NRY, although NGS is improving the known phylogeny all the time (Wang et al. 2014; Wei et al. 2013a). With improved phylogenies the age of a haplotype becomes proportional to the length of its branch. To estimate this age requires knowledge of the mutation rate and this can differ depending on the calibration point used.

Mutation rates were traditionally calculated using chimpanzee and human divergence, but the exact timing of this is not agreed upon; also, generation times may have been different in archaic human ancestors, and the influences of selection and other processes on the separate species are difficult to account for (Wang et al. 2014). More recently, estimates have been made by comparing the number of mutations in haplotypes of individuals within the same family (pedigree analysis; Xue et al. 2009) however, only one haplogroup was used raising questions about whether the rate would also apply to other haplogroups. Also there are some questions about the appropriateness of using short term mutation rates for dating ancient events (Wang et al. 2014).

New calibration points for mutation rate estimates have used archaeologically attested colonisations, such as when humans first arrived in the Americas (Poznik et al. 2013) dated by radio-carbon dating of archaeological sites to about 15,000

years ago (Goebel et al. 2008). There was assumption that NRY haplogroup Q-M3 and Q-L54*(xM3), typical of Native Americans, diverged at the time of American colonisation, which may not have been the case. Indeed, NRY haplogroup Q-M3 had also been found in Siberia, raising the probability that it diverged before the colonisation of America (Malyarchuk et al. 2011). There is also danger of circularity when using demographic events to estimate the mutation rate, if this rate was then used to calculate the age of demographic event events in the same period or region.

The use of aDNA may provide more reliable calibration points with which to improve the accuracy of mutation rate estimates (Wang et al. 2014). Ancient DNA retrieved from skeletal remains of a male infant taken from a Clovis archaeological site in North America, radio-carbon dated to approximately 12,500 years ago, was also found to bear the haplogroup Q-L54*(M3). By comparing this ancient NRY haplogroup with modern variants, the mutation rate was calculated and then used to estimate the divergence of this lineage to 16,900 years ago (95% CI 13,000 to 19,700 years ago; Rasmussen et al. 2014) supporting a pre-colonisation divergence, approximately 2000 years earlier than the dates used by Poznik et al. (2013).

Calibrating the molecular clock using aDNA can be used to improve the parameters of models for estimating the timing of demographic events recorded in the modern gene-pool. Estimates for mutation rates, however, are highly dependent on reliably-dated sources of aDNA from well-attested sites. There are also the usual problems associated with aDNA concerning DNA preservation, finding suitable specimens with extant haplotypes and identifying and avoiding contamination (Wang et al. 2014), there is a greater understanding of the need for protective clothing and careful storage of potential DNA sources amongst

archaeologist, which helps to minimise contamination and post-excavation DNA degradation.

While NGS is increasing the quantity and quality of DNA resources, the ever increasing computer power has made computer modelling faster and far more accessible for individual researchers. As well as an increased use of MCMC and ABC modelling by geneticists (see section 2.3.3), has been a recent renaissance in computer-modelling by archaeologists (Lake 2014; Parker 2011). Archaeologists have had a troubled relationship with computer modelling in the past, leading to its near abandonment in the 1980s (Aldenderfer 1991; Lake 2014). However, there has been a recent resurgence in the use of computer modelling for hypothesis building, particularly in evolutionary archaeology and human evolution, but also for addressing other questions about long-term societal change (Lake 2014).

However, there has been concern that archaeological data provides a poor estimator of the impact of demographic events such as migrations, or changes in population size, whereas genetics has the potential to elucidate these. Likewise, genetic models for estimating demographic and social impacts on populations would benefit from improved estimates for mutation rates and timing parameters which are more reliably inferred from the archaeological record and radio-carbon dates (Steel 2008; Shennan 2008).

Newer developments in post-processual archaeological theory are concerned with breaking down disciplinary boundaries (Lucas 1997). Born out of the original drive for establishing alternative view-points for examining the past, there is awareness that perspectives and skill sets from outside archaeology could help to reduce bias and blind-spots (James, pers. comm). Some geneticists too, are realising the benefits of cross-disciplinary collaborations for a more

holistic approach to analysing the human past.

Research groups which include archaeologists, geneticists, historians and linguists have become more common in the last few years (see, for example http://www2.le.ac.uk/projects/impact-of-diasporas, http://www.arch.ox.ac.uk/SEAL.html ). There are also a growing number of individuals, like myself, who have training in both archaeology and genetics. These initiatives aim to close the knowledge gaps between the fields, giving members greater insight into benefits and limitations of each other's subjects.

In conclusion, the long-awaited synthesis of archaeology and genetics appears to be finally coming into being. Molecular technologies have achieved the required resolution and sophistication to permit the exploration of regional genetic variation. Computing has reached a point where complex historical scenarios suggested by the archaeological record can be tested against genetic data. Perhaps more importantly, within genetics and archaeology, there are a growing number who appreciate the benefits gained by both parties of a more systematic approach to collaboration.

# Chapter Three: The Isle of Man in context

## 3.1 Introduction

The Isle of Man provides a convenient case study for exploring population history, because it has definite boundaries, transcending time and politics. Sea boundaries mean that people would have generally only travelled to the island with a purpose, rather than accidentally wandering over the border. Before the modern advances in transportation, the Isle of Man, could essentially, be considered a self-contained population.

To date, little genetic analysis has been carried out on the Isle of Man itself, and instead research has targeted the surrounding lands and more distant regions which provide the context for this research. Major trends have been identified in the modern gene-pools of the British and Irish Isles, and in ancient DNA from sources in mainland Europe which have implications and provide genetic temporal baselines for the population history of the island.

Historical and archaeological evidence suggest that the island's location in the middle of the Irish Sea may have made it of strategic importance in the past. However, during rough weather and dark winter months the surrounding sea would have acted as a barrier to the casual visitor. During the periods when the climate was unpleasant this sea barrier may have lead to periods of prolonged isolation.

For the sea-faring Vikings the Irish Sea was no barrier, and instead provided a convenient route-way, but although Norwegian Vikings are documented in the island's history, little is known about the genetic impacts of this or possible earlier undocumented migrations, on the Manx gene-pool. However, key

features of the Isle of Man's location, geography, archaeology and genetics make it a useful starting point for exploring the wider genetic landscape of the British and Irish Isles and the Isle of Man's place among the Atlantic populations of Western Europe.

### 3.1.1 Geography and demography


*Figure 3.1: Physical map of the Isle of Man (Mapsof.net 2014)*

The Isle of Man lies in the middle of the Irish Sea almost equidistant from England and Ireland, although its nearest mainland shore is in Scotland (see inset of Figure 3.1). The Island measures 572km$^2$ (221miles$^2$; Isle of Man

government 2014) with a mountain range that divides it into northern and southern regions (Fig. 3.1 main image).

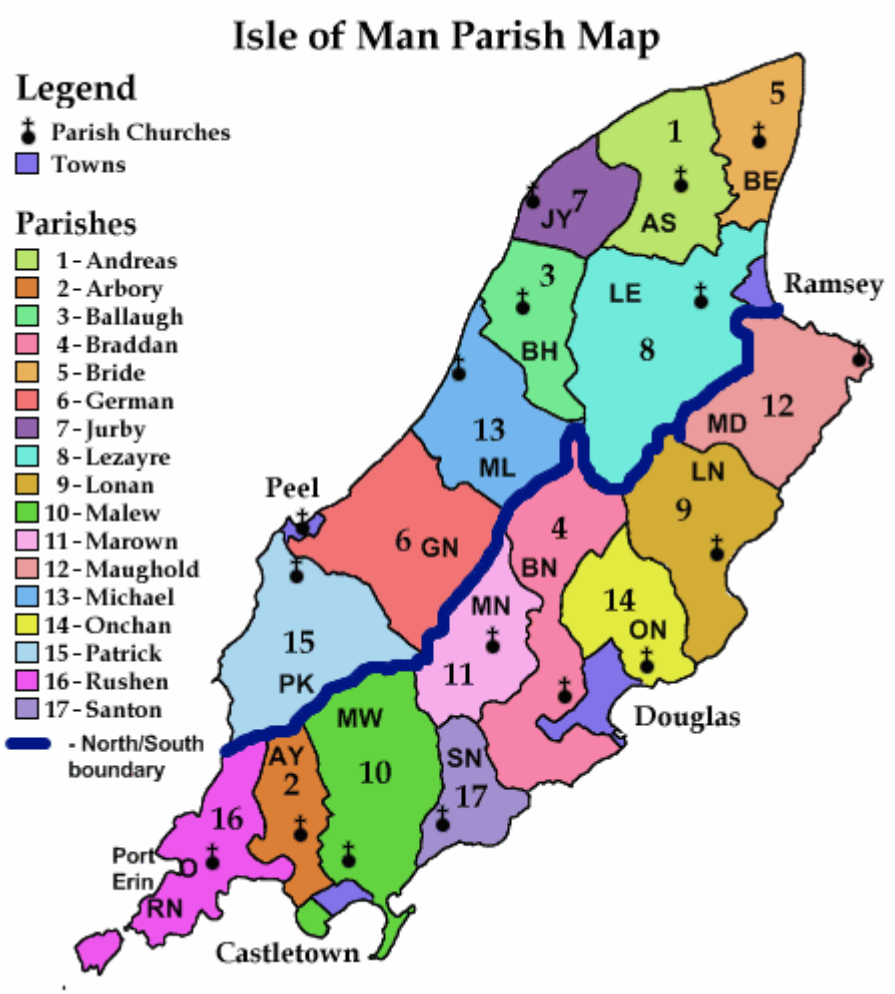



*Figure 3.2: Parish map of the Isle of Man showing the northern and southern regions (amended from (Isle of Man Guide 2014)*

The Island is divided into 24 districts, made up of 4 towns, 4 villages and 16 parishes (Fig. 3.2). According to the 2011 census the population of the Island was just over 84,000, with almost a third (27,938) living in the capital, Douglas. The biggest centres of population outside of Douglas are Onchan (9,172), Ramsey (7,821), Peel (5,083), Port Erin (3,530) and Castletown (3,097). Only 48% of residents were Manx born, and of the remaining 52%, only 7.5% were born outside the British and Irish Isles (Isle of Man Government 2011).

### 3.1.2 Language and Place-names

The Isle of Man lies in the region of the British and Irish Isles often referred to as the 'Celtic Fringe', where Celtic languages have been spoken most recently. Manx Gaelic is part of the Goidelic (*q*-celtic) language family, in common with Irish and Scots Gaelic, differing from Brythonic (*p*-celtic) which includes the languages of Wales, and Brittany, and the extinct language of Cornwall (Kinvig 1975, 16-17).

*Figure 3.3: 5th Century Ogham stone from Rushen (Manx National Heritage 2011). The inscription reads '..of Bivaidu, son of the tribe of Cunava(li)' Cunava is thought to be the early form of the surname Cannell or Connell (Kinvig 1975, 49).*

The main source of early linguistic evidence on the Isle of Man is in the form of Ogham stones. Ogham is the earliest written Irish language found mostly in the form of inscriptions carved into stone and the earliest Ogham stones are found

in Ireland dating to the 4th Century CE (Charles-Edwards 2013). They proliferated during the 5th and 6th centuries and are found all over the western regions of the British Isles known to have been Irish colonies during the time of Irish Diasporas in the 5th century CE. It was Irish settlers called the Scotti who ultimately gave Scotland its name.

Ogham stones are found all over the Isle of Man (Fig. 3.3) indicating the presence of Irish speakers on the island from the 5th century onwards. However, many of the Ogham stones bear both Irish and British Gaelic as well as some Latin inscriptions suggesting that the population was probably bilingual during this period (Charles-Edwards 2013; Jackson 1953).

It is difficult to establish what language was spoken before the 5th century as no written records exist on the island from before this period. However, there is evidence that culturally and linguistically the inhabitants were Brythonic (Phillips 2004). In early 5th century Ogham inscription bearing an Irish name written in a Brythonic accent is considered strong evidence that the language spoken on the Isle of Man was Brythonic (Jackson 1953, 173). A second piece of evidence comes from Brythonic place-names recorded in the Chronicles of Man which have been used to suggesting a strong pre-Goidelic Brythonic influence (Broderick 1999).

It is not clear whether the modern Manx language survived or arose after the Norse period. Whichever is the case, it is most similar to Gaelic spoken by people of Galloway and the Western Isles of Scotland (Charles-Edwards 2013; Phillips 2004). These dialects would have been mutually intelligible during the Early Medieval period, but separated some time in the 14th or 15th centuries CE (Phillips 2004). This connection with Scots Goidelic is reflected in the many Manx surnames, prefixed with the patronymic 'Mac' when recorded in the 16th century

documents. However, the 'Ma' had long since been dropped leaving the hard 'c' sound, hence there are a disproportionate number of modern Manx surnames beginning with C, Q or K. So, for example, at some point McThomas, became Comas and McWilliam became Quilliam (Jones 1998).

Norse is thought to have been introduced to the island during the Viking period (ninth to thirteenth centuries CE) and was probably spoken alongside Irish Gaelic (Phillips 2004). Place-names on the Isle of Man now are mostly of Manx Gaelic or Scandinavian origin, although there are a few English-derived names (Moore 1890). Very few names that can be dated to the pre-Norse period have survived (Fellows-Jensen 1983; Kneen 1925, 9 – 21) perhaps reflecting the longevity and severity of Norse dominance over the island. However, there is evidence that there were more Gaelic names surviving in the thirteenth century, typically for topographical features rather than settlements, which indicates that Gaelic was still spoken throughout the Norse Period (Fellows-Jensen 1983).

## 3.2 Archaeological context

The archaeological record of the Isle of Man shows a rich and complex history of human occupation stretching back 10,000 years. This section will describe the archaeological evidence for each recognised era with reference to the implications for the demographic history of the island. Indications of possible migrations and colonisations provide useful clues about the ancestry of living Manx people, and in turn feed into theories about the population's history.

Archaeological data are particularly useful for determining likely time-frames and possible source populations for the patterns observed in current gene-pools. This information can then be used to design model scenarios to test

theories about admixture using genetic data.

### 3.2.1 Initial colonisations

The last glaciation, called the Midlandian glaciation, lasted from approximately 115,000 to 10,000 years ago, and for much of which the British and Irish Isles were frozen waste lands. During this period, Britain and Ireland were connected to continental Europe by land bridges. During periods of relative warmth, mammoth, deer, bears, wolves and other intrepid mammals were able to range from mainland Europe to as far as Ireland (Waddell 1998).

Human-related archaeological finds in Britain during the later parts of this period are restricted to Wales and England, south of Yorkshire (Barton 2009; Pettitt and White 2012). The scarcity of archaeological sites and artefacts dating from this period seems more consistent with short-term occupation, perhaps indicating hunting parties visiting during milder summers rather than longer or more regular visitations (Pettitt and White 2012). No evidence has been found to suggest that people reached the Isle of Man during this inhospitable period.

By 10,000 years ago the tundra of Ice Age Britain had given way to deciduous woodland of hazel, oak and birch, with red deer and pine martens replacing reindeer and wild horses (Milner and Mithen 2009). However, the land bridge between Britain and Ireland may have been severed between 12,000 and 10,000 years ago. This was before the first human colonists could arrive and also before the arrival of many of the flora and fauna seen in other regions, leaving Ireland much lower in native species than Britain and mainland Europe (Mallory 2013, 30-33). The English channel developed around 10,000 years ago, cutting Britain off from Europe (Bradley 2007) and the Isle of Man became an island sometime after this, although the actual date is unknown (McCartan 2004).


*Figure 3.4: Map of putative Mesolithic coastline (in beige) around Irish Sea and Isle of Man (redrawn from Mallory 2013, 35).*

This rapid rise in atmospheric temperature marked the beginning of the Mesolithic Age (Milner and Mithen 2009) as well as the beginning of the Holocene. Human activity of this period is characterised by small triangular flint tools named microliths, some of which may have been the projectile points of arrow and spears (Milner and Mithen 2009). The first colonists to arrive in Britain are thought to have arrived from the south, possibly from France, but largely

descended from people in the Iberian Ice-Age refugia, who had gradually migrated north as temperatures and the environment became more hospitable (Gamble et al. 2004).

The first colonists to reach the Isle of Man probably arrived by boat sometime before 7600 years ago (Bradley 2007; McCartan 2004) or it is possible that the Isle of Man was still connected to parts of Cumbria and south-west Scotland (Mallory 2013, 58-60). Mallory (2013) argued that the now submerged lands around the Isle of Man may have also held the source population for Ireland. The channel separating the shores would have been narrower (Fig. 3.4), but even now on a clear day it is possible to see Ireland from Man, so given the right impetus, such as hunting grounds increasingly inundated by sea-water, people may have felt compelled to brave the sea to reach new land (Mallory 2013), possibly using small boats like curraghs (Cunliffe 2001, 67).

The earliest sites in Ireland known so far, pre-date the earliest sites on the Isle of Man by about 1000 years, however, the land area surrounding, what was to become the Isle of Man could have been up to 15,000km$^2$ in area, stretching from North Wales to the Solway Firth (Fig. 3.4; Mallory 2013; 65). However, any evidence there may have been for Mesolithic occupation is now submerged beneath the Irish Sea (ibid; 60).

Ireland may have been a relatively impoverished environment for the first human colonists, lacking much of the diversity of food plants and animals seen in Britain (Mallory 2013). Wild pig seems to have been a major food source for Irish Mesolithic hunters, and there is some evidence that they have been introduced by the first colonists to arrive (Mallory 2013, 42-46). Wild pigs also seem to have been an important food source for Mesolithic Manx, lending support to the argument that the Isle of Man was the source population for both humans and

wild pigs (ibid, 58-60).

There are parallel developments of the Mesolithic tool-kits seen in Britain and other parts of Europe until 8000 years ago. After this Mesolithic technologies in Britain and Ireland begin to diverge from mainland Europe and each other, suggestion a period of isolation (Cunliffe 2001, 115). Developments in Late Mesolithic microflint and macrolithic manufacture on Man are very similar to those of the communities in Northern Ireland (McCartan 2004; 2000), suggesting at least some level of contact between the populations.

Population densities are estimated for this period are thought to be between 0.3 and 5 people per $km^2$ based on population densities of modern hunter-gather societies (Cunliffe 2001). Giving an estimated Mesolithic Manx population of between 190 and 3000 people, but probably towards the lower end of this range, given the environment.

### 3.2.2 The transition to a farming economy

The Neolithic Age, marked by the development of agriculture, began in the Near East approximately 8000BCE. The new technologies are thought to have spread through Europe via two distinct routes; across Central Europe and through the Mediterranean, eventually meeting in western France at from around 5000 to 5500BCE (Tresset 2003). It was a time of great cultural change; farming bought a new way of existing on the land with more control over the environment.

Pottery gave people new ways of handling, storing and serving food and there is evidence of feasting (the Atlantic Neolithic is reviewed in Cunliffe 2001; 139 – 158). Beautifully crafted and polished axes made of decorative stone such as jadeite, thought to be quarried in the Alpine regions of continental Europe, are found all over the British and Irish Isles including the Isle of Man. These items

provide evidence that exchange networks may have stretched across Atlantic Europe (Pétrequin et al. 2008; Waddell 1998, 49).

The bones of domesticated animals begin to appear in Irish Mesolithic contexts from around 5500BCE indicating contact between Irish hunter-gathers and continental farmers (Tresset 2003). However, from 4000BCE, agricultural innovations are found almost simultaneously all over the British and Irish Isles (Thomas 2008). The rapidity of the spread has been attributed to multiple introductions via sea-ways, from western France up to the Netherlands (Cunliffe 2011, 137). There is much debate over whether these technological and cultural developments involved colonisation by continental farmers or adoption of practices by indigenous people of the Isles (Sheridan 2003; Thomas 2008; 2006).

When Neolithic site development was modelled using C14 dates as a proxy for population density, results suggested that population growth in Britain during the period was more compatible with immigration than with a peak in the birth rate due to indigenous adoption of farming technologies. Furthermore, the data suggested two points of colonisations, first in southern England, and second, possibly less than a century later, in western Scotland (Collard et al. 2010). Although the study did not include the Isle of Man it seems probable that under this scenario the Manx would have received their Neolithic package during the same Irish Sea-colonising diaspora as Scotland.

The Neolithic package included domesticated animals and plants as well as the first pottery, and was received complete in the British and Irish Isles with few indigenous innovations. Of the domesticates, only pigs were native to the British Isles, so domesticated species other than these must have been introduced from continental Europe. Similarities in pottery-style and tomb design, and layout, as well as the use of quartz crystals in tombs, suggest north-western France as the

possible source for farming technologies, around the Irish Sea, including the Isle of Man (Sheridan 2003).

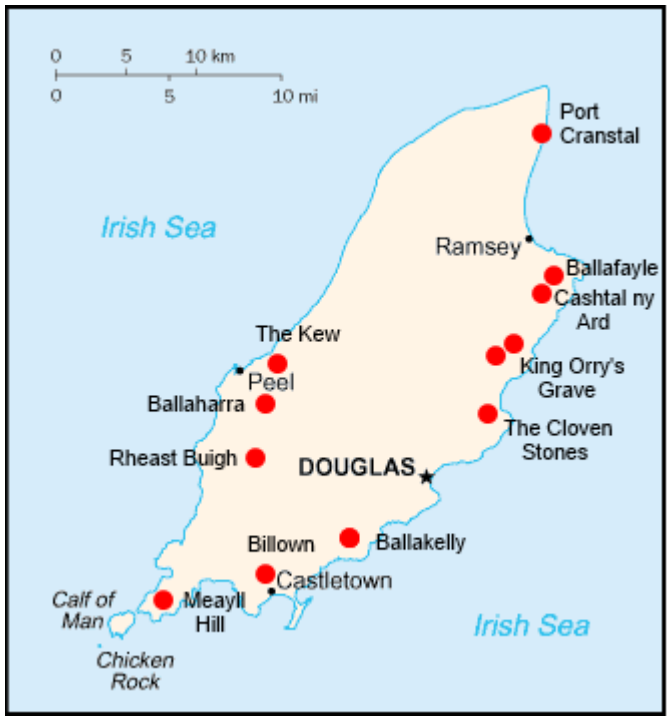



*Figure 3.5: Map showing the twelve remaining Manx Neolithic monuments sites.*

There is no definitive date for the arrival of the Neolithic on the Isle of Man. Carinated pottery, characteristic of the period elsewhere in Britain has never been found on Man. Further, a Mesolithic-type assemblage was found at a site in Rhendhoo dating to between 4150 and 3650BCE, indicating that inhabitants on the island may have maintained a Mesolithic way of life after acquiring farming technologies. However, from 4000BCE there are plenty of middle Neolithic site providing evidence that the Isle of Man was part of the Irish sea Neolithic cultural network (Burrow 1997, 9, 11).

Megalithic structures containing mixed bone assemblages are a feature of this period and are found all over the Atlantic region, although there are regional distinctions. Of the twelve remaining Manx Megalithic tombs (Fig. 3.5), some

have features completely unique to the island most share characteristics in common with other tombs around the Irish Sea region (eg Fig. 3.6). In particular, like many Irish Sea megalithic tombs, all Manx Neolithic monuments are located in elevated positions with views of sea and distant mountains (eg Fig. 3.7), suggesting that connections to seaways and other populations within the region were significant to the builders (Cummings 2004).

*Figure 3.6: Cashtal ny Ard, megalithic tomb in Maughold, Isle of Man (Manx National Heritage 2014)*

Pottery styles also seem to reflect this connection with other Irish Sea communities too (Burrow 1997). In the later Neolithic the Isle of Man became

more insular, demonstrated by the development of the indigenous Ronaldsway culture and pottery traditions, which were unique to the island (Burrow and Darvill 1997), possibly indicating a period of isolation or rejection of neighbouring communities.

*Figure 3.7: Meayll Circle on Meayll Hill above Cregneash on Isle of Man (Manx National Heritage 2014).*

### 3.2.3 Metal Ages and exchange networks

From around 2400BCE the exploitation of metals for tools and weapons became widespread in Western Europe marking the beginning of the local Bronze Age. The production and acquisition of metal objects and weapons was possibly a source of power and status (Champion 2009) and contemporary changes in burial style appear to indicate the rise of the importance of particular individuals over that of the community (Ibid, 143-144). These cultural innovations are accompanied by the use of a new beaker style of pottery (Gibson 1982) and there has been much debate over whether these cultural innovations were spread by migrants from the continent (Pearson 2012).

Isotope analysis of the Amesbury Archer, a particularly rich burial associated with beaker pottery found near Stonehenge, indicated that he spent his childhood in the Alps (Evans et al. 2006). Seven other individuals with isotopic signatures of distant birth-places have been found to date (Pearson 2012) providing evidence that at least some people were mobile during this period. It may be that centres of prestige such as monument buildings, like Stonehenge, were also cosmopolitan places where traders and pilgrims from all over Europe met, exchanging exotic goods, raw materials, ideas, stories, and possibly, genes (Cunliffe 2001).

The later Bronze Age saw the intensification of established coastal sea routes and rivers, as boating technology increased mobility of at least some people (Cunliffe 2012). The western seaways may have been important for linking the communities of Ireland, Wales and western Scotland, and it is easy to imagine that the Isle of Man was an important midway point given its location in the middle of the Irish Sea (Waddell 1998). Perhaps these networks formed a link between the southern Atlantic and more northerly communities.

It is evident that raw materials as well as items made from them were exchanged between communities. These items of prestige were passed along networks likely to have been established during the Neolithic. Gold objects from Ireland are found in Amorica, and Iberian gold neck rings have been found in Britain and Ireland (Taylor 1994). Bronze daggers decorated with gold and apparently made by the same craftsman, have been found in graves in Brittany and Wessex (Cunliffe 2001; 255). Gold cups found in Germany and Cornwall seem to have been made in the same workshop and were possibly inspired by Iberian workmanship (ibid, 256). There is also some speculation that this is the period which saw the development and spread of the Celtic languages, which include

Manx Gaelic, as the lingua franca of the Atlantic region (Cunliffe and Koch 2010).

The Atlantic regions of Europe are particularly rich in metals such as tin and copper, needed to make bronze, and also gold and silver (Cunliffe 2001), and there is strong evidence of metal mining and working are found at sites in Ireland and Wales (Pearson 2009). Ore-smelting first seems to have developed on the Atlantic Fringe in Iberia, and there is speculation that metal prospectors and their families may have moved from there to other regions where the raw materials were to be found, such as western Britain and Ireland, bringing their knowledge of metal working with them (Cunliffe 2012; Roberts 2008).

However, on the Isle of Man few beakers indicative of the period have been found. Although several bronze items have been unearthed on the Island, most were found out of context without organic material necessary for radio-carbon dating (e.g. Figure 3.8; Davey et al. 1999). The island has its own sources of copper, at Langness and Bradda, and hammer-stone tools, putatively dated to 1500-1100BCE and found at Bradda (Pickin and Worthington 1989), indicate that these sources were being exploited, although not everyone agrees on this (Davey et al. 1999).

*Figure 3.8: Late Bronze Age (900BCE – 700BCE) sword, found at Berrag in Jurby (Image from Manx National Heritage 2011)*

Stone tombs and cists of the type found all over the British and Irish Isles are found on the Isle of Man, containing remains from both inhumations and cremations (eg Woodcock 1999a; 1999b), showing developments and changes in common. These commonalities demonstrate some level of contact between the Isle of Man and the wider region over this period.

After 1400BCE cremation burials were the most common burial practice for most of western Britain and Ireland, with south-east England following the European practice of burying cremated remains in urns. However, these burials only accounts for a minority of the population, so may have been alongside archaeologically invisible means of disposing of the dead (Cunliffe 2012). On the Isle of Man burials from this period of any type are rarely found until the Christian era 1000 years later (Manx National Heritage 2011).

At around 800BCE it is thought that Europe went through massive social, political and economic upheaval, marking the beginning of the Iron Age. These changes were thought to be due to the breakdown of ancient exchange networks (Reviewed in Cunliffe 2011; Chapter 8), and world-wide environmental changes making it colder and wetter (Van Geel et al. 1996), as well the growing importance of iron devaluing bronze (Cunliffe 2001, 336 – 365).

Then at the beginning of the Iron Age there is evidence that contact was still maintained between the south of England with the continent, in the form of swords and other items typical of the European Iron Age. These items, although concentrated in Thames Valley archaeological contexts, are also found all over Britain and Ireland as well as France, Belgium and Netherlands, making the Thames Valley the likely exchange route, linking the British and Irish Isles to the continent (Cunliffe 2001; 2012). This contact was probably partially secured by the ancient tradition of British metalworking crafts people.

By 600BCE exchange of metal work between south east England and northern France seems to have diminished to a trickle. There is little evidence of the few European items travelling beyond the Thames Valley (Cunliffe 2001; 322 – 323, Bradley 2007; 233). However, other less archaeological visible items may still have been traded; horses, dogs, furs, and even slaves (Cunliffe 2012). In the

period immediately prior to Roman contact and conquest, there is evidence that links between southern England and France had re-established in the form of shared pottery styles and coins (Cunliffe 2012).

*Figure 3.9: Iron Age fort at Cronk ny Merriu with later Viking Age rectangular building (Image from Manx National Heritage 2011).*

During this period the Irish were very active according to their early literature, which is full of tales about the exploits of their Iron Age heroes (Mallory 2013). Some of these stories feature Manannan mac Lir, meaning 'son of the sea', who was a merchant and a skilled sea-man who lived on the Isle of Man. The Manx name for the island, Ellan Vannin, translates as Manannan's Isle (Chadwick 1970) and is possibly where the English name derives from too.

The shift to iron technology in Ireland seems to have occurred later than in other European contexts, and the majority of weapons recovered from earlier in the period are bronze (Mallory 2013, 159). Horse-related equipment, which was

commonly found in Britain and even more common in Europe (Cunliffe 2001, 320-321), was not found until much later in the period (ibid, 359). All this is taken to indicate that Ireland became insular at least during the first half of the Iron Age. There is some suggestion that Atlantic communities may have been deliberately shunning the new iron technologies, in favour of bronze items (Henderson 2011; Mallory 2013).

The spread of certain art styles in the later Iron Age which were inspired by continental forms demonstrates that knowledge and ideas, were circulating (Cunliffe 2001, 324-328; Mallory 2013). Since ideas require agents to spread, some people must still have been mobile during this period. The skull of a barbary ape, native to North Africa and Gilbraltar, was found in Fort Navan in Northern Ireland, dated to the later Iron Age and is evidence that Ireland may have had some long distance connections (Napier and Jenkins 1997).

The Iron Age may have been a period of relative poverty on the Isle of Man. Some luxury items such as glass beads and brooches (Bersu 1977) have been found on the island, maybe indicators of exchange, as no evidence of insular iron working have ever been found (Bersu 1977). Although this may be due to a paucity of data, it is also possible that changing climate would have resulted in much more blustery conditions affecting surrounding seas, making sea routes more unpleasant, dangerous, and discouraging to outside exchange. Many large roundhouses have been excavated on the island, showing evidence of prolonged occupation over many generations (Bersu 1977). These are largely undefended sites, although there are a handful of forts (Figure 3.9; Kinvig 1975; 31-35) along the coastline, facing out to sea (Frieman 2008).

The Romans, were aware of the Isle of Man, through the writings of Ptolemy (100CE, Book 2, chapter 1, trans. Stevenson 1932) who called it Monaoeda.

However, they never made it part of the empire and there is no evidence that they even visited. Only a handful of Roman coins have been found (English Heritage National Inventory 2014) but it is possible that some trade or maybe raiding may have occurred between Roman-occupied Britain and Man (Kinvig 1975; 32).

### 3.2.4 Christians and Pagans

The Irish Sea is thought have dominated by the Irish in the first millennium CE, and raids on Britain for slaves were common from 360CE onwards. Saint Patrick was a slave taken by the Irish who escaped but later returned to Ireland as a missionary to convert his former captors (Saint Patrick 470CE 1993). It is not certain when the Isle of Man became Christianised, but many of the parishes and churches have been named for the various Christian missionaries who were travelling around in the first centuries of the common era, including Patrick who gives his name to the eponymous parish and Bridget who gives her name to Bride. Over 200 early Christian churches, known as keeils on the island, were built during these early centuries. Many appear to have been built on existing burial sites or sites of spiritual significance (Kinvig 1975, 45-47; Wilson 2008).

The Isle of Man, like Ireland, avoided the influence of the Anglo-Saxons and remained Christian whilst England returned to paganism under Germanic rule. The Christian period of the island is also famous for its stone memorials and crosses. This was a tradition that pre-dates the Christian period, and is considered to have been brought to the island by the Irish some time after the 5$^{th}$ century CE (see Fig. 3.3). The earliest Christian stone slabs (Fig. 3.9) date to about 650CE and the custom of carving stones was to last well into the Viking period with stones bearing Scandinavian art and Norse inscriptions (Cubbon 1971).

*Figure 3.9: Calf of Man, an early Christian stone carving, dating to c. 800CE measuring 65.3 cm x 24.8 cm x 3.5 cm. It pictures Jesus underneath a cross and is thought to be influenced by Irish and Mediterranean art traditions (Manx National Heritage 2014)*

From about 500CE Irish settlements were springing up all over western Britain, forming Irish-speaking colonies in Pembrokeshire, Wales and Galloway in Scotland marked by the appearance of Ogham stones (see Fig. 3.3) and a contemporary Spanish historian recorded that the Isle of Man was under Irish rule in the 5th century CE (Charles-Edwards 2013, 148, 174). Latin and British

inscriptions some stones have made linguists suggest that the population of the island were bilingual (see section 3.1.2, this Chapter).

According to the Anglo-Saxon Chronicles 'heathen men' began raiding along the coasts of Britain, some time between 786CE and 802CE. Small bands of men, initially Norwegians, were targeting richly furnished churches and monasteries in the decades between 750CE and 830CE along both sides of the coast of the North Sea and English Channel (Cróinín 1995; Wormald 1991).

Around 830CE the number and intensity of these raids increased as Danes joined the Norwegians and also began raiding (Cróinín 1995). Raiding parties are thought have been relatively small, just 50 boats, and it seems that the same parties were raiding different regions (Wormald 1991, 145). At around 850CE, the nature of the raiding parties changed, the numbers involved increased three- to five-fold, and the Anglo-Saxon Chronicle records the armies arriving at this time as 'great' (Ibid 145). Up until this time, they had not shown any interest in settling for longer than the winter but at around this time the raiders were after land (Ibid 145).

It is thought that by the time the Vikings were conquering England they already had bases in Orkney, Shetland and the Western Isles (Richards 2009) and established semi-permanent bases in Ireland including Dublin (Valante 2008; Wormald 1991, 144). Viking warlords formed alliances with Irish kings, in some case through marriage; the ruler of Dublin, Olaf, was married to a daughter of the famous Irish ruling Uí Néill clan (Valante 2008) and Irish men joined raiding parties plundering other Irish regions. Despite this general social instability, the Viking presence seems to have helped develop Dublin as an important hub in the Irish Sea trade network (Ibid 54-56).

The alliance between the Vikings and the Irish was not to last, and in 902CE the

Irish king evicted the Norsemen from Dublin, who then settled in parts of Western Britain including the Isle of Man (Valante 2008). The relative poverty of the Isle of Man, as well as difficult surrounding seas, had perhaps protected it from raids during the early Viking period, but its strategic location surely would have made it an important stopping off or overwintering point. However, the earliest archaeological evidence dates from the early 10th century CE in the form of pagan burials, which fits with the timing of the expulsion from Ireland (Wilson 2008, 24 – 27; Valante 2008, 79).

*Figure 3.10: Boat-shaped Viking Burial (in the bottom half of the image) at Balladoole in Arbory; the top of the picture shows the remains of an older keeill- a Manx church, which was founded in the grounds of a much older Iron Age fort (Manx National Heritage 2011, Wilson 2008; 38 – 46).*

There are about 24 identified pagan burial sites that are associated with the Vikings found on Man. These are all inhumations found either in mounds or in flat burial sites. Balladoole in Arbory (Fig. 3.10) has a Viking burial within the Christian cemetery, cutting into earlier graves, although whether this was deliberate or not is uncertain. Distinctive boat-shaped burial sites marked with stones, such as this, are found all over Scandinavia. Unlike Christian burials, they

were often furnished, containing bones of animals, and some even show evidence of human sacrifice (Wilson 2008; 27). Excavation of the grave at Balladoole revealed the remains of a man, buried in a boat along with riding tackle and weapons; indicating that this was an important individual (Wilson 2008).

*Figure 3.11: Beads from the 'Pagan Lady's' Grave, found on St Patrick's Isle, Peel (image from Manx National Heritage 2011, (Freke 2002; 441).*

Another important Viking burial ground was found at St Patrick's Isle at Peel, which is a multi-period site that has had significance since prehistoric times (Freke 2002; 45- 49). Seven pagan, presumably Norse, burials were found at this site, dating from the 10th century; they included two infants and a middle-aged woman of considerable rank (ibid; 441). The 'pagan lady', is interesting not just because of the unique richness of her burial (Figure 3.11), but also because of what she represents. Although her burial is typically Scandinavian in character, her grave contained no specifically Scandinavian objects. Maybe she was native of the island and married into the Norse culture, and therefore buried by pagan relatives, or she could have been an incomer from Norway or perhaps a more local Norse colony, such as Orkney, Iceland or Ireland (Wilson 2008, 48-50, 441). The novelty of this woman's grave also highlights the scarcity of female pagan

burials found on the island, possibly indicating the low status of women in Viking culture. Or perhaps Vikings were marrying local women, either through their own free-will or maybe through coercion. Although as it is possible that Manx women were considered little better than slaves during initial contacts with Norse raiders (McDonald 2008; 29), this was probably not the case within one or two generations when local customs, in particular the Christian style of burial, was re-established (Cubbon 1983). Eight stone inscriptions on the Isle of Man are dedicated to women, at least one of which commemorates the Gaelic wife of a Norse man. These inscriptions imply some level of respect and affection (Hampton 2012). However, it is not clear whether these were Manx women or women from elsewhere in the Irish Sea (Wilson 2008, 89-90).

*Figure 3.12: 10th century Thorwald's cross-slab from Andreas (Manx National Heritage 2011); showing both the Christian cross and the Norse god Odin (Cubbon 1971, 32-33)*

There is evidence that slaves were the object of a number of earlier raids in the region and the Irish too had a long tradition of slave raids. The Irish Annals record that 700 slaves, largely women, were taken from Armagh in one raid (Valante 2008, 87). Slavery is difficult to detect archaeologically, but there are numerous sagas that mentions Scandinavians owning Irish slaves. In this way many people, particularly women, may have been transported around the Irish

Sea, up into Scandinavia and possibly even beyond into Eastern Viking territories, where, it is recorded, there was a considerable market for female slaves for working households or as concubines (ibid 88-89).

After a relatively short time the Norse incomers seemed to adapt to the local customs of the Isle of Man and the practice of furnished pagan burials disappeared as the Vikings became Christianised, rendering them indistinguishable from the locals. They also embraced the local custom of carving memorial stones, something that is not seen elsewhere in the Norse colonies or homeland. These stones have a uniquely Norse flavour, using Scandinavian-influenced art styles of the period and incorporating Norse iconography and mythology alongside the traditional Christian cross as seen in Figure 3.12 (reviewed in Wilson 2008; Chapter 3).

Archaeological evidence for Viking settlements and farmsteads is surprisingly rare. Roundhouses were the pre-Viking architectural style of the Manx, whereas Vikings preferred rectangular long-houses. Braaid in Marown (Fig. 3.13) contains a native roundhouse alongside two later rectangular buildings, one with wide curving walls, built in the Norse style. It is thought that Vikings took over a pre-existing farmstead (Cubbon 1983), a common Viking practice seen on Orkney and in other colonies (Wilson 2008; 96-98).

Only one other farm of Viking age has been definitely identified at Doarlish Cashen in Patrick (Cubbon 1983). What is curious about these two sites is that they were both built on marginal land; Doarlish Cashen is on hill land 200m above sea level and the farm at Braaid was thought to have been eventually abandoned when it became too boggy. So despite the dearth of archaeological evidence for Viking Age farms the location of these settlements is taken as evidence that Norse settlement was dense on the Isle of Man (Cubbon 1983;

Wilson 2008, 98-99).

*Figure 3.13: Viking settlement at Braaid in Marown, Isle of Man*

Some of the Iron Age forts seem to have been taken over by Norse settlers and contain rectangular houses of the Viking Age (e.g. Fig. 3.9), but many of these date to the later Viking period (Wilson 2008; 92-96). If distributions of burial mounds across the island can be taken to indicate the extent of Viking settlement, then Norse settled all over the Island. However, it is possible that the indigenous Manx were copying the Viking style of burial.

In short, archaeological evidence for Norse settlement is frustratingly lacking or inconclusive. The strongest evidence for dense Norse settlement comes from place-names, which have almost completely replaced Gaelic ones. This indicates

that Norse was the dominant language spoken by the ruling elite, and maybe the lingua franca of the period. It is possible that the Gaelic language may have continued to be used throughout the period. Subsequent generations could have been bilingual, having learnt Norse from their fathers and Gaelic from their mothers. This would have been particularly useful given continued close relationships with Gaelic-speaking peoples of Scotland and Ireland (Wilson 2008; 100-103).

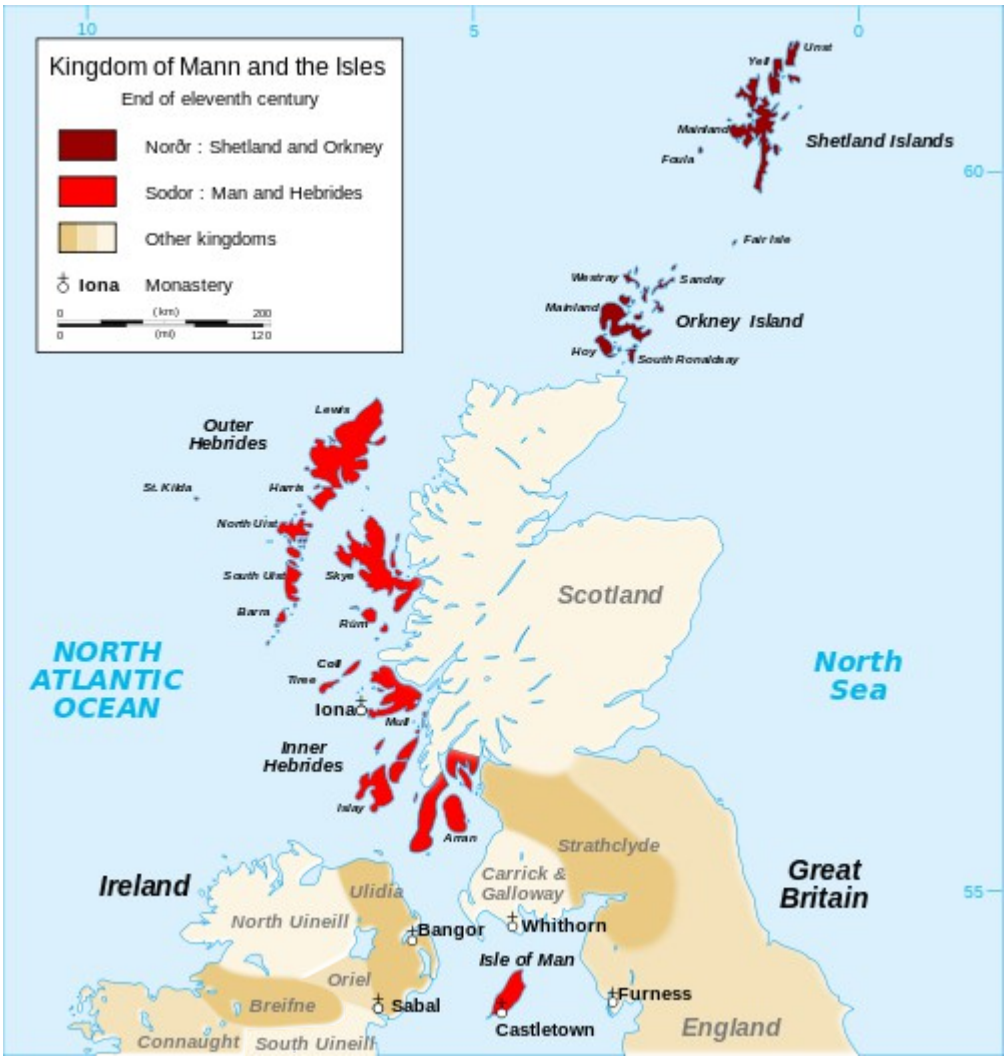



*Figure 3.14: Map showing the extend of the Irish-Norwegian sea Kingdom, circa 1300 CE (Image from WikiCommons 2014).*

The Isle of Man was under the rule of its Nordic colonists until the mid 13th

Century, along with the Western Isles, the Hebrides and the Shetland Isles (Fig. 3.15). Man was considered one of the most important islands in the Irish Sea archipelago Kingdoms, as various Irish, Norse and Scottish warlords and petty kings vied for power and control of the area (McDonald 2008; 26 – 33. Ethnic boundaries in the region became blurred through inter-marriage and fostering, which acted to cement kinship ties.

Fostering between families seems to have been culturally very important to both the Irish and Scandinavians; mentioned in many Icelandic sagas (Steinsland 2011, 7; Valante 2008, 91) and protected by Irish law from an early date (Downham 2014). The implication is that it was common during the Viking period, to the point where it was uncommon for all children to be raised at home (Jochens 2006, 267). Although, fostering of boys was more widely recorded it seems to have involved girls as well, and often foster-mothers would accompany their foster daughter when they married to help her raise her children (ibid, 267).

In a time when infant mortality was high, it would have served to redistribute children amongst families as well as strengthening ties. It seems to have occurred both between social strata (Jochens 2006, 267) and between Celtic and Norse families, as recorded in Manx stone inscriptions (Harding et al. 2014, 57). The people who emerged from these complex relationships and alliances, through marriage and fosterage became known as the Gall-Gaedhil, meaning foreign Gael (McDonald 2008; 26 – 33; Valante 2008, 91).

In the late 12$^{th}$ century Magnus Barelegs, King of Norway, claimed the whole of the Norse-occupied area for Norway with the agreement of the then Scottish king. This agreement between the two major powers of the time demonstrates that the populations of the western-most parts of the Scottish mainland and the Western Isles seemed to have a lot more in common with the peoples of the

Irish Sea and the Isle of Man during this period, than with the rest of Scotland, owing to the geography of the area. It was much easier to sail around these regions and up to Norway than it was to travel overland through the mountains to the power-centres of mainland Scotland (McDonald 2008, 35-37).

In reality, neither king managed to maintain authority in the area and local Norse/Gaelic leaders continued to fight amongst themselves. According to the Manx Chronicles, the island was ruled by Godred Crovan and his descendants from 1079. Eventually a treaty was signed between Scotland and Norway in 1265, giving the islands including Man back to the Scottish king (Cubbon 1983; Kinvig 1975, 58-66). The last of Godred Crovan's descendants, also named Godred, led the Manx in in rebellion against the Scots in 1275, but they were defeated and Godred, son of Magnus, was killed at Ronaldsway, ending the Norse Era (Cubbon 1983). The Norse Era left a legacy on the Isle of Man in its form of government known as the Tynwald which is still used today.

The following 150 years saw the Isle of Man ravaged by skirmishes between the armies of England and Scotland as each side fought for supremacy until 1346, when the island finally came under English rule.

### 3.2.5 The Post-Norse era

From 1346 the Isle of Man was owned by various Earls until finally, Henry IV gifted Man to Sir John Stanley of Derby on condition he pay homage to the Crown (Kinvig 1975, 86-88). The Earls of Derby ruled the Isle of Man from 1405 until 1765 when one of Stanley's descendants sold the Lordship of the island back to the English Crown (Kinvig 1975, 114).

The post-Norse era does not seem to have led to any extensive settlement on the island; the various ruling parties, if they came to the island at all, seem to

have conquered and then left. Even the great Stanley dynasty mostly ruled from afar, only the 2nd and the 7th Earls ever taking any real interest in the island (Kinvig 1975 chapters 6 and 7). The first indication that Man was somewhere that people might want to migrate to was during the 17th century with the rise of the smuggling trade on the island.

During the reign of the Stanleys, trade and commerce were tightly controlled by the governor of the island. Imports exacted a heavy duty and as there was little money in circulation on the island, trade tended to involve bartering. Export was only allowed if the governor deemed it could be spared and strangers could only buy for export under strict conditions. Added to this was the problem of England's Navigation Acts which only allowed goods to be brought into England in English-built and -manned boats and the duty was so high that the Manx people could not afford it anyway (Kinvig 1975, 119-120).

These prohibitive economic conditions led to a rise in illicit trade which was to lead to a flourishing smuggling trade with the Isle of Man at its centre. The late 17th century saw the ports of the Isle of Man in receipt of luxury goods from all over the world with the officials turning a blind eye as the money rolled in. With this trade came merchants and traders from elsewhere keen to make a profit. Eventually the British government tired of this situation and forced the last lord of Man to sell his rights of rule. They then passed the Mischief Act of 1765 and changed the tax laws; acts which were to put an end to this lucrative trade (Kinvig 1975, 114-117, 120-122). Many of the people who had made their home on Man whilst the situation was good, moved away again as the island was plunged into poverty (Ibid 132).

As the people of the Isle of Man rebuilt their economy, rediscovering the fisheries and agriculture industries which had been neglected during the

smuggling years, conditions began to improve. In the nineteenth century, the island became an attractive place to live for people on a low income that could not sustain them in Britain, but gave them a comfortable living on Man. It was popular with retired Napoleonic soldiers and also people on the run from debt in England (Kinvig 1975, 132-133). The later part of the century saw the rise in the Victorian tourist trade from England as the ferry service from Liverpool was introduced. Many of the richer tourists eventually emigrated to the island (Ibid, 150-151).

## 3.3 Genetic context

Focusing on the NRY, a number of research studies published in the early 21st century sought to describe and explain the genetic landscape of the British and Irish Isles. Many went on to try and find likely explanations for the patterns and affinities from archaeology and history, with varying degrees of plausibility. Through this research, the genetic landscape of Great Britain, and some of the associated Isles, has been well characterised. However, how and when these patterns emerged is still little understood.

The first half of this section will describe the genetic landscape of the British and Irish Isles, reviewing the current interpretations for the patterns found there. This will provide the context for beginning the exploration of the genetic diversity of the Isle of Man and where it fits within the British and Irish Isles gene-pools.

Although some regions have been extensively studied for particular time periods, there are obvious gaps in our knowledge due to a lack of of quality data and insufficient sampling. In particular, mtDNA diversity of the British and Irish Isles has received far less attention than that of NRY. In an attempt to fill these

gaps the second section will examine European research which focuses on ancient DNA and what this is revealing about the impact of the Neolithic on modern European genetic diversity. This research has implications for understanding the population history of the British-Irish Isles including the Isle of Man by providing temporal baselines for the formation of specific Europe-wide genetic patterns.

### 3.3.1 British and Irish Isles

Early studies using modern NRY data show that the British and Irish Isles can be divided into three genetically distinct regions; the first accounts for most of England excluding Cornwall and is most similar to mainland Europe, particularly Denmark and Germany. The second region, sometimes refered to as the 'Celtic-fringe' includes the western parts of the British Isles and Ireland, including the Isle of Man (Ellis 2003). These regions appear genetically most similar to south-western Europe, particularly Northern Iberia including the Basque Country. The third, encompassing Orkney and Shetland, is most similar to Norway (Amos et al. 2008; Capelli et al. 2003) Interestingly, in Britain, the distinct genetic regions roughly correspond to the upland and lowland areas.

These patterns and affinities manifest as, firstly, a significant cline in overall genetic diversity decreasing from east to west (Capelli et al. 2003; Weale et al. 2002), and secondly, the distribution of three common haplogroups; R1b, R1a1 and I (Fig. 3.15). R1b-M269 is the most common NRY haplogroup in western Europe. In parts of Ireland and Wales 80-90% of men carry a R1b haplotype, compared to England, which averages about 60%. Haplogroup I is more common in Eastern parts of Britain whereas haplogroup R1a1 is more common in Northern Scotland (Capelli et al. 2003).

Recent genome-wide analysis of the British Isles has used the program fineSTRUCTURE to group people by genetic similarity and compare them to people from mainland European populations. Even at the highest level of structuring, which divided Britain and Northern Ireland into 53 clusters, much of England remained an homogenous group that could not be broken down into smaller regions, whilst the rest of the British Isles and Ireland seemed to form distinct regional clusters (Hellenthal et al. 2014).

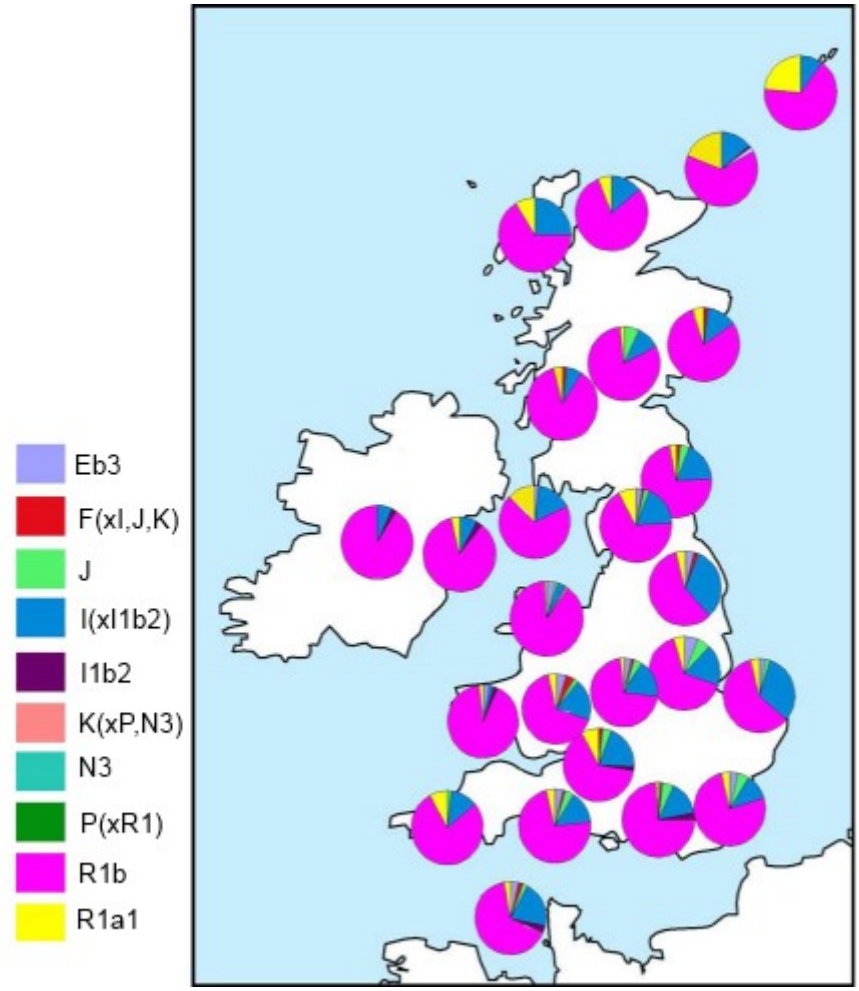



*Figure 3.15: NRY haplotype frequencies for populations of the British Isles. Data from Capelli* et al. *(2003); original image reproduced by permission of M. Jobling.*

Previously, the distinctiveness of most of England from the rest of the British

Isles, which was first detected in NRY haplotype distributions had been attributed to successive waves of immigrant men from genetically similar regions of Denmark and Germany, firstly as Anglo-Saxon and secondly, as Danish Vikings (Capelli et al. 2003; Weale et al. 2002). Continental men were thought to have displaced the native Britons, who were assumed to have been genetically similar to modern Welsh or Cornish populations (Weale et al. 2002). Numbers of immigrants required to produce these patterns were estimated at about 500,000 (Thomas et al. 2006). However, there is no archaeological evidence for migrations of these proportions, and instead archaeologists favour an elite-dominance model for the occupation of England by Germanic tribes during these periods (Hills and O'Connell 2009).

The NRY data were then used to explore the potential effects of social stratification during the Anglo-Saxon period, using computer modelling. The results indicated that restricting intermarriage between incomers and natives, coupled with societal advantage for incomers, could result in the same types of genetic patterns observed within a few generations without requiring a mass migration (Thomas et al. 2006). However, the resolution of the data at the time was insufficient to rule out long-term continental genetic exchange, perhaps dating back to the Neolithic or earlier (Pattison 2008).

Examination of the genome-wide data (Hellenthal et al. submitted) led authors to suggest that although Anglo-Saxon and Danish contributions were detectable in the modern gene-pools of Britain, they seemed to have had a minimal impact on the overall patterning observed. Hellenthal et al. (submitted) concluded that the patterns of differentiation they found in the British and Irish gene-pool, were a result consecutive, early pre-historic immigrations from mainland Europe which followed similar routes. Two routes of initial colonisation were proposed; across

the channel between opposite shores, and through the Irish Sea via coastal routes with France and Denmark seemed to have supplied some of the earliest migrants (ibid).

Regional distinctions were further reinforced by subsequent periodic isolation and historic migrations between specific population, such as the fifth century Scotti migration from Ireland to Scotland, and a twelfth century movement from Flanders in France into Wales. The genetic homogeneity in south-east England was attributed to the Romans breaking down tribal polities in this region which was to last throughout subsequent periods. Linguistic and cultural differences are thought to have provided a genetic barrier to prevent this homogenisation spreading to regions outside of England (Hellenthal et al. submitted).

Hellenthal et al.'s (submitted) analysis of the British Isles and Northern Ireland is the most comprehensive to date. Although it ignores the Isle of Man and the Republic of Ireland, it provides a detailed description of the patterns and affinities seen both within the British and Irish Isles and in comparison to Europe. However, the explanations for the patterns found therein are perhaps over-reliant on finding conveniently matching narratives from archaeology and history. Using specific events from archaeology or history to explain genetic patterns without attempting to ascribe the time-frames needed for these patterns to form, can lead to circularity. The relative order of specific European input was surmised by the amount of geographic diffusion from the putative landing point in Britain or Ireland. While this method provided a crude genetic temporal stratigraphy it largely ignored the effects of genetic drift, social structure or changes in population size, which could impact on patterning in

unpredictable ways.

Norwegian contacts with the British and Irish Isles are perhaps the best studied for both NRY and mtDNA, although results differ between different studies depending on the resolution and methodologies used. Norwegian NRY input for Orkney has been estimated to be between 31% and 55%, with Shetland Isle between 41% and 68% and the Isle of Man between 39% and 58% (Bowden et al. 2008; Capelli et al. 2003; Goodacre et al. 2005).

The estimates for female Norwegian contribution for Orkney are similar to that of Norwegian males, with 37%. The amount of Norwegian contribution further south decreases dramatically to 12% in Skye and the Western Isles (Helgason et al. 2001). The proportions of Norwegian mtDNA in British populations indicates that Norwegian women were also migrating during the Viking period. The largest proportions of female Norwegian lineages are seen in the most northerly populations of the British and Irish Isles, which are thought to be the oldest, most secure Viking colonies. Southerly colonies appear to have more Norwegian male than Norwegian female contributions to their populations, perhaps indicating that less Norse women migrated this far and so Norse men were marrying local women.

Recent ancient mtDNA data from late Iron Age Norwegians show a high degree of lineage sharing particularly with modern Orcadians, Shetland Islanders and ancient Icelanders as well as France and England (Krzewińska et al. 2015) provides support for the involvement of Norwegian women in colonisations and migrations of the Viking period. The genome-wide investigation indicated a much smaller degree of Norwegian input in to these northern gene-pools (24% into Orkney). It is possible that drift has played a role in increasing proportions of Norwegian NRY and mtDNA lineages present in the small island populations.

There is agreement that the similarity between Norwegian and British populations decreased with distance from Norway (Scotland 10/11%, England 3-7%, Wales 5-7%; Hellenthal et al. 2014).

Surprisingly, given the substantial archaeological and historical evidence for long-term interaction with Vikings, no evidence of NRY Scandinavian admixture has been found in Ireland (McEvoy et al. 2006). One study used a number of admixture methods to search for Scandinavian lineages in Ireland, McEvoy et al. (2006). However, they only targeted individuals with surnames of Norse origin, and it is possible that wider sampling would detect a Norwegian signal.

However, there are a number of problems with the use of admixture estimates. Firstly admixture analysis is dependent on the populations chosen as parentals. Historically, putative 'Celtic' populations have been created by mixing mid-Scottish and Irish populations (Bowden et al. 2008; Goodacre et al. 2005), or Irish and Basque populations (Capelli et al. 2003). This technique is problematic as it makes the assumption that through mixing these modern populations an ancient sample of different perhaps unrelated populations can be recreated.

Secondly, admixture methods which simply count number of lineages in modern populations and then assign them to probable parental populations, cannot account for the action of genetic drift (see Section 2.2.1) on the population since the admixture occurred. Genetic drift could have removed any traces of admixture from relatively large populations, such as Ireland, whereas the impact of Norwegian admixture could be exaggerated in smaller island populations. This means that the proportion of putatively Norwegian lineages in modern samples may have little bearing on the number of Norwegian migrants settling during the Viking era (Helgason et al. 2009).

MtDNA diversity of the British and Irish Isles is much less well described but it

exhibits more homogeneity than NRY its distribution across Europe (Fig. 3.16; Richards et al. 1996) and the distinctive patterning seen in NRY distributions in the British-Irish Isles is not apparent for mtDNA (Fig. 3.17; McEvoy et al. 2004; Wilson et al. 2001). In Britain as in most of Europe, the most frequent mtDNA haplogroup is H, carried by around 50% of people, whereas other haplogroups tend to be carried by less than 15% of individuals (Richards et al. 1996). However, exploration of mtDNA and NRY from modern and ancient sources is beginning to illuminate some aspects of British and Irish population history, explored further in the next section.

### 3.3.2 Impact of the Neolithic Transition

New data emerging from ancient DNA sources are providing insight into the development of patterns of genetic diversity within Europe. This is particularly true for mtDNA, but NRY and other nuclear DNA data are also becoming available and are highlighting the impact of the Neolithic transition on modern gene-pools (Pinhasi et al. 2012). Although no aDNA is available from these periods for the British-Irish Isles yet, patterns emerging in Europe can still tell us something about the population history of the Isle of Man.

At the time of writing, 83 mtDNA sequences from 32 Pre-Neolithic populations of Europe had been published (Brandt et al. in press) were dominated by mtDNA haplogroup U, which is only present in about 15% of modern samples. However, Neolithic samples' (238 samples from 30 sites; Brandt et al. in press), haplotype frequencies more closely resemble those of modern day European populations; with higher overall diversity than the Pre-Neolithic samples, higher frequencies of haplogroup H and lower of haplogroup U (Fu et al. 2012; reviewed by Brandt et al. n press).

Evidence from phylogenetic analysis of 1,151 complete mitochondrial genomes indicates that haplogroup U proliferated between 10,000 and 15,000 years ago, probably in response to the end of the Glacial period. Haplogroup H, however, rose in frequency around 7000 years ago, possibly in association with farming innovations (Fu et al. 2012). So far only five Pre-Neolithic NRY samples from two sites (one in Luxembourg and four in Sweden) have been recovered and these all belong to the haplogroup I (Lazaridis et al. 2014).

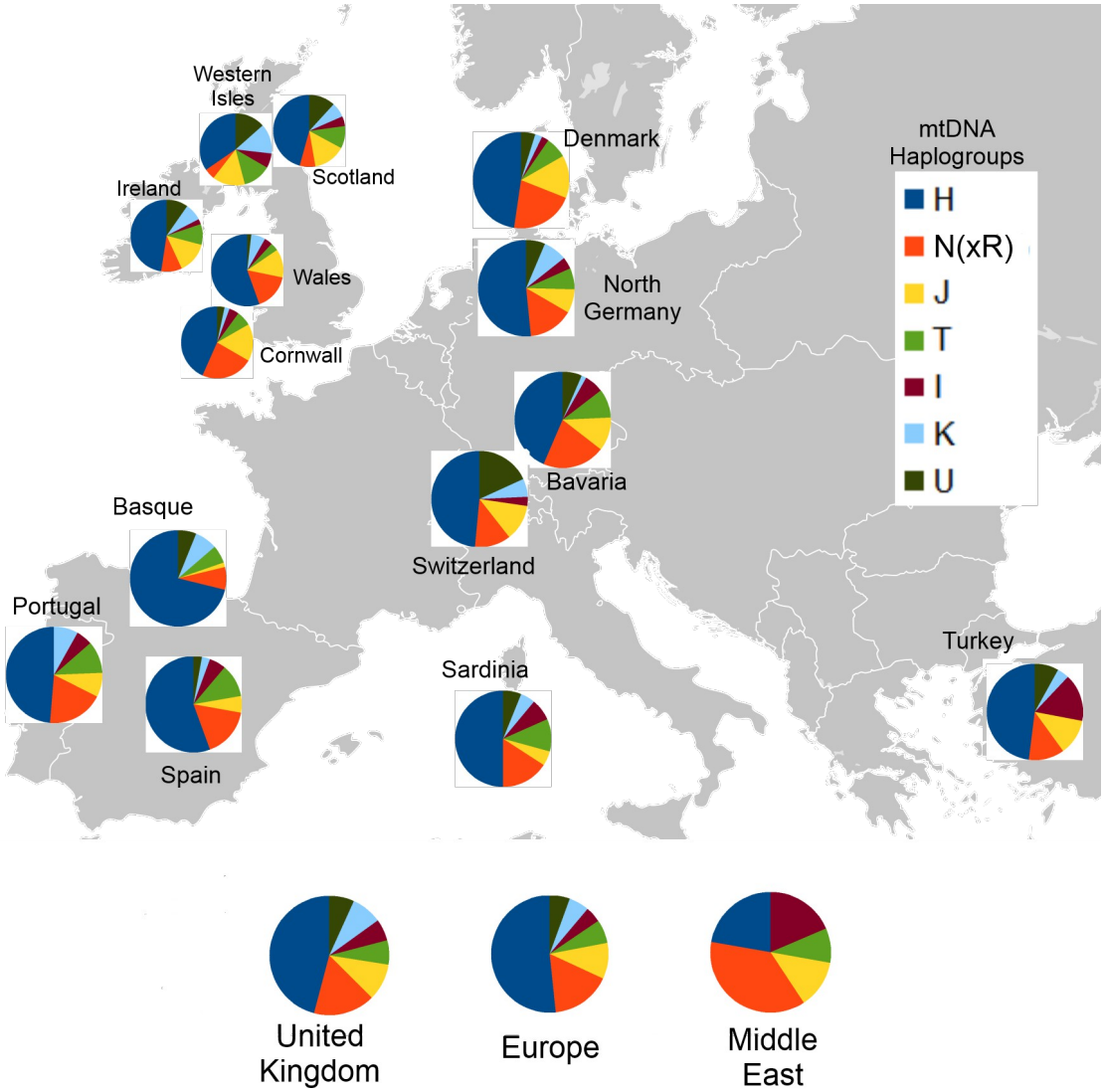



*Figure 3.16: Major mtDNA haplogroup frequencies for populations around Europe. Data taken from Helgason et al 2001 and Richards et al 1996.*

Only 40 sample of NRY have been recovered so far from three Neolithic sites in Germany (Haak et al. 2010), France (Lacan et al. 2011b), Catalonia (Lacan et al. 2011a) and nine from south east Europe. However, 27 of these individuals carried the G2a haplogroup, which is rare in modern population (Lacan et al. 2011b). Its presence at such high frequencies in Neolithic samples from across European sites suggest it was much more common and widespread during the Neolithic. However, the now very common R1b and R1a1, have not yet been found in any ancient assemblages (Brandt et al. in press; Pinhasi et al. 2012).

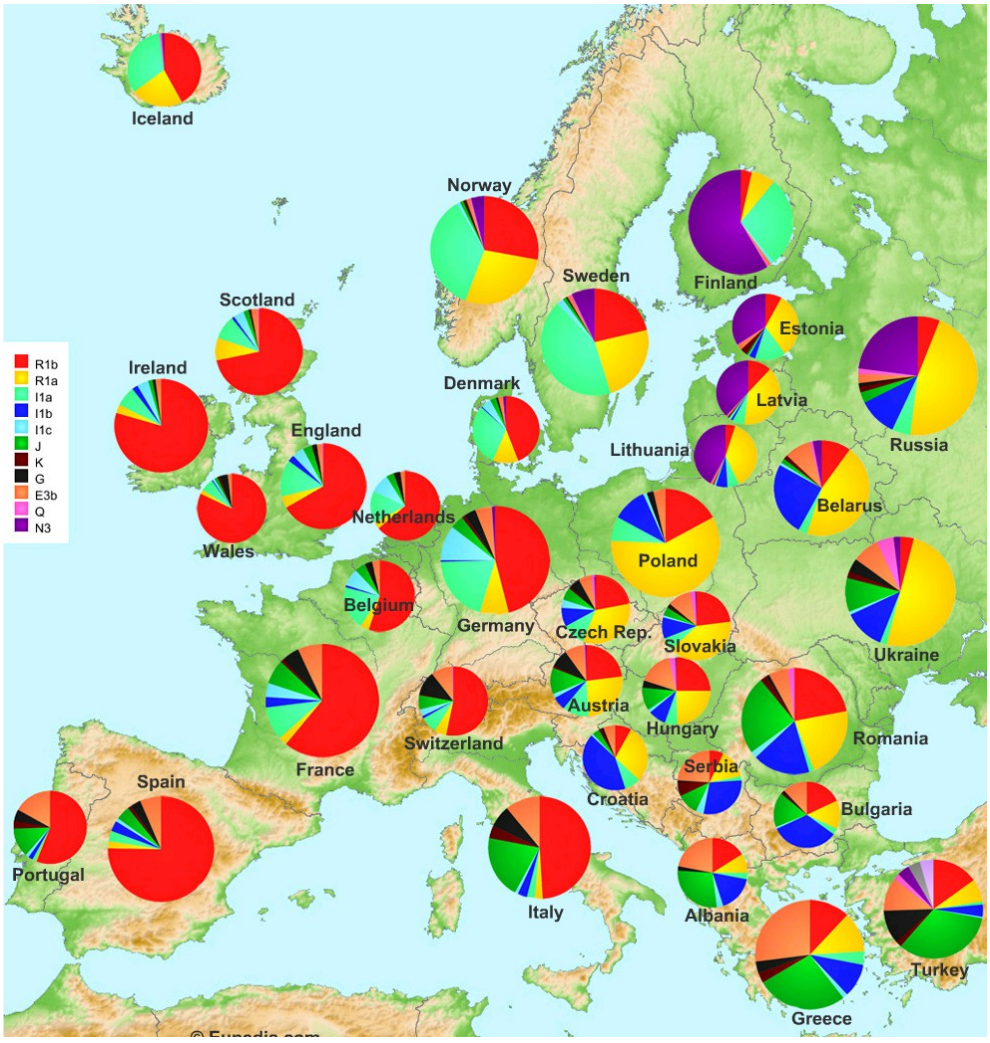



*Figure 3.17: Map of relative NRY Haplotype frequencies for population of Europe (image from Eupedia.com).*

The phylogeny (evolutionary tree) of NRY haplotype R1b is distinctly star shaped, indicative of a population expansion. Although the expansion dates to within the Holocene, there was insufficient data to pinpoint when or where exactly this occurred (Wei et al. 2013). Distributions and frequencies of the haplogroup throughout Europe were thought by some to be the result of Neolithic expansions of farmers out of the Near-East (Balaresque et al. 2010). However, the dating of this research was based on the notoriously unreliable STRs and the gradient of distribution out of the Near-East proposed by Balaresque et al. (2010) disappeared when the extreme western and eastern population of Ireland and Turkey are removed from the analysis (Busby et al. 2012).

Also computer simulations incorporating NRY and mtDNA from both modern and ancient samples found that patterns in both markers were most compatible with gradually decreasing Near-Eastern Neolithic input stretching across Europe, with very few Near-Eastern lineages contributing to the populations of north-western Europe, including the British and Irish Isles (Rasteiro and Chikhi 2013). For mtDNA, this seems to be supported by the relatively high proportions of haplogroup U found in the Western Isles and Scotland (see Fig. 3.17; Helgason et al. 2001).

British-Irish isles and European populations are generally remarkably similar in their mtDNA diversity (Fig. 3.17), suggesting that substantial mixing of females between populations must have occurred since the Neolithic. However, this homogenising effect could simply due to generations of women moving to neighbouring communities when they married (Destro Bisol et al. 2012; Seielstad et al. 1998; Wilkins and Marlowe 2006). The NRY aDNA data-set raises more questions that it answers, where did G2a come from? It was clearly widespread in the Neolithic, but what about before this, and where did it go? Also when did

R1b and R1a1 become so prevalent? More aDNA and better phylogenetic resolution for these haplogroups may help to answer these questions.

## 3.4 Summary and conclusions

First colonised during the Mesolithic, the Isle of Man's location in the middle of the Irish Sea has ensured it received innovations such as farming and metal working, as well as playing its part in periodic exchanges of other valued materials. There is also clear evidence of long-term contacts with other regions around the Irish Sea and particularly Ireland.

These contacts are reflected in the NRY genetic landscape of the region, through significant similarities between Ireland and the western parts of Britain, including the Isle of Man. What is still unclear is when these similarities first arose, and why and how these regions came to be so similar to parts of France and Iberia. Many researchers suspect that these patterns arose as a results of the Neolithic expansion through Europe.

Affinities between mtDNA diversity in the British and Irish Isles and Europe suggests that modern populations are predominantly descended from Neolithic people, who outbred and out-competed the hunter-gatherers in most regions. Although, in western Europe, few of these Neolithic lineages are purported to directly descend from Near-Eastern farmers, there is some evidence that the British and Irish Isles were colonised by mainland European farmers. These colonists were likely to have come from regions adjacent to British coastlines, and in the Irish Sea regions including the Isle of Man this would have been north-western France.

The similarities observed between these populations today, may not indicate a

one-off mass migration with population replacement 6000 years ago. These patterns could have grown steadily similar through repeated and intermittent low-level migrations over the intervening periods, possibly along networks of exchanged that evolved out of these colonising events.

The Isle of Man seems to have been protected from more recent invasions and colonisations affecting England and the east of Britain. Instead its influences were far more local, and it is likely to have been settled by the Irish during the 5th century CE. However, it is unclear whether this settlement was by both men and women. Later, however, its location and island status were no barrier to the sea-faring Norse, and it seems likely that significant numbers of Norwegian and Irish men settled there.

The slave trade was a prominent feature of the both the Early Irish Medieval period and the Viking Era and the majority of these were women. Some women may have been bought in from Scandinavia but most were likely to have been local to the Irish Sea area; some of these may have just been moved around the Scandinavian colonies, but some may have been traded up into Iceland, Norway, and beyond. This may have played some part in the homogenisation of female gene-pools in Atlantic populations.

# CHAPTER FOUR: DATA SOURCES AND METHODOLOGIES

## 4.1 Introduction

In order to carry out this research a source of DNA from the Isle of Man needed to be obtained, extracted and analysed. Before this could happen decisions needed to be made regarding whose DNA was required and how it could be obtained. There also needed to be data to use for comparison; some of this was available from Mark Jobling's laboratory, while the rest was gathered from published research. These comparative data needed to be compatible and equivalent to the Isle of Man data, with information from the same region of the genome, the same SNPs and STRs, and preferably sourced in a similar way to this project. The data needed processing to convert them all to the same format to make them ready for analysis.

This chapter is split into three sections, covering, data collection, DNA purification and typing, and data analysis. The first section describes how DNA was collected from Manx people, including how and why particular individuals were selected for sampling. This section also includes details and sources of comparative populations used in this research.

The second section describes the laboratory methods used to extract, purify and type the Isle of Man samples. The third section covers the methods used to describe and analyse the genetics of the Isle of Man, and to compare it to other populations from Western Europe.

## 4.2 Data Collection

In order to carry out genetic research it is necessary to have a source of DNA. Although the population under consideration was that of the Isle of Man, in order to learn about how the Isle of Man sits within the genetic landscape of Western Europe, other populations from the regions were also required. Comparative data for both NRY and mtDNA were taken from the literature or provided by colleagues in Mark Jobling's laboratory (see Tables 4.2 and 4.3 for details). However, the only published data available from the Isle of Man at the time of writing (Capelli et al. 2003) were of very low resolution by current standards so, as it was not possible to re-analyse those samples, a new sample was required.

Sampling relies on recruiting volunteer donors. Fortunately in western Europe there is no shortage of people willing to donate their DNA for projects if it can provide information about their ancestry. However, a self-selecting sample group may contain biases as the Isle of Man has a long history of attracting migrants from its nearest neighbours. Recent censuses show that only around 50% of the modern Manx population were born on the island, the rest being mostly from within the British and Irish Isles with a minority (~7%) from elsewhere in the EU (Isle of Man Government 2006). A random sample taken on the island was likely to reflect this more recent activity, so a strategy was needed to increase the chances that the gathered data-set was representative of the population of interest.

To overcome the issue of modern migration, population genetics studies generally require DNA donors to be unrelated as far as is known, and to have at least two generations of ancestry from the region under investigation. This study took a different approach to sampling, by using surnames to achieve a deeper

rooted ancestry than the standard two generations. The next section explains how this was achieved.

### 4.2.1 Background to sampling strategy

Over the last two hundred years the Isle of Man received considerable numbers of migrants (see Section 3.2.5). The standard 'two generations of ancestry' criterion bypasses some of the issues of 20th century migration (see above) but does not address the effects of older movements of people. Instead, this project was inspired by the work of Bowden et al. (2008), where volunteers were recruited on the basis of their surnames.

Previous studies had demonstrated that surnames correlated significantly with NRY haplotypes, since both are inherited down the paternal line. Non-paternity events, a son bears the name of someone who is not his biological father, can break this connection, however, studies shows that these events are are infrequent (King et al. 2006). Rare surnames also tend to show strong regionality suggesting that significant numbers of males tend to stay near their place of birth (Cheshire et al. 2010)

By using surnames that appear in the earliest documents from the region of interest, in combination with the two-generations' criterion, Bowden et al. (2008) were able to effectively assemble a sample group representative of the 16th century, thus bypassing migrations of the later millennium triggered by industrialisation. The oldest written documents attesting named residents of the Isle of Man are manorial rolls, dating to 1511 and 1515 (Talbot 1924). Males bearing surnames found in these documents were deemed more likely than a random sample to have deeply-rooted ancestry on the island, extending back to those times. A strict criterion (see following section 4.2.2) was designed to avoid

accidently sampling descendants of later migrants to the island.

It was decided not to recruit a control group without surname ascertainment, since, Bowden et al. (2008) had already shown that surname-ascertained genetic data was an effective way of bypassing the problem of recent historical immigration. It seemed unnecessary to show this again, and collecting a control sample would also increase the expense and amount of time required for the project. However, the existing low-resolution Isle of Man data-set (Capelli et al. 2003) served to investigate differences between samples obtained under alternative criteria without the need for further sampling.

### 4.2.2 Criteria for Surname Selection

A disadvantage of using 16th century documents is that the spelling and pronunciation of many names have changed, sometimes dramatically, in the intervening period. However, Manx surnames are the subject of three books: *The Surnames and Placenames of the Isle of Man* by A. W. Moore (1890) *Personal Names of the Isle of Man* by J. J. Kneen (1925) and *Surnames of the Manks* [sic] by Leslie Quilliam (1989). Using manorial rolls and parish records of baptisms from the 16th and 17th century, authors analysed and categorised Manx surnames, linking old forms to modern forms and spelling variants.

Quilliam (1989) takes care to reconcile the sometimes conflicting interpretations given in the earlier works. He also includes details such as the earliest dates that the names were recorded, and their frequencies in each parish at the time of printing. In appendices he lists names according to their frequency and also includes lists of names that have origins and spelling variants elsewhere (Quilliam 1989). A list of surnames representing families attested on the island for at least 400 years was compiled using all these references.

In order to be included in the study the following criteria were adhered to:

1. Include only surnames from the parochial records (c1500; n=184; Quilliam 1989, appendix F)
2. Include surnames with origins in Man only e.g. Gelling, Kinvig (n=12)
3. Include Manx version of names found elsewhere if they have Manx-specific spelling (n=108; Quilliam 1989, appendix G)
4. Exclude surnames with origins both in Manx and elsewhere, for example Bell, Callow, Garrett (n=23; Quilliam 1989, appendix G)
5. Exclude Manx surnames which originated in Man, but have the same spelling as surnames that have originated elsewhere e.g. Kelly, Lace, Moore (n=20; Quilliam 1989, appendix G)
6. Advertise for recruits from all alternate spelling variants of a surname but include only one from each surname group in the study to avoid bias e.g. Keggen, Keggan (Appendix A1)

For the purposes of this project, surnames that had origins both within and outside Man were excluded, due to the difficulty in distinguishing between descendants of those mentioned in the manorial rolls and more recent incomers. However, names that had spellings specific to the island were included even if there were apparent spelling variants from outside of Man. Often surnames that appeared similar to surnames from elsewhere, were pronounced very differently on the island, suggesting very different roots. Two individuals (Crowe and Cowley), who were passionate about their family genealogy, provided detailed documentary support that they were descended from someone referred to in the documents rather than a later incomer. These particularly individuals were included in the study despite possessing a surname which did not make it onto the list, helping to increase the final sample size.

Using the above criteria 110 surnames were identified as being of ancient Manx origin (Table 4.1). These were divided into 81 surname families of spelling variants (see Appendix A 1 for list of spelling variants). In the Isle of Man 1996 census, some of these names appeared to be extinct or nearing extinction, but they were included anyway in case individuals bearing these names had returned to the island.

*Table 4.1: 110 modern surnames identified as ancient Manx in origin (Kneen 1925; Moore, 1890; Quilliam 1989). Volunteers were required to bear one of these surnames to participate in this study.*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Bridson | Callin | Callister | Cannan | Caren | Carine | Carran | Carroon | Casement |
| Caveen | Christory | Clague | Cleator | Clucas | Cojeen | Collister | Colquitt | Colvin |
| Comaish | Comish | Condra | Cooil | Coole | Corkan | Corkhill | Corkill | Corkish |
| Corlett | Cormode | Corran | Corrin | Corris | Corteen | Costain | Cowen | Cowin |
| Cowle | Crebbin | Creer | Cregeen | Crellin | Crennell | Cretney | Cringle | Crye |
| Cubbin | Cubbon | Curphey | Faragher | Fargher | Fayle | Freer | Gawne | Gelling |
| Joughin | Kaighen | Kaighin | Kaneen | Karran | Kee | Keggan | Keggen | Keggin |
| Keig | Kennaugh | Kennish | Keown | Kermeen | Kermode | Kerruish | Kewin | Killey |
| Killip | Kinley | Kinnish | Kinrade | Kinvig | Kissack | Kneale | Kneen | Lewney |
| Looney | Lowey | Maddrell | Moughtin | Mylcraine | Mylchreest | Mylechreest | Mylrea | Mylroie |
| Quaggan | Quaggin | Qualter | Qualtrough | Quane | Quark | Quaye | Quiggin | Quilleash |
| Quilliam | Quillin | Quine | Shimmin | Skelly | Skillicorn | Taubman | Teare | Vondy |
| Waterson | Watterson | | | | | | | |

The majority of the selected Manx surnames were of Gaelic derivation, although a few names were possibly Norse (Moore, 1890; Quilliam 1989). The merging of Irish Sea and Norse cultures (see Chapter 2) means that bearing a surname of Norse derivation is not necessarily a reliable indicator of Norse ancestry. The following section explains how volunteers bearing surnames from Table 4.1 were recruited to take part in this study.

### 4.2.3 Volunteer Recruitment

In order to recruit people to take part in this project, I designed a website which was then built by Dr. Turi King. The site contained details of the aims of the project and links to connected projects including The Impact of Diasporas on the Making of Britain http://www2.le.ac.uk/projects/impact-of-diasporas as well as links to Prof. Mark Jobling's and Prof. Simon James' sites to lend professional weight to the study.

The website recruitment page contained a list of the selected surnames (Table 4.1) and a form to register a volunteer's interest in taking part. Only one volunteer per surname group was required (see Appendix A1) so the online form featured a drop-down menu containing all the surnames. To ensure one volunteer per name, surnames were removed once a candidate was found for each group.

Volunteers were invited to a public event where their DNA could be collected. The sampling event was held in the Manx Museum in Douglas, Courtesy of Manx National Heritage, on the 19th of February 2011, and included a presentation on what the project hoped to learn about the population history of the Isle of Man. The event was publicised by sending a press release (Appendix A2) to all the local media. This was picked up by Manx Radio, and I was interviewed about the project. The museum also helped to publicise the event by publishing the list of names and directions to the project website on their social media networks, Twitter and Facebook. Details of the location and time of the DNA collection event were also included on the project website.

These methods of volunteer recruitment and face-to-face DNA collection were chosen because they were targeted, and had the advantage of direct interaction with the public, which increased the profile of this and connected projects. It also

seemed more respectful to actually visit the Isle of Man to engage with volunteers, rather than just requesting their DNA from a distance. For people who wanted to take part but who were unable to attend on the day, there was the option to receive a sample kit through the post. As these people were primed to receive the kit, the response rate could be expected to be much higher than through unsolicited recruitment.

Almost 70 potential volunteers were recruited through the website, plus numerous reserves. On the day of the event about 100 people attended the presentation including 35 of the potential volunteers. The event also provided an opportunity for people whose surnames had been excluded because they did not match the criteria, to present documents supporting their ancestry dating back to at least the 16th century. Saliva samples of the attending volunteers were taken using the Oragene® DNA collection kit (see Fig. 4.1), while those who were absent were contacted and had packs sent out to them by post. In total, 67 samples matching the criteria were collected. Full or partial DNA profiles were obtained for 66 of these samples (see Table 4.2, Appendix A3, and also Section 5.2.1, Chapter 5).


*Figure 4.1: Oragene® DNA collection kit, donor spits into the funnel until saliva reaches the line.*

Sampling was done with informed consent following review by the University of

Leicester Research Ethics Committee (Ethical Application Ref: maj4-46d9; 'The Impact of migrations on the making of Britain: genetic approaches to population history (revision)'). Volunteers were provided with details of the project (see Appendix A4) and were required to sign consent forms. Following genetic analysis (see next section) individual results of STR and SNP haplotyping were returned to participants by email or post, depending on their preference, with details of current understandings of their haplotype's origins and history (see Appendix A5).

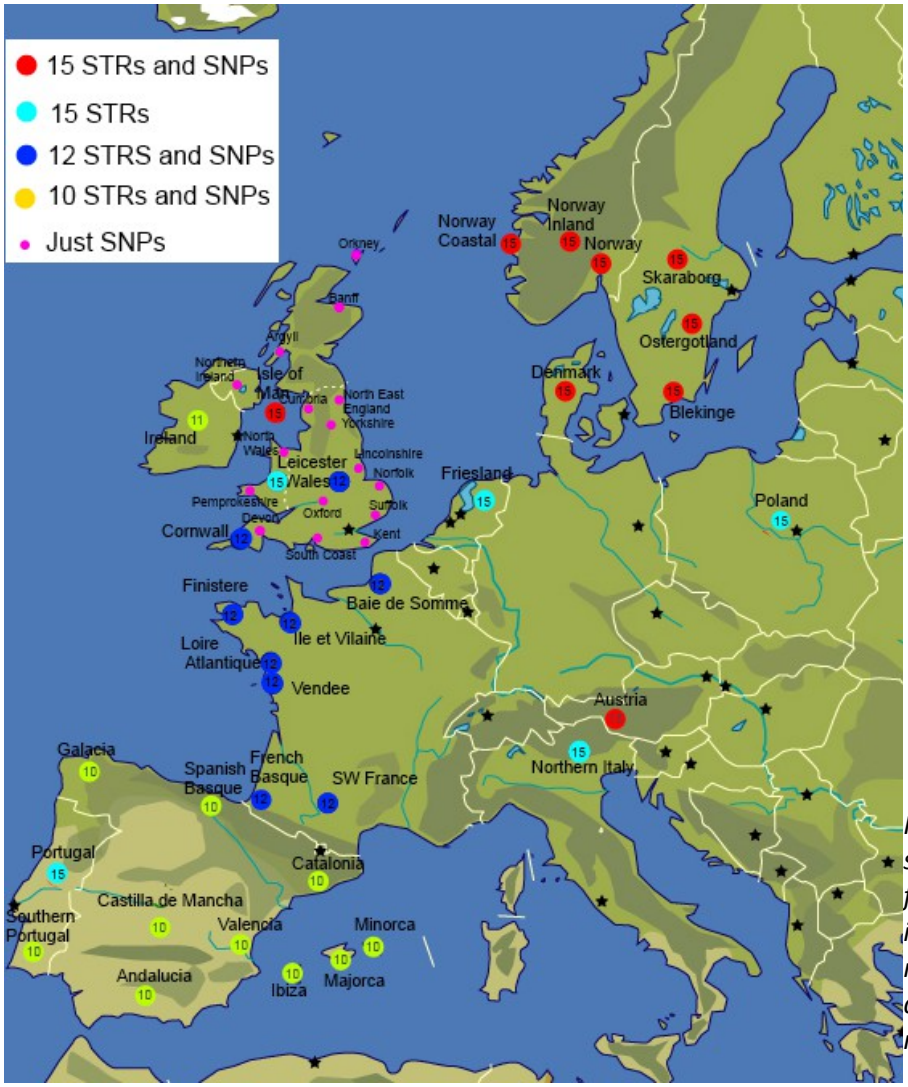



*Figure 4.2: Map showing the regions for NRY data-sets, including details of markers available and levels of STR resolution.*

### 4.2.4 Comparative NRY Data Sets

Comparative data for the Isle of Man NRY data set were taken from the literature or supplied by members of Mark Jobling's research group. To be included they needed to be from populations within the Western European Atlantic region and preferably of high resolution. Only data-sets with at least 10 STRs in common or a high degree of R1b haplogroup sub-division were included in analyses (See Figure 4.1 and Table 4.3 for details of data-sets included in analyses).

### 4.4.2 Comparative mtDNA Data Sets

The MtDNA control region was sequenced for 35 samples from Manx NRY volunteers with two generations of maternal ancestry on the Isle of Man. Of these, usable data was obtained for 33 samples (see Table 4.2 and Appendix B10). This work was carried out in Mark Jobling's laboratory by Lorena Boquete Vilariño (2012) for her Master's dissertation under the supervision of Dr Chiara Batini. Comparative data sets were taken from the literature or supplied by members of Mark Jobling's research groups. To be included they needed to be from populations in Western European and have sequence data from at least hyper-variable regions 1 and 2 (See table 4.4 for details of data sets included in analyses).

*Table 4.2: Breakdown of numbers of samples obtained and used for data analysis. SNP and STR profiles only apply to NRY samples.*

| | Total sampled | Full DNA Profiles | STR only | SNP only | Unusable | Total usable |
|---|---|---|---|---|---|---|
| **NRY** | 67 | 58 | 5 | 3 | 1 | 66 |
| **mtDNA** | 35 | 33 | N/A | N/A | 2 | 33 |

*Table 4.3: details of comparative NRY data sets*

| Country | Region | N | Compatible STRs | SNPs (high/ low resolution) | Reference |
|---|---|---|---|---|---|
| Basque Country | Spanish | 116 | 10 | Yes (low) | (Adams et al. 2008) |
| | French | 60 | 12 | No | Balaresque and Jobling, unpublished |
| Catalonia | | 79 | 10 | Yes (low) | (Adams et al. 2008) |
| Denmark | | 109 | 17 | Yes (high) | Børglum et al. 2007 |
| England | Chippenham | 55 | 17 | No | Wetton, unpublished |
| | Cornwall | 64 | 12 | No | Balaresque and Jobling, unpublished |
| | Cornwall | 41 | 0 | Yes (high) | (Winney et al. 2012) |
| | Cumbria | 123 | 0 | Yes (high) | " " |
| | Devonshire | 51 | 0 | Yes (high) | " " |
| | Kent | 46 | | Yes (high) | " " |
| | Leicester | 46 | 12 | No | Balaresque and Jobling, unpublished |
| | Leicestershire | 36 | 0 | Yes (high) | Winney et al. in print; Wetton, unpublished |
| | Lincolnshire | 72 | 0 | Yes (high) | " " |
| | Norfolk | 66 | 0 | Yes (high) | " " |
| | North East | 122 | 0 | Yes (high) | " " |
| | Oxfordshire | 53 | 0 | Yes (high) | " " |
| | South Coast | 82 | 0 | Yes (high) | " " |
| | Winchester | 30 | 17 | No | Wetton, unpublished |
| | Yorkshire | 102 | 0 | Yes (high) | Winney et al. in print; Wetton, unpublished |
| France | Baie de Somme | 43 | 12 | No | Balaresque and Jobling, unpublished |
| | Finistère | 81 | 12 | No | " " |
| | Ille et Vilaine | 48 | 12 | No | " " |
| | Loire Atlantique | 74 | 12 | No | " " |
| | South west | 57 | 12 | No | " " |
| | Vendée | 50 | 12 | No | " " |

| Country | Region | N | Compatible STRs | SNPs (high/ low resolution) | Reference |
|---|---|---|---|---|---|
| Netherlands | Friesland | 95 | 17 | No | Wetton, unpublished |
| Northern Ireland | | 25 | 0 | Yes (high) | Winney et al. in print; Wetton, unpublished |
| Norway | Coastal | 82 | 17 | Yes (high) | King, unpublished |
| | Inland | 224 | 17 | Yes (high) | " " |
| | Mixed | 72 | 17 | Yes (high) | " " |
| Portugal | South | 77 | 10 | Yes (low) | Adams et al. 2008 |
| Republic of Ireland | | 796 | 10 | No | Ballard et al. 2006 |
| Scotland | Argyll | 20 | 0 | Yes (high) | Winney et al. in print; Wetton, unpublished |
| | Banff | 39 | 0 | Yes (high) | " " |
| | Orkney | 23 | 0 | Yes (high) | " " |
| Spain | Andalucía | 95 | 10 | Yes (low) | Adams et al. 2008 |
| | Castilla la Mancha | 73 | 10 | Yes (low) | " " |
| | Galicia | 62 | 10 | Yes (low) | " " |
| | Ibiza | 88 | 10 | Yes (low) | " " |
| | Majorca | 54 | 10 | Yes (low) | " " |
| | Minorca | 62 | 10 | Yes (low) | " " |
| | Valencia | 36 | 10 | Yes (low) | " " |
| Sweden | Blekinge | 40 | 17 | Yes (high) | Karlsson 2006 |
| | Östergötland | 45 | 17 | Yes (high) | " " |
| | Skaraborg | 40 | 17 | Yes (high) | " " |
| Wales | Abergele | 26 | 17 | No | Wetton and Jobling, unpublished |
| | Llangefni | 32 | 17 | No | Wetton, unpublished |
| | Llanuwchllyn | 30 | 17 | No | " " |
| | North Wales | 44 | 0 | Yes (high) | Winney et al. in print; Wetton, unpublished |
| | Pembrokeshire | 31 | 0 | Yes (high) | " " |
| | Tregaron | 30 | 17 | No | " " |

*Table 4.4: details of comparative mtDNA data-sets*

| Country | Region | N | Genome Region | Reference |
|---|---|---|---|---|
| Basque | Guipúzcoa | 113 | HSV 1&2 | (García et al. 2011) |
| | Viscaya | 90 | HSV 1&2 | " " |
| Denmark | | 201 | HSV 1&2 | (Mikkelsen et al. 2010) |
| England | Cheshire | 54 | Selected genome wide SNPs | (Winney et al. 2012); Wetton, unpublished |
| | Cornwall | 118 | Selected genome wide SNPs | " " |
| | Cumbria | 284 | Selected genome wide SNPs | " " |
| | Devonshire | 99 | Selected genome wide SNPs | " " |
| | Dorset | 46 | Selected genome wide SNPs | " " |
| | Forest of Dean | 60 | Selected genome wide SNPs | " " |
| | Gloucestershire | 49 | Selected genome wide SNPs | " " |
| | Hampshire | 49 | Selected genome wide SNPs | " " |
| | Herefordshire | 32 | Selected genome wide SNPs | Winney et al. In print; Wetton, unpublished |
| | Kent | 110 | Selected genome wide SNPs | " " |
| | Lancashire | 45 | Selected genome wide SNPs | " " |
| | Leicester | 43 | Selected genome wide SNPs | " " |
| | Leicestershire | 91 | Selected genome wide SNPs | " " |
| | Lincolnshire | 159 | Selected genome wide SNPs | " " |
| | North East England | 243 | Selected genome wide SNPs | " " |
| | Norfolk | 128 | Selected genome wide SNPs | " " |
| | Nottinghamshire | 84 | Selected genome wide SNPs | " " |
| | Northamptonshire | 58 | Selected genome wide SNPs | " " |
| | Oxfordshire | 140 | Selected genome wide SNPs | " " |
| | Suffolk | 114 | Selected genome wide SNPs | " " |
| | Sussex | 91 | Selected genome wide SNPs | " " |
| | Worcestershire | 39 | Selected genome wide SNPs | " " |
| | Yorkshire | 234 | Selected genome wide SNPs | " " |
| France | South west | 165 | Control | (Behar et al. 2012) |
| Finland | | 200 | HSV 1&2 | (Hedman et al. 2007) |
| Germany | North | 213 | HSV 1&2 | (Tetzlaff et al. 2007) |
| Northern Ireland | | 69 | Selected genome wide SNPs | Winney et al. In print; Wetton, unpublished |

| Country | Region | N | Genome Region | Reference |
|---|---|---|---|---|
| Norway | | 40 | Control | Jobling's lab. |
| Republic of Ireland | | 50 | Control | " " |
| Scotland | Argyll | 53 | Selected genome wide SNPs | Winney et al. In print; Wetton, unpublished |
| | Banff | 71 | Selected genome wide SNPs | " " |
| | Orkney | 136 | Selected genome wide SNPs | " " |
| Spain | Asturias | 76 | HSV 1&2 | García et al. 2011 |
| | La Rioja | 52 | Control | Behar et al. 2012 |
| Sweden | Blekinge | 38 | Control | (Tillmar et al. 2010) |
| | Gotland | 40 | Control | " " |
| | Östergötland | 40 | Control | " " |
| | Skaraborg | 41 | Control | " " |
| | Uppsala | 54 | Control | " " |
| | Värmland | 42 | Control | " " |
| | Västerbotten | 40 | Control | " " |
| Wales | North | 98 | Chip | Winney et al. In print; Wetton, unpublished |
| | Pembrokeshire | 68 | Chip | " " |

## 4.3 NRY DNA Purification and Typing

### 4.3.1 Introduction

Before carrying out analysis, the DNA must be extracted and purified from the provided samples. For this study, DNA was extracted from buccal (interior cheek) cells suspended in saliva mixed with a buffer. Following DNA extraction and purification, the haplotype of each NRY was determined using binary SNP markers and 17 STR markers. These analyses involved modified versions of polymerase chain reaction (PCR); the method whereby the DNA region of interest is targetted and amplified (Mullis and Faloona 1987; Saiki et al. 1985).

This Nobel-Prize winning technique was developed in the 1980's, and remains central to most forms of molecular genetics research today.

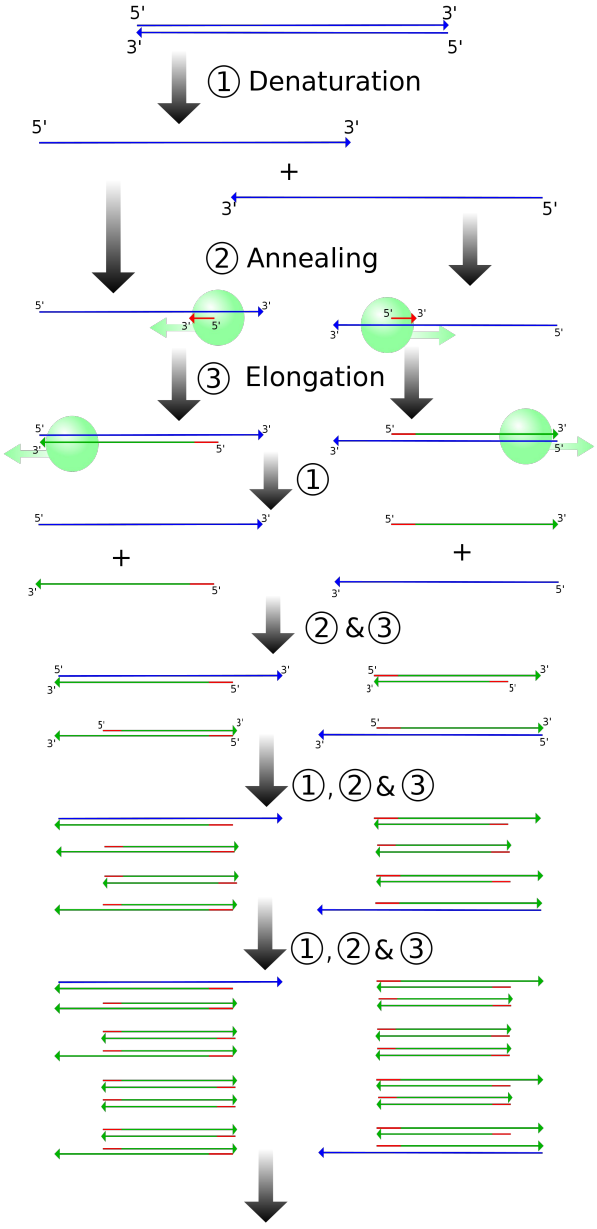



*Figure 4.1: Polymerase Chain Reaction; amplification of DNA region of interest through repeated cycles of heating and cooling. 1. Denaturation, the two DNA strands are separated by heating. 2. Annealing, DNA is cooled, allowing pre-prepared primers, which match sequences close to the region of interest, to attach to respective region of DNA. 3. Elongation, nucleotides attach to the primer replicating the complementary strand of DNA. DNA is then heated again to separate template strands from newly replicated strands. Cooling allows annealing and elongation to occur. The process of heating and cooling is repeated until the DNA region of interest is sufficiently amplified and concentrated. (image by Madeleine Price Ball CC WikiCommons).*

PCR involves cyclically heating and cooling DNA with a heat-tolerant polymerase (DNA-replicating enzyme), and oligonucleotide primers, which are short single-stranded pieces of DNA specially designed to match a section of DNA flanking

the region under investigation. Heating separates the two DNA strands so that primers can bind. Polymerase adds nucleotides (subunits of DNA) to the primers, building new strands of DNA which are complementary to the original strands. In subsequent cycles, new strands separate to act as templates for the next cycle. With each repetition of the cycle, the number of copies of target DNA theoretically doubles (Mullis and Faloona 1987; Saiki et al. 1985; see Fig. 4.1).

This study used modified techniques which allowed STRs and SNPs to be identified by capillary electrophoresis. Capillary electrophoresis is an automated system whereby the negatively charged DNA sequences are drawn through a capillary using an electric field. Sequences are separated by size, as larger molecules move slower than smaller ones. As they move through the capillary different fluorescent labels are detected by a laser (Butler et al. 2004).

### 4.3.2 DNA Extraction

DNA was contained within buccal (interior cheek) cells suspended in saliva and a stabilising solution. Extraction was carried out in Mark Jobling's laboratory by Dr. Turi King and myself using QIAcube QIAamp® DNA Mini QIAcube Kit (240; QIAGEN®). A small number of samples required manual extraction where the automated method had failed, following the Oragene® DNA collection kit manufacturers DNAgenotek protocol (see Appendix A6).

To extract the DNA, the saliva/buffer solution was incubated for an hour at 50º and then vortexed with Oragene•DNA Purifier to the DNA from the cells, deactivate the nucleases which break down the DNA and precipitate out other cellular material. These impurities where then separated from the DNA in solution by centrifuge. The DNA was precipitated out of the supernatant using ethanol. DNA was separated from the liquid by centrifuge and then rehydrated

in DNA buffer TE for storage. The purified DNA solution could then be stored long term at -20ºC.

### 4.3.3 STR Typing

STR haplotypes were ascertained using AmpFℓSTR Yfiler Kit™ PCR Amplification Kit (Applied Biosystems), which allows up to 17 STRs to be typed in one reaction (see Table 4.5 for STRs under investigation). First PCR is performed using the Yfiler Kit which amplifies the NRY regions of interest, then electrophoresis is carried out to identify the NRY haplotype of each sample.

*Table 4.5: showing details of each STR typed in this study.*

| Marker | Repeat unit | Range of repeats |
|---|---|---|
| **DYS19** | TAGA | 10-19 |
| **DYS389 I** | (TCTG) (TCTA) | 9-17 |
| **DYS389 II** | (TCTG) (TCTA) | 24-34 |
| **DYS391** | TCTA | 6-14 |
| **DYS392** | TAT | 6-17 |
| **DYS393** | AGAT | 9-17 |
| **DYS437** | TCTA | 13-17 |
| **DYS438** | TTTTC | 6-14 |
| **DYS439** | AGAT | 9-14 |
| **DYS448** | AGAGAT | 20-26 |
| **GATA_H4** | TAGA | 8-13 (25-30) |
| **DYS390** | (TCTA) (TCTG) | 17-28 |
| **DYS456** | AGAT | 13-18 |
| **DYS385_a,b** | GAAA | 7-28 |
| **DYS458** | GAAA | 13-20 |
| **DYS635** | TCTA or TGTA | 17-27 |

PCR was performed on each of the Isle of Man samples to target and amplify the region of the genome under investigation. The AmpFℓSTR Yfiler kit contained labelled primers which target 17 specific STR sites found on the NRY. One of

each primer pair carried a fluorescent label at its 5´ end, which allowed fluorescent detection of the PCR product after electrophoresis.

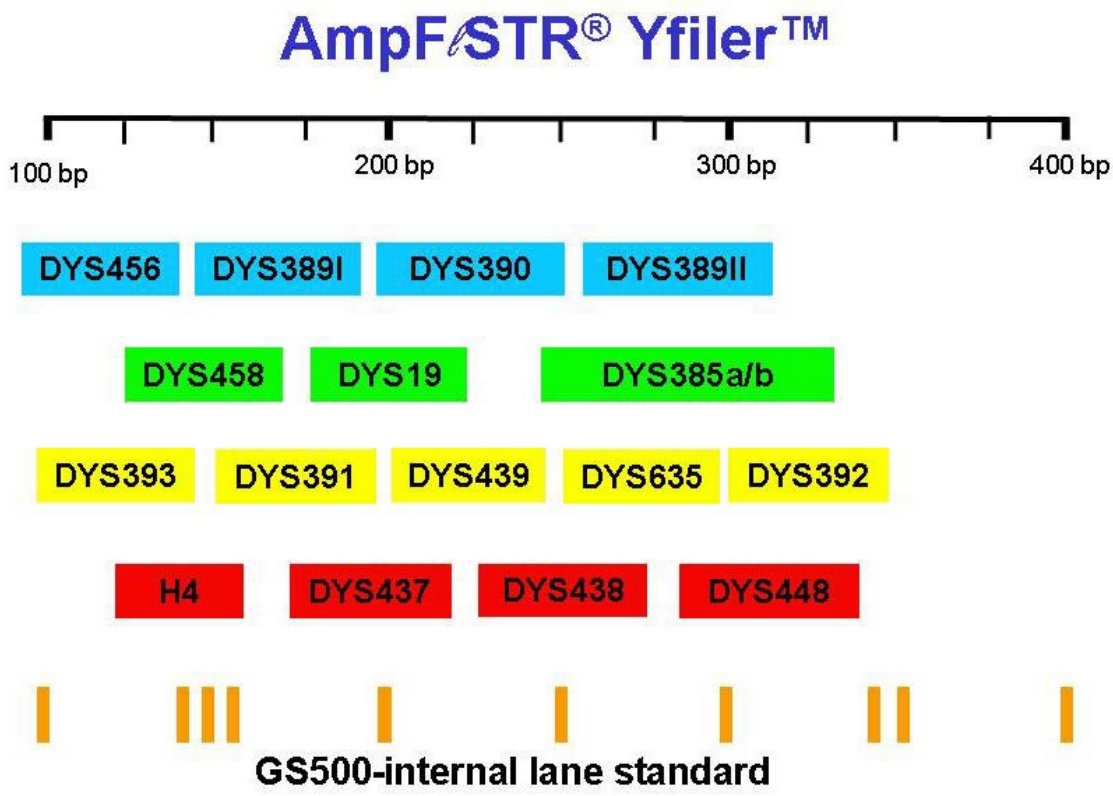



*Figure 4.3: Shows the fluorescent dye label colour and relative PCR product size ranges for the various STR loci present in the AmpFℓSTR Yfiler kit. (image from STRbase 2014)*

A reagent containing PCR reaction mix, primers, and Amplitaq Gold® DNA polymerase was created in a dedicated PCR laminar hood. 5µl of the reagent mix was added to 3µl (approx 1ng) of DNA from each sample, in separate wells of a reaction plate. Two positive controls and one negative control were also included in separate wells. The initial incubation step of PCR ran at 95ºC for 11 minutes. PCR was carried out for 30 cycles of heating and cooling as follows; denature step 94ºC for 1 minute, annealing step 61ºC for 1 minute and extension step 72ºC for 1 minute. For the final extension step the temperature was maintained at 60ºC for 80 minutes before being held at 4ºC until needed.

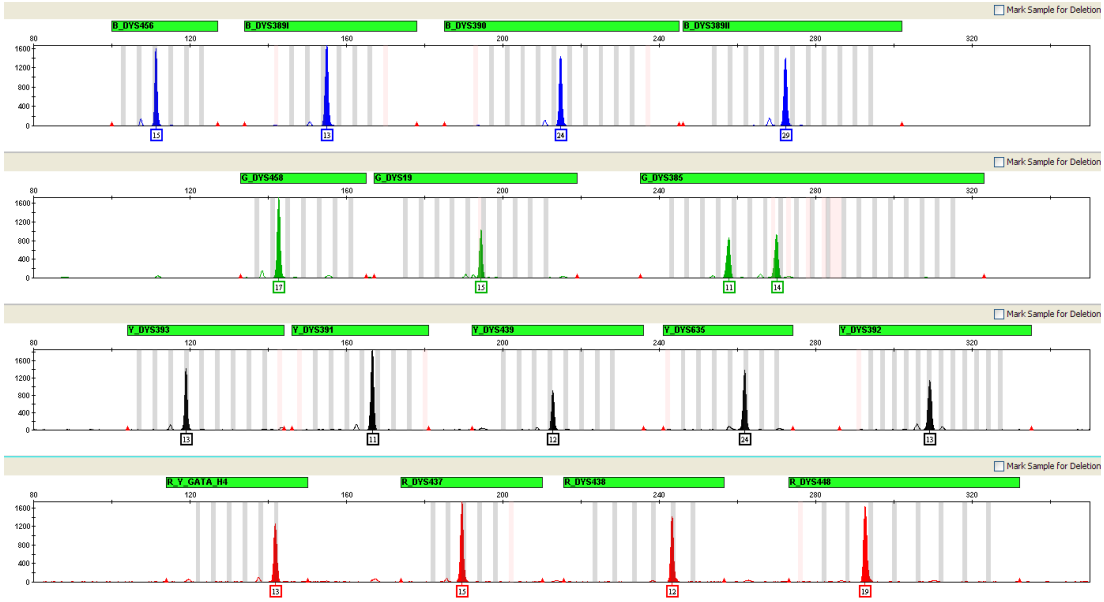



*Figure 4.3: Genemapper® electropherogram for the control sample. The NRY STR loci are identified above each peak. Each peak corresponds to a STR allele and the number of repeats is indicated below the peak. The colour of the peak corresponds to the colour of the fluorescent dye label. (Image taken from the Applied Biosystems AmpFℓSTR Yfiler Kit™ PCR Amplification Kit user guide).*

To prepare for electrophoresis, the PCR products for each sample were transfered to individual wells in a Micro-Amp® Optical 96-well reaction plate. GeneScan™- 500 LIZ® size standard, previously diluted with Hi-Di™ Formamide, and an allelic ladder were added to each sample. The plate was sealed and centrifuged before being incubated for 3 minutes at 95°C to denature the DNA. The plate was then put on ice for another 3 minutes.

Capillary electrophoresis was performed on an ABI PRISM® 3130xl Genetic analyser (Applied Biosystems) to separate and identify the PCR products from the previous step. In every sample, each of the 17 STRs were identified by weight and fluorescent primer (see Fig. 4.2), which was detected as it passed by a laser beam. This information was captured and recorded automatically by GeneMapper® software (Applied Biosystems; see Fig. 4.3). See Appendix A7 for the detailed protocol.

### 4.3.3 SNP Typing

SNPs were identified using single-base extension reaction via the SNaPshot® multiplex system (Applied Biosystems®) which allows 10 SNPs to be typed at a once. 19 SNPs were typed in total, using two multiplexes. The first identified SNPs which define all the major haplogroups, and the second identified 9 SNPs which defined R1b subgroups (see Table 4.6 for details). This work was carried out in Mark Jobling's laboratory by Dr. Turi King and Marta Verdugo, an undergraduate student (Verdugo 2012).

PCR was carried out on each of the Isle of Man samples to target and amplify the SNP loci of interest. Primers matching the flanking regions of the target sites, were affixed at the 5' end with poly-A tails of varying lengths allowing SNPs to be identified by size. The SNaPshot® multiplex system (Applied Biosystems®) contained four types of fluorescent labelled ddNTPs, corresponding to the four different possible nucleotides. The ddNTPs extended and terminated the annealing primer, and the fluorescent label allowed fluorescent detection of the PCR product after electrophoresis (see Fig. 4.4).

Two PCR reactions were required to type SNPs, each followed by a cleaning step to remove excess primers and dNTPs or ddNTPs. The first PCR reaction amplified the regions containing the SNPs. A reaction mix containing AmpliTaq® gold DNA polymerase, dNTPs, and a buffer were mixed with the annealing primers (see Appendix A8 for table containing primers). This mix was then added to each of the Isle of Man DNA samples. Two positive controls carrying known SNP alleles and one negative control with no DNA were also included. The first PCR reaction was carried out beginning with 94ºC for 9 minutes, followed by 35 cycles of 94ºC for 30 seconds, 59ºC for 30 seconds and 72ºC for 1 minute. The reaction was completed with 74ºC for 3 minutes.

*Table 4.6: showing targeted SNPs and the haplogroups they define.*

| SNP ID | hg |
|---|---|
| **M173** | R1 |
| **M170** | I |
| **M253** | I1 |
| **M172** | J2 |
| **M35** | E1b1b1 |
| **M201** | G |
| **M242** | Q |
| **Tat** | N1c |
| **M223** | I1b |
| **M17** | R1a1 |
| **M167** | R1b1b1a1b5 |
| **U106** | R1b1b1a1a |
| **U152** | R1b1b1a1b1 |
| **U198** | R1b1b1a1a1 |
| **M153** | R1b1b1a1b4 |
| **S145** | R1b1b1a1b2 |
| **M222** | R1b1b1a1b2a |
| **S116** | R1b1b1a1b |
| **L11** | R1b1b1a1 |

PCR products were cleaned by mixing with Exonuclease 1 (Exol; New England Biolabs), to digest unused primers, shrimp alkaline phosphatase (SAP; Affymetrix) to dephosphorylate excess dNTPs, and E buffer. This reaction was carried out in a thermal cycler at 37ºC for 2 hours, 80ºC for 15 minutes and finally 4ºC for 15 minutes. The cleaned PCR produced from the first PCR, was mixed with SNaPshot™ multiplex reaction mix containing the fluorescently labelled ddNTPs, AmpliTaq® gold DNA polymerase, buffer and primers.

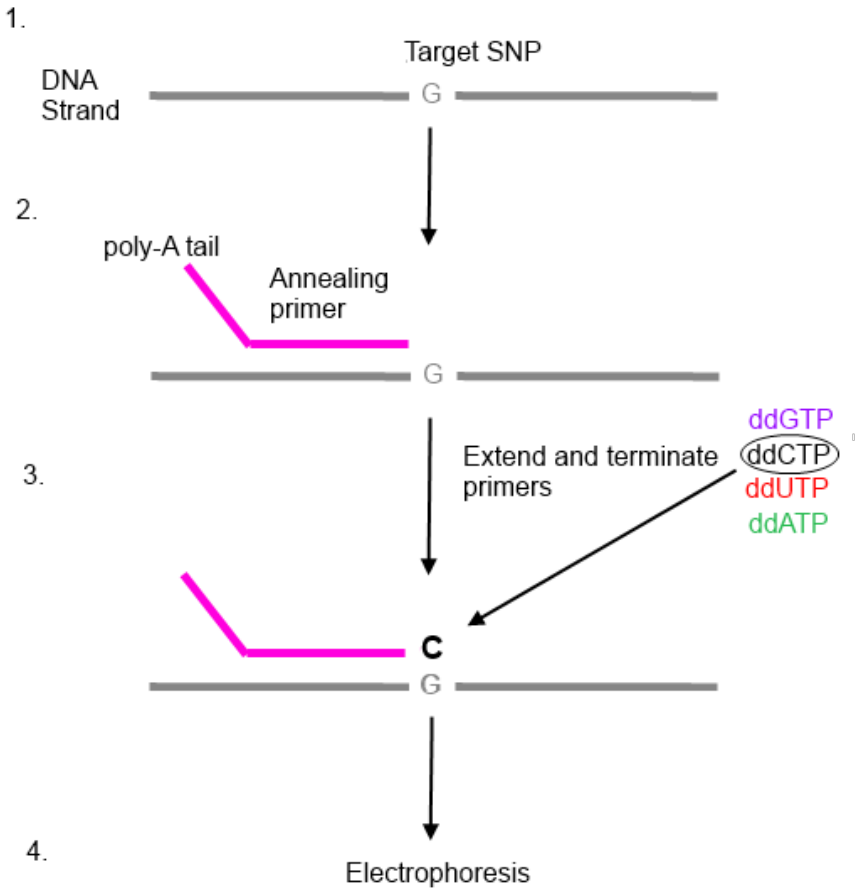



*Figure 4.4: Schematic diagram of SNP typing.1. During the first PCR, the regions containing the SNPs under investigion were amplified. 2. In the second PCR, primers with poly-A tails anneal to the flanking region of the SNP. 3. A labelled terminating ddNTP, complimentary to the SNP, extend the annealling primer by one. PCR product is then identified by electrophoresis on the basis of colour and weight.*

To prepare for electrophoresis, the PCR products for each sample were transferred to individual wells in a Micro-Amp® Optical 96-well reaction plate. GeneScan™- 500 LIZ® size standard, previously diluted with Hi-Di™ Formamide, was added to each sample. The plate was sealed and centrifuged before being incubated for 4 minutes at 95ºC to denature the DNA. The plate was then put on ice for 3 minutes.

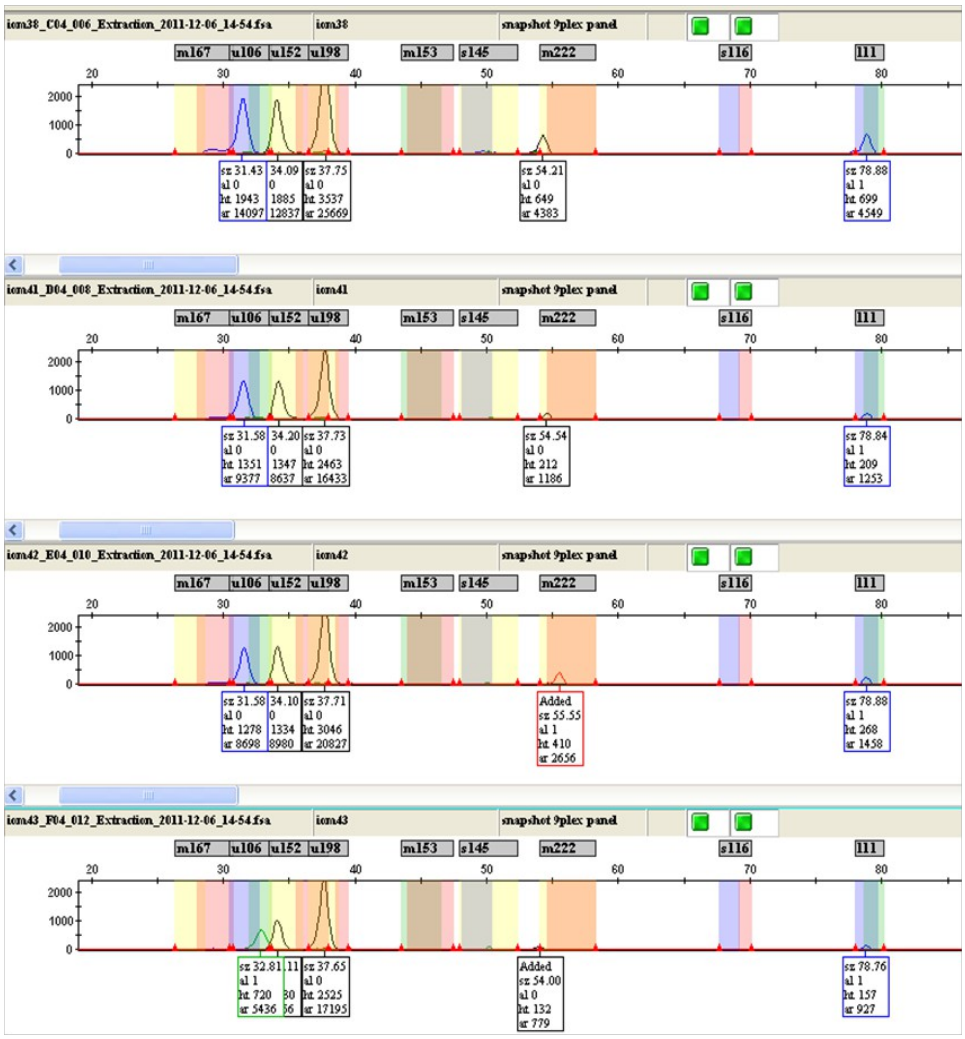



*Figure 4.5: Screenshot of Genemapper® electropherogram showing NRY SNPs for samples IOM38, IOM41, IOM42, and IOM43. SNPs are identified above each peak, and their allelic state is indicated in the box below as either ancestral (0) or derived (1). The colour of the box corresponds to the coloured label added during PCR, respresenting the sequence state. The markers showing no peak are probably a result of inadequate PCR amplification (Image taken from M. Verdugo, unpublished dissertation).*

As with STR typing (see section 4.3.2) capillary electrophoresis was performed on an ABI PRISM® 3130xl Genetic Analyzer (Applied Biosystems) to separate and identify the PCR products. In every sample, each of the SNPs was identified by molecular weight and their state was determined by fluorescently labelled ddNTP, which was detected as it passed by a laser beam. This information was captured and recorded automatically by GeneMapper® software (Applied

Biosystems; See Fig. 4.5).

## 4.4 Data Analysis

### 4.4.1 Introduction

Once the DNA samples had been haplotyped, the resulting dataset and the comparative datasets were all converted to the same format, so that data analysis could begin. For data where haplogroup information was not available, haplogroups were predicted from available STR data using online algorithms and databases (see Table 4.8). To analyse the genetic data of the Isle of Man and surrounding Western European populations, techniques were drawn from the three main classes of genetic methodologies: classical population genetics, phylogeography and computer modelling.

The first two sets of methodologies offered ways to describe and visualise genetic diversity, to provide a sense of the genetic landscape and highlight any interesting features, patterns or affinities between populations. Computer modelling then provided a way to test hypotheses about population history arising from patterns observed in the genetic diversity.

Summary statistics were generated from NRY and mtDNA data to describe the genetic diversity within and between the populations of the region. Comparisons were made between genetic diversity of the Isle of Man and surrounding Western European populations for NRY and mtDNA at sufficient resolutions to include all available data. Then a subset of the populations were analysed at a higher resolution, where data were available.

*Table 4.7: software used for data analysis*

| Name | Function | Reference |
|---|---|---|
| **Arlequin 3.5** | Generates summary statistics from genetic data | (Excoffier and Lischer 2010) |
| **Spider** | Converts input files to different formats | (Lischer and Excoffier 2012) |
| **Phylogenetic Network 4.621 Network Publisher add-on 2.0.0.1** | Calculates and draws phylogenetic networks from haplotype data | (Bandelt et al. 1999; fluxus-engineering.com 2014) |
| **DIYABC version 2.0.3 for WINDOWS.** | User-friendly Approximate Bayesian Computation software. Simulates data for model selection and parameter estimation. | (Cornuet et al. 2014) |
| **SPSS version 20.0** | General statistical analysis | (IBM Corp. 2011) |

Phylogenetic networks were created using pooled datasets which included the Isle of Man and other selected populations to visualise relationships between haplotypes across populations. Networks were also created to explore the geographical distributions of specific haplogroups in detail. Haplotypes that closely matched those found in the Isle of Man data were also searched for in an online database.

*Table 4.8: online tools used for data analysis.*

| Name | Function | Reference |
|---|---|---|
| **Haplogrep** https://haplogrep.uibk.ac.at/ | MtDNA haplogroup predictor | (Kloss-Brandstätter et al. 2011) |
| **YHRD** http://www.yhrd.org/ | Global NRY database | (Willuweit and Roewer 2007) |
| **Whit Athey Haplogroup predictor** http://www.hprg.com/hapest5/hapest5a/hapest5.htm | NRY haplogroup predictor (STRs) | (Athey 2013) |

Finally, a number of historical scenarios were explored using a computer modelling approach. NRY and mtDNA data sets for selected populations were

simulated under various historical demographic conditions to find hypotheses which best explained the data. Parameters relating to timing and impact of potential demographic events were also estimated using this method.

### 4.4.2 Classical Population Genetics

*Table 4.9: definitions of summary statistics used for intra-population calculations.*

| Summary Statistics | Definition |
|---|---|
| **k** | Number of different haplotypes present in the population |
| **S** | Number of polymorphic sites, or segregating sites (mtDNA only) |
| **Haplotype (gene) Diversity** | This is the probability that two haplotypes, randomly chosen from a sample, will be different. A measure of the diversity of within populations (Nei 1987), 180), it reveals how similar, therefore how closely related, haplotypes within a population are. Used for NRY STR haplotypes. |
| **Nucleotide Diversity π** | Measures diversity at each nucleotide along a length of DNA, measuring how much variation there is amongst sequences. Used for mtDNA sequences (Nei 1987, 257). |
| **Mean size variance** | Informative about how STR haplotypes within a population differ and how closely related they are. |
| **Mean Pairwise difference** | Measures amount of genetic diversity within a population by counting the average number of difference between each pair of sequences. |
| **Mean Pairwise Variance** | The variance of the above, informative about how closely related sequences within a population |

MtDNA sequences were edited for analysis by removing poly-C and other repeat sequences that are usually not considered for population analyses due to unavailability of suitable mutation models. Specifically, these edits included:

- 16182C and 16183C corrected back to A
- removal of C inserted around positions 309 and 315
- removal of variations in the CA tract around positions 522-3
- removal of insertions around position 576

The first three edits were taken from the Phylotree website

(http://www.phylotree.org/tree/subtree_D.htm (van Oven and Kayser 2009) and the fourth edit was recommended by Dr. C. Batini (Pers. Comm.). After editing 1123 base pairs (bp) were left for analysis.

mtDNA sequences were assigned to haplogroups using the Haplogrep online prediction tool (Table 4.8). Haplogrep compared sequence polymorphisms with the mtDNA phylogeny and used an algorithm to calculate a list of most the probable haplogroups. Polymorphisms were weighted depending on the number of times they occurred in the phylogenetic tree, so that mutation hotspots were ignored in favour of stable mutations when ranking the most probable haplogroups (Kloss-Brandstätter et al. 2011).

NRY data lacking SNP-derived haplogroup information were assigned to haplogroups on the basis of STR haplotypes, using Whit Athey's haplogroup predictor (Table 4.8). This haplogroup predictor uses a Bayesian approach to calculating the most probable haplotype, based on allele frequencies. Unfortunately STR data alone were insufficient to distinguish between closely-related R1b subgroups, so Whit Athey's haplogroup predictor could only provide the most basic haplogroup information for those data-sets lacking SNP data.

Genetic summary statistics (see Tables 4.9 and 4.10) were generated using Arlequin 3.5 software (Excoffier and Lischer 2010, see Table 4.7). Genetic distances were calculated as $F_{ST}$ and $R_{ST}$ (see Table 4.10 for definitions), and were then converted into 2D scatter-plots, via Multi-Dimensional Scaling (MDS). MDS compresses multi-dimensional data, from a matrix of pairwise figures, into two dimensions whilst preserving the maximum information about relationships between populations. In this way, it allowed genetic similarity between populations to be visualised. S-stress provided an indicator of the goodness-of-fit between original data and MDS plots. MDS was performed through SPSS 20.0

(IBM Corp. 2011, see Table 4.7) using the PROXSCAL option and default settings.

*Table 4.10: definitions of summary statistics used for inter-population calculations.*

| Summary Statistics | Definition |
|---|---|
| $\mathbf{F_{ST}}$ | Used to measure the genetic diversity of a meta-population compared to that of sub-populations and how diversity is distributed, using pairwise comparisons of alleles between populations. Was used as a genetic distance measure between populations for SNP-derived haplogroups, and mtDNA sequences or haplogroups.<br>Gives a figure between 0 and 1 for each pair of populations, where 1 is completely different and 0 is identical. |
| $\mathbf{R_{ST}}$ | Analogous to $F_{ST}$ but used on STR data. Calculations incorporate a Stepwise Mutation Model (SMM), whereby each locus has an equal probability of increasing or decreasing in length by one repeat unit. Provides more detail about how diversity is proportioned within and between populations than is possible for SNP-derived haplogroup data, to reveal any structure.<br>Gives a figure between 0 and 1 for each pair of populations, where 1 is completely different and 0 is identical. |
| $\mathbf{H_{ST}}$ | Similar to Fst, but uses pairwise comparisons between sequences to calculate genetic distance between populations (Hudson et al. 1992). |
| $\mathbf{(Delta\ mu)^2}$ | Another genetic distance measure for use with STRs, which includes a stepwise mutation model (Goldstein et al. 1995) |

### 4.4.3 Phylogeographic Analysis

NRY and mtDNA haplotypes shared between populations were identified using Arlequin 3.5 (Excoffier and Lischer 2010). Further NRY matches were tracked through the online Y Chromosome Haplotype Reference Database (YHRD, see Table 4.8), which contains Y chromosome STR and SNP haplotypes from 126,931 individuals in 176 world-wide populations (http://www.yhrd.org).

Networks were created using Phylogenetic Network Software (Bandelt et al. 1995) following the variance-based weighting system of Qamar et al. (2002). Within-sample variance was calculated for each STR. STRs with highest variances received the lowest weighting (ibid 2002, see Table 4.11). This reduced the

complexity of networks by removing the least likely parallel mutations.

*Table 4.11; weighting given to connections based on the variance of each STR length within the sample. (Qamar* et al. *2002)*

| STR sample variance | Weighting |
|---|---|
| **0 – 0.09** | 5 |
| **0.1-0.19** | 4 |
| **0.2-0.49** | 3 |
| **0.5-0.99** | 2 |
| **=/>1.00** | 1 |

NRY STR haplotype networks were created using a combination of median-joining (MJ; Bandelt et al. 1999) and reduced-median joining (RM) methods (Bandelt et al. 1995) as recommended by the manual. The resultant networks were further simplified using the Maximum Parsimony (MP) option, which removes unnecessary links from the final network (Polzin and Daneshmand 2003).

### 4.4.4 Computer Modelling

Approximate Bayesian Computation (ABC; see Section 2.3.2) was used to estimate the impacts immigration on Manx male and female populations. Through ABC it was possible to select which of a range of hypothetical demographic scenarios best fit with the observed genetic data. Parameters were also estimated, relating to timing of demographic events, such as population divergence and admixture, as well as effective population sizes and rates of admixture where appropriate. The software package DIYABC (version 2.0.3; Cornuet et al. 2014; see Table 4.7) was used to carry out most aspects of this multi-stage analysis.

The process began with designing scenarios incorporating archaeological and historical theories regarding population contacts and the possible timing of

these putative events. Many millions of simulation data-sets were generated based on proposed scenarios using a coalescence simulator incorporated into DIYABC. ABC was then used to select the scenario which produced simulated data-sets with data which most closely resembled the observed data. ABC was also used to estimate parameters of the selected scenario(s) such as effective population size and rates of admixture.

*Table 4.12: displaying details of population data-sets used in simulations for each marker. For NRY, the populations used for simulation 1 were Ireland, the Isle of Man and Norway. For simulation 2, wales was substituted for Ireland. For mtDNA the populations used in Simulation 1 were Ireland, Isle of Man and Norway, but in Simulation 2, Sweden was substituted for Norway.*

| | Resolution | Population | N | Simulation 1 | Simulation 2 |
|---|---|---|---|---|---|
| **NRY** | 10 STRs | Ireland | 796 | pop1 | N/A |
| | | Wales (Tregaron) | 30 | N/A | pop1 |
| | | Isle of Man | 66 | pop2 | pop2 |
| | | Norway (inland) | 224 | pop3 | pop3 |
| **mtDNA** | 1118bp | Ireland | 50 | pop1 | pop1 |
| | | Isle of Man | 33 | pop2 | pop2 |
| | | Norway | 40 | pop3 | N/A |
| | | Sweden | 295 | N/A | pop3 |

ABC was also used to assess the goodness-of-fit between proposed model/parameter combinations and the observed data, as well as indicating confidence in the chosen models, by calculating type I and type II errors. At the time of writing it was not possible to simulate scenarios with periods of continuous or background migration using DIYABC, so it was beyond the scope of this project to investigate scenarios involving background migration.

## Demographic Scenarios

Two sets of scenarios were explored using DIYABC for each genetic marker. The first set of scenarios were identical for both NRY and mtDNA, and explored simple demographic scenarios. Demographic parameters under investigation

were assigned estimated ranges, known as prior distributions. Prior distributions were typically conservative, encompassing a wide range of values that could be reasonably expected, given information drawn from the literature, but narrow enough to keep processing time down. Uniform distributions were used, to give equal probability across the whole range.

Populations were chosen for each marker to represent possible parental Viking and Pre-Viking populations for the Isle of Man. For NRY, Wales and Ireland were the nearest neighbours for which highest resolution data (10 STRs) was available. These were chosen to represent possible pre-Viking source populations. Tregaron in Wales was used as it was the most inland population, so deemed least likely to have been affected by fifth century Irish input (Capper pers. Comm). Norway was used to represent Viking source population, the inland population was used as it was deemed most likely to resemble the ancient population of Norway. Ireland's sample was geographically undifferentiated, so the whole sample was used.

For mtDNA, no suitable British data was available to represent a potential Manx pre-Viking parental population so Sweden was selected to test against Norway as a potential Scandinavian source population for Viking women. As previpous studies had detected no population structure using Swedish mtDNA (Humphreys et al. 2011), the Swedish samples were pooled.

Effective population sizes were estimated from current and historic demographic data sources and were approximated to $^{1}/_{6}$ census population size. Due to uncertainty surrounding calculating effective populations sizes, broad prior distributions surrounding these values were used, after (Bramanti et al. 2009). Estimates for the timing of population divisions and possible admixture events were taken from historic and archaeological sources. Parameters relating to

mutation rates and models were taken from the genetic literature and direct estimates from genetic data.

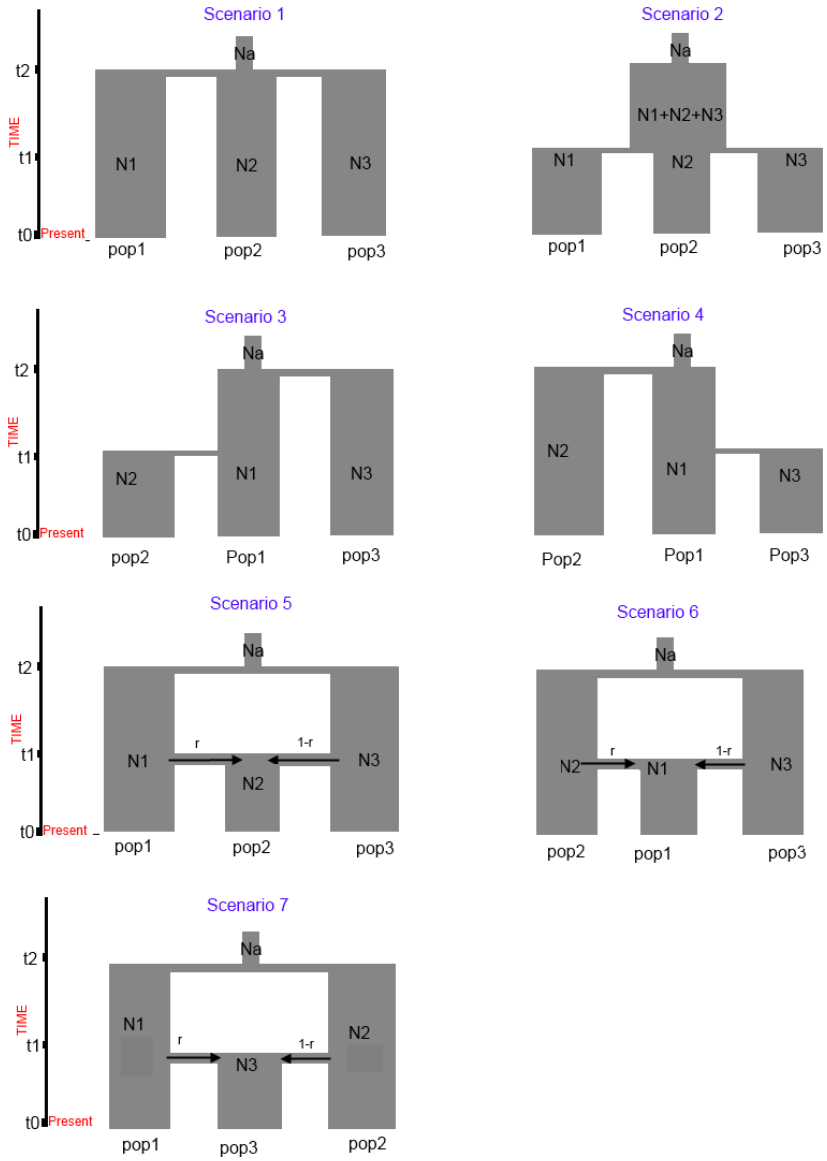



*Figure 4.6: schematic diagrams representing hypothetical demographic scenarios tested using ABC. Scenarios 1 – 4 show permutations of 3 populations diverging from an ancestral population, with no migrations between populations. Scenarios 5-7 show permutations of one population form of admixture between the other 2 populations at a rate of 1-r. Na represented the ancestral effective population size, N1, N2 and N3 represented the effective population sizes of populations 1, 2 and 3 respectively. See Table 4.12 for details populations used for each genetic marker and Tables 4.14 and 4.15 for the prior distributions for shown parameters.*

The admixture rate were calculated for scenarios where two parental populations

contribute to the formation of the third population. The proportion of the third population made up of the first parental population in the scenario is represented by r and is between 0 and 1. In the case where r=1 this means the third population is a daughter population of the first population with no contribution from the second population (1-r). Admixture rates were treated as completely unknown parameters and given ranges to reflect this (see Table 4.13 for NRY and 4.14 for mtDNA).

## Summary Statistics

At the time of writing no formal method existed for choosing statistics that summarise genetic qualities of populations under investigation sufficiently but without too much loss of vital information. Summary statistics were chosen on the basis of their perceived relevance to the parameters and scenarios under investigation (See Table 4.15).

## Generation of simulated data

Simulations were generated using the inbuilt DIYABC coalescence simulator. The simulations generated the same type of genetic data and the same number of samples as the observed data. For NRY this was 10 STRs and for mtDNA this was sequences of 1123 base pairs. There were 1,000,000 simulations per NRY scenario and 100,000 simulations per mtDNA scenario. A smaller number of simulations was generated for mtDNA due to the greater length of time which was required to generate the larger amount of genetic data used for this marker.

Each simulation drew values for scenario and genetic parameters, from prior distributions (see Tables 4.13 and 4.14). This was the most computer-intensive stage, taking several hours to generate data-sets. These were stored on a reference table, to be used for statistical analysis.

*Table 4.13: NRY initial parameter prior distributions, including demographic, historical and archaeological sources for values. The lower boundaries of effective population sizes are set low to reflect lower historical populations sizes. 31 years per male generation after Fenner (2005). Uniform prior distributions were applied to the majority of parameters (U: min-max, with min and max corresponding to the minimum and maximum values of the uniform distribution, respectively).*

| Parameter | Prior distribution | Notes and references |
|---|---|---|
| **N1 (Wales)** | 20 – 10,000 | Effective population size of Wales, based on approximate male census size (Tregaron) of 600 (City Population 2014) |
| **N1 (Eire)** | 5000 – 100,000 | Effective population size of Ireland, based on approximate male census population size of 2 million (Central Statistics Office 2012) |
| **N2 (Isle of Man)** | 500 – 50,000 | Effective population size of the Isle of Man, based on approximate male census population size of 40 000 (Isle of Man Government 2011) |
| **N3 (Norway)** | 5000 – 100,000 | Effective population size of the Norway, based on approximate male census population size of 2.5 milion (Statistisk Sentralbyra 2014) |
| **Na** | 500 - 5000 | Ancestral population size |
| **r1** | 0 – 1 | Admixture rate, unknown parameter |
| **t1** | 20 – 64 gens | ~700 – 2000 years ago. encompassing time-frames for potential Irish (Broderick 1999, 13; Jackson 1953, 157; Phillips 2004, 1) and/or Norwegian genetic input (Richards 2009, 266; Wilson 2008, 24-27; Wormald 1991, 132) including extent of Norwegian occupation in the area (McDonald 2008, 26-33) |
| **t2** | 130 – 355 gens | ~4000 – 11,000 years ago time-frame allowing for population split from ancestral population to have occurred during the Bronze Age, Neolithic Age (Sheridan 2003), or the Mesolithic Age (Mallory 2013, 58-60) |
| **Mutation Model** | Stepwise Mutation Model (SMM) | (Estoup et al. 2002)Excluding insertions and deletions. |
| **Mean Mutation rate (MR)** | 0.001 | (Wetton pers. comm.) Per generation, fixed parameter |
| **Individual locus MR** | 0.001 | (Wetton pers. comm.) Per generation, fixed parameter |

*Table 4.14: mtDNA initial scenario parameter prior distributions, including demographic, historical and archaeological sources for values. The lower boundaries of effective population sizes are set low reflect lower historical populations sizes. 25 years per female generation after (Fenner 2005). Uniform prior distributions were applied to the majority of parameters (U: min-max, with min and max corresponding to the minimum and maximum values of the uniform distribution, respectively).*

| Parameter | Prior distribution | Notes and references |
|---|---|---|
| **N1 (Eire)** | 5000 – 1,000,000 | Effective population size of Ireland, based on approximate female census population size of 2 million (Central Statistics Office 2012) |
| **N2 (Isle of Man)** | 500 – 50,000 | Effective population size of the Isle of Man, based on approximate female census population size of 40,000 (Isle of Man Government 2011) |
| **N3 (Norway)** | 5000 – 1,000,000 | Effective population size of Norway, based on approximate female census population size of 2.5 milion (Statistisk Sentralbyra 2014) |
| **N3 (Sweden)** | 10,000 – 2,000,000 | Effective population size of Sweden, based on approximate female census size of 4.25 million (World Population Review 2014) |
| **Na** | 500 - 5000 | Effective female populations size for ancestral population |
| **t1** | 28 – 80 gens. | 700 – 2000 years ago. Encompassing time-frames for potential Irish (Broderick 1999; Jackson 1953; Phillips 2004)and/or Norwegian genetic input (Richards 2009; Wilson 2008; Wormald 1991)<br>including extent of Norwegian occupation in the area (McDonald 2008), 26-33) |
| **t2** | 160 – 440 gens 25 years per female gen. (Fenner 2005 | 4000 – 11,000 years ago*, time-frame allowing for Irish/Manx population division to have occurred during the Bronze age, Neolithic Age (Sheridan 2003), or Mesolithic Age (Mallory 2013) |
| **r1** | 0 – 1 | Admixture rate, unknown parameter. |
| **Mutation Model** | Kimura | (Kimura 1980) |
| **Mean Mutation rate (MR)** | $1x10^{-6}$ | Per generation, (Jobling et al. 2004), 61). Shape set to 0 so that all loci take the same value (mean) |
| **Individual locus MR** | $1x10^{-6}$ | Per generation, (Jobling et al. 2004, 61). Shape set to 0 so that all loci take the same value (mean) |

*Table 4.15: Summary statistics chosen to represent different aspects of population diversity and genetic distance between populations. from options, given in DIYABC.*

| | Summary statistics | Notes |
|---|---|---|
| **NRY** | Mean Number of Alleles | Indicates historical effective population size and amount of gene-flow population has experienced. Simulation sample sizes match real sample sizes. |
| | Genetic diversity | Used to indicate levels of immigration as well provide clues about long- term effective population size. |
| | $F_{ST}$ | Intra-population measure. Indicates something about recent gene-flow or migrations between populations, or time since the populations diverged |
| | (Delta mu) $^2$ | Inter-population measure. Indicates levels of gene-flow or migrations between populations, or time since populations diverged, increased sensitivity for detecting more ancient events. |
| **mtDNA** | Number of Haplotypes (k) | Indicates long-term effective population size and amount of gene-flow population has experienced. Simulation sample sizes match real sample sizes. |
| | Mean Pairwise differences | Used to indicate levels of immigration as well provide clues about long- term effective population size. |
| | Mean Pairwise Variance | Informative about older immigrations rather than more recent ones. |
| | $H_{ST}$ | Inter-population measure. Indicates levels of gene-flow or migrations between populations, or time since populations diverged |

## Pre-evaluation of scenario/parameter combinations using ABC

Two methods to assess the goodness of fit for both the priors and the scenarios to the observed data through DIYABC which were carried out before the rest of the analysis. The first generated a Principal Components Analysis (PCA) plot showing distances between simulated data-sets for each model and the observed data. Secondly, the software calculated the proportion of summary statistics for the simulated data that fall below the observed value. The summary statistics with very high or very low proportions across all scenarios provided indications of which priors or scenarios did not fit with the observed data, allowing adjustments to be made to scenarios and prior distributions to produce better fits.

## Scenario Choice Using ABC

DIYABC used logistic regression, as proposed by Beaumont (2008), to assess which scenario best described the observed data. This took into account the difference between summary statistics from simulated data-sets within the tolerance range, and those of the observed data. A tolerance of 0.001% was used for NRY and 0.01% for mtDNA to target the 1000 simulations with values closest to the observed data. Posterior probability distributions were displayed as line graphs, point estimates and quantiles.

## Parameter Estimation using ABC

DIYABC used the weighted linear regression step, as recommended by Beaumont (2002), to estimate posterior distributions for parameter values. Parameters were transformed before calculation using the default logit transformation setting. A tolerance of 0.01% was used so the regression step weighted in favour of the 1000 simulations with summary statistics closest to the observed (Beaumont et al. 2002).

Posterior probability distributions were displayed as line graphs, point estimates and quantiles. The range of the posterior distributions for each parameter provided a visual display of the precision of estimates for each parameter.

## ABC Validation

To calculate type I and type II errors, the software chose 500 simulations at random from a selected scenario and used their summary statistics as pseudo-observed data (PODs) of known providence. ABC was used to count how many times the procedure identified the correct scenario verses an incorrect scenario. In this way, it was possible to establish whether, given the available data, ABC could distinguish between scenarios.

The goodness-of-fit between the selected scenario and the observed data was

assessed by generating simulations from the posterior distributions of the parameters for that scenario. A PCA plot was then generated to compare distribution around the observed data. In a good fit, the PODs form a tight cluster surrounding the observed data.

# CHAPTER FIVE: RESULTS

## 5.1 Introduction

This chapter is divided into two parts. The first describes the patterns found in the genetic diversity of the Isle of Man and populations of the European Atlantic fringe. The second details the results of ABC modelling of possible demographic scenarios which could have led to the patterns described in the first section.

## 5.2 The Genetic landscape

This section details patterns and affinities revealed by a number of commonly used classic populations genetics and phylogeographic techniques. The results in this section described the genetic landscape of western Europe. These methods are considered descriptive rather than interpretive.

### 5.2.1 NRY

Of the 67 samples collected from Manx volunteers, complete or partial STR and SNP profiles were obtained from 66. Despite numerous extraction and typing attempts, one sample failed to produce sufficient data to be included in any of the analyses. Eight samples produced incomplete genetic profiles, but contained enough information to be included in either SNP or STR analyses. A full list of Isle of Man NRY STR and SNP haplotypes can be found in Appendix B1. Of the 58 complete genetic profiles obtained only three identical pairs of STR/SNP profiles were found.

SNPs could only be typed for 61 of the 66 samples. Twelve of the 17 possible

SNP haplogroups were present in the Isle of Man sample. In agreement with previous studies (Capelli et al. 2003) the NRY SNP haplogoup profile of the Isle of Man was dominated by the R1b haplogroup (69%). However, unlike previous studies, R1b subgroups were also typed allowing a more detailed investigation of the genetic composition of the island. The most common haplogroup in the Isle of Man was R1b-S145, which represented 43% of R1b and 30% of the total sample (See Figure 5.1). The second most frequent haplogroup found on the Isle of Man was I1 (15%).

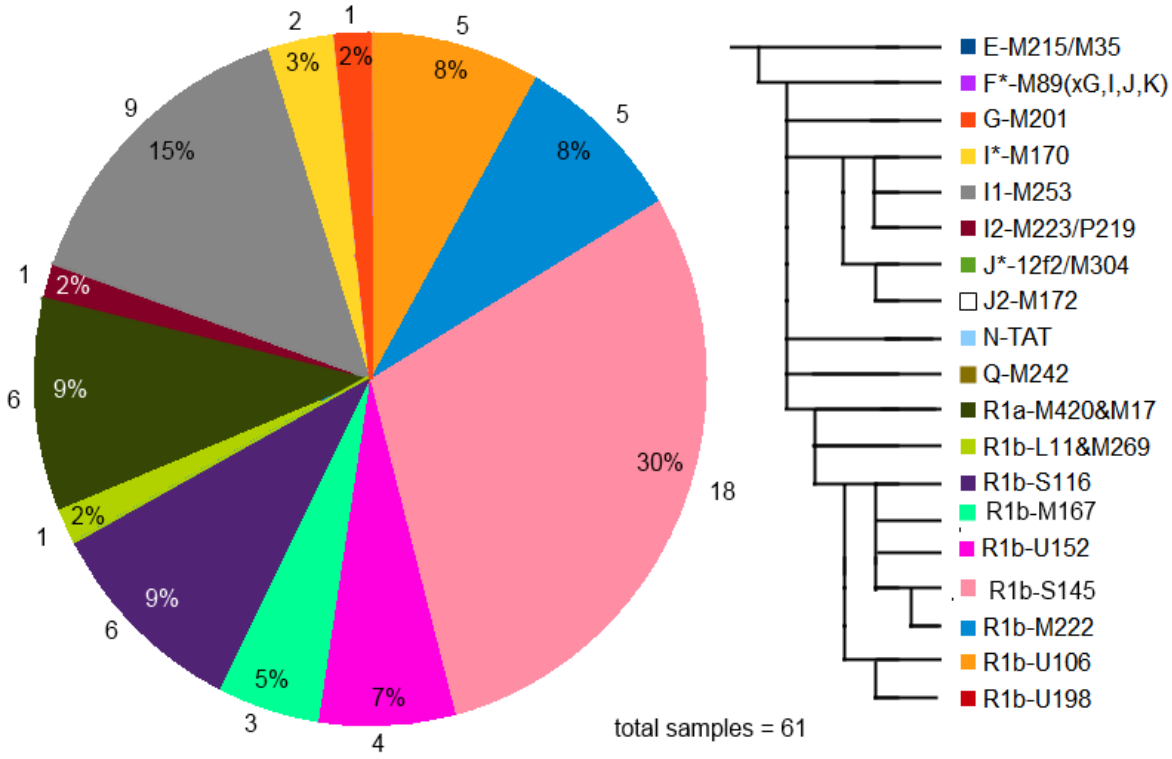



*Figure 5.1: NRY haplogroup frequencies for the Isle of Man. Actual numbers for each haplotype are shown around the outside of the chart.*

## Intra-Population Diversity

Genetic diversity was calculated for SNP haplogroups at various resolutions, to allow comparisons between populations which had been typed for slightly different groups of SNPs. The highest-resolution analysis investigated 17 SNP

haplogroups including six R1b subgroups and three subgroups of the I haplogroup, allowing comparisons between 18 populations in the British Isles and seven in Scandinavia (See Fig. 5.3). Values for R1b-S145 and R1b-M167 were pooled with R1b-S116, as the British Isles data-sets were not typed for these haplogroups.

The lowest resolution, analysing just 8 SNPs, allowed the highest number of Western European populations to be compared (n= 39). A comparison between the haplotype diversity at these two levels of resolution can be seen in Figure 5.2 (see Appendix B2 for table of values).

The Isle of Man was found to have haplogroup diversity close to the average for the whole Atlantic region at both resolutions. The island was more diverse than its nearest neighbours around the Irish Sea region at both resolutions, possibly indicating recent admixture.

At the 8-SNP level, most Scandinavian populations, along with Southern Portugal, Austria and Galicia, showed the highest levels of diversity. Ireland, and western parts of Britain, as well as North-Eastern Iberian populations, including the Basque Country, showed the least diversity (Figure 5.2). This result was typical of results found in previous studies (Capelli et al. 2003), see Appendix B3 for map of relative haplotype frequencies in western Europe).

However, increasing the number of haplogroups within the R1b and I haplogroups revealed much more diversity within British and Irish populations, whilst making little difference to the overall diversity of Scandinavia. At the highest SNP resolution the English populations of Leicestershire, Yorkshire and Kent had the highest levels of diversity, higher than any of the populations in Scandinavia. Irish Sea populations of Wales, Cornwall, Argyll and Northern Ireland remained the least diverse, although the difference between these and

other British populations were less pronounced than at lower resolutions. Data for South-Western Europe was not available at this resolution.

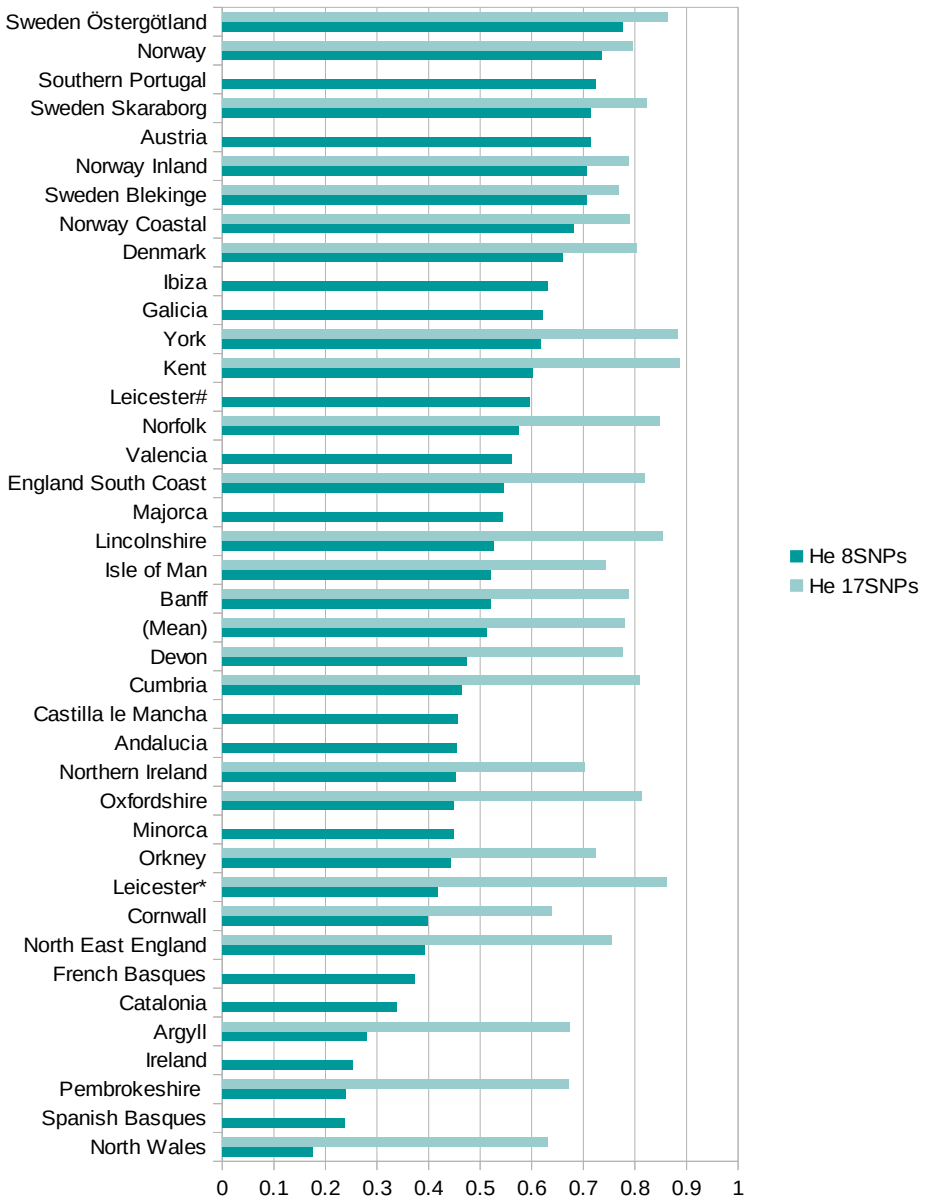



*Figure 5.2: comparison of western European populations' genetic diversity when SNPs are divided into maximums of 8 and 17 haplogroups. Mean is the average values for all samples. *from Hellenthal* et al. *(2014), # From Balaresque and Jobling (unpublished)*

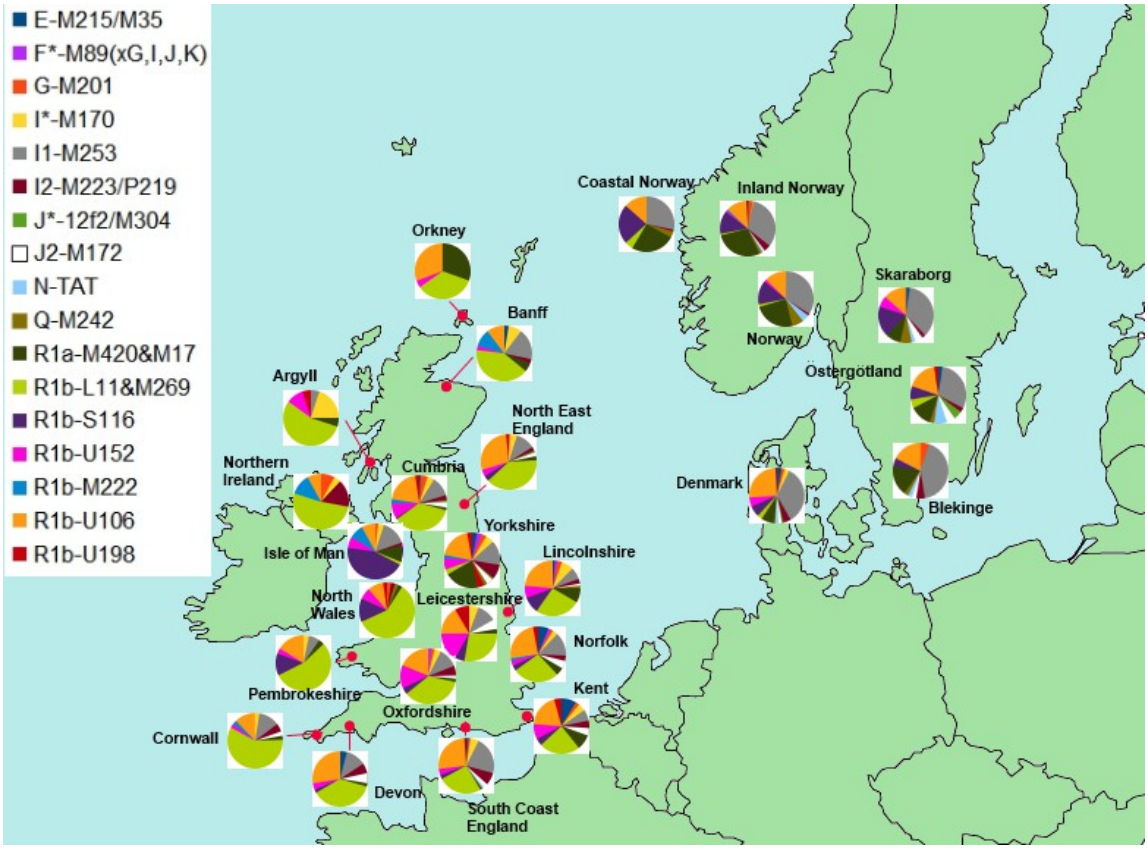



*Figure 5.3: Map showing relative NRY haplogroup frequencies for 17 haplotypes for the British and Irish Isles, and Scandinavia.*

Although parts of Britain showed similar or higher levels of diversity to Scandinavia at the 17-SNP level of resolution, the frequencies of particular haplotypes differed between the regions. Figure 5.3 shows a map highlighting these regional differences in haplotype frequency at the highest resolution. A full list of haplotype frequencies for each population can be found at Appendix B4.

Perhaps most striking was the relatively high frequency of R1b-S116 in the Isle of Man compared to all other populations, but particularly in relation to its nearest neighbours and the rest of the British Isles. Norwegian populations and Skaraborg contained slightly higher frequencies of haplogroup R1b-S116 than was typically found in the British Isles, however, it was not found in the

proportions seen on the Isle of Man.

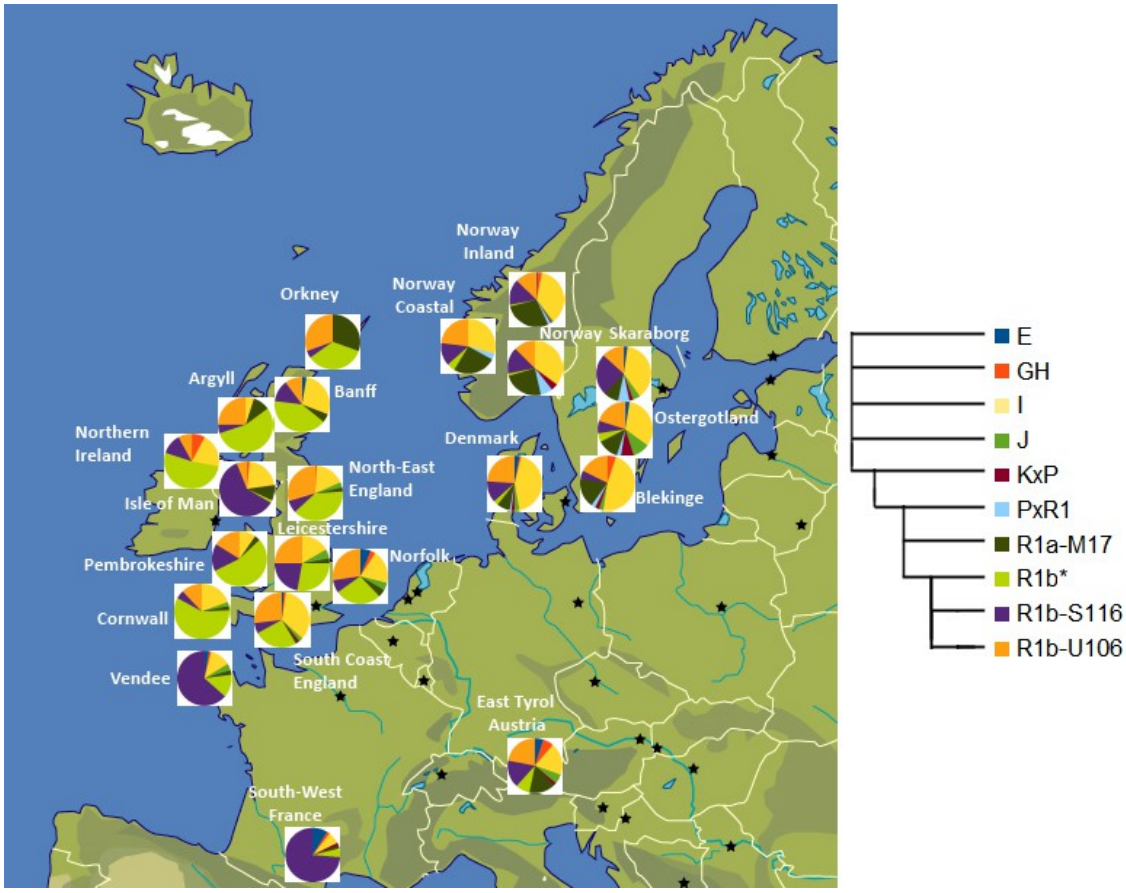



*Figure 5.4: Geographical distribution of relative NRY haplogroup frequencies when the R1b-L11 haplogroup is subdivided by a further two SNPs, U106 and S116. The phylogeny to the right shows the evolutionary relationship between the haplogroups shown.*

To investigate this further, R1b was divided into its three main subgroups R1b-L11, R1b-U106 and R1b-S116, to allow more populations to be included in the analysis (Fig. 5.4). Values for R1b-U152 and R1b-M222, were added to R1b-S116 and values for R1b-M198 were added to R1b-U106. This analysis revealed that the frequency of R1b-S116 found in the Isle of Man was on a comparable scale to populations in France (Figure 5.4). In contrast, the rest of the British Isles populations had higher frequencies of R1b-U106 and R1b-L11. The highest frequencies of R1b-L11 (>50%) were found in the Irish Sea populations (Figures 5.3 and 5.4). Frequencies of R1b-U106 were similar between Scandinavian and

British Isles populations.

Haplotype I1 was found at much higher frequencies all over Scandinavia (~25-40%) than in the British or Irish Isles (<22%), which tended to have a greater range of I haplotypes. Outside of Scandinavia, haplotype R1a1, which is considered to be an indicator of Norwegian contacts, was found in the highest frequencies in Orkney, Yorkshire, Isle of Man, and to a lesser degree, Kent and Lincolnshire. (see Appendix B4 for full list of haplotype frequencies for all populations).

As a comparison, genetic diversity for NRY STRs were also calculated at two resolutions: 10 STRs and a further 5 STRs for a reduced number of populations (see Appendix B5). There were some consistencies between diversity measures for SNP and STR data. Welsh populations showed the least diversity overall, with Llanuwchllyn having almost half the diversity of Portugal and Ostergotland at the top end of the scale. Scandinavian populations were again amongst those with the highest diversity, along with Southern Portugal, Galicia and populations in central Europe (See Appendix B6).

STR data for the British and Irish Isles were only available for Ireland (Eire), four populations in England (Winchester, Chippenham, Leicester and Cornwall) and four populations in Wales (Abergele, Llanwychlin, Llangefni and Tregaron). With the exception of Abergele in Wales, the western regions of the British and Irish Isles were again amongst the populations with the lowest diversity, along with the Basques and Minorca. Winchester, Chippenham and the Isle of Man were around the average for STR diversity, higher than all Irish Sea populations except Abergele. Leicester had the highest STR diversity of all British and Irish populations.

One major difference between STR and SNP diversity was that increasing the

number of STRs used did not have such a dramatic effect on the levels of intra-populations diversity for Irish Sea and northern Iberian populations as was seen when the numbers of SNPs was increased (Fig. 5.2). Closely related R1b haplotypes which are at high frequency in these regions, have very similar STR profiles making them indistinguishable even at the highest STR resolution. This factor resulted in some loss of information when comparing diversity of Atlantic populations using just STRs.

## Inter-population Diversity

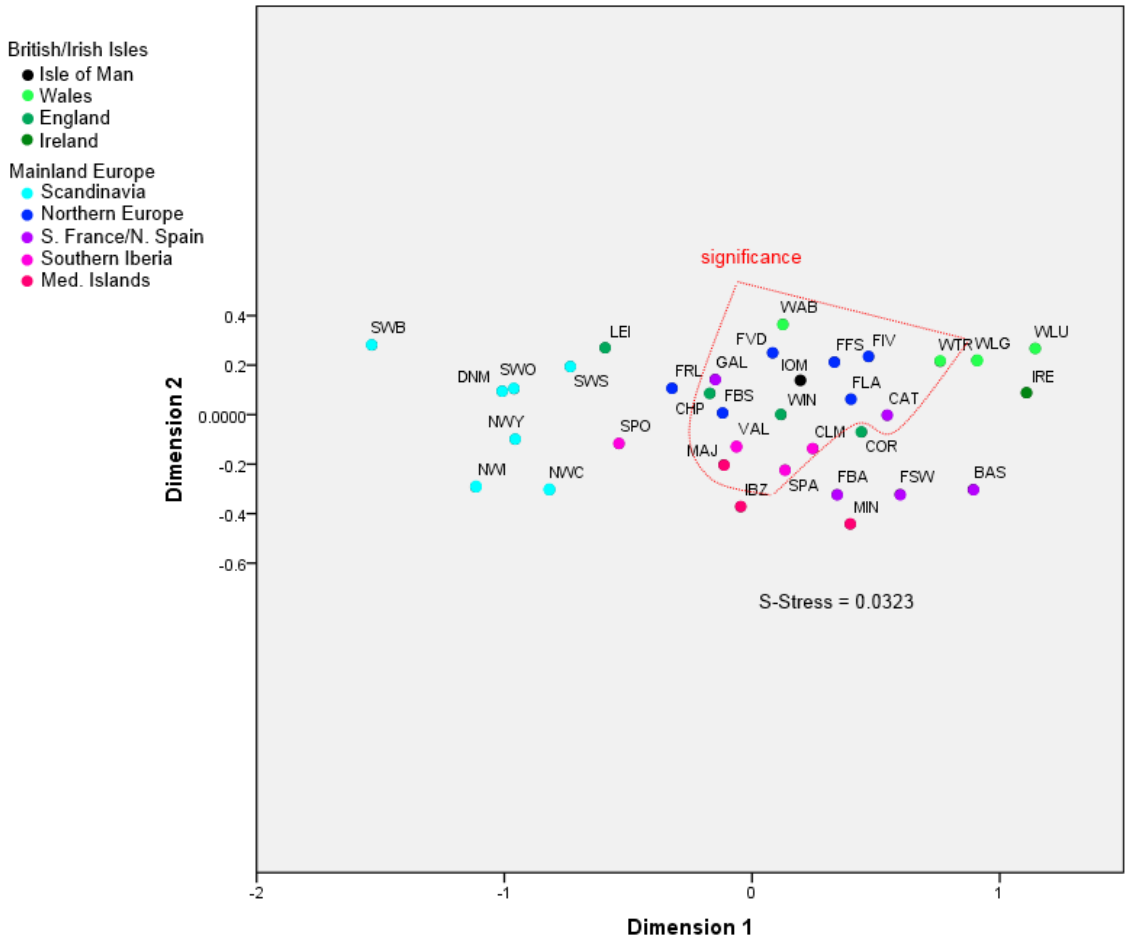



*Figure 5.5: MDS plot of genetic distances between populations based on $R_{ST}$ values for 10-STR haplotypes. Genetic similarity between populations are represented in two dimensional space. Populations within the red line have $R_{ST}$ values which are not significantly different from the Isle of Man (P value = 0.05), meaning they are genetically indistinguishable from the Isle of Man on the basis of $R_{ST}$. S-Stress measures how well the plot fits the data, <0.1 is considered a good fit (Dugard* et al. *2009).*

MDS plots were calculated from inter-population $F_{ST}$ for SNP haplotype variation and $R_{ST}$ for STR variation. These plots were used to describe the genetic relationships between populations. The S-stress value indicated goodness-of-fit between the data and the plot; all plots showed less than 0.1, which was considered to be a good fit (Figures 5.5 and 5.6; Dugard et al. 2009).

Using 10 STRS allowed the maximum number of populations to be displayed, without losing too much information (Appendix B7 for MDS plot using inter-population RST for 15 STRs). The 10 STR plot (Figure 5.5) showed the Scandinavian populations clustered at one side, with Ireland and three Welsh populations on the extreme opposite. Leicester was the British population lying closest to the Scandinavia cluster. In the middle, populations were arranged approximately south to north. The Isle of Man was found in the centre top of the main cluster, but was not significantly different to the French populations or Abergele (Wales). Using $F_{ST}$s based on 17 SNPs (Figure 5.6), the genetic distances between Scandinavia and the British Isles became much more pronounced. The Isle of Man was a clear outlier, significantly different from all other populations, although closer to Scandinavia on one dimension. This was probably due to the high frequency of R1b-S116 (see Figures 5.3 and 5.4). Orkney was also an outlier, although not to the same degree as the Isle of Man, but was found lying below the main British Isles cluster.

The British populations which lay closest to Scandinavia were Yorkshire and the south coast of England, closely followed by Kent and Norfolk. Of the Scandinavian populations, Ostergotland in Sweden lay closest to the British populations, followed by Denmark. Again, the Irish Sea populations were clustered together at the furthest side of the British cluster from the Scandinavian populations.

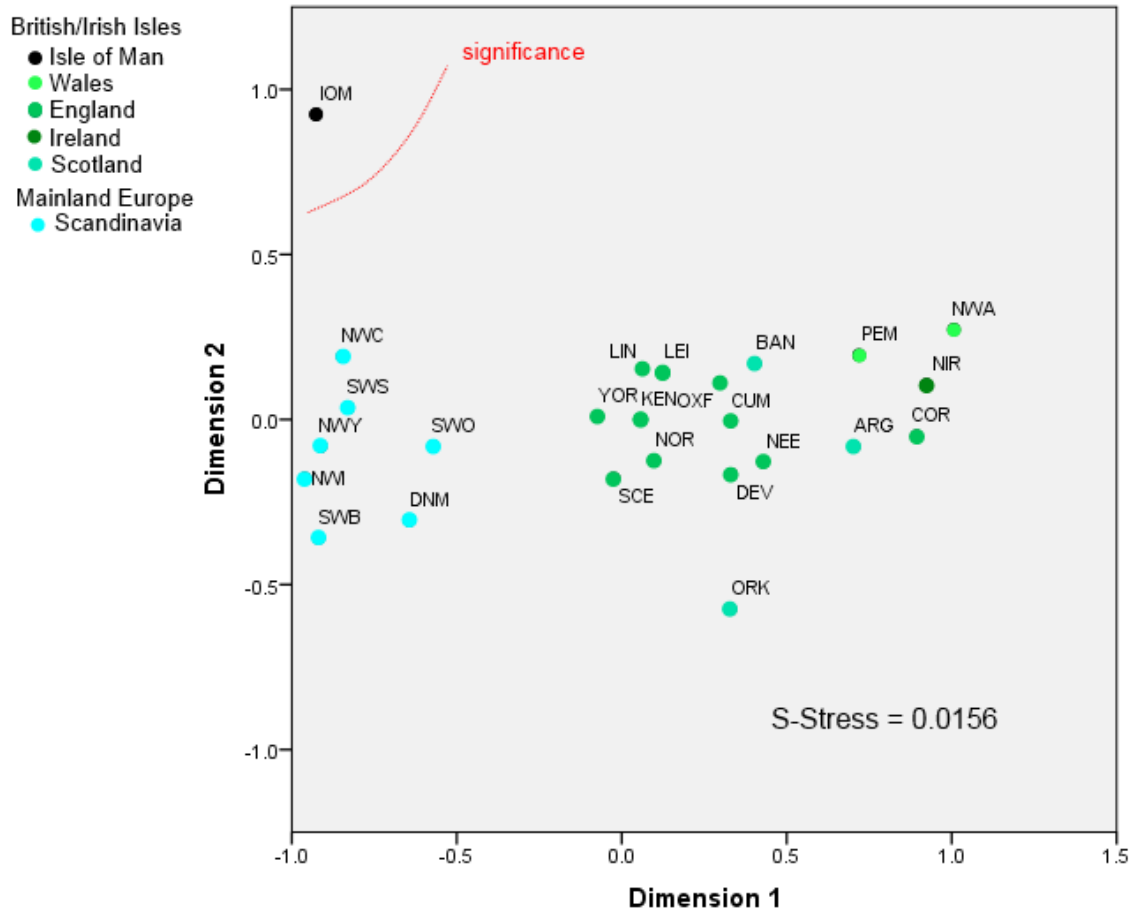



*Figure 5.6: MDS plot of genetic distances between populations based on frequencies of 17 NRY SNP-defined haplogroups. Genetic similarity between populations are represented in two dimensional space. Populations outside of the red line have $F_{ST}$ values which are significantly different from the Isle of Man (P value = 0.05), meaning they are genetically distinct from the Isle of Man. S-Stress measures how well the plot fits the data, <0.1 is considered a good fit (Dugard* et al. *2009).*

## Network analysis

Comparisons between haplotypes of the Isle of Man and other populations revealed that two individuals from the Isle of Man carried SNP and 17 STR haplotypes that exactly matched individuals from Norway. Another individual carried a haplotype which differed at just one STR locus from a Norwegian. Two of these matches belonged to haplogroup I1 (Appendix B8), but the third was haplotype R1b-M222 (see Fig. 5.7), typical of Irish samples (Moore et al. 2006). There were no shared haplotypes between the Isle of Man and any other

population within this study. Neither were matches found between Manx 17 STR/SNP combination haplotypes and individuals from elsewhere using the YHRD world-wide database ( http://yhrd.org/ ).

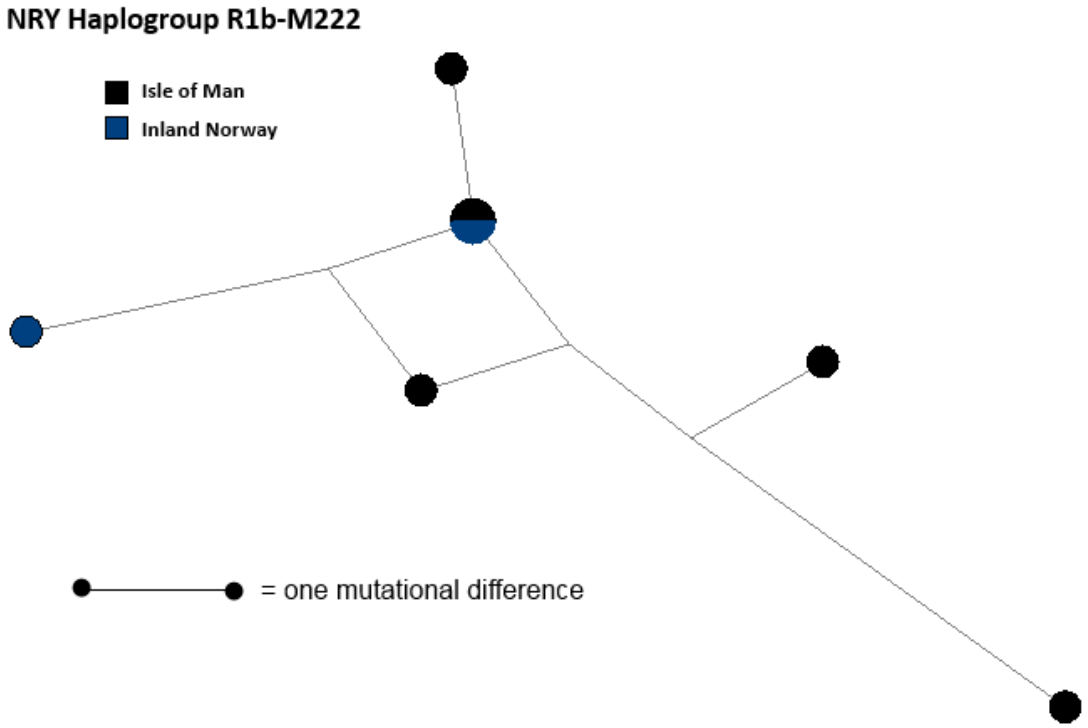



*Figure 5.7: Network diagram showing the genetic relationship between 15-STR haplotypes of all individuals sampled who also carried the NRY R1b-M222 haplotype. The size of the circle is proportional to the number of individuals carrying identical 15-STR haplotypes; the smallest equalling one. The length of the line between haplotypes is proportional to the number of one-step differences separating STR haplotypes. NWI = Inland Norway, IOM = Isle of Man.*

Median-joining network analysis (Bandelt et al. 1995) was used to identify potentially closely related individuals between populations and to look for interesting features which could indicate demographic events. The Irish R1b subgroup, M222, was used to examine the potential relationship between the two Norwegians and five Manx individuals belonging to this lineage. The network showed that the haplotype shared by a Norwegian and a Manx was one mutational step from a third individual also from the Isle of Man. Given STR mutation rates of 0.001 per generation, it is likely that these three individuals

shared a common ancestor within the last 1000 years (Fig. 5.7). None of the other individuals carrying R1b-M222 look closely related.

The I1 haplogroup network (Appendix B8) appeared to have several star-like clusters, with no clear core. This seemed in agreement with previous studies that have detected traces of ancient population expansions within this haplogroup (Wei et al. 2013a). However, no geographic structuring could be detected and the Isle of Man individuals were scattered about the network with no discernible pattern.

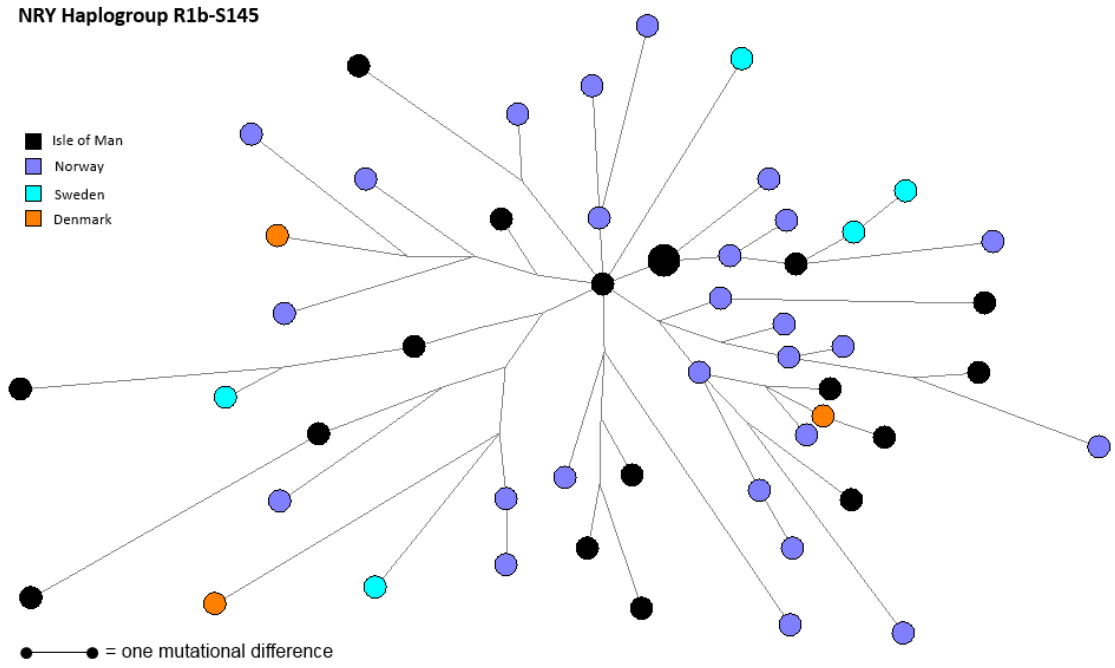



*Figure 5.8: Network diagram showing the genetic relationship between 15-STR haplotypes of all individuals sampled who also carried the NRY R1b-S145 haplotype. The size of the circle is proportional to the number of individuals carrying identical 15-STR haplotypes; the smallest equalling one. The size of the line between haplotypes is proportional to the number of one-step differences separating STR haplotypes.*

The R1b-S145 network had the most striking star-like pattern, shown in Figure 5.8. However, there was no obvious geographic structuring, and the Isle of Man samples were scattered throughout the network including at the core. The R1b-U106 network was also star-shaped although more complex, probably due to

the greater number of samples (Appendix B9). Again there was no obvious structuring and the Isle of Man samples were scattered but in this case, far from the core.

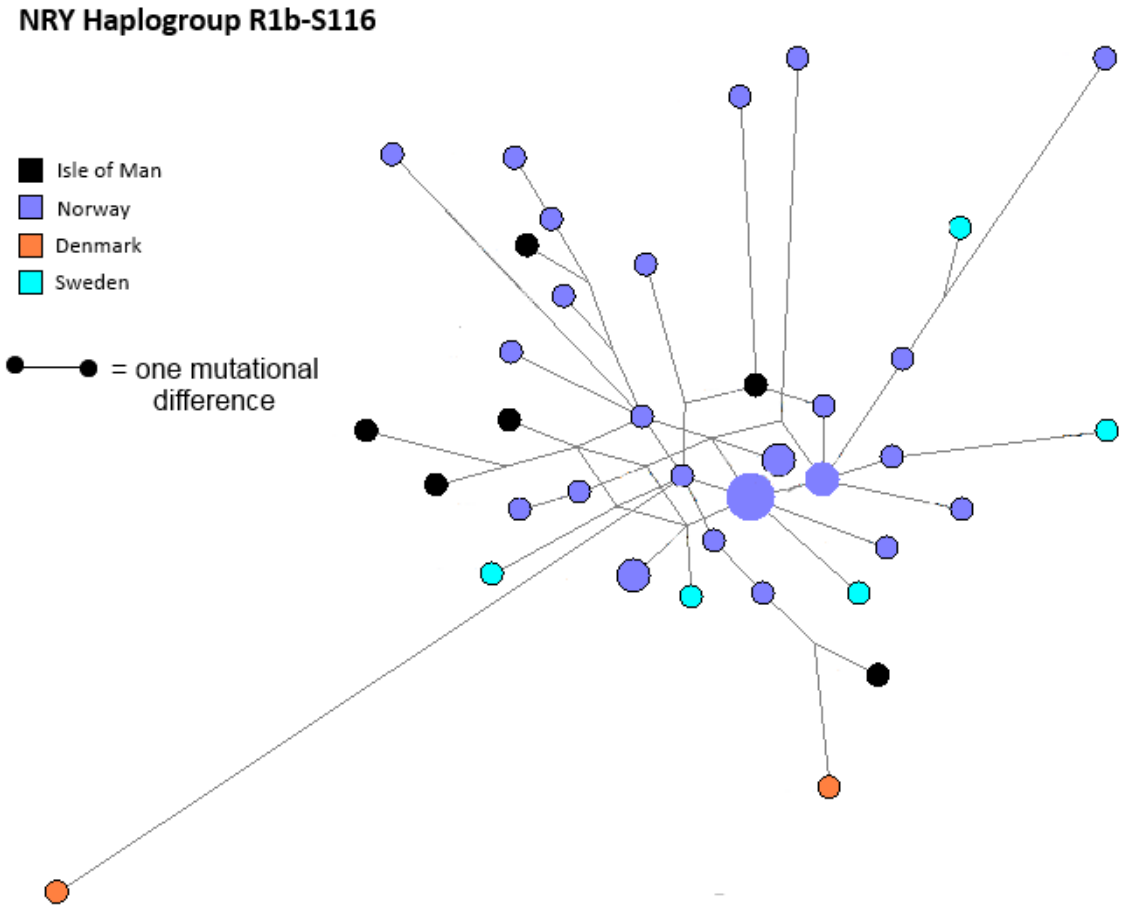



*Figure 5.9: Network diagram showing the genetic relationship between 15-STR haplotypes of all individuals sampled who also carried the NRY R1b-S116 haplotype. The size of the circle is proportional to the number of individuals carrying identical 15-STR haplotypes; the smallest equalling one. The size of the line between haplotypes is proportional to the number of one-step differences separating STR haplotypes*

Examination of the R1b-S116 network did not reveal an obvious star-like pattern (Figure 5.9). However, there did seem to be a cluster composed exclusively of Manx individuals, suggesting some geographical structuring within the haplogroup.

### 5.2.2 mtDNA

A total of 33 mtDNA sequences were obtained from the Isle of Man samples. These were taken from individuals within the NRY data-set that had at least two generations of maternal ancestry on the island. A maximum of 1136 bp were sequenced encompassing the entire control region (16024-576) of the mitochondrial DNA (see Figure 2.4, in Chapter 2).

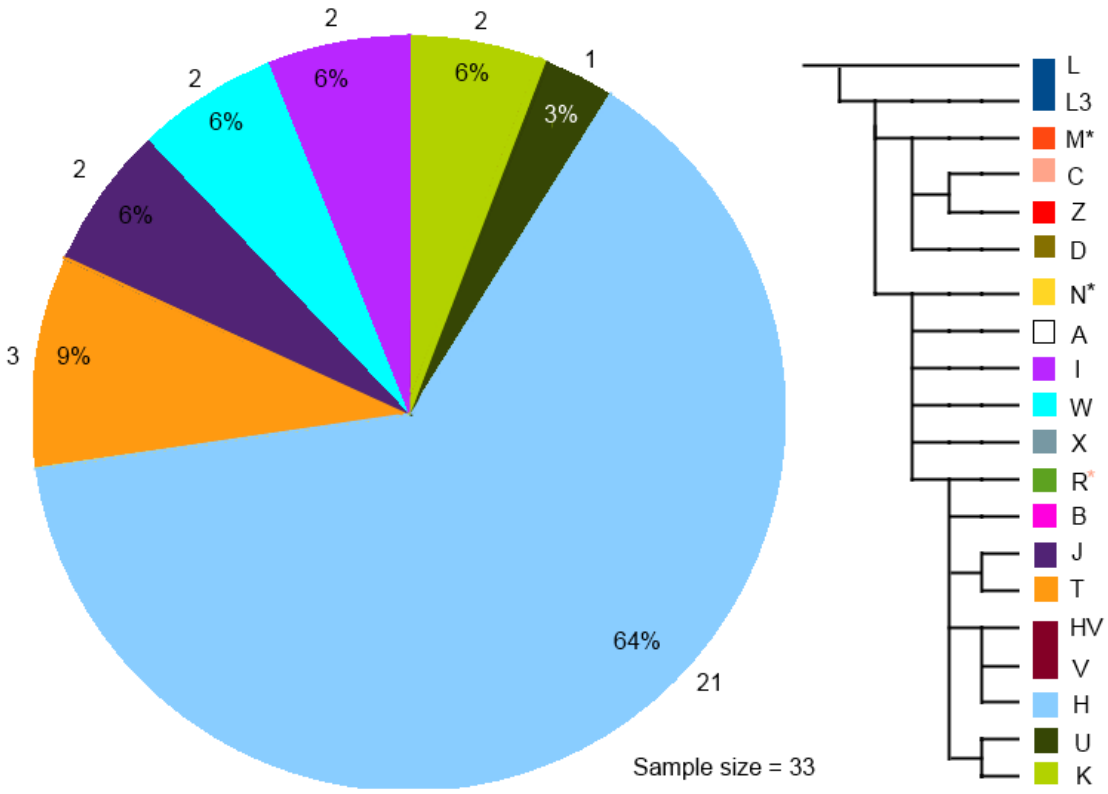



*Figure 5.10: mtDNA haplogroup frequencies for the Isle of Man. Actual numbers for each haplotype are shown around the outside of the chart.*

Sequences were assigned to a haplotype using Haplogrep online prediction tool (see Section 4.4.2, Chapter 4). The Isle of Man was found to be dominated by the H haplogroup (64%, see Figure 5.10), with the remaining sequences almost equally distributed between haplotypes T (9%), J (6%), W (6%), K (6%) and U (3%). This was the first study to explore the mtDNA haplotype frequencies of the

Isle of Man.

Of the 32 sequences obtained, four sequences were identical and a further six pairs of identical sequences were found. The four identical sequences belonged to haplogroup H, as did four of the matching pairs of sequences. The remaining pairs belonged to haplogroup W and J (see Appendix B10 for list of Isle of Man samples).

## Intra-Population Diversity

As well as the Isle of Man, sequence data were available for the control region (1123 bp) of another twelve populations of western Europe, while a further six populations had sequence data available for just HVS 1 and 2 (634 bp). Nucleotide diversity was calculated using Arlequin for all populations at both resolutions where available (See Appendix B11 for list of nucleotide diversity values). Increasing the number of base pairs under investigation, decreased the overall nucleotide diversity, suggesting that HVS 1 and 2 are the most diverse regions of the sequences, so including the rest of the control region acts to dilute the diversity (see Fig. 5.11).

The Isle of Man had just below average diversity for the region, with similar levels of diversity to Skaraborg in Sweden. Northern Spain, including the Basque Country, had the lowest diversity, in agreement with previous research (Bertranpetit et al. 1995). The Swedish populations of Gotland and Östergötland, as well as Leicester in England, and Northern Germany, had the highest diversity at both resolutions.

Sequence data were not available for most British Isles data-sets. Instead haplotypes had been inferred from combinations of mitochondria whole genome polymorphic sites. To enable comparisons between the Isle of Man and all comparative data, haplotypes were inferred from control region or HVS1 and

2 polymorphisms for the 19 populations using the online haplotype predictor Haplogrep (http://haplogrep.uibk.ac.at ).

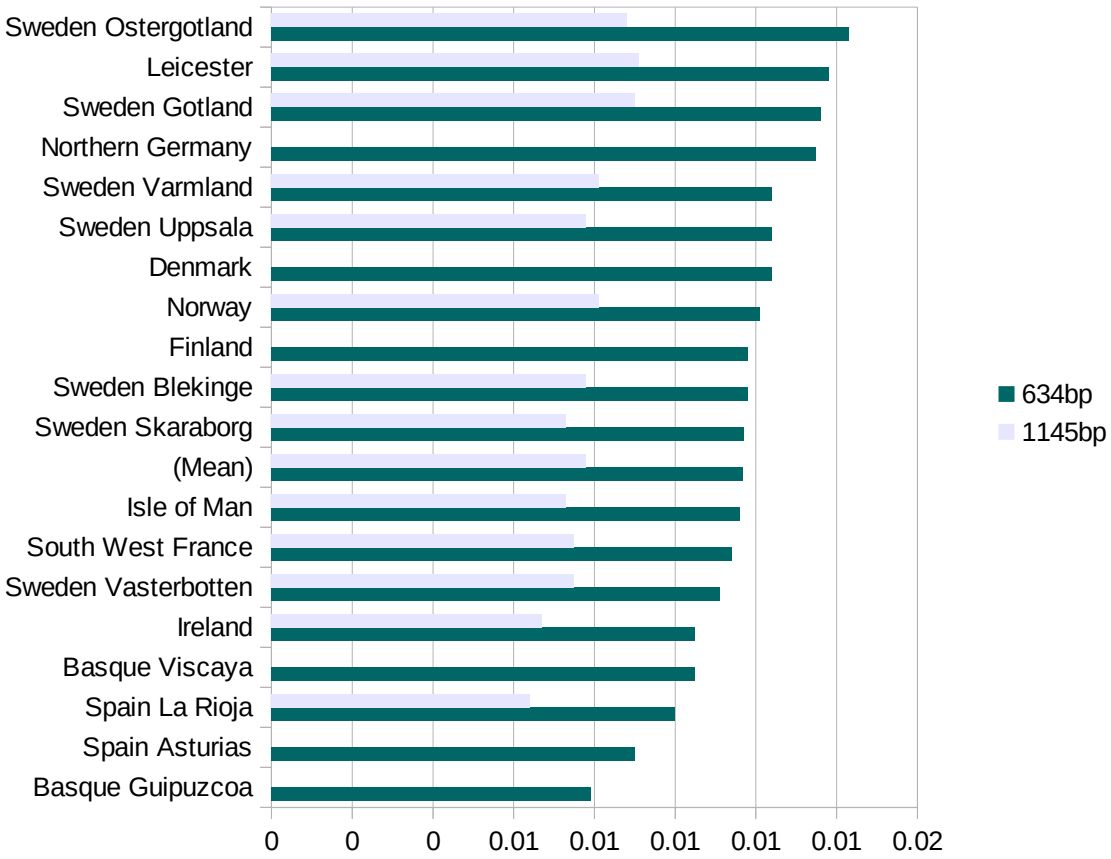



*Figure 5.11: comparison of western European population mtDNA nucleotide diversity for sequences of 634 base pairs and 1145 base pairs. Mean indicates the average value calculated for the region*

Sequences were grouped into either ten or nineteen haplogroups to give two levels of resolution. However, very little difference was found in the overall levels of diversity between these two resolutions (Appendix B12). For some populations, changing the type of diversity measure had little impact on their

relative diversity. La Rioja (Spain) and Guipúzcoa (Basque Country), were amongst the populations with the lowest diversity whether looking at sequence diversity or haplogroup frequency. Likewise Östergötland (Sweden) and North Germany were amongst those with the highest diversity for both measures.

However, for some of the populations, looking at haplogroup diversity rather than directly at sequence diversity altered the amount of diversity they appeared to have. For example, on the basis of haplogroups, the Isle of Man and Norway changed from being somewhere near the mean of the group, to the least and most diverse populations, respectively. Leicestershire and the Republic of Ireland also dramatically changed positions in the ranking.

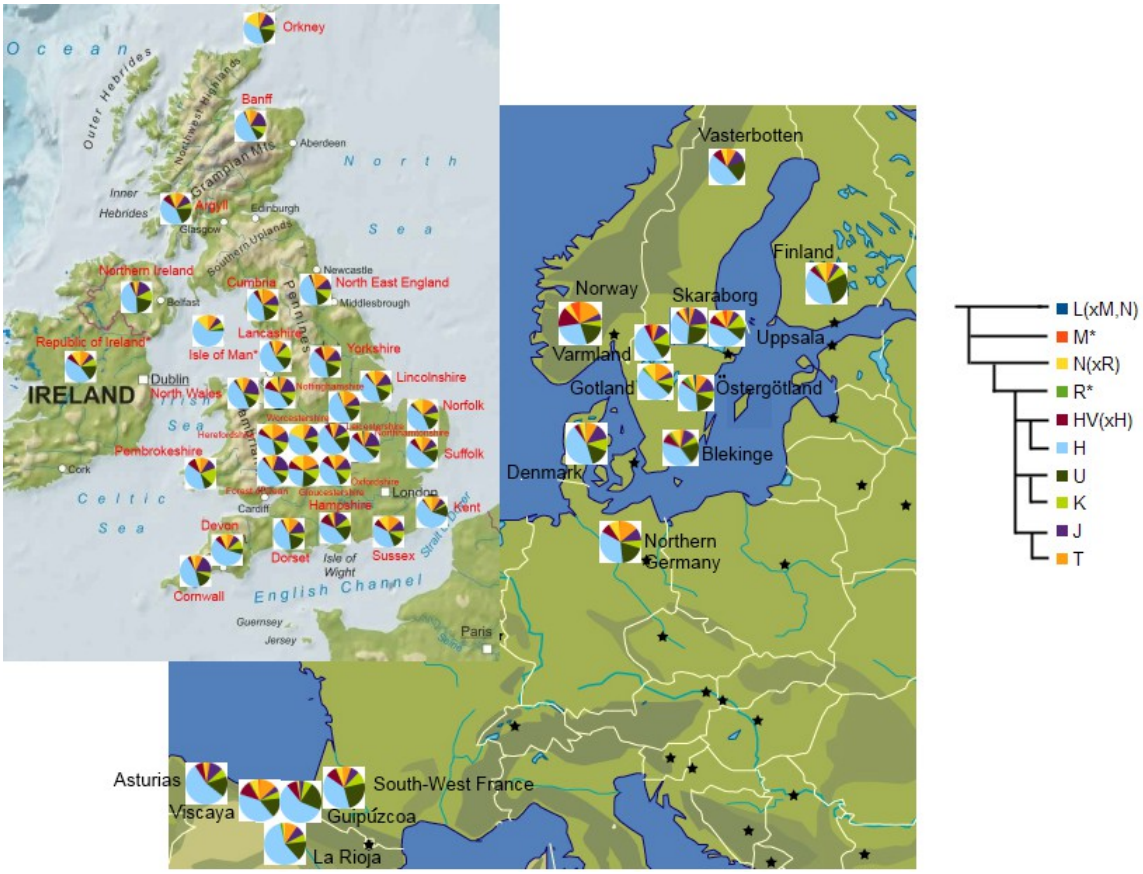



*Figure 5.12: Map showing the frequencies and distributions of the ten major mtDNA haplogroups in populations of western Europe. The British and Irish Isles are shown larger for clarity. The phylogeny shows the evolutionary relationship between haplogroups, shown on the map.*

The discrepancies between the two types of diversity measure were probably related to the high frequencies of the H haplogroup in western Europe, which could not be sub-divided using the data available. Figure 5.12, which shows haplogroup frequencies in the populations under investigation, confirmed that Isle of Man had the highest frequency of haplogroup H and Norway had the lowest, which would have impacted on their relative diversities (see Appendix B13 for full list of haplogroup frequencies).

## Inter-population Diversity

Multi-dimensional scaling was used to examine the relationships between the Isle of Man and other populations around north-western Europe. $F_{ST}$ values of the pairwise differences for the two resolutions of haplogroup diversity and sequence diversity. Although sequence data might have given a better comparison of population affinities, more populations could be included for analysis when just using haplogroups (see Fig. 5.13).

Even at the highest (19) haplogroup resolution the Isle of Man was not significantly different from Ireland (Eire) and many British populations including North Wales. Many Scandinavian populations were also not significantly different from the Isle of Man, including Denmark and four Swedish populations including Västerbotten which was the most northern population of Sweden used in this research. However, it was significantly different from some of its other nearest neighbours, such as Argyll, Northern Ireland and Cumbria.

When exploring genetic distances using the perhaps more reliable sequence data (Fig. 5.14), the Isle of Man was not significantly different to four Swedish populations whist showing a significant difference from the geographically much closer Ireland. In contrast to the haplogroup MDS plot (Fig. 5.13), there was also no significant difference between Norway and the Isle of Man.

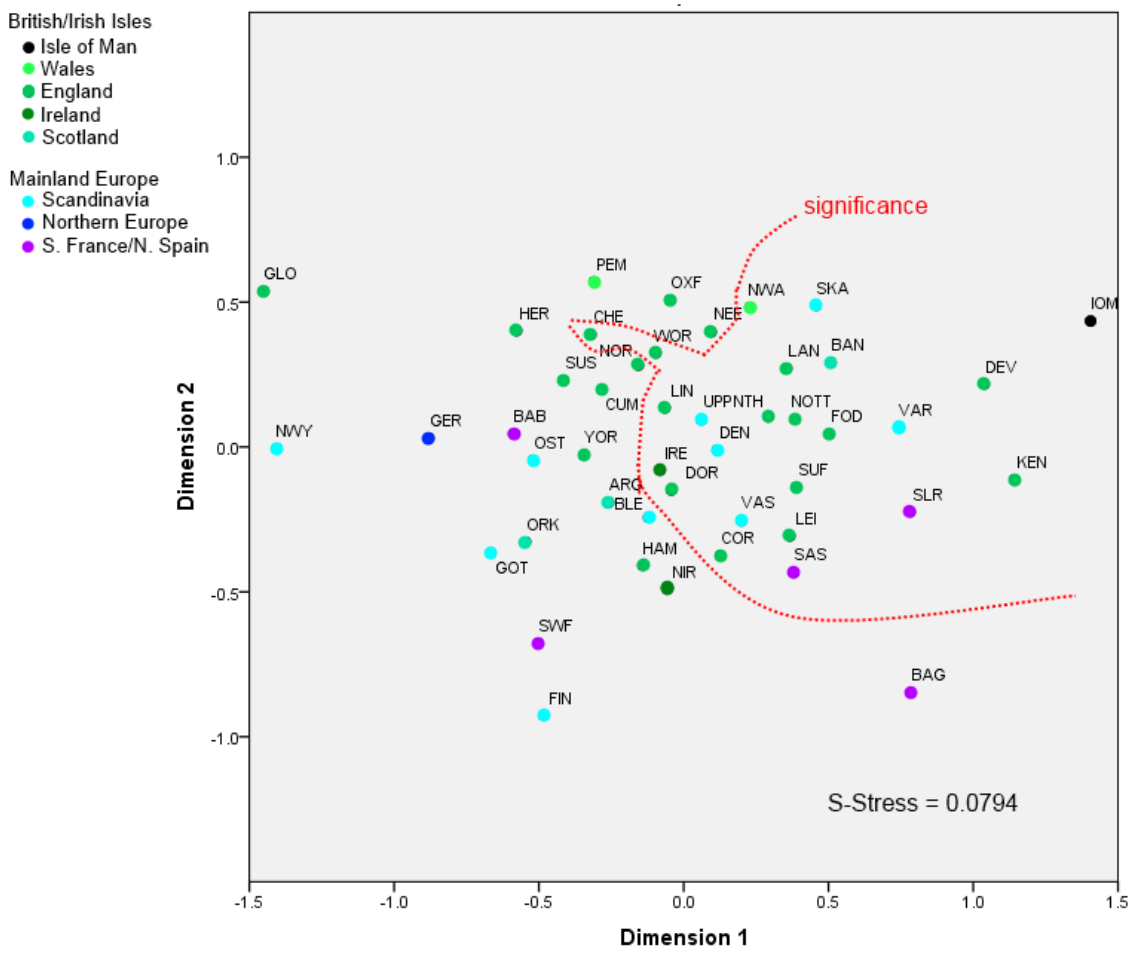



*Figure 5.13: MDS plot of genetic distances between populations based on frequencies of 19 mtDNA haplogroups. Genetic similarity between populations are represented in two dimensional space. Populations to the right of the red line have $R_{ST}$ values which are not significantly different from the Isle of Man (P value = 0.05), meaning they are genetically indistinguishable from the Isle of Man on the basis of $R_{ST}$. S-Stress measures how well the plot fits the data, <0.1 is considered a good fit (Dugard* et al. *2009).*

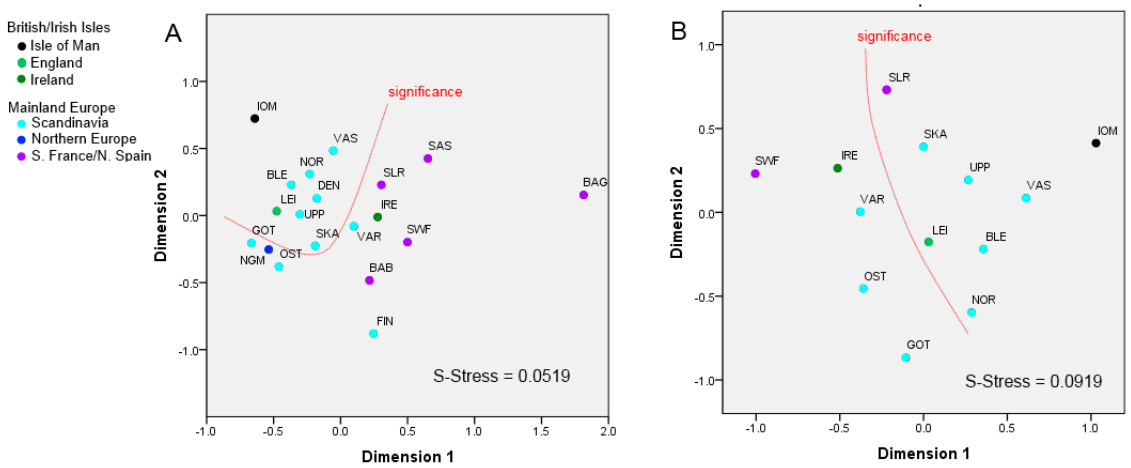



*Figure 5.14: MDS plot of genetic distances between populations based on mtDNA sequences lengths of (A) 634 bp and (B) 1123 bp. Genetic similarity between populations are represented in two dimensional space. Populations on the same side of the red line as the Isle of Man have $F_{ST}$ values which are not significantly different from that of the Isle of Man (P value = 0.05). meaning that they are genetically indistinguishable from the Isle of Man on the basis of $F_{ST}$. S-Stress measures how well the plot fits the data, <0.1 is considered a good fit (Dugard* et al. *2009).*

## Network Analysis

Matches among 1123 bp sequences were found between the Isle of Man and all twelve of the populations where these data were available. One of these sequences was found in 31 individuals across 12 of 13 populations, (it was not found in Leicester). Another sequence was carried by 14 individuals across nine populations. Both these sequences belonged to haplogroup H.

A further sequence belonging to the H haplogroup was carried by seven individuals but was found only in the Isle of Man, south-west France and La Rioja. South-west France shared the greatest number of sequences with the Isle of Man, with six sequences in common. However, this was probably due to south-west France's sample size (n=164) which was at least triple that of all the other populations.

One sequence, belonging to the T haplogroup, was found in six individuals across Sweden, Ireland and the Isle of Man. However, network analysis of this

haplogoup T revealed no obvious geographic structuring, although there were some star-shaped clusters possibly indicating population growth but with closely related sequences dispersed throughout western Europe (Fig. 5.15).

Network analysis of mtDNA control sequences belonging to the H haplogroups showed distinctive star-shaped clusters indicative of the ancient population expansion that has previously described in the literature (see Appendix B15; Fu et al. 2012). However, no obvious geographical structuring of lineages was apparent, instead closely related linages were found dispersed all over the whole region from southern Europe to northern Scandinavia. Similar star clusters with closely-related but distantly-dispersed lineages were also found for haplogroups J, K and I (see Fig. 5.15).

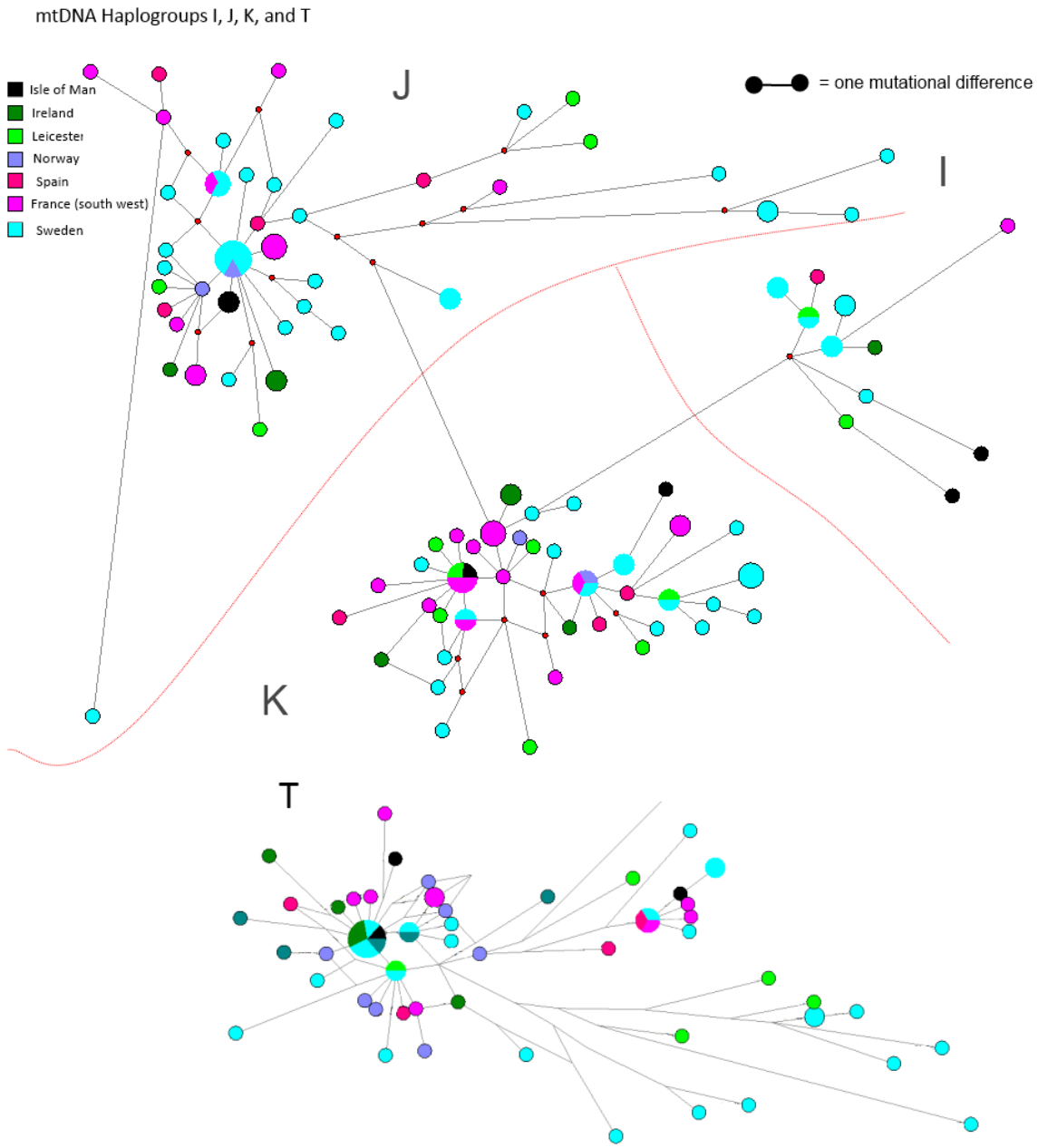



*Figure 5.15: Network diagram showing the genetic relationship between mtDNA sequences of individuals within the I, J, K and T haplogroups. The size of the circle is proportional to the number of individuals carrying identical sequences; the smallest equalling one. The size of the line between haplotypes is proportional to the number of single mutational differences separating the sequences.*

## 5.3 Hypothesis Building

In this section some the patterns revealed by the descriptive genetics methodologies were used to inform scenarios which were tested against one another using the computer modelling method, approximate Bayesian computation (ABC). NRY and mtDNA were tested separately using data from four populations from around the Irish Sea and Scandinavia, to look for demographic events specific to the Viking Age Norwegian diaspora.

### 5.3.1 NRY

In each set of simulations only three populations were used; one set of seven scenarios were modelled with Ireland, Isle of Man and Norway, and another set used Wales, Isle of Man and Norway. In the set of scenarios which included Ireland, the scenario with the highest posterior probability, was where the Isle of Man was the admixed population (0.9714; 95% CI: 0.9475 – 0.9952; see Fig. 5.16). The results indicate that Ireland's population was comprised of approximately 30% Irish and 70% Norwegian (see Fig. 5.18). However, the confidence in this model was only 30% indicating that the ABC process could not distinguish between scenarios with a high degree of certainty (Appendix B16 for type I and type II errors).

When the same scenario's were modelled with Wales instead of Ireland, two scenarios had equal posterior probabilities (Fig. 5.17). One showed an ancient split of Norway from Wales and a more recent one of Isle of Man from Wales (0.5257; 95% CI: 0.4165 – 0.6349). The other had the Isle of Man as the admixed population (0.4743; 95% CI: 0.3651 – 0.5834), with a rate of approximately 60% Welsh to 40% Norwegians (see Fig. 5.18). However, the ABC was more confident in the first scenario (80%) than in the second (47%; Appendix B16 for errors).

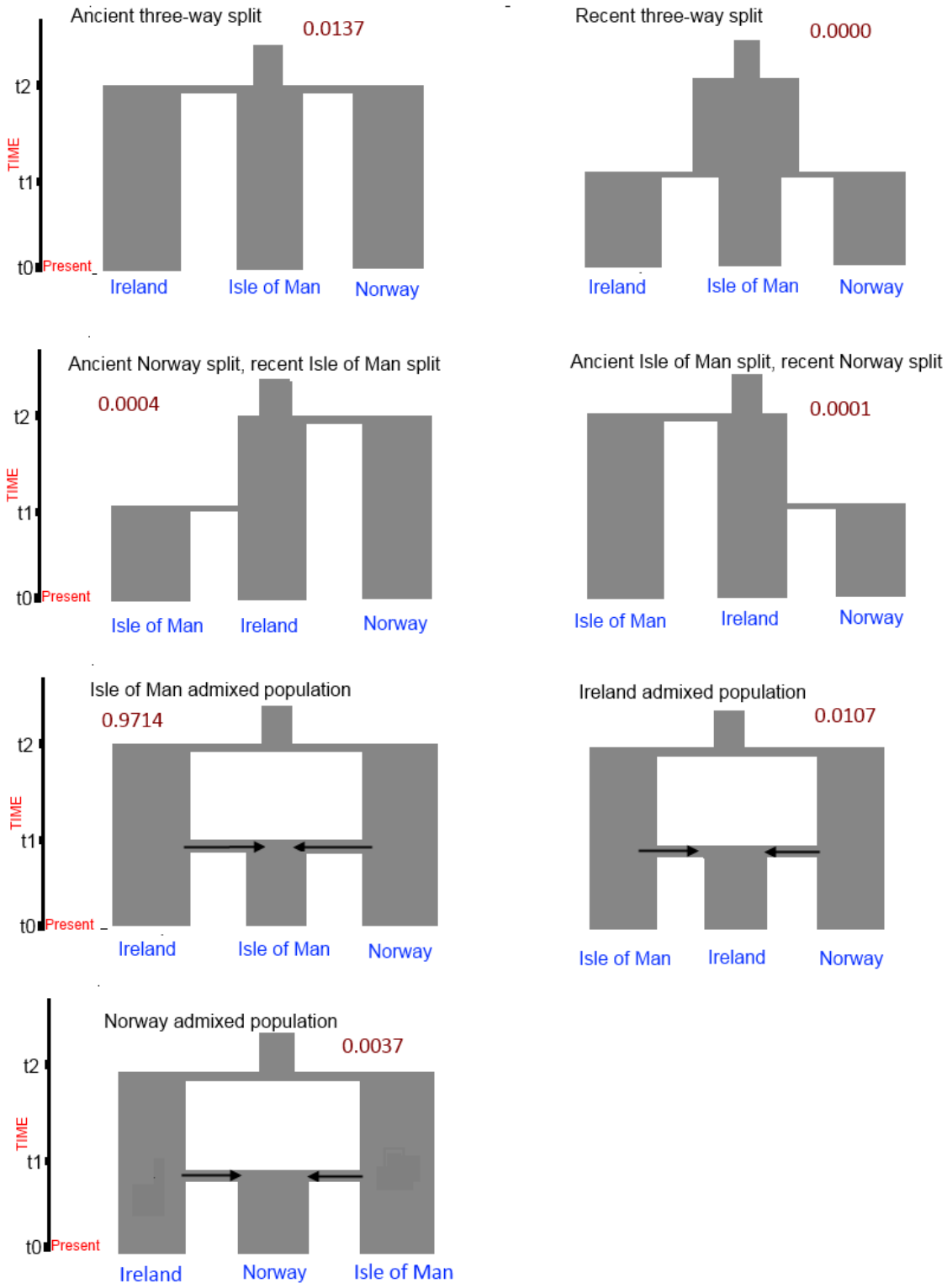



*Figure 5.16: NRY scenarios simulated with Ireland, Isle of Man and Norway. The numbers in brown represent the posterior probability of each scenario.*

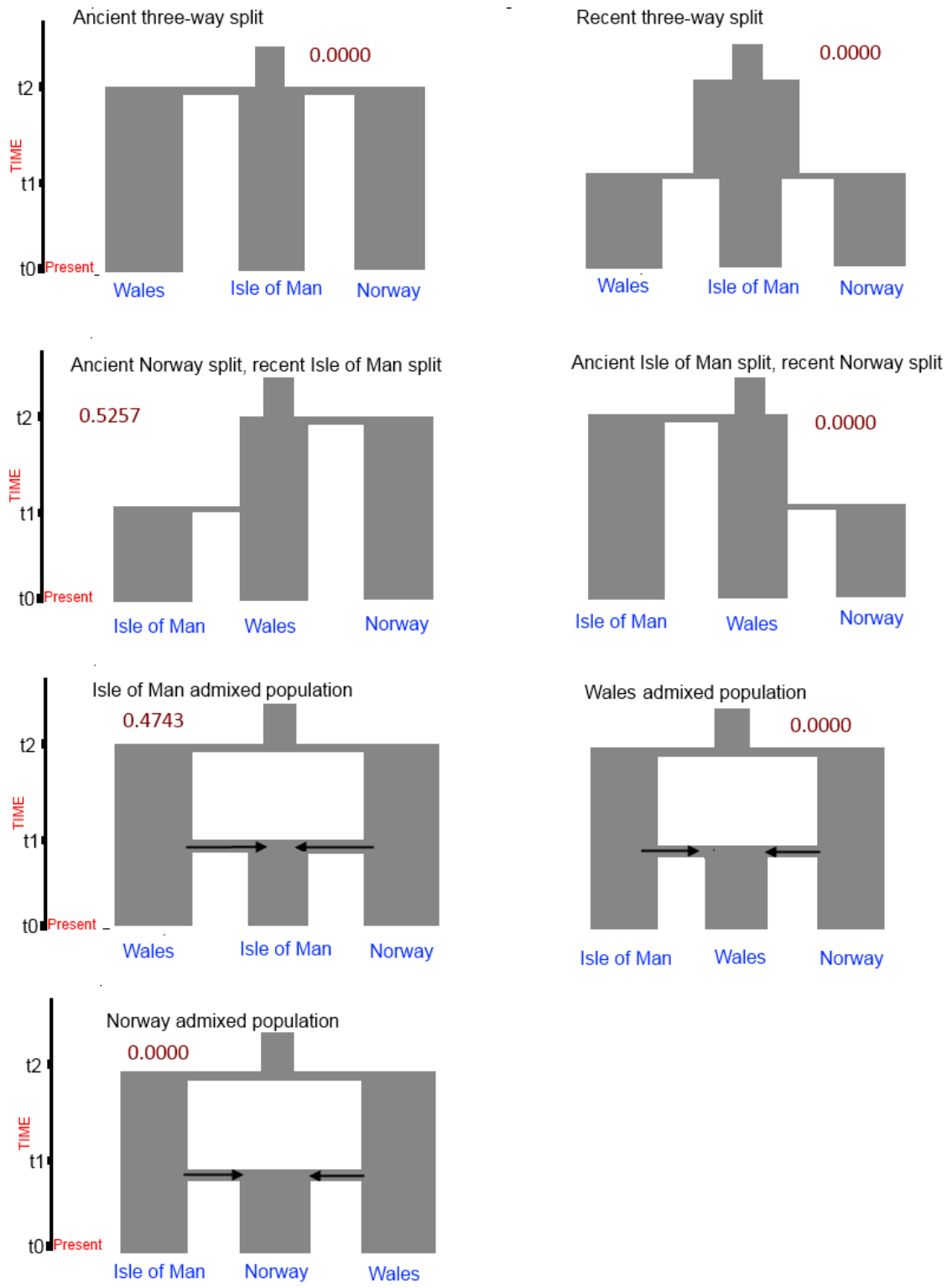



*Figure 5.17: NRY scenarios simulated with Wales, Isle of Man and Norway. The numbers in brown represent the posterior probability of each scenario.*

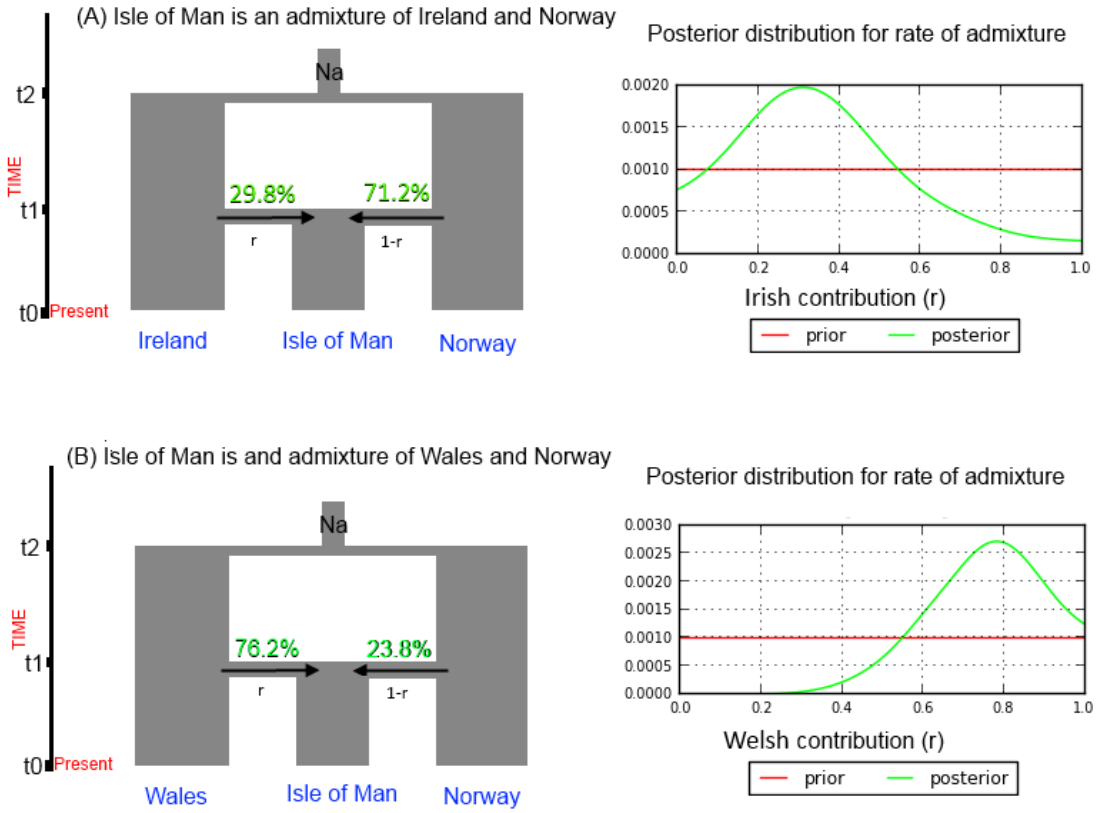



*Figure 5.18: Estimated admixture rates of parental population contribution to the Isle of Man, for two sets of simulations; (A) simulations with Ireland, Isle of Man and Norway, (B) simulations with Wales instead of Ireland. The contribution of Wales or Ireland is represented by r, and the contribution of Norway is 1-r. The graphs to the right show the posterior probability distribution for admixture rates.*

The estimates for the rate of admixture, seemed to indicate a larger contribution of Welsh lineages to the Isle of Man, than of Irish. The simulations created using Wales, also seemed to form a better fit with the observed data than those produced by incorporating Ireland (Appendix B17). However, the scenario which produced data most like the observed, with the most confidence associated with it, was the one where there was no admixture, and the Isle of Man had recently diverged from Wales.

*Table 5.1: Showing parameter estimates for the scenario where Isle of Man is the admixed population and Ireland and Norway are the parental populations (Posterior probability of scenario – 0.9714).*

| | Description | Prior | Mode | 95% Credible Intervals |
|---|---|---|---|---|
| **t1** | Time of admixture (generations) | 20 – 64 | 29.5 | 21– 62.2 |
| **t2** | Time of divergence (generations) | 100 – 355 | 100 | 100 – 173 |
| **N1** | Irish effective population size | 5000 – 1,000,000 | 957,000 | 49,100 – 985,000 |
| **N2** | Manx effective population size | 10 – 50,000 | 9,100 | 2690 – 47,900 |
| **N3** | Norwegian effective population size | 10 – 100,000 | 1,450 | 736 – 2,550 |
| **Na** | Ancestral effective population size | 10 – 5000 | 38 | 13 – 238 |
| **r** | Irish contribution | 0.001 – 0.999 | 0.298 | 0.0559 – 0.822 |

*Table 5.2: Showing estimates for parameters generated from scenario simulated where (A) Norway and Wales diverged first and the Isle of Man and Wales diverged later (Posterior probability – 0.5257). (B) Isle of Man is the admixed population and Wales and Norway are the parental populations (Posterior probability – 0.4743).*

| **(A)** | **Description** | **Prior** | **Mode** | **95% Credible Intervals** |
|---|---|---|---|---|
| **t1** | Time since Wales/Isle of Man diverged | 20 – 64 (gens) | 22.2 | 20.4 – 58.9 |
| **t2** | Time since Wales/Norway diverged | 100 – 355 (gens) | 100 | 100 – 130 |
| **N1** | Welsh effective population size | 10 – 10,000 | 515 | 277 – 1180 |
| **N2** | Manx effective population size | 10 – 500,000 | 47,500 | 1,590 – 48,500 |
| **N3** | Norwegian effective population size | 10 – 1,000,000 | 30,100 | 17,700 – 954,000 |
| **Na** | Ancestral effective population size | 10 – 5000 | 189 | 32 – 1,310 |
| **(B)** | **Description** | **Prior** | **Mode** | **95% Credible Intervals** |
| **t1** | Time of admixture (generations) | 20 – 64 | 22.6 | 20.4 – 61.4 |
| **t2** | Time of divergence (generations) | 100 – 355 | 105 | 101 – 152 |
| **N1** | Welsh effective population size | 10 – 10,000 | 590 | 279 – 1,300 |
| **N2** | Manx effective population size | 10 – 500,000 | 105,000 | 13,700 – 486,000 |
| **N3** | Norwegian effective population size | 10 – 1,000,000 | 3360 | 2,210 – 810,000 |
| **Na** | Ancestral effective population size | 10 – 5000 | 159 | 28 – 1,020 |
| **r** | Welsh contribution | 0.001 – 0.999 | 0.762 | 0.461 – 0.966 |

Estimates for the effective population sizes of Norway, the Isle of Man and Ireland, could not be calculated with accuracy from this genetic data, as reflected

in the wide 95% credible intervals (Table 5.1 and 5.2). The effective population size of Norway varied depending on the demographic scenario. For the admixture scenario where Norway and Ireland were the parentals, the effective population size of Noway was significantly smaller than both Ireland and the Isle of Man. Perhaps this was a reflection of the size of the Norwegian founder population in the Irish Sea region.

Estimates for the effective population size of Wales were much narrower and it was consistently and significantly the smallest population. The ancestral populations from which all four of these populations diverged was also consistently very small across all selected scenarios, with no more than 1310 men.

The estimates for the timing of demographic events also seemed to be in agreement across all the selected scenarios. However, according to these estimates, both admixture and divergence events happened much more recently than would perhaps be expected based on archaeological or historical sources. Admixture was calculated to have occurred within the last 700 years, based on a 31 year per male generation time (Fenner 2005). Likewise according to this scenario the divergence of Norway and Wales or Ireland happened within the last 4000 years. However, using STRs is known to produce underestimates for the timing of demographic events, because back mutations reduce the number of visible changes, which leads to errors when using the molecular clock (see Section 3.2.2; Wei et al. 2013).

Interestingly, of the summary statistics generated from the simulations (Appendix B18), delta mu squared produced the worst fit with the observed data. The majority of simulations producing smaller values for this summary statistic than seen for the observed data. This indicates the simulations produced

populations that were more closely related than the observed populations, perhaps confirming that earlier divergence times are needed than are estimated by these scenarios.

### 5.3.2 mtDNA scenarios

Like NRY, two sets of scenarios were simulated using three populations. For the first set, the populations Ireland, Isle of Man and Norway were used. However, for the second set, Sweden was substituted for Norway as no suitable British mtDNA data-set was available.

For the first set, none of the scenarios were chosen outright. The four scenarios with the highest posterior probabilities which were not significantly different from each other. These were the two with different divergence timings from Ireland, and two admixture scenarios, one where Ireland was the admixed population and the other where Norway was the admixed population (see Fig. 5.19). The estimated admixture rates for Ireland were approximately 55% Manx to 45% Norwegian, and 55.4% Irish into Norway (see Fig. 5.20).

The posterior probabilities for these scenarios were not significantly different from each other (95% CI ranging between: 0.1609 – 0.2085 and 0.1196 – 0.1641) but they were higher than for the remaining scenarios. Confidences in scenario choice were not very high; ranging from 20% to 41%, the highest being for the scenario where Isle of Man diverged from Ireland more recently (see Appendix B21 for errors). However, all four of these scenarios produced a good fit with the observed data, but the two without admixture were perhaps slightly better (Appendix B20).

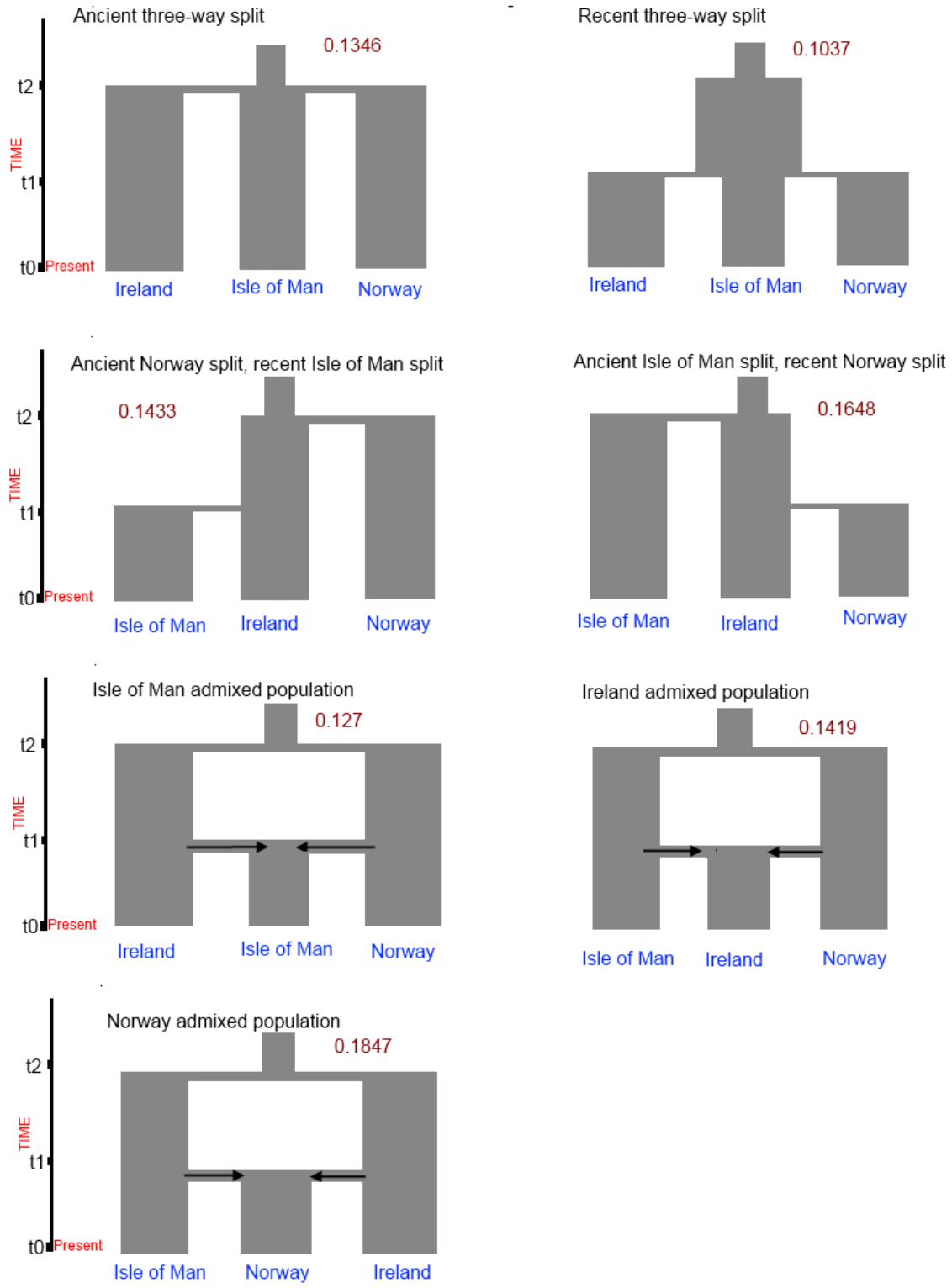



*Figure 5.19: mtDNA scenarios simulated with Ireland, Isle of Man and Norway. The numbers in brown represent the posterior probability of each scenario.*

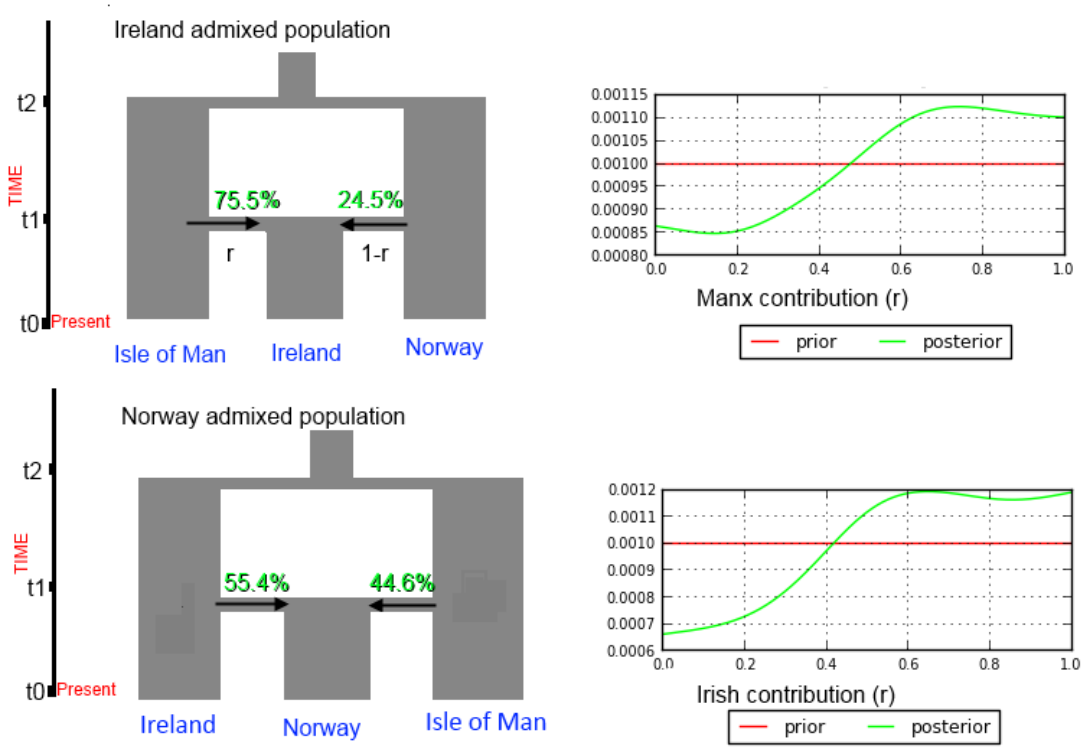



*Figure 5.20: Estimated admixture rates for two admixture scenarios with the highest posterior probabilities from the set of simulations using Ireland, Isle of Man and Norway. The contribution of the first parental population to the daughter population is represented by r, so the contribution of the second parental population is 1-r. The graphs to the right show the posterior probability distribution for admixture rates.*

Parameters estimated from all of these scenarios (Appendix B22), agreed that Norway had a much bigger effective population size than either Ireland or Isle of Man, with peak estimates of between 1,200,000 and 6,320,000. Ireland tended towards a bigger model estimate than the Isle of Man, but they were not significantly different for any of the scenarios.

The range for all effective population sizes was very wide for all populations in all scenarios, indicating the difficulty in making these estimations from the available genetic data. However, like the NRY, the ancestral population was indicated to be very small, with modal values of between 121 and 413. Again these were not significantly different between scenarios.

It was difficult to interpret the timings for demographic events, because the patterns of events in each scenario contradict each other particularly in the scenarios with divergence without admixture, but despite this, the timings for the first and second events are not significantly different between any of the scenarios. They also have wide 95% credible intervals, indicating the difficulty that ABC had in recovering dates from the data. The timings of the divergences were also more recent that would be expected, given data from archaeological and historical sources, possible partly due to the occurrence of back mutations in the HVS of mtDNA, which leads to underestimates of the number of mutations, in turn leading to under-estimates of time-frames.

The simulations with Sweden produced much clearer results. The highest posterior probability with 0.6192 (95% CI: 0.5631 – 0.6752), was for the scenario with an ancient divergence of Ireland and Sweden and a recent divergence of Ireland and Isle of Man (Fig. 5.21). However, there was only 38.4% confidence in this scenario (Appendix B20), and the fit between the simulations and the observed data, was not as good as for any of the scenarios which had included Norway (Appendix B21).

Interestingly, two admixture models were not completely ruled out, the ones where the Isle of Man or Ireland were the admixed populations. Both these scenarios had approximately 16-19% posterior probability, perhaps indicating some Swedish genetic influence on the Irish Sea gene-pool.

Estimates for the effective population size for the Isle of Man did not differ significantly between NRY and mtDNA. In populations that practice patrilocality, female mobility rates are generally higher, which leads to higher effective population sizes. This is not the case for the Isle of Man perhaps indicating that Manx men have been much more mobile than women in recent times.

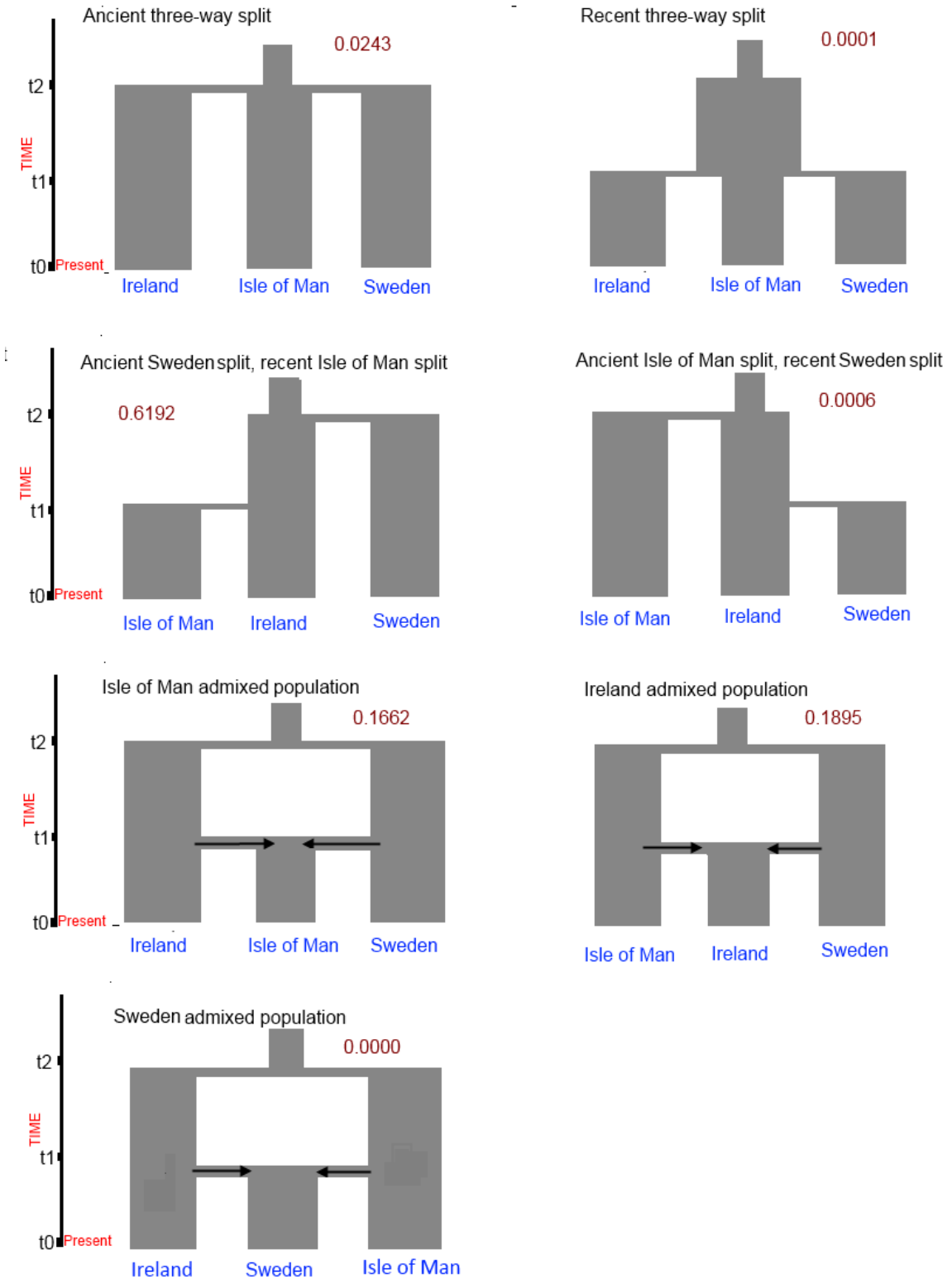



*Figure 5.21: mtDNA scenarios simulated with Ireland, Isle of Man and Sweden. The numbers in brown represent the posterior probability of each scenario.*

*Table 5.2: Showing estimates for parameters generated from scenario simulated where Sweden and Ireland diverged first and the Isle of Man and Ireland diverged later.*

| | Description | Prior | Mode | 95% Credible Intervals |
|---|---|---|---|---|
| **t1** | Time since Ireland/Isle of Man divergence | 28 – 159 (gens) | 28 | 28.8 – 150 |
| **t2** | Time since Ireland/Sweden divergence | 160 – 1,600 (gens) | 834 | 257 – 1,560 |
| **N1** | Irish effective population size | 10 – 100,000 | 889 | 309 – 35,500 |
| **N2** | Manx effective population size | 10 – 50,000 | 31,100 | 1,550 – 48,700 |
| **N3** | Swedish effective population size | 10 – 10,000,000 | 5,690 | 813 - 6,300,000 |
| **Na** | Ancestral effective population size | 10 – 50,000 | 5,130 | 1,930 – 30,000 |

The timing of demographic events also had large credible interval ranges again indicating the unsuitability of these data, for making accurate dating assessments. Nevertheless, the estimate for the timing of the divergence of Sweden and Ireland, was 20,850. Given that it is expected that this is an under-estimation of the actual timing, this could indicate that populations diverged before the last Ice-Age. If the populations are descended from different palaeolithic populations, this could perhaps explain the larger ancestral effective population size, than seen for the simulations which included Norway.

# CHAPTER SIX: DISCUSSION AND CONCLUSIONS

## 6.1 Introduction

This last chapter contains a discussion of what the results mean particularly in relation to what genetic affinities with Scandinavia can tell us about the Viking Diaspora in the Irish Sea region. This section includes a description of where the Isle of Man fits into the wider genetic landscape of the British and Irish Isles and Atlantic Europe, and what this potentially implies about wider connections between the Irish Sea and beyond. I will also consider how genetic analysis of modern data can be better integrated into the archaeological tool kit, particularly for exploring demographic histories of regional contexts.

The conclusions summarise the most important findings of this work, for both the demographic history of the Isle of Man and for the use of genetic data for the exploration of the human past. Finally, the last section will look at potential future projects which could expand or improve on this research, both for the Isle of Man, and for other populations in the Irish Sea region, but also how this type of regional study could be applied elsewhere.

## 6.2 Discussion

This discussion is formed of four parts. The first part discusses the placing of the Isle of Man within the wider genetic landscape of the British and Irish Isles and Atlantic Europe, highlighting possible implications for its ancient and more recent demographic history. The second section discusses the results of computer modelling and what can be inferred about the demographic history of

the Isle of Man and its nearest neighbours. The third section focuses on evidence for Scandinavian contacts with the Isle of Man and how this can inform our thinking about the Viking migrations. The final part discusses the use of modern genetic data and how the Isle of Man case-study provides insights into how such data can be better integrated into archaeological theories about our understanding of the past.

### 6.2.1 The genetic landscape of Atlantic Europe

This research has shown that the Irish Sea populations have some of the lowest male and female diversity values in the whole of the Atlantic region. Diversity is of a similar level to that of the better studied Basque Country and other Northern Iberian populations. This lack of diversity is possibly due to its location at the very edge of Europe and is probably, at least in part, linked to the general lack of diversity found in the biosphere of the region.

During the Mesolithic, Ireland and its nearest neighbours were amongst the last regions in Atlantic Europe to be colonised. It is likely that the founding populations had already experienced several founder events along their way to the region over the several generations it would have taken to get from the putative ancestral origin in the Iberian glacial refugia. These initial colonists would have most likely arrived via France into the British and Irish Isles.

Colonists may have moved along the coast up into Wales, to arrive in the now-submerged area, containing what would later become the Isle of Man, and then possibly, on into Ireland. Alternatively, the Isle of Man may have been colonised by people moving north through England, via Cumbria or Merseyside. It is difficult to ascertain the routes the initial colonist took from archaeological evidence, because any coastal Mesolithic sites in the Irish Sea region are now

under water (see Fig. 3.4).

Likewise, there is an absence of aDNA data from the Mesolithic period of this region to provide a temporal baseline for the general patterns and affinities seen in the Irish Sea populations. Furthermore, if immigration from the south had occurred during the subsequent periods following the initial colonisation, the archaeological evidence indicates that it is mostly likely to have been from the same related populations as colonised the Irish Sea in the first place, via similar routes (Sheridan 2003).

Some aspects of the genetic data seem to support older connections between the Irish Sea and southern populations, in both male and female lineages. For males, there is more genetic affinity between these populations and France and Iberia than with England and Scandinavia (Fig. 5.5 and 5.6). These genetic similarities could indicate either that these populations have the same ancestral root or that there were periods of prolonged or extreme gene-flow between these regions. Alternatively, these patterns could be the result of a combination of all these factors.

The high frequency of the male NRY R1b-S116 haplogroup in the Isle of Man which is also found at high frequencies in France, also seems to indicate contacts between these regions. However, the Isle of Man is unique in the British and Irish Isles in having high frequencies of this haplogroup. Further, this haplogroup is also found in relatively high frequencies in parts of Scandinavia (Fig. 5.4), so while it is possible that this is evidence of ancient connections between France and the Isle of Man, R1b-S116 it is more likely to be explained by Norwegian immigration.

The exceptionally high frequency of NRY R1b-S116 compared to the rest of the British and Irish Isles, may be caused by the action of genetic drift which can

increase the frequency of rare haplogroups in small island populations even when they have disappeared from the much larger populations of their neighbours. The impact of the Viking Diaspora on this region involved the influx of people potentially from different glacial refugia with a different mix of incomers in the intervening period. This period had the potential for a much greater impact on the gene-pools of the Irish Sea region by introducing a completely new mix of haplotypes into the region so dramatically increasing the populations diversity.

It is also worth noting that the network of R1b-S116 daughter lineage R1b-S145, also found in high frequency in the Isle of Man showed a distinctive star shaped pattern indicative of a population expansion. Although no date has been attributed to this expansion yet, it is thought to be relatively recent, occurring some time since the Neolithic. It could be a signal of an expansion in Atlantic Europe triggered by farming technologies, or possibly some later dramatic demographic event. It could have spread to the Isle of Man at the time of this expansion, or at some later point via the south or from the north. As yet it is is not possible to say (Batini Pers. Comm).

For the female lineages the patterns are less dramatic; the Irish mtDNA sample clustered slightly more closely with French and Spanish populations than most of the British or Scandinavian populations (Fig. 5.14). Also the Manx mtDNA sample shared the highest proportion of identical mtDNA sequences with La Rioja in northern Spain (19%; Appendix B14). Again it is unclear when these patterns emerged, and whether they are indications of old contacts between regions, more recent dispersal through patrilocality or perhaps even the long-distance movement of women through the Viking slave trade.

It is also important to remember that male-related STR and female-related HVS

haplotypes, are fast evolving regions of the genome, which means that similarities between haplotypes could arise over time by convergence rather than because they share a common ancestor. Higher resolution data, utilising slower evolving regions of the mtDNA genome, alongside SNP-STR combination haplotypes would give a more reliable indication of how closely-related these populations really are.

The simulations were not used to explore these southern contacts and affinities directly. Instead any impacts that these may have had on the Irish Sea region were treated as part of the general composition of the Welsh and Irish populations. The subtle distinctions between the Irish Sea populations, which may have developed through different outside contacts and separate demographic histories, were exploited to explore affinities of these neighbouring populations with the Isle of Man.

### 6.2.2 Demographic History of Irish Sea populations

Modelling seemed to indicate that the modern Manx NRY population contains Welsh, Irish and Norwegian components (Section 5.3.1). The location of the Isle of Man means that there is reasonable expectation that gene-flow has also occurred between the island and Cumbria, the Wirral, and Scotland at points in its history. These populations may even have contained the original sources for the Manx populations. It is important to note here, that in this context, Wales, Norway and Ireland act as proxies for all the populations which may have contributed to the Manx population. In other words, the Welsh population in the context of these simulations, represents the British mainland, as no other suitable British population was available.

Another study (Hellenthal et al. in press) has noted the genetic similarities

between northern parts of Ireland and western Scotland. So, it is not clear how much of the Irish or the Welsh contribution actually represents Scotland. As both Cumbria and Scotland are known to have experienced admixture during Viking and Irish diasporas respectively, it seems likely that the inland Welsh population may provide a good proxy for the ancient western British populations, with minimal admixture in this instance. However, since Wales and Ireland were modelled separately, in each case they represented all British and Irish contributions to the Manx gene-pool.

Interestingly, the results indicated a larger Welsh than Irish contribution to the Isle of Man. Also, this contribution is estimated to have occurred more recently (c.1300CE; 95% CI: 1391 – 100CE; Table 5.2B) than the Irish admixture (c.1100CE; 95% CI: 1350 – 100CE; Table 5.1). The timings estimated using NRY STRs are expected to underestimate the amount of time since an event, making these dates too recent. This, alongside the wide credible intervals, shows that these dates are not particularly reliable indications of when these contacts occurred. While it is safe to surmise that contacts with Britain have had a greater impact on the modern Manx NRY population than any migrations from Ireland, this does not allow for any inference about the composition of the population prior to these influences.

Less can be inferred about the original source of the female Manx population from the ABC modelling, as no suitable neighbouring British populations were available. For the female simulations, Ireland was used to represent all nearby populations. In the models, Ireland and the Isle of Man were estimated to have diverged around 700 years ago (c. 1300CE; 95% CI: 1300CE – 1750BCE). Again, mtDNA HVSs are expected to underestimate the timing since events, so this, and the wide credible intervals makes this estimate unhelpful. These and other

ambiguities in the results may be an indication that Ireland was not a suitable source population for Manx females.

Conversely, the results indicate the Isle of Man as a potential source of females for Ireland, via migrations some 1800 years ago (c. 200CE; 95% CI: 1250CE – 15CE; Appendix B22). Again, these estimates are not reliable indicators for when Manx women were moving into Ireland. It is also not clear if the Isle of Man was the actual source or whether in this context, the Manx females represent another British population not sampled here. Wales or Scotland are known to have been in close contact with Irish populations in the past, particularly during the fifth century Irish diaspora but also perhaps during the Bronze or Iron Ages (Section 3.2.3). It may be these populations, rather than the Manx, which supplied Ireland with a source of brides or female slaves, which is being picked up by the modelling.

Estimates for the ancestral effective population sizes for both men and women in this study indicate that populations in this region were very small in the past, and possibly for much of the region's past. The effective size of a population is thought to be about a quarter to a sixth of the census population size, although the actual relationship between census size and population size is still little understood. The effective population size however, is known to have an impact on the amount of drift experienced by a population (Section 2.2.2). So, small population sizes may at least partially explain the general lack of diversity seen in some of the Irish Sea populations (see above).

In particular, Ireland and the Isle of Man are likely to have had smaller carrying capacities compared to Britain and Europe as they were both cut off from the mainland before all of the flora and fauna seen in Britain were able to colonise the regions. Equally it might be expected that Britain would have a lower

carrying capacity than Europe for the same reason. There is archaeological evidence that food mammals had to be introduced to Ireland, even before the knowledge and technologies of the Neolithic Age made the management of animals common practice in the British and Irish Isles (Bradley 2007, 8; Mallory 2013, 42-46).

Long-term isolation is another factor which impacts on effective population size. It seems likely that the Irish Sea populations would have remained relatively isolated until boats became a reliable form of transport for everyone, rather than just the desperate or intrepid. The climate in these regions is much harsher than in southern regions, and wetter than the more sheltered eastern parts of Britain. This makes for more extreme winters which may have impacted on mortality rates and long-term population sizes in the past. The unpredictable weather and occasionally dangerous seas probably would have required specialised knowledge and skills for survival which may have also discouraged immigrants in any large quantity.

The Isle of Man is the exception to the low diversity generally seen in the region, as it has relatively high diversity of both female and male markers for the region (see Figures 5.2 and 5.11). This is unexpected because both its relatively small population size and island-status mean that drift would have been strongly acting on the population, removing diversity much quicker than larger neighbouring populations. This provides some evidence that the island must have experienced considerable gene-flow to both male and female gene-pools, recently enough and of a significant size, for genetic drift to not have had time to remove.

### 6.2.3 The Scandinavian Diaspora and the Isle of Man

Unfortunately due to the imprecise nature of estimating times since events using the available genetic data, it was not possible to reliably attribute any of the Scandinavian influences found in the Isle of Man directly to the Viking period. Dates for male Norwegian genetic input were estimated to between 1300 and 1100 with wide credible intervals (Tables 5.1 and 5.2B; admixture with Wales c.1300CE; 95% CI: 1391 – 100CE; admixture with Ireland c.1100CE; 95% CI: 1350 – 100CE). Also these are expected to be underestimates, placing the actual data at some unspecified earlier time. While these dates do correspond to the end of the Norwegian occupation in the region and therefore could indicate Viking Age immigration to the island, it does not indicate it with any certainty.

Despite this uncertainty, it is clear there was significant mixing of both Manx males and females with Norwegians. For males there is clear evidence of Norwegian genetic input in the Manx populations. Although the genetic data precluded an exact estimate of the amount of Norwegian admixture, it seems that there were more Norwegian migrants than Irish, but less Norwegian than British (represented by the Welsh sample, see above; Fig. 5.18). The modelling seemed to indicate that the Manx NRY gene-pool looks more like the Welsh (or British) than the Irish gene-pool giving a 24% Norwegian contribution to the population (95% CI: 46.1% - 96.6% Welsh contribution).

This is the estimated proportion of Norwegians arriving on the island at the time and is much lower than 31 – 58% previously described using other methods to calculate admixture. However, these authors used methods which calculated the proportion of Norwegian lineages in the modern Manx population, without taking into account how drift may have impacted on the relative frequencies in the intervening years. Whereas in this project, the modelling estimated the

proportion of Norwegian linages entering the Manx gene-pool at the time.

This discrepancy between proportion of Norwegian lineages in the gene-pool now, compared to at the time of admixture, perhaps indicates that Norwegian lineages in the Manx population have increased in the intervening period through chance (ie. drift). It is possible that men of Norwegian ancestry had social advantages which made them more likely to pass on their haplotype to the next generation, thereby increasing the frequency of Norwegian lineages relative to local lineages over time. However, the difference could simply be due to the different methodologies used.

Recent aDNA research in Europe (reviewed in Pinhasi et al. 2012 and Brandt et al. in press) indicates that the complete turnover of NRY gene-pools may have been more common and have taken much less time than previously imagined. The results seems to support this, demonstrating that migrations of sufficient size or social impact, can also shape the modern gene-pool. It may be because drift has not had sufficient time to remove all traces of the immigrants but also in some cases, drift and social factors may act to increase the incoming lineages obscuring older patterns.

The mtDNA of the populations explored seems to be much less clearly defined than the NRY. When scenarios using Norway, Ireland and the Isle of Man were modelled, ABC was unable to choose which scenario fitted the data best, indicating that the mtDNA gene-pool of the region is very homogeneous. The fact that ABC had little problem in choosing a scenario when Sweden was substituted for Norway, suggests that this mixing is indicative of the movement of women during the Norwegian occupation of the area, and that women were moving between all three populations significantly enough for no clear and discernible pattern to have formed.

The low but significant probability that Sweden may have contributed to Ireland and Manx gene-pools however, may be an indicator of the general background level of female mobility between the Scandinavian populations, which is then showing up in the Irish Sea gene-pool. Interestingly, the scenarios with the low but significant probabilities, were the admixture scenarios where either Ireland or the Isle of Man was the admixed population (Fig. 5.21). This provides further evidence that these populations were difficult to distinguish due to the substantial mixing of females. However, there was no support for Irish Sea women ever reaching Sweden.

This difference in pattern seen between male and female genetics perhaps suggests that groups of, possibly related, men migrated the longer distances between Norway and the Irish Sea, but when they got there seemed to stay put. Whereas the movement of women around the Irish Sea and up into Norway was much more prevalent and diffuse, so related women would end up scattered throughout the region.

This homogenisation of female gene-pools is generally considered to be the product of patrilocality, which is practised by most populations (~70%). In these populations female mobility rates are generally higher, which, in turn leads to higher effective population sizes (see Section 2.2.3; Wilkins and Marlowe 2006). So women marry into different communities and therefore are much more mobile in general. This is often reflected in higher effective population sizes for women than men.

However, estimates for the effective population size for the Isle of Man did not differ significantly between NRY and mtDNA. Norwegian and Irish effective population size was significantly higher for men than for females. These results are interesting because they suggest that something else is happening. Perhaps

these are a result of men being much more mobile than women during the Early Medieval period. In Ireland, the difference between the effective population sizes is more dramatic than in the Isle of Man, possibly as a result of longer periods of Irish male mobility over greater distances.

During the Irish diaspora and into the Viking period, if the men who migrated were mainly marrying local women they would effectively have been practising matrilocality. It is not clear whether the Irish and Norwegian tradition of fosterage could have had an impact on relative population sizes. The term 'fostering' implies that they would have eventually returned to their birth-place. However, as it was apparently common and widespread, before, during and after the Viking period, the implication for the gene-pool should be examined.

Whatever the reason for this unusual pattern seen in the ratio of male to female effective population sizes, it is seen to a lesser degree in the Manx population, than in Norwegian and Irish populations. This indicates that the Manx may have had different traditions from those of the Irish and Norwegians. As the Manx population seems to be largely descended from British populations, perhaps the differences seen between relative population sizes is a reflection differing cultural influence. No comparative female Welsh data was available to explore this possibility.

However, there also remains the possibility that these patterns are an artefact of the genetic data used, or perhaps of the types of scenarios modelled. It would be interesting to see if these patterns emerge from whole mtDNA genome data when compared with high resolution NRY SNP data, for exploring the same and different scenarios.

### 6.2.4 Genetic data for exploring the past

Assessing the populations history of such a small geographic region using uni-parental markers was never going to be easy. The main problem is that the populations involved are likely to be very closely related and possibly have experienced extensive mixing throughout time. Another problem is the type of genetic data which was used here; the hyper-variable regions of mtDNA and STRs of NRY, are both highly mutable. On one hand this means they acquire diversity quickly, but on the other hand the problem of back mutation can lead to blurring of the patterns of divergence as well as underestimate of timing since events occurred.

Tests using ABC also indicated that the HVS of mtDNA and STRs of NRY are not useful for estimating effective population sizes, particularly of the larger populations. However, it was able to provide evidence of relative populations sizes and provide a way to compare male and female effective population sizes. However, although there was some difficulty in estimating the exact parameters of the scenarios, or in the case of mtDNA establishing which scenario explains the patterns of diversity in the region, modelling still provided some clues about the relationships between populations in this region as well as highlighting differences in male and female population histories.

The wider aim of this thesis was to bring the fields of archaeology and genetics closer together, firstly by integrating archaeological data into genetics models in an attempt to illuminate patterns found in gene-pools, and secondly by trying to provide a way for archaeologists to evaluate the quality of hypotheses built upon genetic data. Historically there has been a cultural divide between the fields borne out of different approaches towards the treatment and assessment of data. Both fields have been through a period of introspection, where crossing

disciplinary boundaries has been perhaps both undesirable and impractical. However, I think ultimately both fields have been striving to understand what can be objectively known about the past given the limitations of the methodologies and data available. The main advantage of using ABC in this setting is that it provides a way of assessing the confidence in predictions made using the genetic data. By using ABC it was possible to quantify the level of certainty in the results and the interpretation providing a measure of the limits of the data.

Human population geneticists and archaeologists are both interested in the past although perhaps in the beginning their interest in the past was formed of different motivations. When the patterns in the modern gene-pool were first discovered geneticists looked to the past and archaeology as a way to help to explain the patterns that they saw. Archaeologists, on the other hand, saw that the patterns in the gene-pool, in a similar way to patterns in material culture, could be used to uncover how people in the past behaved and interacted. This led to circularity and to disappointment and frustration, as both sides began to realise that the other did not hold obvious answers to the questions that they wished to answer. Genetic patterns like patterns in material culture are subject to interpretation and bias, and archaeological theories about the past are fluid and much debated.

The important developments over the last two or three decades in genetics have been a deeper understanding of the precise way that processes have acted together to shape the gene-pool. Through advances in molecular and computational technology. Existing theories about how natural selection, mutation, drift, migration and recombination interact to shape genomes and gene-pools can finally be tested against real data. This has led to modifications

of models used to measure these processes. Archaeological data provides another set of parameters to test genetic data against. Ancient DNA provides a temporal aspect which has been missing from genetic research of the human past.

Historically, the infancy of genetic methodologies, a lack of high resolution data from sufficient populations has caused difficulty when interpreting genetic patterns. However, even in the time since beginning this project, developments in genetics have made much of this work obsolete. It is far more possible now to look at a wide range of NRY SNPs as well as whole mtDNA genomes. Using these data-sets for modelling has the potential to increase the accuracy of parameter estimates. Adding aDNA would also help provide temporal resolution to make the dating of demographic events more accurate.

However, it is probably still beyond the budget of the average archaeological research unit to obtain this type of genetic material, whereas the type of genetic data used in this project is abundant and comparatively cheap. Also, I hope this research has shown that it can provide a new perspective on the human past, particularly in relation to different demographic histories of men and women. Data-sets collected from small geographically-defined areas can be compared, within a modelling framework, to the genetic data already available in the literature, including aDNA as and when it becomes available. This can then provide fresh insights and new theories to explore, about the human past of specific regions.

## 6.3 Conclusions

The wider aim of this thesis was to try to better integrate genetic data into an

archaeological framework. When geneticists first started to show an interest in the past I think they were more interested in what it could say about the patterns seen in modern populations, to provide clues about the underlying processes which have interacted to form the patterns seen in the gene-pool. Whereas, archaeologists are more interested about what patterns could say about the past. Geneticist used the past as a tool to understand the present patterns, whereas archaeologists saw patterns in the archaeological record as a way to understand the past.

With this in mind, I have attempted to use the genetic data to provide another perspective on a particular period of the past. In particular, to see if it was possible to detect the specific early tenth century Viking migrations to the Isle of Man. Unfortunately, the short answer to this question is, no. Although the genetics data indicates a male Norwegian contribution to Isle of Man gene-pool, the data available did not provide sufficient temporal resolution to enable it to be pinned the Viking period. It seems likely that much of the Scandinavia signal, as well as the homogenisation of the mtDNA gene-pool, occurred in the post-Viking period, when people of mixed Norwegian and Celtic heritage roamed the area. To date these patterns would require either aDNA from the region, dating to these periods, or higher resolution data from the more stable regions of the NRY and mtDNA

Another aim of this research was to demonstrate that genetic data can be used to explore and compare different sex-related behaviour, and in this it was much more successful. This research was able to demonstrate considerable differences in male and female gene-pools in the region and to link these with the Norwegian diaspora. There was significant evidence of male Norwegian admixture in the Manx gene-pool, although the estimate for initial input was

much lower than previous estimates for proportion in the modern population. This difference could indicate the higher social standing of those of Norwegian descent in the Irish Sea area, however, more testing would be required to support this.

More compelling is the evidence for the dispersal of women around the Irish Sea, in the form of brides, slaves or both. This mixing was only found between Ireland, the Isle of Man and Norway but not to any significant degree with Sweden. This is interesting because while past studies have recognised the general homogenisation of the mtDNA gene-pools of Europe and attributed it to higher mobility of women through patrilocality, here we seem to see the linkage of a specific community of populations through the distribution of its women.

Overall, this research has demonstrated that genetic data can be used to examine human past on the regional scale. Genetic data is particularly useful for examining demographic factors relating to human populations and for exploring the difference social behaviours of men and women. This work has also, once again, demonstrated the difficulty of using highly variable genetic data to provide a temporal framework for these patterns. However, with time, these problems will be overcome through technological advances improving the resolution of genetic data and increasing the availability of aDNA. The ABC, in the meanwhile provides a useful way of building and testing hypotheses from patterns in the available genetic data.

## 6.4 Future prospects

While the attempts to elucidate the impact of Viking diaspora in the Irish Sea

were only partially successful, this research does open up other potential areas for future exploration. There are clear indications of Norwegian genetic input into the Irish Sea gene-pool, but while this research was unable to fix dates to these patterns, there is potential for further explorations using more precise data, such as high resolution NRY SNPs and genome-wide mtDNA data. This type of genetic data is likely to give more reliable dating estimates, but by far the best way to increase the accuracy would be to include aDNA data in the computer models.

This research has also revealed unusual patterns in the ratios of effective population sizes of men to women. However, more testing is needed to explain these patterns and it cannot necessarily be attributed to the Norwegian diaspora as it does not affect all the populations involved in the same way. Exploration of this phenomenon using the types of genetic data mentioned above, may reveal something more about past Norwegian and Irish Sea societies and provide a way of testing that these patterns are not just an artefact of the data used in this study.

As well as these more recent Scandinavian migrations into the Irish Sea, there are the older migrations and colonisations of the region which have not yet been explored in detail using genetic data. Comprehensive sampling of geographically-defined populations around the Irish Sea and down into France could be used to explore the earlier contacts. Older migrations are expected to have the most impact on a population's genetic composition because they occurred when populations were small. Comparisons between the Irish Sea populations and other British populations may also reveal something of the relationships between the regions of the British and Irish Isles.

It seems that small-scale migrations were perhaps a common feature of human

pre-history as well as more recent history. Women in particular, seem have been exchanged between populations, binding communities together, even across large distances. In the case of diasporas, this may have served to form kinship bonds between incomers and locals, as a means of maintaining links with the source communities, and for forging links between source and host populations. Patterns of mtDNA dispersal and regions of mtDNA homogeneity may provide a useful markers for these ancient community ties.

The mobility of women throughout time still needs to be explored in more depth. A detailed examination of Medieval literature about the use of marriage and fosterage to cement relationships, may provide insights into how this type small-scale mobility functioned in these societies. Both fosterage and marriage could be also be modelled to provide clues about the kind of impact they were likely to have on genetic patterns.

This work would also benefit from analysis of the literature on the anthropology of migrations and diasporas; push and pull factors, which motivate people to move and colonise new region. This would provide a deeper understanding of migrant behaviour, particularly in relation to integration between host and migrant populations, and the forging of links between host and source populations.

# APPENDIX A

## Appendix A1: List of Manx surnames which matched the criteria for inclusion in this study and their spelling variants.

| | | | |
|---|---|---|---|
| Carroon | Cretney | Keig | Kissack |
| Christory | Cringle | Kennaugh | Kneen |
| Cojeen | Crye | Kennish/Kinnish | Quilliam |
| Colquitt | Creer/Freer | Kewin/Keown | Quine |
| Condra | Moughtin | Killip | Callister/Collister |
| Quillin | Mylcraine | Kinley | Clague |
| Callin | Mylroie | Kinrade | Corkill/Corkhill |
| Casement | Quaggin/Quaggan | Looney/Lewney | Corlett |
| Caveen | Watterson/Qualtrough/ Waterson/Qualter | Maddrell | Cubbon/Cubbin |
| Colvin | Quane | Mychreest/MyleChreest | Faragher/Fargher |
| Kaneen | Quiggin | Mylrea | Kneale |
| Keggin/Keggen/Keggan | Quilleash | Skillicorn | Shimmin |
| Kermeen | Skelly | Bridson | Teare |
| Kinvig | Taubman | Cannan | Crellin/Crennell |
| Lowey | Cooil/Coole/Cowle | Karran/Carine/Carran/ Caren | Quaye/Kee |
| Comish/Comaish | Cleator | Clucus | Killey |
| Corkan | Corris/Kerruish | Corrin/Corran | Cregeen |
| Corkish/Quark | Crebbin | Cowin/Cowen | Kaighin/Kaighen |
| Kermode/Cormode | Fayle | Curphey | Gelling |
| Costain/Corteen | Joughin | Gawne | |

# Appendix A2: Press release which was circulated around Manx local media

**To the Editor:**

26th January 2011

Dear Ms Kenny,

**Seeking Viking ancestry on the Isle of Man**

We write to draw your listeners' attention to a new research project on Viking ancestry on the Isle of Man. The arrival of the Vikings in the British Isles around a thousand years ago was a dramatic event that has left a lasting legacy on our language, landscape and place-names. But did the Vikings leave their genes behind as well? And can we disentangle the different influences of the Norse and Danish Vikings?

Our study focuses on the Y chromosome, part of our DNA that is passed down from fathers to sons. I want to recruit male volunteers with local surnames, because these are passed down the generations in the same way as the Y chromosome, and provide us with a stronger link to the past.

Samples will be collected at a one day event being held at the Manx Museum in Douglas on the 19th February 2011 from 11am till 2pm with the kind support of Manx National Heritage. There will also be a short presentation explaining the project held in the lecture theatre at 11.30am. Sampling involves simply brushing the inside of the cheek, and it is free to take part. In return for participating, volunteers will receive a description of their own Y-chromosome type when the work is completed. Men interested in taking part are asked to register online at www.leicestersurnamesproject.org.uk where they can either indicate whether they will attend on the day or if they wish to have a sample pack sent out by post.

Yours sincerely,

Prof Mark A. Jobling
*Wellcome Trust Senior Research Fellow and Professor of Genetics*

Dr Simon James
*Reader in Archaeology*

Ms Hayley Dunn
*Doctoral Research Student*

# Appendix A3: List of volunteers, their birthplace within the Isle of Man, and birthplace of two generations of ancestors (IOM = Isle of Man, Unk = Unknown).

| ID | Surname | Birthplace | Father's birthplace | Mother's birthplace | Birthplace of Father's father | Birthplace of Father's Mother | Birthplace of Mother's father | Birthplace of Mother's father |
|---|---|---|---|---|---|---|---|---|
| IOM 1 | Taubman | Castletown IOM | IOM | IOM | IOM | IOM | IOM | UNK |
| IOM 2 | Vondy | Douglas IOM | Ramsey IOM | Douglas IOM | Ramsey IOM | Ramsey IOM | Tipperary Ireland | Ramsey IOM |
| IOM 3 | Crellin | Ramsey IOM | Leeds | Manchester | Ramsey IOM | Leeds | Manchester | Manchester |
| IOM 4 | Kneale | IOM | IOM | IOM | IOM | England? | IOM | IOM |
| IOM 5 | Kerruish | Douglas IOM | Douglas IOM | South Africa | Ramsey IOM | Bradford | Wirral | Santon IOM |
| IOM 6 | Kinvig | Douglas IOM | IOM | IOM | IOM | IOM | IOM | IOM |
| IOM 7 | Clague | Peel IOM | Peel IOM | Liverpool | Rushen IOM | Malew IOM | Lhergydhoo IOM | Peel IOM |
| IOM 8 | Kewin | Douglas IOM | Douglas IOM | Stockport | Douglas IOM | Unk | Stockport | Unk |
| IOM 9 | Comish | Douglas IOM | IOM | IOM | IOM | IOM | IOM | Suffolk |
| IOM 10 | Bridson | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM |
| IOM 11 | Kaneen | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | IOM | IOM |
| IOM 12 | Clucas | Watford | Douglas IOM | Onchan IOM | Cornaa IOM | UNK | Douglas IOM | Douglas IOM |
| IOM 13 | Craine | Douglas IOM | Ballaugh IOM | Norwich | Ballaugh IOM | Douglas IOM | Norwich | Norwich |
| IOM 14 | Cleator | Douglas IOM | Jurby IOM | Maughold IOM | Jurby IOM | Andreas IOM | Arbory IOM | Lezayre IOM |
| IOM 15 | Creer | IOM | IOM | IOM | IOM | IOM | IOM | IOM |
| IOM 16 | Costain | Widnes | Colby IOM | Ballakilpheric IOM | Colby IOM | Surby IOM | Ballakilpheric IOM | Castletown IOM |
| IOM 17 | Curphey | Hull | Sulby IOM | Hull | Sulby IOM | Sulby IOM | Hull | Hull |
| IOM 18 | Quilleash | Douglas IOM | IOM | IOM | IOM | IOM | IOM | IOM |
| IOM 19 | Kinley | IOM | IOM | IOM | IOM | Manchester | IOM | IOM |
| IOM 20 | Colvin | IOM | IOM | Stratford | IOM | IOM | Stratford | Stratford |
| IOM 21 | Killip | Cardiff | Liverpool | Ormskirk | IOM | IOM | Lancashire | Lancashire |
| IOM 22 | Quaye | IOM | IOM | IOM | IOM | IOM | IOM | IOM |
| IOM 23 | Mylchreest | IOM | IOM | IOM | IOM | IOM | IOM | IOM |
| IOM 24 | Kissack | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM |
| IOM 25 | Shimmin | Douglas IOM | Foxdale IOM | Liverpool | Foxdale IOM | Patrick IOM | Braddan IOM | German IOM |
| IOM 26 | Lowey | Arbory IOM | Rushen IOM | Arbory IOM | Rushen IOM | Rushen IOM | Douglas IOM | Arbory IOM |
| IOM 27 | Teare | IOM | IOM | IOM | IOM | IOM | IOM | IOM |
| IOM 28 | Kneen | Port St. Mary IOM | Port St. Mary IOM | Michael IOM | Port St Mary IOM | Wirrel | Micheal IOM | Colby IOM |
| IOM 29 | Cannon | Douglas IOM | Douglas IOM | Douglas IOM | Ramsey IOM | Douglas IOM | Liverpool | Douglas IOM |
| IOM 30 | Gelling | IOM | IOM | Liverpool | IOM | IOM | IOM | IOM |
| IOM 31 | Keggin | IOM | IOM | IOM | IOM | IOM | IOM | London |
| IOM 32 | Cannan | Maughold IOM | Jurby IOM | Toronto | Bride IOM | Bride IOM | Lezarye IOM | Maughold |
| IOM 33 | Faragher | Peel IOM | Peel IOM | Douglas IOM | Peel IOM | Peel IOM | Douglas IOM | Peel IOM |
| IOM 34 | Cowin | IOM | IOM | Lincolnshire | IOM | IOM | England | England |
| IOM 35 | Kermeen | Douglas IOM | Dalby IOM | Peel IOM | Dalby IOM | Dalby IOM | Liverpool | Colby IOM |
| IOM 36 | Cooil | Douglas IOM | Douglas IOM | Colby IOM | Castletown IOM | Castletown IOM | Jurby IOM | Colby IOM |
| IOM 37 | Cubbon | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Rhyll Wales |
| IOM 38 | Qualtrough | Port St. Mary IOM | Port St. Mary IOM | Port St. Mary IOM | Colby IOM | Rushen IOM | Rushen IOM | Port st. Mary IOM |
| IOM 39 | Kaighin | Douglas IOM | Ramsey IOM | Ballaugh IOM | Jurby IOM | Ballaugh IOM | Ballaugh IOM | Santon IOM |
| IOM 40 | Fayle | Douglas IOM | Lonan IOM | Peel IOM | Lonan IOM | Lonan IOM | Peel IOM | Lancashire |
| IOM 41 | Kermode | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Liverpool |
| IOM 42 | Cringle | IOM | IOM | Ireland | IOM | IOM | Uk | Ireland |
| IOM 43 | Skelly | Douglas IOM | Douglas IOM | IOM | IOM | IOM | IOM | IOM |
| IOM 44 | Keown | IOM | IOM | Scotland | IOM | IOM | Scotland | Scotland |
| IOM 45 | Kennaugh | Liverpool | Castletown IOM | Douglas IOM | Unk | Unk | Unk | Unk |
| IOM 46 | Lewney | IOM | IOM | Norfolk | IOM | IOM | Norfolk | Norfolk |
| IOM 47 | Corlett | Winchester | IOM | Winchester | IOM | IOM | UK | India |
| IOM 48 | Crebbin | Douglas IOM | IOM | Unk | IOM | IOM | Unk | Unk |
| IOM 49 | Quilliam | Douglas IOM | Douglas IOM | Laxey IOM | Douglas IOM | Peel IOM | Ireland | Liverpool |
| IOM 50 | Kinrade | IOM | IOM | IOM | IOM | IOM | IOM | Austria |
| IOM 51 | Colquitt | IOM | IOM | Scotland | IOM | IOM | Unk | Scotland |
| IOM 52 | Kinley | Douglas IOM | Peel IOM | Douglas IOM | Peel IOM | Peel IOM | Douglas IOM | Douglas IOM |
| IOM 53 | Quane | Douglas IOM | Dalby IOM | Glanmaye IOM | Dalby IOM | IOM | Peel IOM | Arbory IOM |
| IOM 54 | Collister | Castletown IOM | Castletown IOM | Liverpool | unk | unk | unk | unk |
| IOM 55 | Bridson | Douglas IOM | Douglas IOM | Laxey IOM | IOM | IOM | IOM | IOM |
| IOM 56 | Skillicorn | IOM | IOM | Stockport | IOM | IOM | Stockport | Stockport |
| IOM 57 | Karran | Cheshire | Wrexham | Chatham | Manx ship | Castletown IOM | Leicester | Castletown IOM |
| IOM 58 | Condra | Douglas IOM | Laxey IOM | Laxey IOM | Lonan IOM | Dundee | Laxey IOM | Laxey IOM |
| IOM 59 | Cregeen | Douglas IOM | Castletown IOM | Birkenhead | Castletown IOM | Rushen IOM | Stafford | Bridge North |
| IOM 60 | Corkhill | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Liverpool | Douglas IOM | Northern Ireland |
| IOM 61 | Mylrea | Douglas IOM | Douglas IOM | Douglas IOM | Douglas IOM | Manchester | Douglas IOM | Douglas IOM |
| IOM 62 | Cretney | Douglas IOM | Douglas IOM | Liverpool | Andreas IOM | IOM | Liverpool | Liverpool |
| IOM 63 | Quaggin | Douglas IOM | Douglas IOM | Leith Scotland | Douglas IOM | Maldon | Leith | Calcutta |
| IOM 64 | Maddrell | Douglas IOM | Douglas IOM | Liverpool | unk | unk | Unk | unk |
| IOM 65 | Casement | IOM | IOM | IOM | IOM | england | IOM | IOM |
| IOM 66 | Gawne | Peel IOM | Douglas IOM | Douglas IOM | Malew IOM | Maughold IOM | Douglas IOM | Peel IOM |
| IOM 68 | Quiggin | IOM | IOM | IOM | IOM | unk | IOM | IOM |
| IOM 69 | Cowley | Douglas IOM | Glasgow | Peel IOM | Micheal IOM | Glasgow | Peel IOM | Micheal IOM |
| IOM 70 | Crowe | Micheal IOM | Micheal IOM | Baldwin IOM | Micheal IOM | Micheal IOM | Baldwin IOM | Baldwin IOM |
| IOM 71 | Joughin | IOM | IOM | IOM | IOM | england | England | England |
| IOM 72 | Cannell | Douglas IOM | IOM | Glasgow | IOM | IOM | Glasgow | Glasgow |
| IOM 73 | Keig | Kingston | Douglas IOM | London | Douglas IOM | Port Erin IOM | Lithuania | Lithuania |
| IOM 75 | Corrin | UK | IOM | UK | IOM | IOM | UK | India |

**Appendix A4: Volunteer pack paperwork sent to each participant with the DNA collection kit. Pack contained; welcome letter, two consent forms (just one shown), genealogical information form, project information sheet, and participant motivation questionnaire.**

The Isle of Man
Viking ancestry project

Researchers: Prof Mark Jobling, Dr. Simon James & Hayley Dunn

University of Leicester

**University Road**
Leicester LE1 7RH · UK
Tel: 0116 252 3377
Mobile: 07512 586 493
Email: surnames@le.ac.uk
web: www.le.ac.uk/genetics/maj4

Dear Participant,

Thank you for expressing interest in participating in our study of genetics and the history of the people of Isle of Man.

In this study we aim to look at the proportion of people with Viking ancestry on the Isle of Man. The only criterion for participating is that you are a man whose father's father was born on the Isle of Man. Simply giving a saliva sample provides us with all we need.

We are carrying out this work as a research project and thus it is absolutely free for you to take part. In return, we will provide you with a summary of the results, designed for a layperson, at the end of the study in 2013. In addition we will send you a copy of your Y chromosome genetic fingerprint and an explanation sheet designed for the layperson.

In this study we wish to study only unrelated men, so if you know that a male-line relative (e.g. a son or brother) has taken part, unfortunately we will not be able to analyse your DNA.

During this project we are looking at normal variation only, and no targeted tests of any medical consequence are done. However, while analysing Y-chromosomal variation it can be found, in very rare cases, that a man has lost part of his Y chromosome which is related to fertility. Therefore, if you concerned about the risks of detecting infertility, we would suggest that you do not take part.

Please find enclosed everything you need to take part in the study. Follow the kit instructions for providing the saliva sample, then write your name on the white sticky label and stick it to the sample tube. If you are returning the sample to us postally, place this in the padded envelope with the completed questionnaire and consent form by May 1st 2011. Use the enclosed stamp and sticky address label to return it to us.

Yours sincerely,
Hayley Dunn
Doctoral Research Student

# The Isle of Man Viking ancestry project

Researchers: Prof Mark Jobling, Dr. Simon James & Hayley Dunn

University of Leicester

DEPARTMENT OF GENETICS

University Road
Leicester LE1 7RH · UK
Tel: +44 (0) 116 252 3377/3427
Fax: +44 (0) 116 252 3378
Email: surnames@le.ac.uk
www:http://www.le.ac.uk/genetics/

## CONSENT FORM

Please **initial** box

1. I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. ☐
3. I consent to the storage of my sample. ☐
4. I agree to take part in the above study. ☐

| | | |
|---|---|---|
| _____ | _____ | _____ |
| Name of Participant | Date | Signature |
| _____ | _____ | _____ |
| Researcher | Date | Signature |

## Please retain this copy for your records.

When completed, 1 for participant; 1 for researcher site file;
Please retain one copy for yourself and return the other copy to Hayley Dunn, Department of Genetics, University of Leicester, University Road, Leicester, LE1 7RH, with your sample.

# The Isle of Man Viking ancestry project

Researchers: Prof Mark Jobling, Dr. Simon James & Hayley Dunn

University of Leicester

DEPARTMENT OF GENETICS

University Road
Leicester LE1 7RH · UK
Tel: +44 (0) 116 252 3377/3427
Fax: +44 (0) 116 252 3378
Email: surnames@le.ac.uk
www:http://www.le.ac.uk/genetics/

## Genealogy Questionnaire

Below we ask some questions about the origins of your family; we include your maternal ancestry because we want to have as much information as possible about the patterns of both male and female descent in our samples.

| | Place of birth | First Language |
|---|---|---|
| Me | | |
| My father | | |
| My mother | | |
| My father's father | | |
| My father's mother | | |
| My mother's father | | |
| My mother's mother | | |

**In order to ensure that we avoid sampling close male relatives, we would be grateful if you if could list your male line relatives (living or deceased) who share your surname, and who are closer than second cousin (e.g. father, grandfather, sons, brothers, uncles, cousins).**

| **Relationship** | **Name** | **Current or last place of residence (City, Town, or Village, and County)** |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

**Has a close relative of yours taken part in the study?** Yes/No/Don't know

**Your full name in BLOCK LETTERS**: .....

Signature of participant: ..... DATE: .....

**My place of permanent residence (Town, County):**.....

**Were you or your father adopted, or have you changed your surname?** YES /NO

*If YES, please give the original surname, if known*:……

**Current age:** .....

• I am interested to receive a description of the results of the study YES /NO

If YES, please give your email/postal address: ..... ..... ..... ..... .....

..... ..... ..... ..... ..... .....

..... ..... ..... ..... ..... .....

**ETHNIC ORIGIN**
*Please tick the box that you feel most accurately describes your ethnic origin.*

**WHITE**

English A1 ☐
Scottish A2 ☐
Welsh A3 ☐
Other British A4 ☐
*(Please specify)* ☐
Irish A5 ☐
Any other White background A6 ☐
*(Please specify)* ☐

White and Black Caribbean B1 ☐
White and Black African B2 ☐
White and Asian B3 ☐
White and Chinese B4 ☐
Any other Mixed background B5 ☐
*SPECIFY)* ☐

**OTHER ETHNIC GROUP**

Any other Ethnic group not above F1 ☐
*(Please specify)* ☐

**ASIAN,** *INCLUDING*
**Asian English, Asian Scottish, Asian Welsh, Asian British**

Indian C1 ☐
Pakistani C2 ☐
Bangladeshi C3 ☐
Any other Asian background C4 ☐
*(Please specify)* ☐

**BLAC**
*INCLUD*
**Black English, Black Scottish, Black Welsh, Black British**

***Caribbean*** ***D1*** ☐
African D2 ☐
Any other Black background D3 ☐
*(Please specify)* ☐

**CHINESE,** *including*
**Chinese English, Chinese Scottish, Chinese Welsh, Chinese British**

Chinese E1 ☐
Any other Chinese background E2 ☐
*(PLEAS*
*SPECIF* ☐

## Appendix A4 Replace this sheet with PDF of Project information sheet (1 page)

**Research Participant Motivation Questionnaire**

**BIOGRAPHICAL**

**Surname** *(please print)* **……**

**Age**

**18 – 29 [ ] 30 – 39 [ ] 40 – 49 [ ] 50 – 59 [ ] 60 – 69 [ ] 70 – 79 [ ] 80+ [ ]**

**Parish……..**

**MOTIVATIONS**

**1. How strongly are you motivated by the following reasons to take part in this study?** *(Please mark a cross on the line to represent how strongly you feel for each answer)*

**a. An interest in your own or your family's history**

**not strongly** **very strongly**

---

**b. An interest in Manx history**

**not strongly** **very strongly**

---

**c. A desire or interest in being part of a scientific study**

**not strongly** **very strongly**

---

**c. Any other reasons:** *please state***……..**

**not strongly** **very strongly**

---

**IDENTITY**

**2. Which nationality do you most identify with?** *(please tick one)*

**a. [ ] Manx**

**b. [ ] Irish**

**c. [ ] Scottish**

**d. [ ] English**

**e. [ ] British**

**f. [ ] all of the above**

**g. other:** *please state*…….

**h. [ ] I don't know**

**3. Who do you feel are most likely to be ancestors?** *(please tick one)*

**a.[ ] Gaelic peoples**

**b.[ ] Norse (Viking)** *(go to Q5)*

**c.[ ] Anglo Saxon (English)**

**d.[ ] other:** *please state*…….

**e.[ ] I don't know**

**VIKING ANCESTRY**

**4. Do you think you have any Viking ancestry?**

**a. [ ] Yes** *(go to Q5)*

**b. [ ] No** *(go to Q6)*

**c. [ ] I Don't know** *(If you have also answered 'I don't know' to question 3, please go to Q10)*

**5. Why do you think this?** *(Tick as many as you want)*

**a. [ ] Physical appearance (Height, build, hair or eye colour)**

**b. [ ] Family story**

**c. [ ] Gut feeling**

**d. [ ] Surname**

**e. [ ] Other:** *please state*…….

**6. How surprised will you be if the results do not match your expectations?** *Please mark a cross on the line below*:

**Not surprised very surprised**

**<_____>**

**7. Why would you be surprised?.....**

**……..**

**……..**

**8. How disappointed will you be if the results do not match your expectations?** *Please mark a cross on the line below*:

**Pleased not bothered very disappointed**

**<_____>**

**9. Why would you be pleased/disappointed?.....**

**……..**

**……..**

**RELATED INTERESTS**

**10. Have you ever considered tracing your family tree, history genealogy?** *(Please tick one box)*

**a.[ ] Yes I have or am in the process of doing this**

**b.[ ] Yes, I have considered it, but never done it**

**c.[ ] No, because someone else in my family is doing this.**

**d. [ ] No, I have never considered doing it, but I might consider it in the future.**

**e.[ ] No, it does not really interest me**

**11. Which box best describes your level of understandings of genetics:**

**a. [ ] no understanding whatsoever – what is genetics?**

**b. [ ] some knowledge of the basics or fundamentals**

**c. [ ] interested/self-educated layperson**

**d. [ ] some level of biological/medical training or education**

**e. [ ] some level of genetics training or education**

**COMMENTS**

**Please feel free to share more information about why you have volunteered for this project and your thoughts and feelings about this and similar projects, here**

**……**

**……**

**……**

**……**

**……..**

**……**

**……**

……

……

……

……

……

**Thank you for taking the time to answer these questions. Your participation is greatly appreciated.**

## Appendix A5: Volunteer Results Pack, contained letter of thanks, personalised results certificate and a brief explanation of what their results could indicate.

Dear Mr

Please find attached your Y chromosome results and an explanation of what it means. I'm sorry this has taken so much longer than I thought it would and I thank you for your patience. I would also like to take this opportunity to thank you again for your participation in this study. I will return to the Isle of Man to present my results to all those who took part once my PhD is completed. I will be sure to invite you and I hope that you will be able to attend. If you have any problems opening the documents do not hesitate to contact me.

Yours sincerely,

Ms Hayley Dunn

Postgraduate Research Student
School of Archaeology and Ancient History
Department of Genetics
University of Leicester

**replace this sheet with results certificate**

**Replace this sheet with info sheet 1**

**Replace this sheet with info sheet 2**

## Appendix A6: DNA Extraction and Purification Protocol

- Micro-centrifuge capable of 13,000rpm (15,000 x *g*)
- Water incubator (50ºC)
- Ethanol (95-100%) at room temperature
- DNA buffer: TE (10mM Tris-HCL, 1mM EDTA, pH 8.0)
- Ethanol (70%) at room temperature

| STEPS | NOTES |
|---|---|
| 1. Invert and shake vial | To ensure that saliva and suspension fluid is well mixed |
| 2. Incubate 500µL of each sample at 50ºC in a water incubator for a minimum of 1 hour | Heat treatment is necessary to ensure that DNA is released from the cells and the complete deactivation of nucleases; enzymes that break down DNA. Remainder of sample is stable and can be stored at room temperature. |
| 3. Transfer to micro-centrifuge tubes and add 20µL of Oragene•DNA Purifier (OG-L2P) then mix by vortexing for a few seconds | This ensures cell lysis and precipitates cellular material such as proteins and other impurities, so that the solution becomes turbid. |
| 4. Incubate on ice for 10 minutes. | |
| 5. Centrifuge for 5 minutes at 13,000rpm. | To condense suspended impurities into a pellet, leaving the DNA in solution. |
| 6. Transfer clear supernatant into a fresh micro-centrifuge tube, taking care not to disturb the pellet | This pellet contains impurities and can be discarded. The clear liquid contains DNA. |
| 7. To the supernatant, add an equal volume of ethanol (95-100%) and invert gently to mix. | DNA is not soluble in alcohol so precipitates out and may be visible as a fine fibres in the liquid. |
| 8. Stand at room temperature for 10 minutes. | |
| 9. Centrifuge for 2 minutes at 13,000rpm | The resultant pellet contains DNA |
| 10. Remove the supernatant with a pipette tip and discard it, being careful not to disturb the pellet. | |
| 11. Carefully add 250µL of 70% ethanol and let it stand at room temperature for 1 minute. Completely remove the ethanol without disturbing the pellet. | This helps to remove any residual impurities. |
| 12. Add 100µL of DNA buffer TE to dissolve the DNA pellet and vortex for at least 5 seconds. | Rehydrates DNA into a solution |
| 13. Can be stored long term at -20 | May cause precipitation and require rehydration. |
| 14. Rehydrate for further analysis by either<br>a. vigorous pipetting and vortexing | The expected concentration of DNA is between 20 and 200ng/µL |

b. incubating at 50°C for 1 hour or
c. Incubating for 1-2 days at room temperature

## Appendix A7: STRs Identification Protocols

- Laminar Flow Hood
- Disposable tips
- AmpFℓSTR Yfiler kit (Applied Biosystems) containing;
  - AmpFℓSTR Yfiler PCR reaction mix
  - AmpFℓSTR Yfiler Primer mix with 6-FAM™ (blue), VIC® (green), NED™ (yellow>black), and PET® (red) dye-labelled and unlabelled primers
  - Amplitaq Gold® DNA Polymerase
  - 007 control DNA of known genotype
- Centrifuge
- MicroAmp® Optical 96-well reaction plate
- Laboratory male control (MAJ)
- MJ Research Peltier Thermal PTC-200 cycler

| STEPS | NOTES |
|---|---|
| 1. In laminar hood using dedicated PCR pipettes and disposable tips, create master mix containing, per sample, 3.06µl PCR reaction mix, 1.67µl primer mix, 0.27µl Amplitaq Gold® DNA polymerase | To minimise contamination.<br>Remember to add extra for 3 controls |
| 2. Vortex and centrifuge reagent mix | |
| 3. Add 5µl to each well which will contain a sample. | |
| 4. Add 3µl of DNA from each sample into assigned well. Add 3µl each of 007 control, Lab control and no DNA control into assigned wells. Seal plate firmly | Draw a plate map showing the location of each sample and control.<br>Leave a well empty for allelic ladder |
| 5. Centrifuge to remove bubbles, load into thermal cycler. | Following steps (6-11) were programmed to occur automatically by Dr Turi King |
| 6. Initial incubation step runs at 95ºC for 11 minutes. | |
| 7. Denature step 94ºC for 1 minute | Repeat steps 7-9 for 30 cycles |
| 8. Annealing step 61ºC for 1 minute | |
| 9. Extension step 72ºC for 1 minute | |
| 10. Final extension step 60ºC for 80 minutes | |
| 11. Store at 4ºC until needed | |

- MicroAmp® Optical 96-well reaction plate and septa
- Amplified DNA
- GeneScan™ – 500 LIZ® Size Standard plus orange LIZ® Fluorophore dye
- Hi-Di™ Formamide
- AmpFℓSTR Yfiler Allelic ladder
- Centrifuge
- MJ Research Peltier Thermal PTC-200 cycler
- Ice
- ABI PRISM® 3130xl Genetic Analyzer (Applied Biosystems)

| STEPS | NOTES |
|---|---|
| 1.Dilute 35µl 500 LIZ® size standard with 1ml formamide. Into a new 96-well plate, add 10µl of this mix into each well which will contain a sample. | 500 LIZ® was designed to provide a size standard for fragments of DNA between 35 and 500bp long. Each fragment was labelled to give single peaks in the readout at 35, 50, 75, 100, 139, 150, 160 200, 250, 300, 340, 350, 400, 450, 490 and 500 bases. |
| 2. Add 1µl of each post-PCR sample to assigned well. Add allelic ladder to a spare well. | Allelic Ladder contains amplified and labelled alleles. |
| 3. Seal with septa and centrifuge | Each plate must be given a name and a layout map |
| 4. Load into thermal cycler for 3 minutes at 95ºC | Denatures all DNA. |
| 5. Put on ice | |
| 6. Load into ABI genetic analyzer | By passing an electric field across capillaries containing polymer POP-7 (Applied Biosystems), fragments are separated by weight. A laser beam causes labels attached to primers attached to each STR to fluoresce at a different frequency. The fluorescence is picked up by an optical detector and converted into digital data by data collection software. |
| 7. Results processed by GeneMapper® software (Applied Biosystems) | GeneMapper® converts the digital readout into to genotype data |

## Appendix A8: List of primers used to target specific SNP sites.

| SNP ID | Forward Primer (5'-3')<br>Reverse primer (3'-5') | hg | SNP | Ref ID | Reference |
|---|---|---|---|---|---|
| **M173** | 'AGAGGGAGCAATGAGGACA'<br>'AGGTGTATCTGGCATCCGTTA' | R1 | A>T | rs2032624 | (Underhill et al. 2001) |
| **M170** | 'GTTTGTTCAAATAATTGCAGCTC'<br>'TGAGACACAACCCACACTGAA' | I | A>C | rs2032597 | (Paracchini et al.2002) |
| **M253** | 'GCAACAATGAGGGTTTTTTTG'<br>'CAGCTCCACCTCTATGCAGTTT' | I1 | C>T | rs9341296 | (Cinnioğlu et al. 2004) |
| **M172** | 'TAAAGAGGCCAGCTTTGTGC'<br>'CACTCCATGTTGGTTTGGAAC' | J2 | T>G | rs2032604 | paracchini |
| **M35** | 'TAAGCCTAAAGAGCAGTCAGAG'<br>'AGAGGGAGCAATGAGGACA' | E1b1b1 | G>C | N/A | underhill |
| **M201** | 'TTGTGTGTGTATGCATTTGTTGA'<br>'CATCATGGTGTGACGAACG' | G | G>T | rs2032636 | (Bosch et al. 2006) |
| **M242** | 'AACTCTTGATAAACCGTGCTG'<br>'TCCAATCTCAATTCATGCCTC' | Q | C>T | rs8179021 | cinnioglu |
| **Tat** | 'GACTCTGAGTGTAGACTTGTGA'<br>'GAAGGTGCCGTAAAGTGTGAA' | N1c | T>C | rs34442126 | Zerjal 1997 |
| **M223** | 'TTCAGCAAGAGTAAGCAAGAGG'<br>'CCTTTTTGGATCATGGTTCTT' | I1b | G>A | N/A | Underhill 2001 |
| **M17** | 'CTGGTCATAACACTGGAAATC'<br>'TGAACCTACAAATGTGAAACT' | R1a1 | ->G | rs3908 | " " |
| **M167** | 'AACAGGAGGAGGTGTCGATG'<br>'GAGGCTGGGCCAAGTTAAG' | R1b1b1a1b5 | C>A | rs1800865 | M. Jobling's laboratory |
| **U106** | 'GCAAATCCCAAAGCTCCACG'<br>'TGTGTGTGCACACCTGTGG' | R1b1b1a1a | C>T | rs16981293 | " " |
| **U152** | 'CTTAGCTATACAGCCTCTTTTTGG'<br>'AACATTCCACGCTTGAGGATAA' | R1b1b1a1b1 | C>T | rs1236440 | " " |
| **U198** | 'TAGGTTCTATGGTGATTTGAAC'<br>'CTTAATCAGAACAAGACATTCC' | R1b1b1a1a1 | G>A | rs17222279 | " " |
| **M153** | 'GAGCACGCTATCCCGTTAGACTTGTGTATGCCTTCCGATTT'<br>'CGCTGCCAACTACCGCACATGTTCTCAGACACCAATGGTCCT' | R1b1b1a1b4 | T>A | N/A | " " |
| **S145** | 'CCAAGTCTTTGATGTGCTGTC'<br>'TCAAGGAGGTTCTTGATTTATGC' | R1b1b1a1b2 | C>G | rs11799226 | " " |
| **M222** | 'CATTCAAGATCCCAGAACTGTC'<br>'GGTGATGGATGAGGAGTAAAAA' | R1b1b1a1b2a | G>A | rs20321 | " " |

| | | | | | |
|---|---|---|---|---|---|
| **S116** | 'AACCTGCAGCCATAAGTCTC' 'CAAGGGAGTGAGGCACTTAG' | R1b1b1a1b | C>A | rs342763000 | " " |
| **L11** | 'GGTTTTTTTATGCTGCTGCA' 'ACTCTTTTGCCTAAATTGCTTGT' | R1b1b1a1 | T>C | rs9786076 | " " |

# APPENDIX B

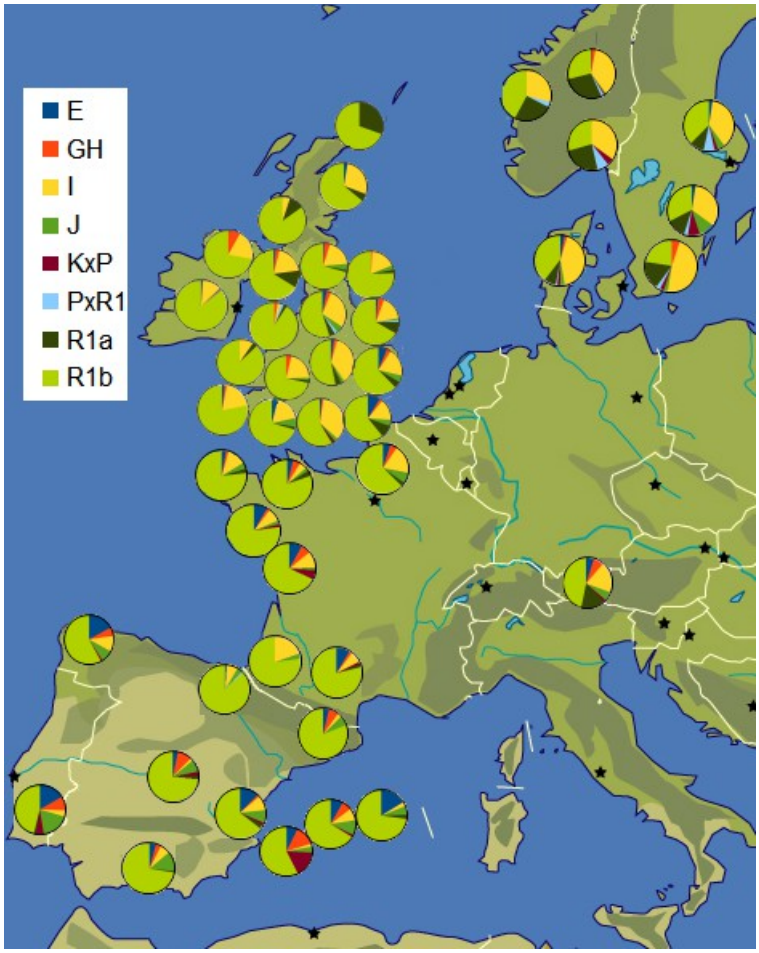
## Appendix B1: List of Isle of Man NRY STR and SNP haplotypes

| Sample | DYS19 | DYS389 | DYS389 | DYS391 | DYS392 | DYS393 | DYS437 | DYS438 | DYS439 | DYS448 | GATA_H | DYS390 | DYS456 | DYS385 | DYS385 | DYS458 | DYS635 | SNP | Notes |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IoM 1 | 14 | 13 | 29 | 11 | 13 | 13 | 15 | 12 | 11 | 19 | 12 | 25 | 16 | 11 | 14 | 17 | 23 | R1b (not U106) | STR only |
| IoM 2 | 14 | 13 | 29 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 12 | 24 | 16 | 11 | 14 | 17 | 23 | R1b-S145 | |
| IoM 3 | 14 | 13 | 29 | 10 | 14 | 13 | 15 | 12 | 11 | 17 | 12 | 25 | 17 | 11 | 13 | 17 | 23 | R1b-M222 | |
| IoM 4 | 14 | 12 | 28 | 10 | 11 | 13 | 16 | 10 | 13 | 19 | 11 | 23 | 14 | 13 | 14 | 15 | 21 | I1-M253 | |
| IoM 5 | 16 | 13 | 29 | 11 | 11 | 13 | 14 | 11 | 10 | 19 | 13 | 25 | 16 | 11 | 14 | 15 | 24 | R1a1 | |
| IoM 6 | 15 | 13 | 29 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 12 | 24 | 18 | 12 | 14 | 17 | 23 | R1b-S145 | |
| IoM 7 | 14 | 13 | 29 | 12 | 13 | 13 | 15 | 12 | 13 | 19 | 12 | 23 | 16 | 11 | 14 | 20 | 24 | R1b-U106 | |
| IoM 8 | 15 | 13 | 29 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 12 | 24 | 18 | 12 | 14 | 17 | 23 | R1b-M167 | |
| IoM 9 | 15 | 13 | 29 | 11 | 11 | 13 | 14 | 11 | 10 | 20 | 12 | 25 | 17 | 11 | 13 | 15 | 23 | R1a1 | |
| IoM 10 | 14 | 14 | 30 | 11 | 13 | 13 | 15 | 12 | 12 | 18 | 12 | 24 | 16 | 11 | 14 | 17 | 23 | R1b- S145 | |
| IoM 11 | 14 | 14 | 30 | 11 | 13 | 13 | 15 | 12 | 12 | 19 | 13 | 24 | 15 | 11 | 15 | 17 | 23 | R1b- S145 | |
| IoM 12 | 14 | 13 | 29 | 10 | 13 | 13 | 15 | 12 | 13 | 19 | 11 | 24 | 16 | 11 | 14 | 20 | 23 | R1b- S145 | |
| IoM 14 | 14 | 13 | 29 | 11 | 13 | 13 | 14 | 12 | 11 | 19 | 12 | 24 | 16 | 11 | 14 | 16 | 23 | R1b-S145 | |
| IoM 15 | | | | 11 | | 13 | | 12 | | | 11 | | 15 | | | 17 | | UNK | discard |
| IoM 16 | 14 | 14 | 30 | 11 | 14 | 13 | 15 | 12 | 12 | 18 | 12 | 25 | 17 | 10 | 13 | 17 | 23 | R1b-M167 | |
| IoM 17 | 16 | 14 | 31 | 10 | 11 | 13 | 14 | 11 | 10 | 20 | 13 | 25 | 15 | 11 | 15 | 16 | 24 | R1a1 | |
| IoM 18 | 14 | 12 | 29 | 10 | 11 | 13 | 16 | 10 | 11 | 20 | 11 | 22 | 14 | 14 | 15 | 16 | 22 | I1-M253 | |
| IoM 19 | 15 | 13 | 30 | 11 | 11 | 13 | 14 | 10 | 11 | 20 | 11 | 24 | 17 | 13 | 16 | 18 | 21 | I-M170 | |
| IoM 20 | 15 | 12 | 28 | 10 | 11 | 12 | 14 | 9 | 12 | 19 | 12 | 24 | 15 | 16 | 20 | 16 | 21 | H (Not M82) | |
| IoM 21 | 16 | 13 | 29 | 11 | 13 | 13 | 14 | 12 | 12 | 19 | 12 | 24 | 15 | 10 | 14 | 16 | 23 | R1b-S116 | |
| IoM 22 | 14 | 12 | 29 | 10 | 13 | 13 | 15 | 13 | 12 | 19 | 11 | 24 | 16 | 11 | 14 | 17 | 23 | R1b-M167 | |
| IoM 23 | 14 | 14 | 31 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 12 | 24 | 16 | 11 | 10 | 17 | 23 | R1b-S116 | |
| IoM 24 | 14 | 13 | 29 | 11 | 13 | 13 | 15 | 12 | 13 | 19 | 12 | 25 | 16 | 11 | 15 | 17 | 23 | R1b-S116? | STR only |
| IoM 25 | 15 | 14 | 32 | 10 | 12 | 14 | 14 | 10 | 11 | 20 | 11 | 23 | 15 | 14 | 15 | 13 | 20 | I2b | |
| IoM 26 | 15 | 13 | 29 | 10 | 9 | 12 | 16 | 10 | 11 | 20 | 11 | 22 | 15 | 13 | 13 | 16 | 22 | UNK | STR only |
| IoM 27 | 14 | 13 | 29 | 11 | 14 | 13 | 15 | 12 | 12 | 18 | 12 | 25 | 17 | 11 | 13 | 17 | 23 | R1b-M222 | |
| IoM 28 | 13 | 13 | 29 | 11 | x | 13 | 14 | 12 | 12 | 19 | 11 | 23 | 16 | 11 | 15 | 18 | 25 | R1b-U106 | SNP only |
| IoM 30 | 14 | 13 | 29 | 11 | 13 | 13 | 15 | 12 | 12 | 19 | 12 | 24 | 15 | 11 | 14 | 17 | 24 | R1b-U152 | |
| IoM 31 | 15 | 13 | 29 | 9 | 13 | 13 | 15 | 12 | 14 | 20 | 12 | 24 | 15 | 11 | 15 | 17 | 23 | R1b-U106 | |
| IoM 32 | 13 | 12 | 28 | 11 | 13 | 13 | 16 | 12 | 12 | 19 | 12 | 24 | 15 | 11 | 14 | 16 | 23 | R1b-U152 | |
| IoM 33 | 13 | 13 | 29 | 11 | 13 | 13 | 15 | 12 | 12 | 19 | 13 | 24 | 15 | 10 | 15 | 17 | 23 | R1b-U152 | |
| IoM 34 | 14 | 13 | 30 | 12 | 13 | 13 | 15 | 12 | 12 | 19 | 11 | 24 | 16 | 12 | 14 | 16 | 25 | R1b-S145 | |
| IoM 35 | 16 | 13 | 29 | 11 | 11 | 13 | 14 | 11 | 10 | 20 | 13 | 25 | 16 | 11 | 14 | 15 | 23 | R1a1 | |
| IoM 36 | 14 | 12 | 28 | 10 | 11 | 13 | 16 | 10 | 11 | 20 | 11 | 23 | 14 | 14 | 15 | 14 | 22 | I1-M253 | |
| IoM 37 | 14 | 14 | 30 | 10 | 13 | 13 | 15 | 12 | 11 | 19 | 12 | 24 | 16 | 11 | 14 | 17 | 23 | R1b- S145 | |
| IoM 38 | 14 | 14 | 31 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 11 | 24 | 16 | 11 | 14 | 17 | 23 | R1b-S145 | |
| IoM 40 | 16 | 13 | 29 | 11 | 11 | 13 | 14 | 11 | 10 | deleted | 11 | 25 | 17 | 11 | 14 | 15 | 23 | R1a1-M17 | SNP only |
| IoM 41 | 15 | 13 | 29 | 11 | 13 | 13 | 15 | 12 | 12 | 19 | 12 | 24 | 17 | 12 | 14 | 17 | 23 | R1b-S116 | |
| IoM 42 | 14 | 13 | 29 | 10 | 14 | 14 | 15 | 12 | 11 | 18 | 12 | 26 | 16 | 11 | 12 | 17 | 23 | R1b-M222 | |
| IoM 43 | 14 | 13 | 28 | 11 | 13 | 13 | 15 | 12 | 11 | 19 | 12 | 24 | 15 | 10 | 14 | 16 | 23 | R1b-U106 | |
| IoM 45 | 14 | 14 | 31 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 12 | 24 | 15 | 11 | 16 | 17 | 23 | R1b-S116 | |
| IoM 46 | 14 | 13 | 29 | 10 | 13 | 13 | 15 | 12 | 12 | 18 | 12 | 24 | 16 | 11 | 14 | 17 | 23 | R1b- S145 | |
| IoM 47 | 15 | 14 | 30 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 11 | 24 | 17 | 11 | 14 | 16 | 23 | R1b- S145 | |
| IoM 48 | 15 | 12 | 28 | 10 | 13 | 13 | 16 | 12 | 13 | 19 | 12 | 23 | 16 | 11 | 14 | 18 | 23 | R1b-L11 | |
| IoM 49 | 14 | 13 | 29 | 11 | 13 | 14 | 15 | 12 | 12 | 18 | 12 | 23 | 16 | 11 | 14 | 15 | 23 | R1b- S145 | |
| IoM 50 | 14 | 12 | 28 | 10 | 11 | 14 | 16 | 10 | 12 | 20 | 10 | 23 | 14 | 14 | 14 | 15 | 21 | I1 | |
| IoM 51 | 14 | 15 | 32 | 11 | 13 | 12 | 15 | 12 | 12 | 19 | 12 | 24 | 15 | 11 | 14 | 17 | 23 | R1b-U152 | |
| IoM 53 | 14 | 13 | 28 | 11 | 13 | 13 | 15 | 12 | 11 | 19 | 12 | 24 | 15 | 10 | 14 | 16 | 23 | R1b-U106 | |
| IoM 54 | 14 | 13 | 29 | 11 | 14 | 13 | 14 | 12 | 12 | 18 | 12 | 25 | 17 | 11 | 13 | 17 | 23 | R1b-M222 | |
| IoM 56 | 14 | 12 | 28 | 10 | 11 | 13 | 16 | 10 | 12 | 20 | 10 | 22 | 15 | 13 | 14 | 15 | 22 | I1 | |
| IoM 57 | 14 | 12 | 29 | 10 | 11 | 13 | 16 | 10 | 11 | 21 | 11 | 22 | 14 | 14 | 14 | 14 | 21 | I1 | |
| IoM 58 | 14 | 13 | 29 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 11 | 24 | 16 | 11 | 14 | 19 | 23 | R1b-S116 | |
| IoM 59 | 14 | 13 | 29 | 10 | 11 | 13 | 16 | 10 | 11 | 20 | 11 | 22 | 14 | 13 | 14 | 13 | 23 | I1 | |
| IoM 60 | 15 | 13 | 29 | 11 | 11 | 13 | 14 | 10 | 11 | 20 | 11 | 24 | 17 | 13 | 16 | 17 | 22 | Not-R1a1 | STR only |
| IoM 61 | 14 | 13 | 30 | 11 | 13 | 13 | 15 | 12 | 13 | 19 | 12 | 24 | 15 | 11 | 15 | 16 | 23 | R1b-S145 | |
| IoM 62 | 16 | 12 | 26 | 10 | 11 | 13 | 14 | 11 | 10 | 20 | 12 | 25 | 15 | 11 | 15 | 16 | 23 | R1a1-M17 | |
| IoM 63 | 16 | 13 | 30 | 10 | 11 | 13 | 14 | 11 | 10 | 20 | 12 | 26 | 16 | 11 | 14 | 15 | 23 | Not R1a1 | STR only |
| IoM 64 | 14 | 13 | 30 | 10 | 13 | 13 | 15 | 12 | 12 | 19 | 12 | 24 | 15 | 11 | 14 | 17 | 23 | R1b-S116 | |
| IoM 65 | 14 | 12 | 30 | 10 | 11 | 13 | 16 | 10 | 11 | 20 | 11 | 23 | 15 | 13 | 14 | 15 | 21 | I1-M253 | |
| IoM 66 | 14 | 13 | 29 | 10 | 13 | 13 | 15 | 12 | 12 | 18 | 12 | 24 | 16 | 11 | 14 | 17 | 23 | R1b-S145 | |
| IoM 68 | 14 | 13 | 29 | 10 | 13 | 13 | 14 | 13 | 13 | 19 | 11 | 24 | 15 | 11 | 14 | 20 | 23 | R1b-S145 | |
| IoM 69 | 14 | 12 | 28 | 10 | 11 | 13 | 16 | 10 | 13 | 19 | 11 | 23 | 14 | 14 | 15 | 15 | 22 | I1-M253 | |
| IoM 70 | 14 | 14 | 30 | 11 | 13 | 13 | 15 | 12 | 12 | 19 | 11 | 25 | 15 | 11 | 15 | 17 | 23 | R1b-S145 | |
| IoM 71 | 14 | 13 | 29 | 11 | 13 | 14 | 15 | 12 | 12 | 19 | 11 | 24 | 15 | 11 | 14 | 18 | 23 | R1b-S145 | |
| IoM 72 | 15 | 14 | 30 | 11 | 11 | 13 | 15 | 10 | 11 | 20 | 11 | 25 | 17 | 11 | 17 | 18 | 22/24 | I-M170 | SNP only |
| IoM 73 | 14 | 13 | 29 | 10 | 14 | 13 | 15 | 12 | 12 | 18 | 12 | 25 | 17 | 11 | 13 | 18 | 23 | R1b-M222 | |
| IoM 75 | 15 | 13 | 29 | 11 | 13 | 13 | 15 | 12 | 12 | 19 | 11 | 24 | 17 | 12 | 14 | 17 | 23 | R1b-S145 | |

# Appendix B2 NRY haplotype diversity at two resolutions for Western European populations

| Populations | ID | N | He 8SNPs | He 17SNPs |
|---|---|---|---|---|
| **North Wales** | NWA | 44 | 0.17548 | 0.63108 |
| **Spanish Basques** | BAS | 116 | 0.23658 | |
| **Pembrokeshire** | PEM | 31 | 0.23871 | 0.67097 |
| **Ireland** | IRE | 796 | 0.254 | |
| **Argyll** | ARG | 20 | 0.27895 | 0.67368 |
| **Catalonia** | CAT | 79 | 0.33788 | |
| **French Basques** | FBA | 61 | 0.37268 | |
| **North East England** | NEE | 125 | 0.3929 | 0.7551 |
| **Cornwall** | PCO | 41 | 0.39878 | 0.6378 |
| **Leicester*** | PLE | 36 | 0.41746 | 0.8619 |
| **Orkney** | ORK | 23 | 0.44269 | 0.72332 |
| **Minorca** | MIN | 37 | 0.44895 | |
| **Oxfordshire** | OXF | 53 | 0.4492 | 0.8135 |
| **Northern Ireland** | NIR | 25 | 0.45333 | 0.70333 |
| **Andalucia** | SPA | 95 | 0.45442 | |
| **Castilla le Mancha** | CLM | 63 | 0.45622 | |
| **Cumbria** | CUM | 123 | 0.46488 | 0.80928 |
| **Devon** | DEV | 51 | 0.47451 | 0.77647 |
| **Banff** | BAN | 39 | 0.51957 | 0.78812 |
| **Isle of Man** | IOM | 74 | 0.52018 | 0.74359 |
| **Lincolnshire** | LIN | 72 | 0.52582 | 0.85329 |
| **Majorca** | MAJ | 62 | 0.54469 | |
| **England South Coast** | SCE | 82 | 0.54622 | 0.81963 |
| **Valencia** | VAL | 73 | 0.56088 | |
| **Norfolk** | NFK | 66 | 0.57529 | 0.84802 |
| **Leicester#** | LEI | 47 | 0.59574 | |
| **Kent** | KEN | 46 | 0.60193 | 0.88696 |
| **York** | YOR | 102 | 0.61774 | 0.88391 |
| **Galicia** | GAL | 88 | 0.622 | |
| **Ibiza** | IBZ | 54 | 0.63033 | |
| **Denmark** | DEN | 106 | 0.66092 | 0.80341 |
| **Norway Coastal** | NWC | 83 | 0.68146 | 0.78982 |
| **Sweden Blekinge** | SWB | 40 | 0.70641 | 0.76795 |
| **Norway Inland** | NWI | 225 | 0.70726 | 0.78738 |
| **Austria** | AET | 141 | 0.71499 | |
| **Sweden Skaraborg** | SWS | 45 | 0.71515 | 0.82222 |
| **Southern Portugal** | SPO | 78 | 0.72394 | |
| **Norway** | NWY | 71 | 0.73481 | 0.79638 |
| **Sweden Östergötland** | SWO | 40 | 0.77564 | 0.82929 |
| **Total/Mean** | | 3453 | 0.51355 | 0.77906 |

## Appendix B3: Map showing relative NRY haplogroup frequencies for populations of Western Europe at the lowest resolution haplogroup resolution (n=8)



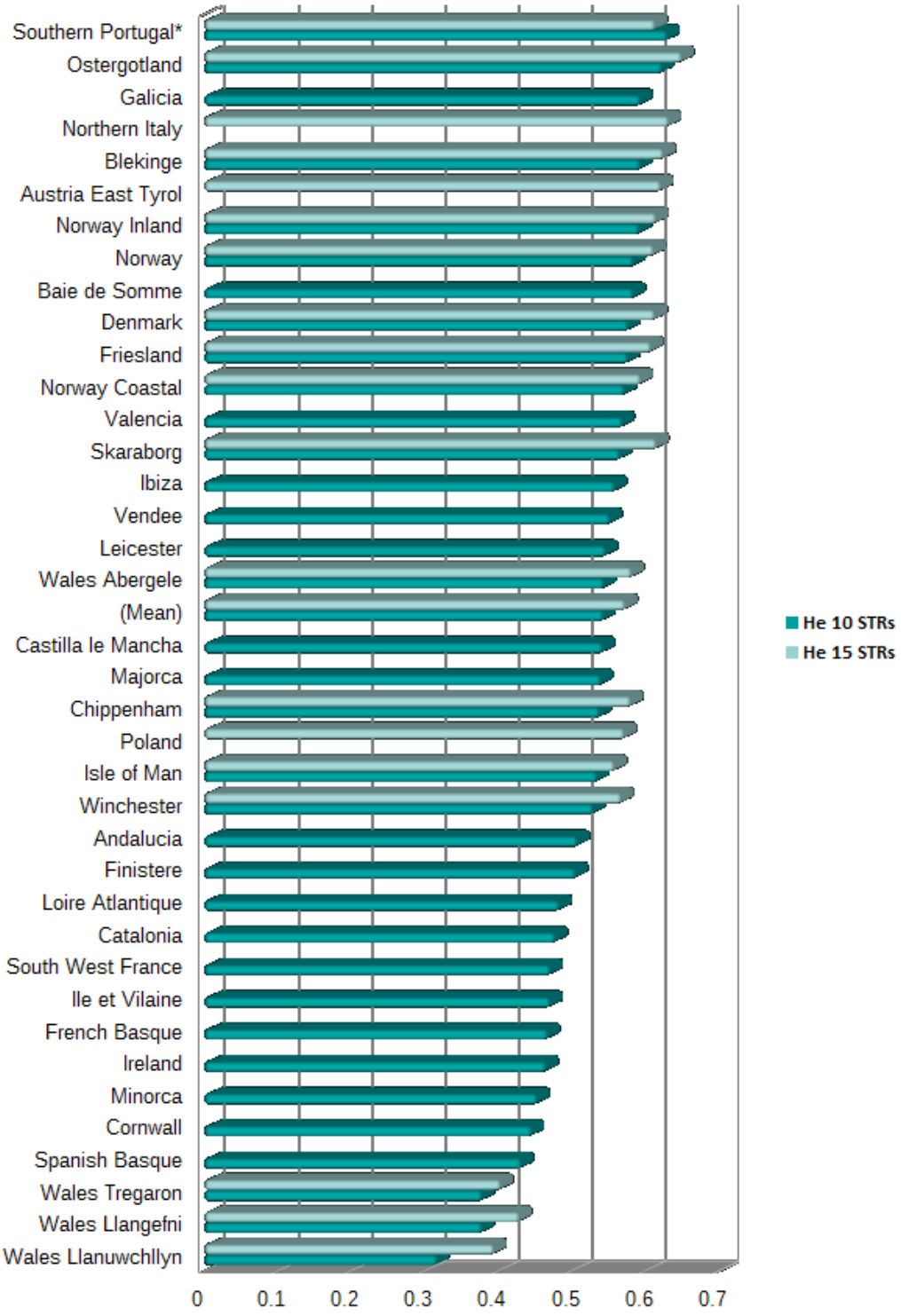
# Appendix B4: Number of each NRY Haplogroup for the British and Irish Isles and Scandinavia at the highest resolution (n=17)

| | E | F*(xG,I,J,K) | G | I* | I1 | I2 | J* | J2 | N | Q | R1a | R1b | R1b | R1b | R1b | R1b | R1b | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | M215/M35 | M89 | M201 | M170 | M253 | M223/P219 | I2f2/M304 | M172 | TAT | M242 | M420/M17 | M269 | S116 | U152 | M222 | U106 | U198 | N |
| Isle of Man | 0 | 0 | 1 | 2 | 9 | 1 | 0 | 0 | 0 | 0 | 6 | 1 | 27 | 4 | 5 | 5 | 0 | 61 |
| Pembrokeshire | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 17 | 4 | 1 | 0 | 5 | 0 | 31 |
| Cornwall | 0 | 0 | 0 | 1 | 5 | 2 | 0 | 1 | 0 | 0 | 1 | 24 | 0 | 1 | 1 | 5 | 0 | 41 |
| Leicestershire | 0 | 0 | 0 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 1 | 10 | 2 | 6 | 0 | 6 | 3 | 36 |
| N. Ireland | 0 | 0 | 2 | 1 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 13 | 0 | 0 | 3 | 2 | 0 | 25 |
| Orkney | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 8 | 0 | 1 | 0 | 7 | 0 | 23 |
| Argyll | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 11 | 0 | 1 | 0 | 4 | 1 | 20 |
| Oxfordshire | 0 | 1 | 1 | 2 | 6 | 3 | 0 | 1 | 0 | 0 | 1 | 19 | 2 | 7 | 0 | 10 | 0 | 53 |
| S. Coast England | 1 | 0 | 1 | 4 | 18 | 6 | 0 | 2 | 1 | 0 | 1 | 21 | 2 | 3 | 0 | 21 | 1 | 82 |
| Banff | 1 | 0 | 0 | 3 | 7 | 1 | 0 | 0 | 0 | 0 | 2 | 16 | 0 | 1 | 4 | 4 | 0 | 39 |
| North Wales | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 26 | 6 | 3 | 0 | 4 | 1 | 44 |
| Devon | 2 | 0 | 0 | 0 | 6 | 3 | 0 | 3 | 0 | 0 | 1 | 19 | 1 | 2 | 0 | 14 | 0 | 51 |
| Kent | 4 | 0 | 1 | 2 | 3 | 2 | 0 | 2 | 0 | 0 | 4 | 11 | 2 | 4 | 0 | 9 | 2 | 46 |
| Lincolnshire | 1 | 1 | 2 | 5 | 5 | 2 | 0 | 1 | 0 | 1 | 6 | 19 | 7 | 5 | 0 | 17 | 0 | 72 |
| Yorkshire | 3 | 0 | 3 | 5 | 15 | 9 | 1 | 4 | 0 | 2 | 4 | 21 | 3 | 7 | 2 | 20 | 3 | 102 |
| Norfolk | 4 | 1 | 1 | 2 | 9 | 2 | 0 | 3 | 0 | 0 | 3 | 18 | 2 | 2 | 1 | 16 | 2 | 66 |
| N.E. England | 0 | 0 | 1 | 5 | 14 | 3 | 1 | 3 | 0 | 0 | 3 | 49 | 4 | 5 | 0 | 35 | 2 | 125 |
| Cumbria | 1 | 0 | 5 | 5 | 14 | 4 | 1 | 4 | 0 | 0 | 2 | 44 | 1 | 11 | 3 | 25 | 3 | 123 |
| Norway Coastal | 0 | 0 | 0 | 0 | 23 | 1 | 0 | 0 | 0 | 3 | 21 | 4 | 19 | 0 | 0 | 11 | 0 | 82 |
| Norway Inland | 1 | 0 | 6 | 1 | 72 | 9 | 0 | 2 | 2 | 4 | 63 | 2 | 30 | 2 | 2 | 27 | 2 | 225 |
| Norway | 0 | 0 | 0 | 0 | 24 | 1 | 0 | 0 | 3 | 5 | 17 | 1 | 10 | 1 | 0 | 9 | 0 | 71 |
| Denmark | 3 | 0 | 2 | 3 | 37 | 5 | 1 | 2 | 2 | 1 | 8 | 3 | 7 | 6 | 0 | 27 | 1 | 108 |
| Blekinge | 0 | 0 | 2 | 0 | 17 | 2 | 0 | 1 | 1 | 1 | 7 | 0 | 2 | 0 | 0 | 7 | 0 | 40 |
| Skaraborg | 1 | 0 | 0 | 0 | 16 | 1 | 0 | 2 | 1 | 3 | 4 | 0 | 8 | 3 | 0 | 6 | 0 | 45 |
| Ostergotland | 1 | 0 | 0 | 0 | 12 | 1 | 2 | 2 | 3 | 1 | 5 | 5 | 13 | 0 | 0 | 0 | 0 | 45 |

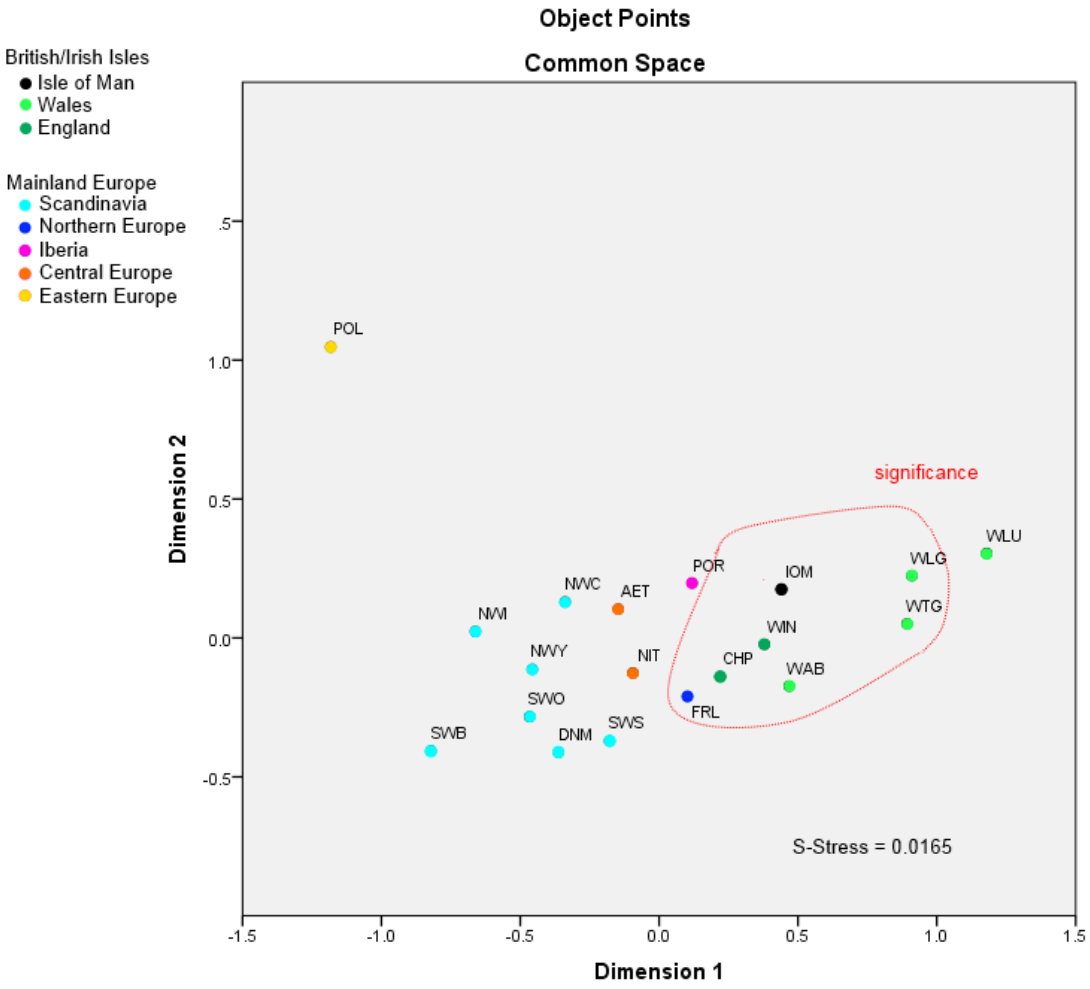
## Appendix B5: List of NRY STR diversity values.

| Population | ID | N | 10 STRS | | 15 STRs | |
|---|---|---|---|---|---|---|
| | | | k | $H_e$ | k | $H_e$ |
| **Isle of Man** | IOM | 64 | 50 | 0.53254 | 59 | 0.55643 |
| **Andalucia** | SPA | 95 | 78 | 0.50611 | | |
| **Friesland** | FRL | 95 | 75 | 0.57474 | 90 | 0.60618 |
| **Spanish Basque** | BAS | 116 | 74 | 0.42982 | | |
| **Castilla le Mancha** | CLM | 62 | 54 | 0.5404 | | |
| **Catalonia** | CAT | 79 | 66 | 0.47567 | | |
| **Denmark** | DNM | 109 | 89 | 0.57506 | 103 | 0.61137 |
| **Leicester** | LEI | 46 | 37 | 0.54435 | | |
| **Cornwall** | COR | 64 | 44 | 0.4437 | | |
| **Baie de Somme** | FBS | 43 | 41 | 0.58184 | | |
| **Ile et Vilaine** | FIV | 81 | 62 | 0.46898 | | |
| **Loire Atlantique** | FLA | 48 | 41 | 0.48147 | | |
| **Finistere** | FFS | 74 | 63 | 0.50337 | | |
| **South West France** | FSW | 57 | 49 | 0.46911 | | |
| **Vendee** | FVD | 50 | 38 | 0.55029 | | |
| **French Basque** | FBA | 60 | 48 | 0.46559 | | |
| **Ireland** | IRE | 796 | 507 | 0.46327 | | |
| **Norway Coastal** | NWC | 82 | 58 | 0.5701 | 73 | 0.59203 |
| **Norway Inland** | NWI | 224 | 146 | 0.59011 | 195 | 0.6131 |
| **Norway** | NWY | 72 | 62 | 0.58192 | 70 | 0.60999 |
| **Ostergotland** | SWO | 40 | 40 | 0.6209 | | |
| **Skaraborg** | SWS | 45 | 39 | 0.56374 | 42 | 0.6136 |
| **Blekinge** | SWB | 40 | 29 | 0.59128 | 36 | 0.62316 |
| **Galicia** | GAL | 88 | 75 | 0.59148 | | |
| **Ibiza** | IBZ | 54 | 31 | 0.55667 | | |
| **Majorca** | MAJ | 62 | 51 | 0.53797 | | |
| **Minorca** | MIN | 36 | 28 | 0.45143 | | |
| **Southern Portugal** | SPO | 77 | 71 | 0.62816 | 223 | 0.6118 |
| **Valencia** | VAL | 73 | 65 | 0.56693 | | |
| **Wales Abergele** | WAB | 26 | 25 | 0.54215 | 24 | 0.57978 |
| **Wales Llangefni** | WLG | 32 | 25 | 0.37641 | 32 | 0.42729 |
| **Wales Llanuwchllyn** | WLU | 30 | 23 | 0.3154 | 29 | 0.39264 |
| **Wales Tregaron** | WTR | 30 | 21 | 0.37655 | 27 | 0.40199 |
| **Chippenham** | CHP | 55 | 48 | 0.53616 | 52 | 0.57863 |
| **Winchester** | WIN | 35 | 30 | 0.52739 | 35 | 0.56571 |
| **Austria East Tyrol** | AET | 137 | 103 | | 103 | 0.61954 |
| **Northern Italy** | NIT | 154 | 151 | | 151 | 0.63035 |
| **Poland** | POL | 255 | 243 | | 243 | 0.56991 |
| **Total/Mean** | | 3586 | | 0.54128 | | 0.57112 |

## Appendix B6: Graph showing comparison of STR genetic Diversity for populations of Western Europe.

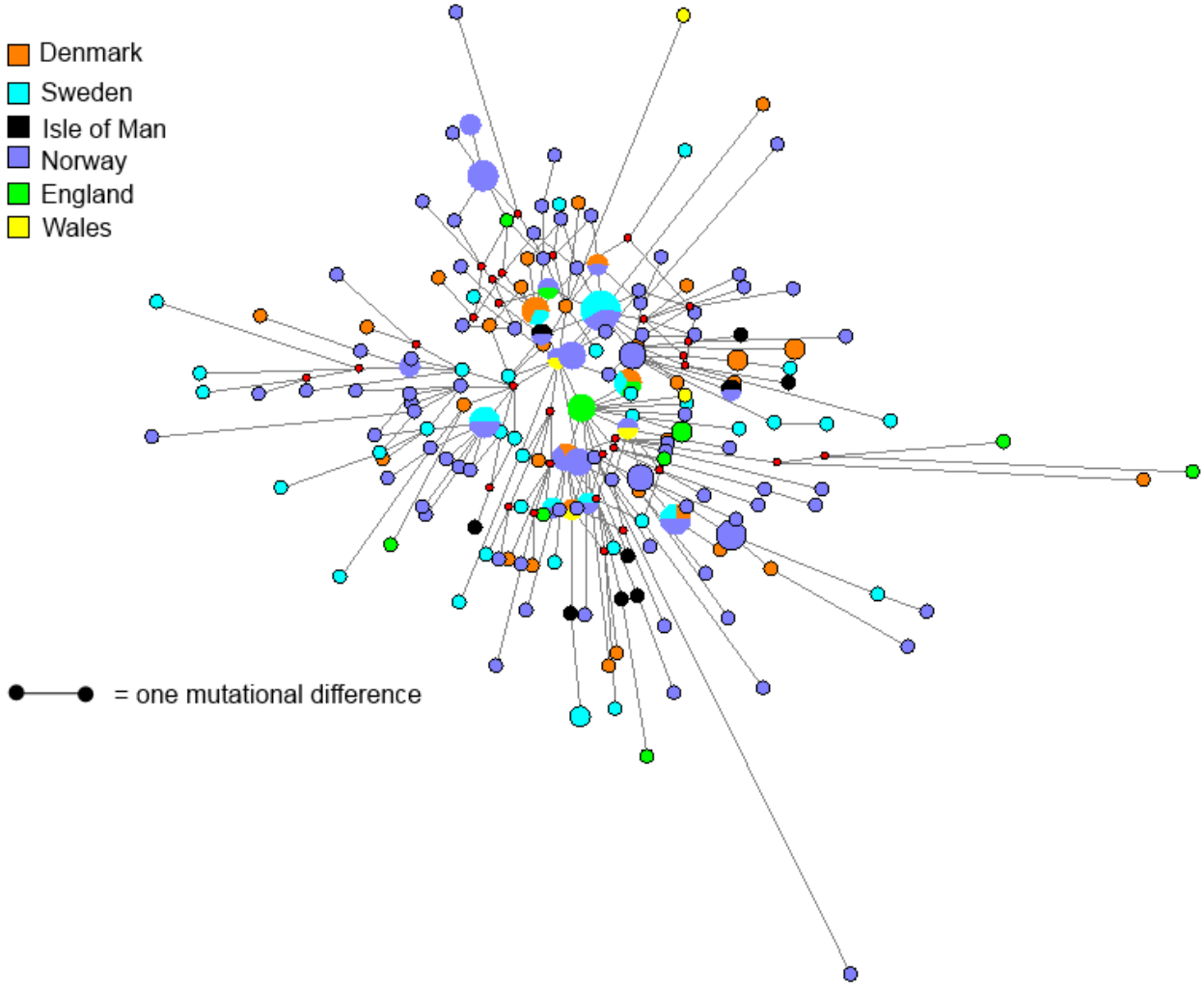


## Appendix B7: MDS plot for NRY 15 STRs based on RST values for 19 European populations.

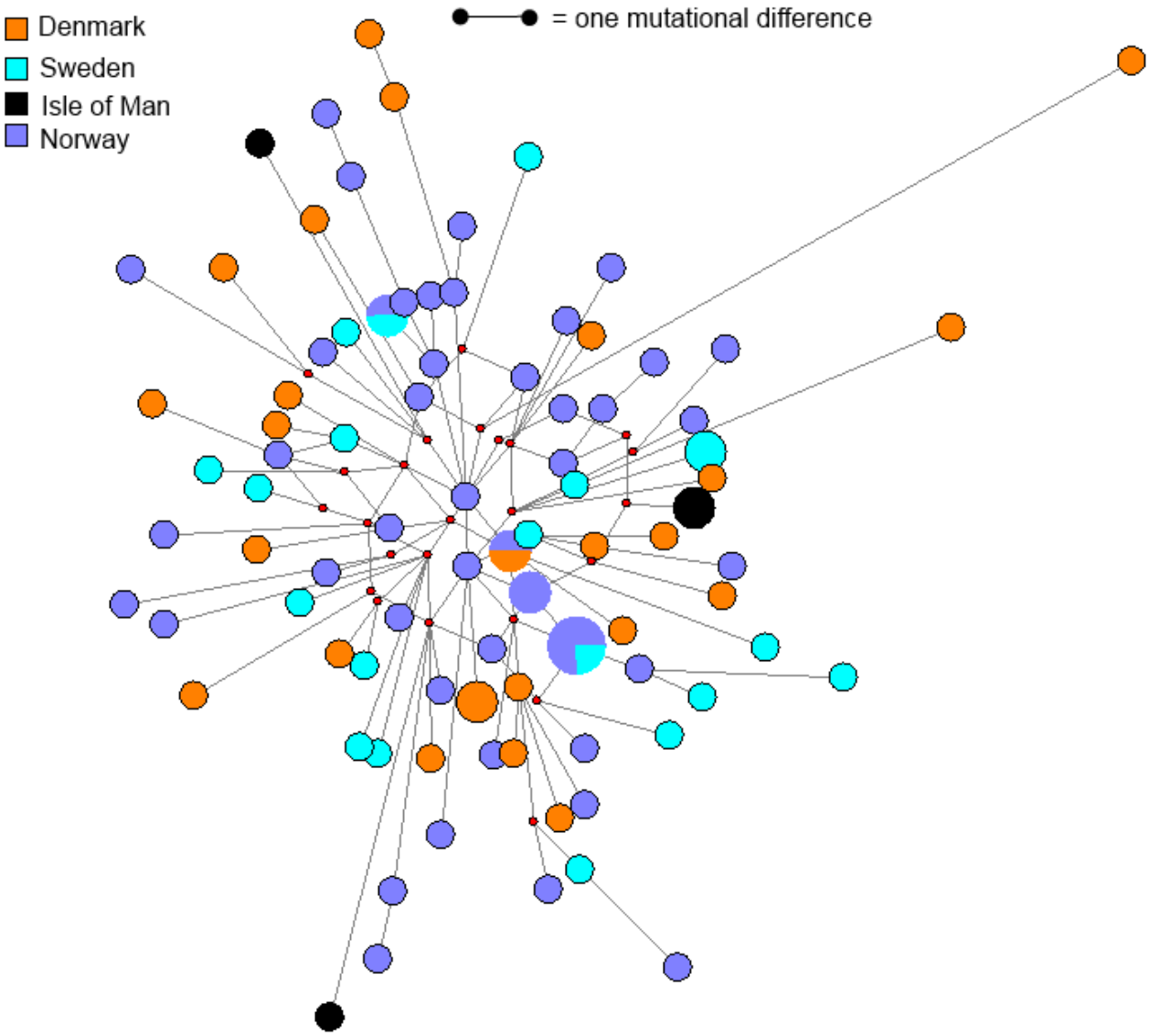



Populations contained within the red line indicate those which are not significantly different from the Isle of Man (P value = 0.05), meaning they are genetically indistinguishable from the Isle of Man. S-Stress indicates how well the plot fits the data, < 0.1 is considered a good fit (Dugard et al. 2009)

## Appendix B8: Network for NRY haplotypes within haplogroup I1-M253.



## Appendix B9: Network for NRY haplogroup R1b-U106



# Appendix B10: List of Isle of Man mtDNA samples and sequence data

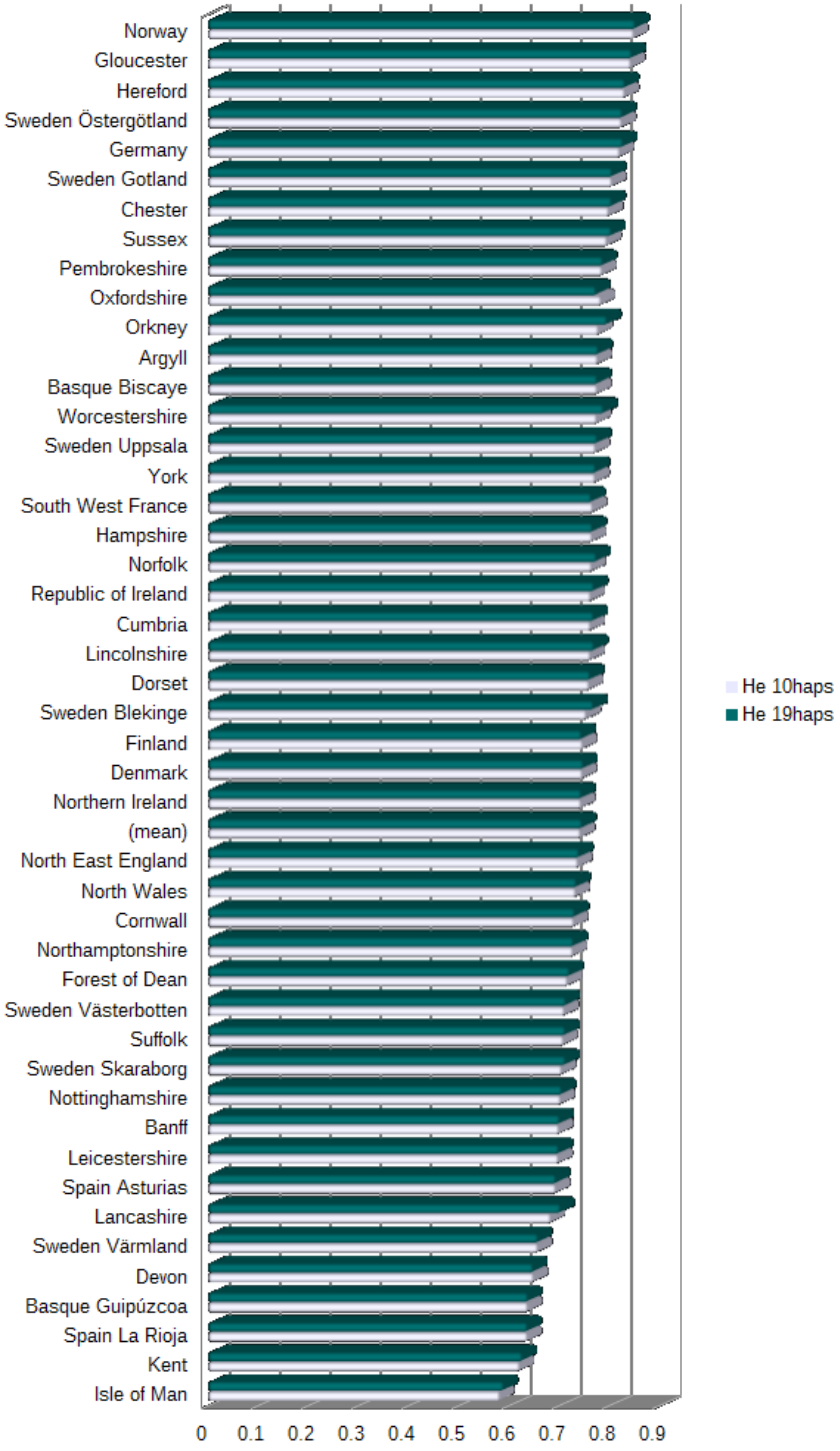
| ID | Hap. | Mutational differences from Cambridge Reference Sequences (CRS) |
|---|---|---|
| **IoM 02** | H | 16291T 16519C 263G 315.1C |
| **IoM 04** | H | 16311C 16519C 152C 263G 309.1C 315.1C |
| **IoM 05** | K | 16224C 16311C 16519C 73G 263G 315.1C 497T |
| **IoM 06** | W | 16223T 16278T 16292T 16295T 16519C 73G 118C 119C 189G 195C 198T 204C 207A 263G 315.1C |
| **IoM 07** | U | 16270T 16292T 16362C 73G 150T 263G 315.1C 517G |
| **IoM 10** | T | 16126C 16163G 16186T 16189C 16294T 16519C 73G 152C 183G 195C 263G 315.1C |
| **IoM 11** | H | 16286T 16519C 152C 263G 315.1C 523d 524d |
| **IoM 12** | H | 16092C 16140C 16293G 16311C 195C 263G 309.1C 315.1C |
| **IoM 14** | I | 16086C 16129A 16223T 16287T 16319A 16391A 16519C 73G 152C 199C 204C 207A 239C 250C 263G 309.1C 315.1C 524.1A 524.2C524.3A 524.4C 573.1C 573.2C 573.3C 573.4C 573.5C 573.6C |
| **IoM 15** | H | 16274A 16519C 146C 263G 309.1C 309.2C 315.1C 524.1A 524.2C |
| **IoM 16** | H | 16256T 16311C 16519C 152C 263G 309.1C 315.1C |
| **IoM 18** | H | 16519C 152C 263G 315.1C |
| **IoM 19** | H | 16519C 152C 263G 315.1C 524.1A 524.2C |
| **IoM 22** | H | 16248T 16362C 41T 239C 263G 297G 315.1C |
| **IoM 23** | I | 16129A 16172C 16223T 16311C 16391A 16519C 73G 199C 203A 204C 250C 263G 315.1C 455.1T 573.1C 573.2C 573.3C 573.4C 573.5C573.6C |
| **IoM 24** | H | 16519C 263G 315.1C |
| **IoM 25** | H | 263G 309.1C 315.1C |
| **IoM 26** | H | 309.1C 315.1C |
| **IoM 27** | H | 16311C 16519C 152C 263G 309.1C 315.1C |
| **IoM 29** | H | 16235G 16291T 16293G 16400T 263G 309.1C 315.1C |
| **IoM 32** | J | 16069T 16126C 73G 152C 185A 228A 263G 295T 309.1C 315.1C 462T 489C |
| **IoM 35** | H | 16086C 73G 263G 309.1C 315.1C 523d 524d |
| **IoM 38** | K | 16311C 16320T 16519C 73G 146C 152C 263G 315.1C 498d 573.1C573.2C |
| **IoM 39** | H | 16311C 16519C 152C 263G 309.1C 315.1C |
| **IoM 43** | H | 16192T 16519C 182T 195C 263G 315.1C |
| **IoM 52** | J | 16069T 16126C 73G 152C 185A 228A 263G 295T 309.1C 315.1C 462T 489C |
| **IoM 53** | H | 16311C 16519C 152C 263G 309.1C 315.1C |
| **IoM 55** | T | 16093C 16126C 16294T 16296T 16304C 16519C 73G 199C 263G309.1C 315.1C |
| **IoM 57** | T | 16126C 16294T 16296T 16304C 16519C 73G 263G 309.1C 315.1C |
| **IoM 58** | W | 16223T 16278T 16292T 16295T 16519C 73G 118C 119C 189G 195C 198T 204C 207A 263G 315.1C |
| **IoM 61** | H | 16274A 16519C 146C 263G 309.1C 309.2C 315.1C 524.1A 524.2C |
| **IoM 65** | H | 16291T 16519C 263G 315.1C |
| **IoM 66** | H | 16286T 16519C 152C 263G 315.1C 523d 524d |

Converted from sequences using mtDNA profiler (Yang et al. 2013; Yonsei DNA Profiling Group 2013)

# Appendix B11: Table of mtDNA genetic diversity and nucleotide diversity values

| | | | 10 haplotypes | | 19 haplotypes | | Nucleotide Diversity | | |
|---|---|---|---|---|---|---|---|---|---|
| | ID | N | k | He | k | He | k | 634bp | 1145bp |
| Isle of Man | IoM | 33 | 6 | 0.5814 | 7 | 0.589 | 55 | 0.0116 | 0.0073 |
| Argyll | ARG | 53 | 8 | 0.7765 | 9 | 0.7779 | | | |
| Banff | BAN | 71 | 7 | 0.6994 | 8 | 0.7006 | | | |
| Chester | CHE | 54 | 8 | 0.7987 | 10 | 0.8043 | | | |
| Cornwall | COR | 118 | 7 | 0.7291 | 9 | 0.7307 | | | |
| Cumbria | CUM | 286 | 9 | 0.7622 | 12 | 0.7684 | | | |
| Devon | DEV | 99 | 7 | 0.6489 | 9 | 0.6477 | | | |
| Dorset | DOR | 46 | 7 | 0.7594 | 8 | 0.7614 | | | |
| Forest of Dean | FOD | 60 | 7 | 0.7164 | 9 | 0.7203 | | | |
| Gloucester | GLO | 49 | 7 | 0.8427 | 7 | 0.8444 | | | |
| Hampshire | HAM | 49 | 8 | 0.7653 | 9 | 0.7662 | | | |
| Hereford | HER | 32 | 7 | 0.8306 | 7 | 0.8306 | | | |
| Kent | KEN | 110 | 7 | 0.6209 | 9 | 0.6262 | | | |
| Lancashire | LAN | 45 | 7 | 0.6828 | 8 | 0.7021 | | | |
| Leicestershire | LEI | 91 | 7 | 0.6982 | 8 | 0.6989 | | | |
| Lincolnshire | LIN | 160 | 9 | 0.762 | 13 | 0.7694 | | | |
| North East England | NEE | 243 | 8 | 0.7373 | 10 | 0.7394 | | | |
| Northern Ireland | NIR | 69 | 8 | 0.7434 | 9 | 0.7438 | | | |
| Norfolk | NOR | 128 | 8 | 0.7644 | 11 | 0.7739 | | | |
| Nottinghamshire | NOTT | 84 | 7 | 0.7025 | 9 | 0.7048 | | | |
| Pembrokeshire | PEM | 68 | 7 | 0.7853 | 8 | 0.7867 | | | |
| Northamptonshire | NTH | 58 | 8 | 0.7272 | 9 | 0.7284 | | | |
| North Wales | NWA | 98 | 8 | 0.7328 | 9 | 0.7341 | | | |
| Orkney | ORK | 136 | 7 | 0.778 | 9 | 0.795 | | | |
| Oxfordshire | OXF | 140 | 7 | 0.7831 | 9 | 0.773 | | | |
| Suffolk | SUF | 114 | 7 | 0.7083 | 9 | 0.712 | | | |
| Sussex | SUS | 91 | 7 | 0.7951 | 10 | 0.8037 | | | |
| Worcestershire | WOR | 39 | 6 | 0.7746 | 7 | 0.7881 | | | |
| York | YOR | 234 | 8 | 0.7714 | 11 | 0.7724 | | | |
| Republic of Ireland | IRE | 50 | 7 | 0.7624 | 9 | 0.769 | 38 | 0.0105 | 0.0067 |
| Norway | NWY | 40 | 8 | 0.8487 | 9 | 0.8513 | 43 | 0.0121 | 0.0081 |
| South West France | SWF | 164 | 9 | 0.7666 | 11 | 0.7636 | 106 | 0.0114 | 0.0075 |
| Spain La Rioja | SLR | 52 | 7 | 0.6365 | 7 | 0.6365 | 39 | 0.01 | 0.0064 |
| Sweden Blekinge | BLE | 38 | 8 | 0.7539 | 9 | 0.7679 | 33 | 0.0118 | 0.0078 |
| Sweden Gotland | GOT | 40 | 9 | 0.8051 | 9 | 0.8051 | 36 | 0.0136 | 0.009 |
| Sweden Uppsala | UPP | 54 | 9 | 0.7722 | 11 | 0.7757 | 48 | 0.0124 | 0.0078 |
| Sweden Östergötland | OST | 40 | 9 | 0.8244 | 10 | 0.8256 | 38 | 0.0143 | 0.0088 |
| Sweden Skaraborg | SKA | 41 | 7 | 0.7049 | 8 | 0.7122 | 31 | 0.0117 | 0.0073 |
| Sweden Värmland | VAR | 41 | 7 | 0.6573 | 7 | 0.6573 | 37 | 0.0124 | 0.0081 |
| Sweden Västerbotten | VAS | 40 | 6 | 0.7103 | 7 | 0.7128 | 34 | 0.0111 | 0.0075 |
| Basque Biscaye | BAB | 90 | 9 | 0.7755 | 9 | 0.7755 | 64 | 0.0105 | |
| Basque Guipúzcoa | BAG | 113 | 8 | 0.6372 | 9 | 0.6373 | 48 | 0.0079 | |
| Spain Asturias | SAS | 76 | 6 | 0.693 | 8 | 0.6933 | 61 | 0.009 | |
| Denmark | DEN | 201 | 9 | 0.7461 | 14 | 0.7478 | 156 | 0.0124 | |
| Germany | GER | 213 | 9 | 0.8205 | 15 | 0.8271 | 163 | 0.0135 | |
| Finland | FIN | 200 | 8 | 0.7467 | 11 | 0.7443 | 120 | 0.0118 | |
| Leicester | LEI | 43 | | | | | 42 | 0.0138 | 0.0091 |
| Mean | | | | **0.7422** | | **0.7455** | | **0.01167** | **0.0078** |

## Appendix B12: Graph showing mtDNA haplogroup diversity at two resolutions.



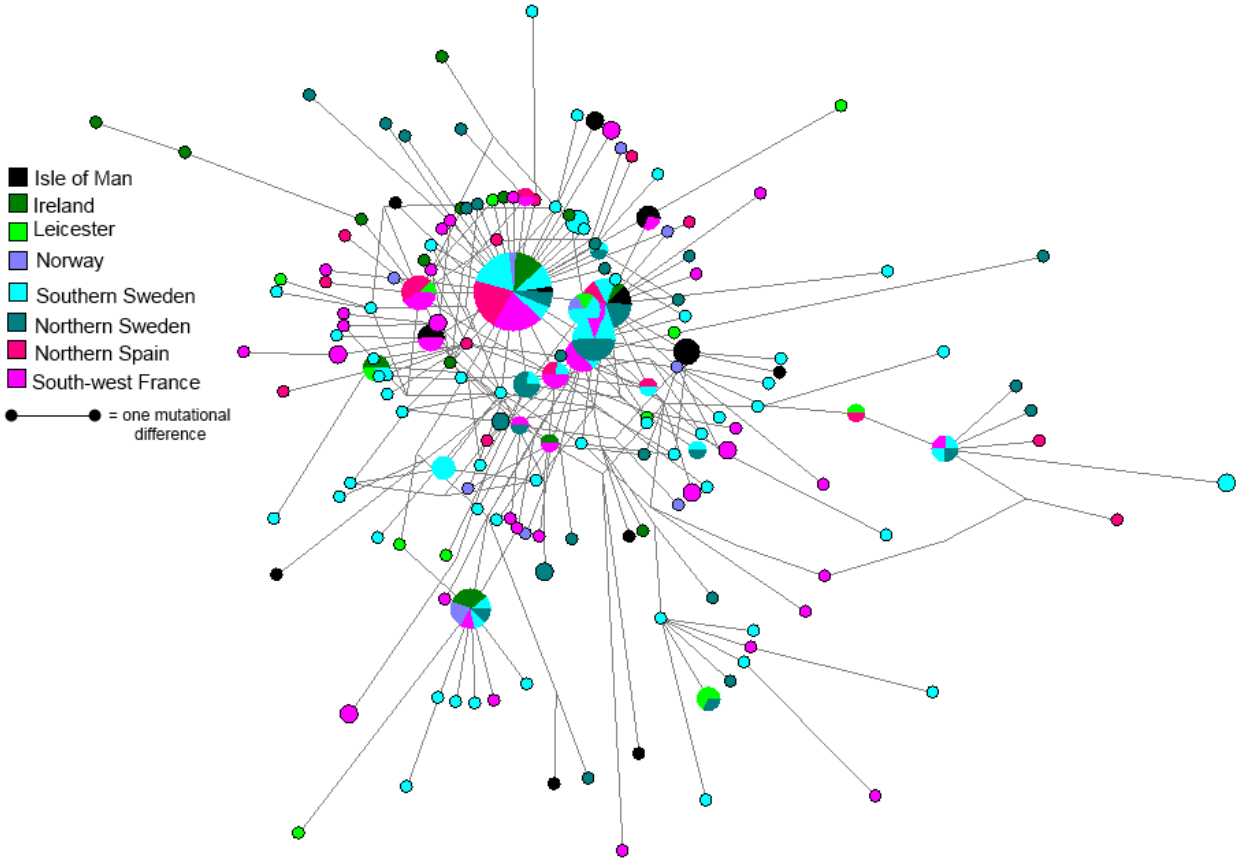
# Appendix B13: Table of mtDNA haplogroup counts for each population.

| | L | M* | C | Z | D | G | N* | A | I | W | X | R* | B* | J | T | HV | H | U | K | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IOM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 0 | 2 | 3 | 0 | 21 | 1 | 2 | 33 |
| IRE | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 0 | 0 | 3 | 5 | 3 | 22 | 8 | 4 | 50 |
| NOR | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 7 | 6 | 11 | 7 | 2 | 40 |
| SWF | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 7 | 1 | 0 | 11 | 8 | 14 | 64 | 40 | 15 | 15 | 178 |
| SLR | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 4 | 5 | 0 | 30 | 8 | 3 | 1 | 53 |
| BLE | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 1 | 5 | 2 | 3 | 17 | 7 | 1 | 39 |
| GOT | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 7 | 2 | 2 | 14 | 9 | 2 | 2 | 42 |
| UPP | 2 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 3 | 3 | 4 | 24 | 6 | 6 | 54 |
| OST | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 5 | 5 | 1 | 14 | 7 | 3 | 40 |
| SKA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 3 | 1 | 0 | 2 | 6 | 0 | 21 | 3 | 3 | 41 |
| VAR | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 2 | 1 | 23 | 4 | 6 | 41 |
| VAS | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 4 | 3 | 3 | 20 | 7 | 0 | 0 | 40 |
| ARG | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 0 | 5 | 5 | 4 | 22 | 10 | 3 | 53 |
| BAN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 0 | 0 | 12 | 6 | 1 | 36 | 6 | 6 | 71 |
| CHE | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 1 | 0 | 0 | 9 | 4 | 5 | 21 | 4 | 5 | 54 |
| COR | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 3 | 0 | 0 | 22 | 9 | 4 | 54 | 17 | 6 | 118 |
| CUM | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 5 | 4 | 4 | 2 | 1 | 31 | 43 | 13 | 119 | 41 | 22 | 286 |
| DEV | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 3 | 0 | 0 | 7 | 7 | 6 | 57 | 6 | 9 | 99 |
| DOR | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 6 | 5 | 1 | 20 | 8 | 3 | 46 |
| FOD | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 1 | 0 | 0 | 12 | 3 | 2 | 29 | 5 | 4 | 60 |
| GLO | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 2 | 9 | 5 | 13 | 8 | 6 | 49 |
| HAM | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 7 | 0 | 5 | 21 | 8 | 4 | 49 |
| HER | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 4 | 5 | 2 | 11 | 3 | 3 | 32 |
| KEN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 4 | 2 | 0 | 0 | 7 | 10 | 2 | 65 | 13 | 3 | 110 |
| LAN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 2 | 6 | 1 | 24 | 6 | 6 | 47 |
| LEI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 1 | 0 | 0 | 0 | 10 | 5 | 5 | 46 | 16 | 5 | 91 |
| LIN | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 6 | 2 | 4 | 1 | 0 | 20 | 17 | 4 | 68 | 21 | 14 | 160 |
| NEE | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 3 | 4 | 0 | 0 | 28 | 33 | 6 | 110 | 30 | 25 | 243 |
| NIR | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 10 | 2 | 2 | 30 | 14 | 8 | 69 |
| NOR | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 5 | 7 | 1 | 0 | 2 | 11 | 17 | 3 | 54 | 17 | 10 | 128 |
| NOT | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 2 | 0 | 0 | 8 | 12 | 1 | 42 | 12 | 4 | 84 |
| PEM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 0 | 0 | 7 | 6 | 6 | 27 | 6 | 12 | 68 |
| NTH | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 10 | 4 | 4 | 28 | 5 | 3 | 58 |
| NWA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 1 | 0 | 1 | 19 | 7 | 2 | 45 | 10 | 7 | 98 |
| ORK | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 1 | 12 | 0 | 0 | 18 | 10 | 0 | 51 | 24 | 9 | 136 |
| OXF | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 3 | 3 | 0 | 0 | 17 | 16 | 11 | 59 | 11 | 16 | 140 |
| SUF | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 4 | 1 | 0 | 0 | 10 | 11 | 7 | 57 | 16 | 4 | 114 |
| SUS | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 4 | 2 | 0 | 0 | 10 | 12 | 6 | 35 | 10 | 8 | 91 |
| WOR | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 2 | 0 | 0 | 5 | 3 | 0 | 16 | 5 | 3 | 39 |
| YOR | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 5 | 2 | 1 | 0 | 0 | 29 | 27 | 19 | 95 | 37 | 17 | 234 |
| SAS | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 7 | 3 | 7 | 39 | 13 | 5 | 76 |
| BAB | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 1 | 4 | 12 | 10 | 37 | 14 | 5 | 90 |
| BAG | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 3 | 4 | 9 | 63 | 24 | 7 | 113 |
| DEN | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 5 | 1 | 2 | 1 | 1 | 25 | 17 | 7 | 90 | 29 | 20 | 201 |
| FIN | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 4 | 5 | 5 | 5 | 0 | 9 | 12 | 9 | 80 | 55 | 13 | 200 |
| GER | 0 | 3 | 0 | 0 | 1 | 1 | 4 | 0 | 3 | 8 | 2 | 2 | 6 | 20 | 29 | 14 | 70 | 38 | 13 | 214 |
| Total | 12 | 8 | 2 | 2 | 5 | 1 | 17 | 5 | 116 | 79 | 78 | 15 | 41 | 450 | 423 | 329 | 1851 | 605 | 333 | 4372 |

# Appendix B14: Table of 1123bp sequences shared between Isle of Man and populations of western Europe.

| Isle of Man | (n=32) | IoM5 | IoM2/65 | IoM11/66 | IoM18/19 | IoM24 | IoM25 | IoM26 | IoM57 | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| **Leicester** | (n=43) | 1 | | | | | | 1 | | 2 |
| **Ireland** | (n=50) | | | | 1 | 4 | | | 2 | 7 |
| **Blekinge** | (n=40) | | | | | 2 | | | 1 | 3 |
| **Östergötland** | (n=40) | | | | | 2 | | 2 | 1 | 5 |
| **Gotland** | (n=40) | | | | | 1 | | | | 1 |
| **Skaraborg** | (n=41) | | | | 1 | 3 | | | | 4 |
| **Uppsala** | (n=54) | | | | 1 | 2 | | | | 3 |
| **Varmland** | (n=42) | | | | 1 | 1 | | | 1 | 3 |
| **Vasterbotten** | (n=40) | | | | 1 | 1 | | | | 2 |
| **Norway** | (n=40) | | | | | 1 | | | | 1 |
| **La Rioja** | (n=52) | | | | 2 | 7 | 1 | | | 10 |
| **S.W. France** | (n=164) | 2 | 2 | 1 | 4 | 7 | 4 | | | 20 |
| **Total** | (n=677) | 3 | 2 | 1 | 14 | 31 | 5 | 3 | 5 | 61 |

**Appendix B15: Network of mtDNA H haplogroup for western European samples, showing distinctive star shaped cluster.**



## Appendix B16: Type I and Type II errors for NRY scenarios selected by ABC.

***Scenarios with Ireland; Scenario 5, where Isle of Man is the admixed population (0.9714; 95% CI: 0.9475 – 0.9952)***

| Scenarios | 1 | 2 | 3 | 4 | 6 | 7 | total |
|---|---|---|---|---|---|---|---|
| **Type I** | 0.08 | 0.16 | 0.14 | 0.08 | 0.09 | 0.15 | 0.7 |
| **Type II** | 0.12 | 0.09 | 0.16 | 0.02 | 0.03 | 0.11 | 0.53 |

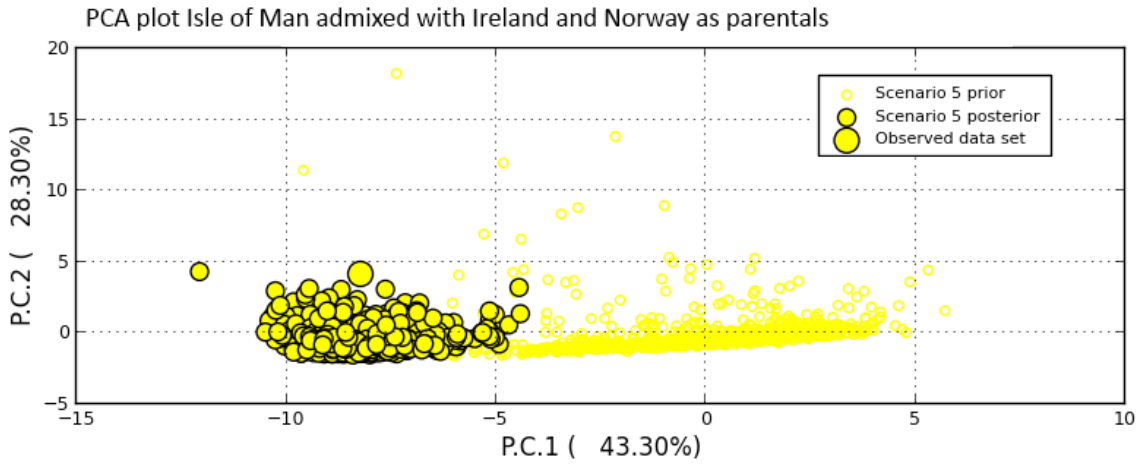
***Scenarios with Wales; Scenario 3, ancient divergence of Wales and Norway, recent divergence of Wales and Isle of Man (0.5257; 95% CI: 0.4165 – 0.6349).***

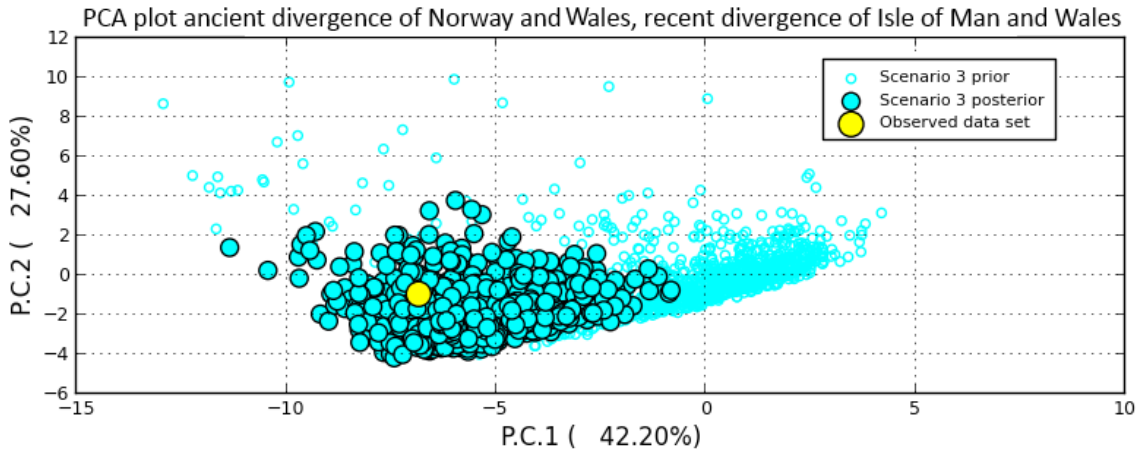
| Scenarios | 1 | 2 | 4 | 5 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|
| **Type I** | 0.01 | 0.01 | 0.01 | 0.12 | 0.04 | 0.01 | 0.2 |
| **Type II** | 0.02 | 0.01 | 0.01 | 0.23 | 0.03 | 0.01 | 0.31 |

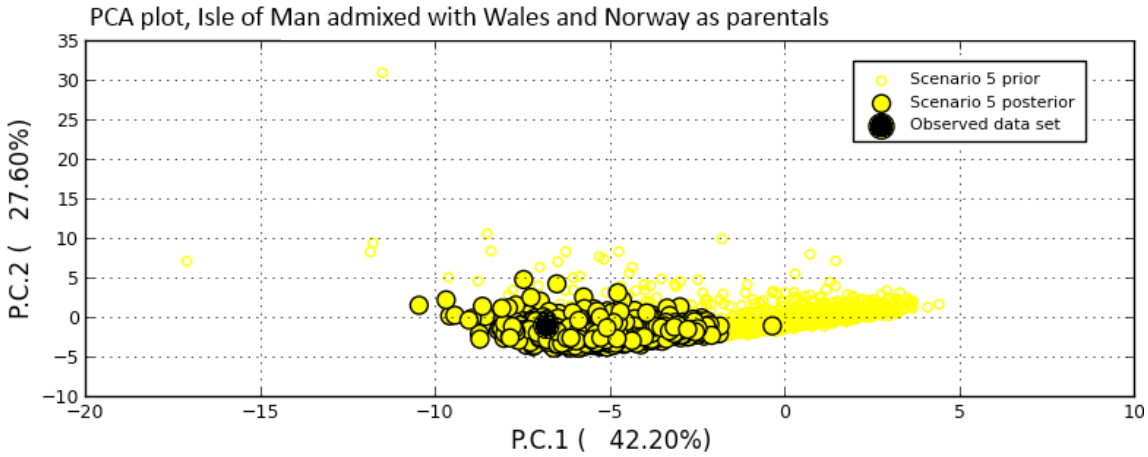
***Scenario 5, where Isle of Man is the admixed population (0.4743; 95% CI: 0.3651 – 0.5834)***

| Scenarios | 1 | 2 | 3 | 4 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|
| **Type I** | 0.15 | 0.07 | 0.23 | 0.01 | 0.06 | 0.01 | 0.53 |
| **Type II** | 0.12 | 0.07 | 0.12 | 0.01 | 0.08 | 0.02 | 0.42 |

## Appendix B17: Goodness of fit between observed NRY data and pseudo-observed data generated derived from selected scenarios.



# Appendix B18: comparisons between observed NRY Summary statistics and those generated from simulations for the three selected scenarios.

### *Isle of Man is an admixture of Ireland and Norway*

| Summary statistic | Observed Value | Proportion Sim<Obs |
|---|---|---|
| Number unique haplotypes: Ireland | 6 | 0.447 |
| Number of unique haplotypes: Isle of Man | 4.3 | 0.6635 |
| Number of unique haplotypes: Norway | 5 | 0.98 |
| Genetic Diversity: Ireland | 0.464 | 0.2395 |
| Genetic Diversity: Isle of Man | 0.5408 | 0.827 |
| Genetic Diversity: Norway | 0.5895 | 0.979 |
| $F_{ST}$ Genetic distance: Ireland/Isle of Man | 0.0232 | 0.918 |
| $F_{ST}$ Genetic distance: Ireland/Norway | 0.1572 | 0.999 |
| $F_{ST}$ Genetic distance: Norway/Isle of Man | 0.0590 | 0.9165 |
| $Dm^2$ STR genetic distance: Ireland/Isle of Man | 0.2882 | 1 |
| $Dm^2$ STR genetic distance: Ireland/Norway | 1.6128 | 1 |
| $Dm^2$ STR genetic distance: Norway/Isle of Man | 0.6224 | 1 |

### *Ancient divergence of Wales and Norway, recent divergence of Wales and Isle of Man*

| Summary statistic | Observed Value | Proportion Sim<Obs |
|---|---|---|
| Number unique haplotypes: Wales | 2.9 | 0.2450 |
| Number of unique haplotypes: Isle of Man | 4.3 | 0.4765 |
| Number of unique haplotypes: Norway | 5 | 0.026 |
| Genetic Diversity: Wales | 0.3766 | 0.122 |
| Genetic Diversity: Isle of Man | 0.5408 | 0.484 |
| Genetic Diversity: Norway | 0.5895 | 0.343 |
| $F_{ST}$ Genetic distance: Wales/Isle of Man | 0.0377 | 0.5285 |
| $F_{ST}$ Genetic distance: Wales/Norway | 0.1705 | 0.922 |
| $F_{ST}$ Genetic distance: Norway/Isle of Man | 0.0590 | 0.642 |
| $Dm^2$ STR genetic distance: Wales/Isle of Man | 0.2010 | 0.851 |
| $Dm^2$ STR genetic distance: Wales/Norway | 1.2898 | 0.959 |
| $Dm^2$ STR genetic distance: Norway/Isle of Man | 0.6224 | 0.937 |

### *Isle of Man is an admixture of Wales and Norway*

| Summary statistic | Observed Value | Proportion Sim<Obs |
|---|---|---|
| Number unique haplotypes: Wales | 2.9 | 0.222 |
| Number of unique haplotypes: Isle of Man | 4.3 | 0.2495 |
| Number of unique haplotypes: Norway | 5 | 0.0630 |
| Genetic Diversity: Wales | 0.3766 | 0.119 |
| Genetic Diversity: Isle of Man | 0.5408 | 0.3365 |
| Genetic Diversity: Norway | 0.5895 | 0.429 |
| $F_{ST}$ Genetic distance: Wales/Isle of Man | 0.0377 | 0.4575 |
| $F_{ST}$ Genetic distance: Wales/Norway | 0.1705 | 0.923 |
| $F_{ST}$ Genetic distance: Norway/Isle of Man | 0.0590 | 0.895 |
| $Dm^2$ STR genetic distance: Wales/Isle of Man | 0.2010 | 0.834 |
| $Dm^2$ STR genetic distance: Wales/Norway | 1.2898 | 0.98 |
| $Dm^2$ STR genetic distance: Norway/Isle of Man | 0.6224 | 0.986 |

## Appendix B19: Type I and Type II errors for mtDNA scenarios selected by ABC.

***Norway scenario 3: Ancient divergence of Ireland and Norway, recent divergence of Ireland and Isle of Man***

| Scenarios | 1 | 2 | 4 | 5 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|
| **Type I** | 0.05 | 0.28 | 0.1 | 0.13 | 0.13 | 0.09 | 0.78 |
| **Type II** | 0.09 | 0.09 | 0.04 | 0.14 | 0.14 | 0.05 | 0.55 |

***Norway scenario 4: Ancient divergence of Isle of Man and Ireland, recent divergence of Norway and Ireland***

| Scenarios | 1 | 2 | 3 | 5 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|
| **Type I** | 0.09 | 0.16 | 0.04 | 0.04 | 0.1 | 0.16 | 0.59 |
| **Type II** | 0.12 | 0.09 | 0.1 | 0.11 | 0.15 | 0.2 | 0.51 |

***Norway Scenario 6: Ireland is admixture of Isle of Man and Norway***

| Scenarios | 1 | 2 | 3 | 4 | 5 | 7 | Total |
|---|---|---|---|---|---|---|---|
| **Type I** | 0.11 | 0.24 | 0.14 | 0.15 | 0.07 | 0.04 | 0.76 |
| **Type II** | 0.1 | 0.1 | 0.13 | 0.1 | 0.9 | 0.04 | 0.29 |

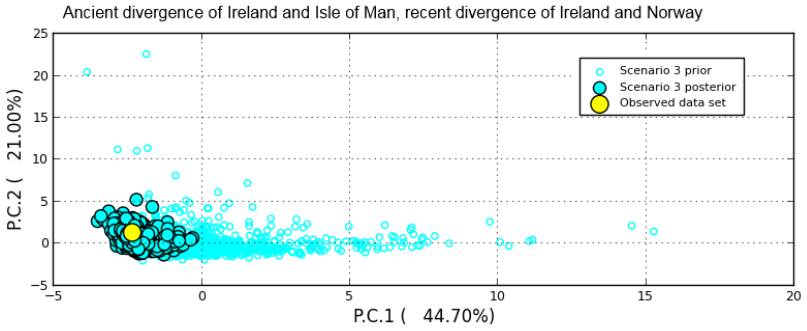
***Norway scenario 7: Norway is admixture of Ireland and Isle of Man***

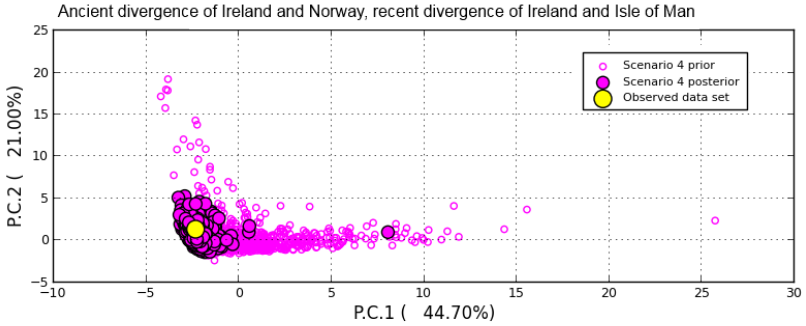
| Scenarios | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| **Type I** | 0.1 | 0.19 | 0.05 | 0.2 | 0.07 | 0.04 | 0.64 |
| **Type II** | 0.11 | 0.1 | 0.09 | 0.16 | 0.08 | 0.04 | 0.49 |

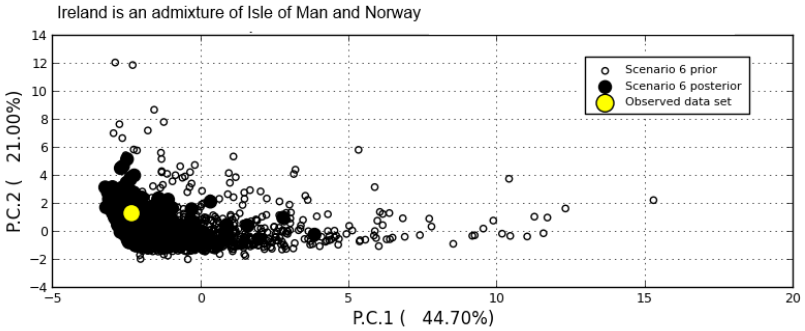
***Swedish scenario 4: Ancient divergence of Sweden and Ireland, recent Divergence of Isle of Man and Ireland***

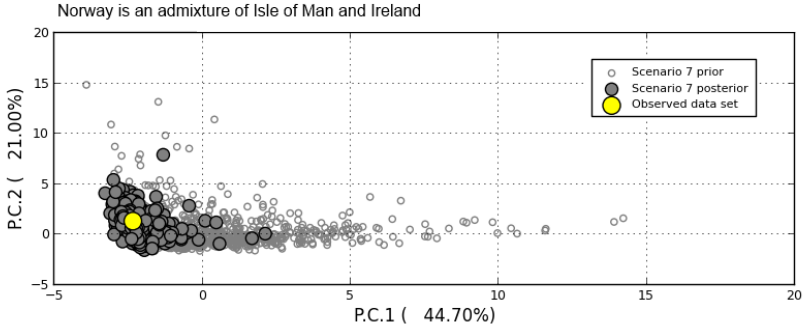
| Scenarios | 1 | 2 | 4 | 5 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|
| **Type I** | 0.05 | 0.17 | 0.05 | 0.14 | 0.16 | 0.05 | 0.62 |
| **Type II** | 0.03 | 0.05 | 0.02 | 0.18 | 0.18 | 0.16 | 0.62 |

# Appendix B20: Goodness of fit between observed mtDNA data and pseudo-observed data generated derived from selected scenarios which included Norway

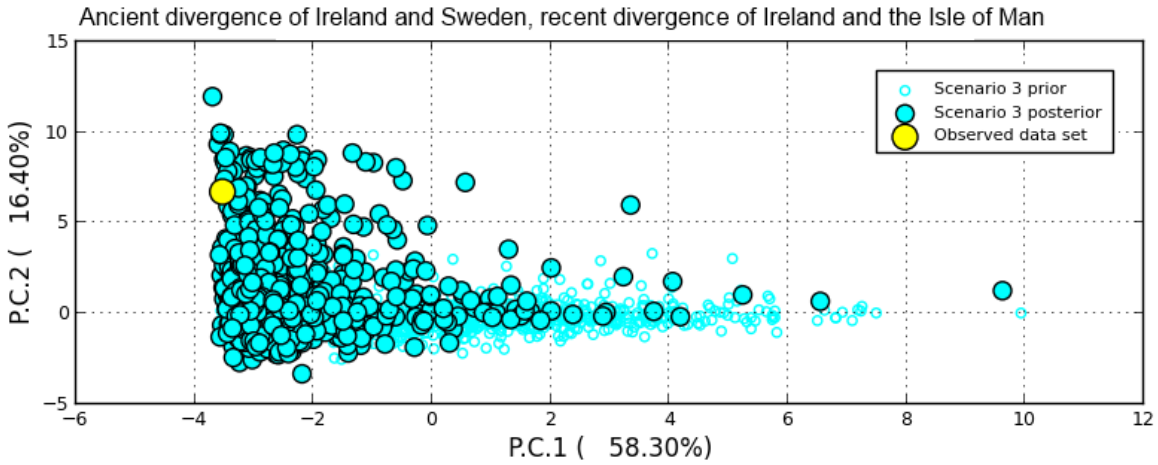


**Appendix B21: Goodness of fit between observed mtDNA data and pseudo-observed data generated derived from selected scenario which included Sweden.**



# Appendix B22: mtDNA parameter values estimated from the scenarios which had the highest posterior probabilities.

### *Norway Scenario 3: Ancient divergence of Norway and Ireland, recent divergence of Isle of Man and Ireland.*

| | Description | Prior | Mode | 95% Credible Intervals |
|---|---|---|---|---|
| **t1** | Time since Ireland/Isle of Man divergence | 28 – 80 (gens) | 30.2 | 29.1 – 78.8 |
| **t2** | Time since Ireland/Norway divergence | 100 – 440 (gens) | 121 | 103 – 218 |
| **N1** | Irish effective population size | 10 – 100,000 | 15,600 | 6,880 – 94,700 |
| **N2** | Manx effective population size | 10 – 50,000 | 2,760 | 1,210 – 43,300 |
| **N3** | Norwegian effective population size | 10 – 10,000,000 | 6,320,000 | 221,000 – 9,720,000 |
| **Na** | Ancestral effective population size | 10 – 50,000 | 121 | 102 – 211 |

### *Norway Scenario 4: Ancient divergence of Isle of Man and Ireland, recent divergence of Norway and Ireland.*

| | Description | Prior | Mode | 95% Credible Intervals |
|---|---|---|---|---|
| **t1** | Time since Ireland/Norway divergence | 28 – 80 (gens) | 59.5 | 31.5- 79.1 |
| **t2** | Time since Ireland/Isle of Man divergence | 100 – 440 (gens) | 117 | 103 – 332 |
| **N1** | Irish effective population size | 10 – 100,000 | 8,290 | 4,890 – 90,800 |
| **N2** | Manx effective population size | 10 – 50,000 | 4,520 | 1,890 – 43,200 |
| **N3** | Norwegian effective population size | 10 – 10,000,000 | 5,770,000 | 287,000 – 9,740,000 |
| **Na** | Ancestral effective population size | 10 – 50,000 | 289 | 54 – 1,320 |

### *Norway Scenario 6: Ireland is admixture of Isle of Man and Norway*

| | Description | Prior | Mode | 95% Credible Intervals |
|---|---|---|---|---|
| **t1** | Time since admixture | 28 – 80 (gens) | 72 | 29.7 – 79.4 |
| **t2** | Time since Isle of Man/Norway divergence | 100 – 440 (gens) | 105 | 101 – 210 |
| **N1** | Irish effective population size | 10 – 100,000 | 15,400 | 7,060 – 95,200 |
| **N2** | Manx effective population size | 10 – 50,000 | 3,600 | 1,810 – 42,900 |
| **N3** | Norwegian effective population size | 10 – 10,000,000 | 1,850,000 | 185,000 – 9,740,000 |
| **Na** | Ancestral effective population size | 10 – 50,000 | 413 | 60 – 1,650 |
| **r** | Admixture, Manx contribution to Ireland | 0.001 – 0.999 | 0.755 | 0.264 – 0.978 |

### *Norway scenario 7: Norway is admixture of Ireland and Isle of Man*

| | Description | Prior | Mode | 95% Credible Intervals |
|---|---|---|---|---|
| **t1** | Time since admixture | 28 – 80 (gens) | 43.9 | 29.5 – 78.8 |
| **t2** | Time since Ireland/Isle of Man divergence | 100 – 440 (gens) | 128 | 103 – 330 |
| **N1** | Irish effective population size | 10 – 100,000 | 8,620 | 4,430 – 88,700 |
| **N2** | Manx effective population size | 10 – 50,000 | 5,180 | 2,090 – 43,300 |
| **N3** | Norwegian effective population size | 10 – 10,000,000 | 1,200,000 | 212,000 – 9,760,000 |
| **Na** | Ancestral effective population size | 10 – 50,000 | 304 | 62 – 1,330 |
| **r** | Admixture, Irish contribution to Norway | 0.001 – 0.999 | 0.554 | 0.0414 – 0.979 |

# GLOSSARY

**Admixture** The mixing of two or more populations to form a hybrid population.

**Allele** Alternative forms of a gene or haplotype.

**Autosome** All of the chromosome which are not involved in sex determination

**Base-pair A** single unit of DNA; a pair of nucleotides bonded across in the double-helix which constitute the building blocks of DNA and spell out the genetic code.

**Cline** A statistically significant gradient of gene frequencies over a geographical region.

**Coalescent Theory** Mathematical theory which considers lineages backwards in time, allowing extinct lineages to be ignored, thus simplifying models.

**Demic Diffusion** Demographic population growth which expands over a geographical area as the population increases.

**Haplogroup** A set of closely related Y chromosome or mtDNA haplotypes defined by a relatively slow and stable mutational marker.

**Haplotype** A particular combination of mutational states on a single DNA molecule or chromosome.

**Holocene** Geological epoch beginning approximately 10,000 years ago.

**Homoplasy** Mutations shared between lineages which is not derived from a common ancestor, but has arisen on separate occasions. Sometimes called convergence

**Hypergamy** The practise of marrying into a higher social group.

**Kilo-base (kb)** One thousand base-pairs.

**Mega-base (Mb)** One million base-pairs.

**Mesolithic Age** Meaning 'middle stone age', the period of human pre-history beginning after the last ice age, also marks the beginning of the Holocene.

**Mitochondrion** The cellular organelle concerned with energy generation, inherited down the female line.

**Mitochondrial DNA (mtDNA)** The circular genome of mitochondria.

**Most Recent Common Ancestor (MRCA)** The single ancestor of all extant lineages.

**Natural Selection** The differential contribution of individuals of one generation to the next on the basis of their ability to survive or reproduce.

**Neolithic Age** Meaning 'new stone age', this is the period of human pre-history when agriculture was first adopted, dates for the beginning of this period differ between regions as they are dependant on the uptake of the practice.

**Neolithic Package** New technologies and materials adopted by people during the Neolithic Age, includes pottery, domesticated animals and plants, as well as new styles of stone tools.

**Next Generation Sequencing (NGS)** New methods of DNA sequencing which supersedes traditional Sanger methods for establishing the order of the four nucleotides bases; adenine, cytocine, guanine and thymine, which make up the genetic code. The target sequence is broken down into small fragments and which are then sequenced in parallel making it much faster and cheaper than traditional methods. The small fragment length required by these methods make it

particularly suitable for sequencing small fragments of ancient DNA.

**Non-Recombining portion of the Y chromosome (NRY)** The majority of the Y chromosome which does not undergo recombination with its sister X chromosome during germ-line cell division.

**Palaeolithic Age** Meaning 'old stone age', this is period of pre-history, encompassing all the earliest phases of human evolution when stone tools were used, up until approximately 10,000 years ago

**Patrilocality** The practice of males staying near to their birthplace whilst females move to the home of their partners. Opposite to matrilocality.

**Phylogeny** A tree-like structure representing the evolutionary relationship of a set of organisms, haplotypes or haplogroups

**Phylogeography** A set of methodologies which analyse the geographical distribution of different branches within a phylogeny

**Polymerase Chain Reaction (PCR)** The exponential amplification of a specific region of DNA through a cyclical process of heating and cooling.

**Polymorphic/polymorphism** A gene or sequence of DNA having two or more forms at significant levels within a population.

**Principal Components Analysis (PCA)** A type of analysis which allows multidimensional data to displayed visually, with minimal loss of information.

**Recombination** The exchange of DNA between homologous chromosomes during germ-line cell division.

**Y chromosome** One of the sex chromosomes, only found in men (together with one X chromosome), and male-determining.

# BIBLIOGRAPHY

Adams, S.M., Bosch, E., Balaresque, P.L., Ballereau, S.J., Lee, A.C., Arroyo, E., López-Parra, A.M., Aler, M., Grifo, M.S.G., Brion, M., Carracedo, A., Lavinha, J., Martínez-Jarreta, B., Quintana-Murci, L., Picornell, A., Ramon, M., Skorecki, K., Behar, D.M., Calafell, F., Jobling, M.A., 2008. The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *American Journal of Human Genetics* 83, 725–736.

Aeschbacher, S., Beaumont, M.A., Futschik, A., 2012. A Novel Approach for Choosing Summary Statistics in Approximate Bayesian Computation. *Genetics* 192, 1027–1047.

Ahern, D., 2005. Dáil Debates Official Report. *Parliamentary Debates Official Report.* 606.

Aldenderfer, M., 1991. The analytical engine: Computer simulation and archaeological research. *Archaeological Method and Theory* 3, 195–247.

Ammerman, A., 1989. On the Neolithic transition in Europe: a comment on Zvelebil & Zvelebil (1988). *Antiquity* 63(283), 162-165.

Ammerman, A.J., Cavalli-Sforza, L.L., 1971. Measuring the Rate of Spread of Early Farming in Europe. *Man* 6, 674.

Amorim, A., 1999. Archaeogenetics. Journal of Iberian Archaeology 1, 15–25.

Amos, W., Jow, H., Burrows, N., 2008. 'A Combined Markov Chain Monte Carlo and Simulation-based Approach for Inferring the Colonization History of the British Isles.', in: Matsumura, S., Forster, P., Renfrew, C. (Eds.), *Simulations, Genetics and Human Prehistory*. Cambridge: McDonald Institute for Archaeological Research, pp. 49-57.

Athey, W., 2013. Whit Athey's Haplogroup Predictor. Available at http://www.hprg.com/hapest5/hapest5a/hapest5.htm?order=num (accessed 6.5.14).

Balaresque, P., Bowden, G.R., Adams, S.M., Leung, H.-Y., King, T.E., Rosser, Z.H., Goodwin, J., Moisan, J.-P., Richard, C., Millward, A., Demaine, A.G., Barbujani, G., Previderè, C., Wilson, I.J., Tyler-Smith, C., Jobling, M.A., 2010. A predominantly neolithic origin for European paternal lineages. *PLoS Biology* 8 (1), e1000285.

Ballantyne, K.N., Goedbloed, M., Fang, R., Schaap, O., Lao, O., Wollstein, A., Choi, Y., van Duijn, K., Vermeulen, M., Brauer, S., Decorte, R., Poetsch, M., von Wurmb-Schwark, N., de Knijff, P., Labuda, D., Vézina, H., Knoblauch, H., Lessig, R., Roewer, L., Ploski, R., Dobosz, T., Henke, L., Henke, J., Furtado, M.R., Kayser, M., 2010. Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *American Journal of Human Genetics.* 87, 341–353.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Bilogy and Evolution*. 16, 37–48.

Bandelt, H.J., Forster, P., Sykes, B.C., Richards, M.B., 1995. Mitochondrial portraits of human populations using median networks. *Genetics* 141, 743–753.

Barbujani, G., Bertorelle, G., Chikhi, L., 1998. Evidence for Paleolithic and Neolithic gene flow in Europe. *American Journal of Human Genetics* 62, 488–492.

Barbujani, G., Chikhi, L., 2006. DNAs from the European Neolithic. *Heredity* 97, 84–85.

Barbujani, G., Goldstein, D.B., 2004. Africans and Asians abroad: genetic diversity in Europe.

*Annual Review Genomics and Human Genetics* 5, 119–150.

Barbujani, G., Sokal, R.R., Oden, N.L., 1995. Indo-European origins: a computer-simulation test of five hypotheses. *American Journal of Physical Anthropology* 96, 109–132

Barton, N., 2009. The Lateglacial or Latest Palaeolithic Occupation of Britain., in: Hunter, J., Ralston, I. (Eds.), *The Archaeology of Britain: An Introduction from Earliest Times to the Twenty-First Century.* London; New York: Routledge, pp. 18-52

Beaumont, M., 2008. Joint determination of topology, divergence time, and immigration in population trees., in: Shuichi Matsumura, Forster, P., Renfrew, C. (Eds.), *Simulations, Genetics and Human Prehistory.* Cambridge: McDonald Institute for Archaeological Research, pp. 135–154.

Beaumont, M.A., Zhang, W., Balding, D.J., 2002. Approximate Bayesian computation in population genetics. *Genetics* 162, 2025–2035.

Behar, D.M., Harmant, C., Manry, J., van Oven, M., Haak, W., Martinez-Cruz, B., Salaberria, J., Oyharçabal, B., Bauduer, F., Comas, D., Quintana-Murci, L., 2012. The Basque Paradigm: Genetic Evidence of a Maternal Continuity in the Franco-Cantabrian Region since Pre-Neolithic Times. *The American Journal of Human Genetics* 90, 486–493.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., Haussler, D., 2004. Ultraconserved elements in the human genome. *Science* 304, 1321–1325.

Bellwood, P., Renfrew, C. (Eds.), 2002. *Examining the farming/language dispersal hypothesis*. Cambridge: McDonalds Institute for Archaeological Research

Bersu, G., 1977. *Three Iron Age round houses in the Isle of Man.* Douglas, Isle of Man: Manx Museum and National Trust

Bersu, G., Wilson, D.M., 1966. *Three Viking graves in the Isle of Man*. Leeds: Society for Medieval Archaeology

Bertorelle, G., Benazzo, A., Mona, S., 2010. ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*. 19, 2609–2625.

Bertranpetit, J., 2000. Genome, diversity, and origins: The Y chromosome as a storyteller. *Proceedings of the National Academy of Sciences* 97, 6927–6929.

Bertranpetit, J., Sala, J., Calafell, F., Underhill, P.A., Moral, P., Comas, D., 1995. Human mitochondrial DNA variation and the origin of Basques. *Annals of Human Genetics* 59, 63–81.

Binford, L.R., 1962. Archaeology as Anthropology. *American Antiquity* 28, 217–225.

Birky, C.W., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences* 92, 11331–11338.

Birky, C.W., Fuerst, P., Maruyama, T., 1989. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* 121, 613–627.

Boquete Vilariño, L., 2012 The mitochondrial control region in the study of human populations: the Isle of Man as a case study. *MSc Molecular Genetics dissertation*, University of Leicester, UK.

Bosch, E., Calafell, F., González-Neira, A., Flaiz, C., Mateu, E., Scheil, H.-G., Huckenbeck, W., Efremovska, L., Mikerezi, I., Xirotiris, N., Grasa, C., Schmidt, H., Comas, D., 2006. Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. *Annals of Human Genetics* 70, 459–487.

Bowden, G.R., Balaresque, P., King, T.E., Hansen, Z., Lee, A.C., Pergl-Wilson, G., Hurley, E., Roberts, S.J., Waite, P., Jesch, J., Jones, A.L., Thomas, M.G., Harding, S.E., Jobling, M.A., 2008. Excavating past population structures by surname-based sampling: the genetic legacy of the Vikings in northwest England. *Molecular Bioliology and Evolution.* 25, 301–309.

Bradley, R., 2007. *The prehistory of Britain and Ireland*. Cambridge; New York: Cambridge University Press.

Bramanti, B., Thomas, M.G., Haak, W., Unterlaender, M., Jores, P., Tambets, K., Antanaitis-Jacobs, I., Haidle, M.N., Jankauskas, R., Kind, C.-J., Lueth, F., Terberger, T., Hiller, J., Matsumura, S., Forster, P., Burger, J., 2009. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science* 326, 137–140.

Brandt, G., Szécsényi-Nagy, A., Roth, C., Alt, K.W., Haak, W., in press. Human paleogenetics of Europe – The known knowns and the known unknowns. *Journal of Human Evolution*.

Broderick, G., 1999. *Language Death in the Isle of Man: An Investigation into the Decline and Extinction of Manx Gaelic as a Community Language in the Isle of Man*. Tübingen: Max Niemeyer Verlag GmbH & Co KG

Brown, K.A., Pluciennik, M., 2001. Archaeology and human genetics: Lessons for both. *Antiquity* 75, 101–106.

Burrow, S., 1997. *The Neolithic culture of the Isle of Man : a study of the sites and pottery*. Oxford: Archaeopress,

Burrow, S., Darvill, T., 1997. AMS dating of the Manx Ronaldsway Neolithic. *Antiquity* 71(272), 412-412

Burr, T., Skurikhin, A., 2013. Selecting Summary Statistics in Approximate Bayesian Computation for Calibrating Stochastic Models. *Biomed Research International* 2013.

Busby, G.B.J., Brisighelli, F., Sánchez-Diz, P., Ramos-Luis, E., Martinez-Cadenas, C., Thomas, M.G., Bradley, D.G., Gusmão, L., Winney, B., Bodmer, W., Vennemann, M., Coia, V., Scarnicci, F., Tofanelli, S., Vona, G., Ploski, R., Vecchiotti, C., Zemunik, T., Rudan, I., Karachanak, S., Toncheva, D., Anagnostou, P., Ferri, G., Rapone, C., Hervig, T., Moen, T., Wilson, J.F., Capelli, C., 2012. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. *Proceedings of the Royal Society B: Biological sciences*. 279, 884–892.

Butler, J.M., Buel, E., Crivellente, F., McCord, B.R., 2004. Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis* 25, 1397–1412.

Caballero, A., 1995. On the effective size of populations with separate sexes, with particular reference to sex-linked genes. *Genetics* 139, 1007–1011.

Cann, R.L., Stoneking, M., Wilson, A.C., 1987. Mitochondrial DNA and human evolution. *Nature* 325, 31–36.

Capelli, C., Redhead, N., Abernethy, J.K., Gratrix, F., Wilson, J.F., Moen, T., Hervig, T., Richards, M., Stumpf, M.P.H., Underhill, P.A., Bradshaw, P., Shaha, A., Thomas, M.G., Bradman, N., Goldstein, D.B., 2003. A Y chromosome census of the British Isles. *Current Biology* 13, 979–984.

Cavalli-Sforza, L.L., Menozzi, P., Piazza, A., 1996. *The History and Geography of Human Genes*:, Abridged. ed. New Jersey: Princeton University Press.

Cavalli-Sforza, L.L., Minch, E., 1997. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. *American Journal of Human Genetics* 61, 247.

Central Statistics Office, 2012. This is Ireland - Highlights from Census 2011, Part 1: Available at http://cso.ie/en/newsandevents/pressreleases/2012pressreleases/pressreleasethisisireland-highlightsfromcensus2011part1/ (accessed 7.26.14).

Chadwick, N.K., 1970. 'Early Literary contacts between Wales and Ireland', in: Donald Moore (Ed.), *The Irish Sea Province in Archaeology and History.* Aberystwyth c/o National Library of Wales: Cambrian Archaeological Association, pp. 66–77.

Chakravarti, A., 1998. It's raining SNPs, hallelujah? *Nature Genetics* 19, 216–217

Chakravarti, A., 1999. Population genetics-making sense out of sequence. *Nature Genetics*. 21, 56–60.

Champion, T., 1991. 'Theoretical Archaeology in Britain', in: Hodder, I. (Ed.), *Archaeological Theory in Europe: The Last Three Decades.* London; New York: Routledge, pp. 129 – 160.

Champion, T., 2009. 'The later Bronze Age', in: Hunter, J., Ralston, I. (Eds.), *The Archaeology of Britain: An Introduction from Earliest Times to the Twenty-First Century*. London; New York: Routledge, pp. 126–148.

Charles-Edwards, T.M., 2013. *Wales and the Britons, 350-1064*. Oxford: Oxford University Press

Charlesworth, B., Charlesworth, D., 2009. Darwin and genetics. *Genetics* 183, 757–766.

Cheshire, J.A., Longley, P.A., Singleton, A.D., 2010. The surname regions of Great Britain. *Journal of Maps* 6, 401 – 409.

Chikhi, L., 2009. Update to Chikhi et al.'s "Clinal variation in the nuclear DNA of Europeans" (1998): genetic data and storytelling-from archaeogenetics to astrologenetics? *Human Biology*. 81, 639–643.

Chikhi, L., Destro-Bisol, G., Pascali, V., Baravelli, V., Dobosz, M., Barbujani, G., 1998. Clinal variation in the nuclear DNA of Europeans. *Human Biology*. 70, 643–657.

Chikhi, L., Nichols, R.A., Barbujani, G., Beaumont, M.A., 2002. Y genetic data support the Neolithic demic diffusion model. *Proceedings of National Academy of Science*. U.S.A. 99, 11008–11013.

Cinnioğlu, C., King, R., Kivisild, T., Kalfoğlu, E., Atasoy, S., Cavalleri, G.L., Lillie, A.S., Roseman, C.C., Lin, A.A., Prince, K., Oefner, P.J., Shen, P., Semino, O., Cavalli-Sforza, L.L., Underhill, P.A., 2004. Excavating Y-chromosome haplotype strata in Anatolia. *Human Genetics* 114, 127–148.

City Population, 2014. Tregaron (Ceredigion, Wales / Cymru, United Kingdom) - population statistics, map and location. Available at http://www.citypopulation.de/php/uk-wales.php?cityid=W37000260 (accessed 7.28.14).

Clarke, D.L., 1968. *Analytical archaeology*. London: Methuen

Collard, M., Edinborough, K., Shennan, S., Thomas, M.G., 2010. Radiocarbon evidence indicates that migrants introduced farming to Britain. *Journal of Archaeological Science* 37, 866–870

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., Estoup, A., 2014. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30, 1187–1189.

Cox, M.P., Woerner, A.E., Wall, J.D., Hammer, M.F., 2008. Intergenic DNA sequences from the human X chromosome reveal high rates of global gene flow. *BMC Genetics*. 9, 76.

Crawford, B.E., 1987. *Scandinavian Scotland*. Leicester: Leicester University Press

Cróinín, D.O., 1995. *Early medieval Ireland, 400-1200*. London: Longman

Crow, J.F., 2000. The origins, patterns and implications of human spontaneous mutation. *Nature Review Genetic* 1, 40–47.

Csilléry, K., Blum, M.G.B., Gaggiotti, O.E., François, O., 2010. Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution* 25, 410–418.

Cubbon, A.M., 1971. *The art of the Manx crosses: a selection of photographs with notes.* Douglas, Isle of Man: Manx Museum and National Trust

Cubbon, M., 1983. 'The archaeology of the Vikings in the Isle of Man', in: Fell, C., Foote, P., Graham-Campbell, J., Thompson, R. (Eds.), *Viking Age in the Isle of Man - Select Papers 9th Viking Congress, Isle of Man, 1981.* London: Viking Society for Northern Research, pp. 13 – 24.

Cummings, V., 2004. 'The setting and form of Manx chambered cairns: Cultural comparisons and social interpretations', in: Cummings, V., Fowler, C. (Eds.), *The Neolithic of the Irish Sea: Materiality and Traditions of Practice.* Oxford: Oxbow, pp. 113 – 128.

Cunliffe, B., 2001. *Facing the Ocean: The Atlantic and Its Peoples, 8000 BC to AD 1500*. Oxford: Oxford University Press.

Cunliffe, B., 2011. *Europe Between the Oceans: 9000 BC-AD 1000*. New haven, Connecticut: Yale University Press.

Cunliffe, B., 2012. *Britain Begins*. Oxford: Oxford University Press

Cunliffe, B.W., Koch, J.T., 2010. Introduction, in: Cunliffe, B.W., Koch, J.T. (Eds.), *Celtic from the West: Alternative Perspectives from Archaeology, Genetics, Language, and Literature*. Oxford; Oakville, CT: Oxbow Books, pp.

Currat, M., Excoffier, L., Maddison, W., Otto, S.P., Ray, N., Whitlock, M.C., Yeaman, S., 2006. Comment on "Ongoing Adaptive Evolution of ASPM, a Brain Size Determinant in Homo sapiens" and "Microcephalin, a Gene Regulating Brain Size, Continues to Evolve Adaptively in Humans." *Science* 313, 172a–172a.

Davey, P.J., Northover, P., O'Conner, B., Woodcock, J.J., 1999. 'Bronze age Metallurgy on the Isle of Man: a symposium', in: Davey, P., White, S.D. (Eds.), *Recent Archaeological Research on the Isle of Man*. Oxford: Archaeopress, pp. 39–62.

Davies, N., 1999. *The Isles: A history*. London: Macmilan Publishers

Deguilloux, M.-F., Soler, L., Pemonge, M.-H., Scarre, C., Joussaume, R., Laporte, L., 2011. News from the west: Ancient DNA from a French megalithic burial chamber. *American Journal of Physical Anthropology* 144, 108–118.

Destro Bisol, G., Capocasa, M., Anagnostou, P., 2012. When gender matters: new insights into the relationships between social systems and the genetic structure of human populations. *Molecular Ecology* 21, 4917–4920.

Downham, C., 2014. Religious and Cultural Boundaries between Vikings and Irish: The Evidence of Conversion. *academia.edu*. Available at http://www.academia.edu/2020479 (accessed 12.30.14).

Dugard, P., Todman, J., Staines, H., 2009. *Approaching Multivariate Analysis: A Practical Introduction.* 2 edition. ed. London ; New York, NY: Routledge,

Dupanloup, I., Pereira, L., Bertorelle, G., Calafell, F., Prata, M.J., Amorim, A., Barbujani, G., 2003. A

recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. *Journal of Molecular Evolution*. 57, 85–97.

Ellis, S.G., 2003. Why the History of "the Celtic Fringe" Remains Unwritten. *European Review of History: Revue européenne d'histoire* 10, 221–231.

English Heritage National Inventory, 2014. Archaeology Data Service: Faceted Browsing. Available at http://archaeologydataservice.ac.uk/archsearch/browser.jsf (accessed 12.19.14).

Estoup, A., Jarne, P., Cornuet, J.-M., 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* 11, 1591–1604.

Evans, J.A., Chenery, C.A., Fitzpatrick, A.P., 2006. Bronze Age Childhood Migration of Individuals Near Stonehenge, Revealed by Strontium and Oxygen Isotope Tooth Enamel Analysis. *Archaeometry* 48, 309–321.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10, 564–567.

Excoffier, L., Ray, N., 2008. Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecological Evolution* 23, 347–351.

Fellows-Jensen, G., 1983. 'Scandinavian settlement in the Isle of man and North-west England: The place-name evidence'., in: Fell, C., Foote, P., Graham-Campbell, J., Thompson, R. (Eds.), *Viking Age in the Isle of Man - Select Papers 9th Viking Congress, Isle of Man, 1981.* London: Viking Society for Northern Research, pp. 37–52.

Fenner, J.N., 2005. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *American Journal Physical Anthropology* 128, 415–423.

Fisher, R.A., 1930. *The Genetical Theory Of Natural Selection*. Gloucester: Clarendon Press.

Fitch, W.M., 1971. Toward Defining the Course of Evolution: Minimum Change for a Specific Tree Topology. *Systematic Zoology* 20, 406–416.

fluxus-engineering.com, 2014. Free Phylogenetic Network Software. Available at: URL http://www.fluxus-engineering.com/netwinfo.htm (accessed 6.26.14).

François, O., Currat, M., Ray, N., Han, E., Excoffier, L., Novembre, J., 2010. Principal component analysis under population genetic models of range expansion and admixture. *Molecular Biology Evolution* 27, 1257–1268.

Freke, D., 2002. *Excavation on St Patrick's Isle, Peel, Isle of Man.* Liverpool: Liverpool University Press.

Frieman, C., 2008. Islandscapes and islandness: the prehistoric Isle of Man in the Irish Sea-scape. *Oxford Journal of Archaeology* 27, 135 – 151.

Fu, Q., Rudan, P., Pääbo, S., Krause, J., 2012. Complete mitochondrial genomes reveal neolithic expansion into Europe. *PLoS ONE* 7, e32473.

Gamble, C., Davies, W., Pettitt, P., Richards, M., 2004. Climate change and evolving human diversity in Europe during the last glacial. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 359, 243–254.

García, O., Fregel, R., Larruga, J.M., Álvarez, V., Yurrebaso, I., Cabrera, V.M., González, A.M., 2011. Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge. *Heredity (Edinburgh)* 106, 37–45

Gibson, A.M., 1982. *Beaker domestic sites: a study of the domestic pottery of the late third and early second millennia B.C. in the British Isles.* Oxford: British Archaeological Reports

Goebel, T., Waters, M.R., O'Rourke, D.H., 2008. The Late Pleistocene Dispersal of Modern Humans in the Americas. *Science* 319, 1497–1502.

Goldstein, D.B., Chikhi, L., 2002. Human migrations and population structure: what we know and why it matters. *Annual Review Genomics and Human Genetics* 3, 129–152.

Goldstein, D.B., Ruiz Linares, A., Cavalli-Sforza, L.L., Feldman, M.W., 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139, 463–471.

Goodacre, S., Helgason, A., Nicholson, J., Southam, L., Ferguson, L., Hickey, E., Vega, E., Stefánsson, K., Ward, R., Sykes, B., 2005. Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity* 95, 129–135.

Haak, W., Balanovsky, O., Sanchez, J.J., Koshel, S., Zaporozhchenko, V., Adler, C.J., Der Sarkissian, C.S.I., Brandt, G., Schwarz, C., Nicklisch, N., Dresely, V., Fritsch, B., Balanovska, E., Villems, R., Meller, H., Alt, K.W., Cooper, A., Members of the Genographic Consortium, 2010. Ancient DNA from European early neolithic farmers reveals their near eastern affinities. *PLoS Biology*. 8, e1000536.

Hallast, P., Batini, C., Zadik, D., Maisano Delser, P., Wetton, J.H., Arroyo-Pardo, E., Cavalleri, G.L., de Knijff, P., Destro Bisol, G., Dupuy, B.M., Eriksen, H.A., Jorde, L.B., King, T.E., Larmuseau, M.H., Lopez de Munain, A., Lopez-Parra, A.M., Loutradis, A., Milasin, J., Novelletto, A., Pamjav, H., Sajantila, A., Schempp, W., Sears, M., Tolun, A., Tyler-Smith, C., Van Geystelen, A., Watkins, S., Winney, B., Jobling, M.A., 2014. The Y-Chromosome Tree Bursts into Leaf: 13,000 High-Confidence SNPs Covering the Majority of Known Clades. *Molecular Biology and Evolution.* Available online.

Hampton, V.D., 2012. Viking Women in the Isle of Man. *The Medievalists.net.* Available at http://www.medievalists.net/2012/09/16/viking-women-in-the-isle-of-man/ (accessed 5.29.14).

Harding, S.E., Griffiths, D., Royles, E., 2014. *In Search of Vikings: Interdisciplinary Approaches to the Scandinavian Heritage of North-West England*. Florida, USA: CRC Press.

Hayden, E.C., 2014. Technology: The $1,000 genome. *Nature* 507, 294–295.

Hedman, M., Brandstätter, A., Pimenoff, V., Sistonen, P., Palo, J.U., Parson, W., Sajantila, A., 2007. Finnish mitochondrial DNA HVS-I and HVS-II population data. *Forensic Science International* 172, 171–178.

Helgason, A., Hickey, E., Goodacre, S., Bosnes, V., Stefánsson, K., Ward, R., Sykes, B., 2001. mtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *American Journal Human Genetics* 68, 723–737.

Helgason, A., Lalueza-Fox, C., Ghosh, S., Sigurðardóttir, S., Sampietro, M.L., Gigli, E., Baker, A., Bertranpetit, J., Árnadóttir, L., Þorsteinsdottir, U., Stefánsson, K., 2009. Sequences From First Settlers Reveal Rapid Evolution in Icelandic mtDNA Pool. *PLoS Genetics* 5, e1000343.

Hellenthal, G., Leslie, S., Winney, B., Davison, D., Boumertit, A., Day, T., Hutnik, K., Royrvik, E.C., Cunliffe, B., Wellcome Trust Case Control Consortium, Internation Multiple Sclerosis Genetics Consortium, Lawson, D.J., Falush, D., Freeman, C., Pirinen, M., Myers, S., Robinson, M., Donnelly, P., Bodmer, W., 2014. Fine-scale genetic structure of the British population. *In print*.

Henderson, J., 2011. *The Atlantic Iron Age: Settlement and Identity in the First Millennium BC*, Reprint edition. ed. London: Routledge

Hills, C.M., O'Connell, T.C., 2009. New light on the Anglo-Saxon succession: two cemeteries and their dates. *Antiquity* 83, 1096–1108.

Hodgson, J.A., Bergey, C.M., Disotell, T.R., 2010. Neandertal Genome: The Ins and Outs of African Genetic Diversity. *Current Biology* 20, R517–R519.

Hudson, R.R., Slatkin, M., Maddison, W.P., 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* 132, 583–589.

Humphreys, K., Grankvist, A., Leu, M., Hall, P., Liu, J., Ripatti, S., Rehnström, K., Groop, L., Klareskog, L., Ding, B., Grönberg, H., Xu, J., Pedersen, N.L., Lichtenstein, P., Mattingsdal, M., Andreassen, O.A., O'Dushlaine, C., Purcell, S.M., Sklar, P., Sullivan, P.F., Hultman, C.M., Palmgren, J., Magnusson, P.K.E., 2011. The Genetic Structure of the Swedish Population. PLoS One 6 (8), e22547

Hunley, K.L., Healy, M.E., Long, J.C., 2009. The global pattern of gene identity variation reveals a history of long-range migrations, bottlenecks, and local mate exchange: implications for biological race. *American Journal Physical Anthropology* 139, 35–46.

Hurst, G.D.D., Jiggins, F.M., 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society of London B: Biological Sciences* 272, 1525–1534.

IBM Corp., 2011. IBM SPSS Statistics. Armonk, New York: IBM Corp.

Isle of Man Government, 2011. Isle of Man Government - Census (Census). Economic Affairs Division- Isle of Man Government Treasury, Isle of Man. Available at http://www.gov.im/census/ (accessed 10.12.2014)

Isle of Man government, 2014. Isle of Man Government - Island Facts. *Official Isle of Man Government Website*. Available at http://www.gov.im/categories/business-and-industries/iom-key-facts-guide/island-facts/ (accessed 2.27.14).

Isle of Man Guide, 2014. Isle of Man Guide - MAPS, Parish Map. Available at http://www.iomguide.com/parishmap.php (accessed 12.11.14).

Jackson, K.H., 1953. *Language and history in early Britain: a chronological survey of the Brittonic languages, 1st to 12th c. A.D.* Edinburgh: Edinburgh University Press,

James Steel, 2008. 'Population structure and diversity indices., in: Matsumura', S., Peter Forster, Colin Renfrew (Eds.), *Simulations, Genetics and Human Prehistory*. Oxford: McDonald Institute for Archaeological Research, pp. 173–176.

Jobling, M.A., 2012. The impact of recent events on human genetic diversity. *Philosophical Transactions of Royal Society B* 367, 793–799.

Jobling, M.A., Tyler-Smith, C., 2003. The human Y chromosome: an evolutionary marker comes of age. Nature Review Genetics 4, 598–612.

Jobling, M., Hollox, E., Hurles, M., Kivislid, T., Tyler-Smith, C., 2013. *Human evolutionary genetics: origins, peoples and disease,* 2nd ed. New York: Garland Science,

Jobling, M., Hurles, M., Tyler-Smith, C., 2004. *Human evolutionary genetics: origins, peoples and disease*, 1st ed. New York: Garland Science,

Jochens, J., 2006. Fosterage, in: Schaus, M. (Ed.), *Women and Gender in Medieval Europe: An Encyclopedia.* Taylor & Francis, pp. 269–267.

Jones, H.R., 1998. *Manx Names in the Early 16th Century*. Available at http://medievalscotland.org/manxnames/jonesmanx16.shtml (accessed 6.2.14).

Karafet, T.M., Mendez, F.L., Meilerman, M.B., Underhill, P.A., Zegura, S.L., Hammer, M.F., 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research* 18, 830–838.

Kayser, M., Roewer, L., Hedman, M., Henke, L., Henke, J., Brauer, S., Krüger, C., Krawczak, M., Nagy, M., Dobosz, T., Szibor, R., de Knijff, P., Stoneking, M., Sajantila, A., 2000. Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *American Journal of Human Genetics* 66, 1580–1588.

Kermode, P.M.C., 1994. *Manx Crosses.* Facsimile of 1907 edition .Balgavies, Angus, Scotland: Pinkfoot Press

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evololution* 16, 111–120.

Kingman, J.F.C., 1982. On the Genealogy of Large Populations. *Journal of Applied Probability* 19, 27–43. doi:10.2307/3213548

King, T.E., Ballereau, S.J., Schürer, K.E., Jobling, M.A., 2006. Genetic signatures of coancestry within surnames. *Current Biology* 16, 384–388.

Kinvig, R.H., 1975. *The Isle of Man: A Social, Cultural and Political History,* 2nd Revised edition. Liverpool: Liverpool University Press.

Kloss-Brandstätter, A., Pacher, D., Schönherr, S., Weissensteiner, H., Binna, R., Specht, G., Kronenberg, F., 2011. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Human Mutation* 32, 25–32

Kneen, J.J., 1925. *Index - Manx Place-Names.* Isle of Man: Yn Cheshaght Ghailckagh (The Manx Society).

Krzewińska, M., Bjørnstad, G., Skoglund, P., Olason, P.I., Bill, J., Götherström, A., Hagelberg, E., 2015. Mitochondrial DNA variation in the Viking age population of Norway. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 370, 20130384.

Kumar, S., 2005. Molecular clocks: four decades of evolution. *Nature Review Genetics* 6, 654–662.

Kunstader, P., Buhler, R., Stephan, F.F., Westoff, C.F., 1963. Demographic variability and Preferential Marriage Paterrns. *American Journal of Physical Anthropology*. 21, 511–519.

Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Duranthon, F., Guilaine, J., Crubézy, E., Ludes, B., 2011a. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *Proceedings of National Academy of Science U.S.A.* 108, 9788–9791

Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Tarrús, J., Bosch, A., Guilaine, J., Crubézy, E., Ludes, B., 2011b. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proceeding National Academy Science U.S.A*. 108, 18255–18259.

Lake, M.W., 2014. Trends in Archaeological Simulation. *Journal of Archaeological Methods and Theory* 21, 258–287.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., Berger, B., Economou, C., Bollongino, R., Fu, Q., Bos, K.I., Nordenfelt, S., Li, H., de Filippo, C., Prüfer, K., Sawyer, S., Posth, C., Haak, W.,

Hallgren, F., Fornander, E., Rohland, N., Delsate, D., Francken, M., Guinet, J.-M., Wahl, J., Ayodo, G., Babiker, H.A., Bailliet, G., Balanovska, E., Balanovsky, O., Barrantes, R., Bedoya, G., Ben-Ami, H., Bene, J., Berrada, F., Bravi, C.M., Brisighelli, F., Busby, G.B.J., Cali, F., Churnosov, M., Cole, D.E.C., Corach, D., Damba, L., van Driem, G., Dryomov, S., Dugoujon, J.-M., Fedorova, S.A., Gallego Romero, I., Gubina, M., Hammer, M., Henn, B.M., Hervig, T., Hodoglugil, U., Jha, A.R., Karachanak-Yankova, S., Khusainova, R., Khusnutdinova, E., Kittles, R., Kivisild, T., Klitz, W., Kučinskas, V., Kushniarevich, A., Laredj, L., Litvinov, S., Loukidis, T., Mahley, R.W., Melegh, B., Metspalu, E., Molina, J., Mountain, J., Näkkäläjärvi, K., Nesheva, D., Nyambo, T., Osipova, L., Parik, J., Platonov, F., Posukh, O., Romano, V., Rothhammer, F., Rudan, I., Ruizbakiev, R., Sahakyan, H., Sajantila, A., Salas, A., Starikovskaya, E.B., Tarekegn, A., Toncheva, D., Turdikulova, S., Uktveryte, I., Utevska, O., Vasquez, R., Villena, M., Voevoda, M., Winkler, C.A., Yepiskoposyan, L., Zalloua, P., Zemunik, T., Cooper, A., Capelli, C., Thomas, M.G., Ruiz-Linares, A., Tishkoff, S.A., Singh, L., Thangaraj, K., Villems, R., Comas, D., Sukernik, R., Metspalu, M., Meyer, M., Eichler, E.E., Burger, J., Slatkin, M., Pääbo, S., Kelso, J., Reich, D., Krause, J., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513, 409–413

Lercher, M.J., Hurst, L.D., 2002. Human SNP variability and mutation rate are higher in regions of high recombination. *Trends in Genetics* 18, 337–340.

Lewontin, R.C., Dunn, L.C., 1960. The Evolutionary Dynamics of a Polymorphismm in the House Mouse. *Genetics* 45, 705–722.

Li, J.Z., Absher, D.M., Tang, H., Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsh, G.S., Feldman, M., Cavalli-Sforza, L.L., Myers, R.M., 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 319, 1100–1104.

Lischer, H.E.L., Excoffier, L., 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28, 298–299

Li, W.H., Sadler, L.A., 1991. Low nucleotide diversity in man. Genetics 129, 513–523.

Lucas, G., 1997. 'Interpretation in comtemporary archaeology: Some philosophical issues.', in: Hodder, I., Shanks, M., Alexandri, A., Buchli, V., Carman, J., Last, J., Lucas, G. (Eds.), *Interpreting Archaeology: Finding Meaning in the Past*. London; New York: Routledge, pp. 37–44.

Mallory, J.P., 2013. *The origins of the Irish.* London: Thames & Hudson.

Malyarchuk, B., Derenko, M., Denisova, G., Maksimov, A., Wozniak, M., Grzybowski, T., Dambueva, I., Zakharov, I., 2011. Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a. Journal of Human Genetics. 56, 583–588.

Manx National Heritage, 2014. Manx National Heritage: iMuseum: Explore Collections [WWW Document]. URL http://www.imuseum.im/Collections/View.mth?EntryID=2526722&issourcerecord=true# (accessed 12.15.14).

Mapsof.net, 2014. Map of Isle of Man [WWW Document]. URL http://mapsof.net/isle-of-man (accessed 12.11.14).

Marks, S.J., Levy, H., Martinez-Cadenas, C., Montinaro, F., Capelli, C., 2012. Migration distance rather than migration rate explains genetic diversity in human patrilocal groups. *Molecular Ecology* 21, 4958–4969.

Mary Clayton, 2002. 'An Edition of Ælfric's Letter to Brother Edward', in: Treharne, E.M., Rosser, S. (Eds.), *Early Medieval English Texts and Interpretations: Studies Presented to Donald G.*

*Scragg.*, Tempe, AZ: Arizona Center for Medieval and Renaissance Studies pp. 286–284.

Matsmura, S., Forster, P., 2008. Introduction, in: *Simulations, Genetics and Human Prehistory.* Cambrige: McDonald Institute for Archaeological Research, pp. 1–6.

McCartan, S.B., 2000. 'The Utilisation of Island Environments in the Irish Mesolithic: Agendas for Rathlin Island.', in: Desmond, A., Anderson, L. (Eds.), *New Agendas in Irish Prehistory: Papers in Commemoration of Liz Anderson*. Bray, Co. Wicklow: Wordwell,pp. 15 – 50.

McCartan, S.B., 2004.' The Mesolithic in the Isle of Man: an island perspective.', in: Saville A., (Ed.), *Mesolithic Scotland and Its Neighbours : The Early Holocene Prehistory of Scotland, Its British and Irish Context, and Some Northern European Perspectives*. Society of Antiquaries of Scotland, Edinburgh, pp. 271–283.

McDonald, R.A., 2008. *Kingdom of the isles: Scotland's western seaboard, c.1100-c.1336*. Edinburgh: John Donald.

McEvoy, B., Brady, C., Moore, L.T., Bradley, D.G., 2006. The scale and nature of Viking settlement in Ireland from Y-chromosome admixture analysis. *European Journal of Human Genetics* 14, 1288–1294.

McEvoy, B., Richards, M., Forster, P., Bradley, D.G., 2004. The Longue Durée of Genetic Ancestry: Multiple Genetic Marker Systems and Celtic Origins on the Atlantic Facade of Europe. *The American Journal of Human Genetics* 75, 693–702.

Menozzi, P., Piazza, A., Cavalli-Sforza, L., 1978. Synthetic maps of human gene frequencies in Europeans. *Science* 201, 786–792.

Mikkelsen, M., Sørensen, E., Rasmussen, E.M., Morling, N., 2010. Mitochondrial DNA HV1 and HV2 variation in Danes. Forensic Science International: Genetics 4, e87–e88.

Milner, N., Mithen, S., 2009. Hunter-gatherers of the Mesolithic., in: Hunter, J., Ralston, I. (Eds.), The Archaeology of Britain: An Introduction from Earliest Times to the Twenty-First Century. Routledge, London; New York.

Mishmar, D., Ruiz-Pesini, E., Golik, P., Macaulay, V., Clark, A.G., Hosseini, S., Brandon, M., Easley, K., Chen, E., Brown, M.D., Sukernik, R.I., Olckers, A., Wallace, D.C., 2003. Natural selection shaped regional mtDNA variation in humans. Proc. Natl. Acad. Sci. U.S.A. 100, 171–176.

Moore, A.W., 1890. *The Surnames and Placenames of the Isle of Man. With an Introduction by Professor Rhys.*, 2nd ed, revd edition. ed. Elliot Stock, London.

Moore, L.T., McEvoy, B., Cape, E., Simms, K., Bradley, D.G., 2006. A Y-chromosome signature of hegemony in Gaelic Ireland. *American Journal Human Genetics* 78, 334–338.

Mullis, K.B., Faloona, F.A., 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* 155, 335–350.

Munch, P.A., 1874. *The Chronicle of Man and the Sudreys: With historical notes*. Isle of Man: Manx Society Publications.

Myres, N.M., Rootsi, S., Lin, A.A., Järve, M., King, R.J., Kutuev, I., Cabrera, V.M., Khusnutdinova, E.K., Pshenichnov, A., Yunusbayev, B., Balanovsky, O., Balanovska, E., Rudan, P., Baldovic, M., Herrera, R.J., Chiaroni, J., Di Cristofaro, J., Villems, R., Kivisild, T., Underhill, P.A., 2011. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur.opean Journal of Human Genetics* 19, 95–101.

Nachman, M.W., Crowell, S.L., 2000. Estimate of the Mutation Rate per Nucleotide in Humans. *Genetics* 156, 297–304.

Napier, P., Jenkins, P.D., 1997. 'Monkey skull and mandible of the Iron Age found at Site B.', in:

Waterman, D.M., Lynn, C.J., Hamlin, A. (Eds.), *Excavations at Navan Fort 1961-71, County Armagh* . Belfast: Stationery Office, pp. 121–124.

National Human Genome Research Intitute, 2010. Human Genome Project Completion: Frequently Asked Questions [WWW Document]. URL http://www.genome.gov/11006943 (accessed 5.12.14).

Nei, M., 1987. *Molecular Evolutionary Genetics*. New York USA: Columbia University Press.

Nickerson, D.A., Taylor, S.L., Weiss, K.M., Clark, A.G., Hutchinson, R.G., Stengård, J., Salomaa, V., Vartiainen, E., Boerwinkle, E., Sing, C.F., 1998. DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. *Nature Genetics* 19, 233–240.

Nielsen, R., Beaumont, M.A., 2009. Statistical inferences in phylogeography. *Molecular Ecology* 18, 1034–1047.

Novembre, J., Stephens, M., 2008. Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics* 40, 646–649.

OED Online, 1924. u'nited, adj. OED Online.

OED Online, 2008. British Isles, n. OED Online.

Oppenheimer, S., 2007. *The Origins of the British: A Genetic Detective Story*, Robinson Publishing, London.

Oxford Dictionaries Online, 2014. United Kingdom: definition of United Kingdom Available at http://www.oxforddictionaries.com/definition/english/United-Kingdom (accessed 11.29.14).

Paracchini, S., Arredi, B., Chalk, R., Tyler-Smith, C., 2002. Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. *Nucleic Acids Research* 30, e27.

Parker, D., 2011. The Complexity of Lithic Simplicity - Computer Simulation of Lithic Assemblage Formation in Western New South Wales, Australia. *Academia.edu,* Available at http://www.academia.edu/3724416/The_Complexity_of_Lithic_Simplicity_-_Computer_Simulation_of_Lithic_Assemblage_Formation_in_Western_New_South_Wales_Australia (accesed 1.3.15)

Pascual, M., Chapuis, M.P., Mestres, F., Balanyà, J., Huey, R.B., Gilchrist, G.W., Serra, L., Estoup, A., 2007. Introduction history of Drosophila subobscura in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* 16, 3069–3083.

Pattison, J.E., 2008. Is it necessary to assume an apartheid-like social structure in Early Anglo-Saxon England? *Proceedings of the Royal Society B: Biological Sciences* 275, 2423–2429; discussion 2419–2421.

Pattison, J.E., 2011. Integration versus apartheid in post-Roman Britain: a response to Thomas et Al. (2008). *Human Biology* 83, 715–733.

Pearson, M.P., 2009. The Earlier Bronze Age., in: Hunter, J., Ralston, I. (Eds.), *The Archaeology of Britain: An Introduction from Earliest Times to the Twenty-First Century*. London; New York: Routledge.

Pearson, M.P., 2012. *Stonehenge exploring the greatest Stone Age mystery*. London; New York; Sydney: Simon & Schuster,

Pétrequin, P., Sheridan, A., Cassen, S., Errera, M., Gauthier, E., Klassen, L., Le Maux, N., Pailler, Y., 2008. 'Neolithic Alpine axeheads, from the Continent to Great Britain, the Isle of Man and Ireland', in: Fokkens, H., Coles, B., van Gijn, A., Kleijne, J., Hedwig Ponjee, Corijanne

Slappendel (Eds.), *Between Foraging and Farming*. Leiden: Faculty of Archaeology, pp. 261– 280.

Pettitt, P., White, M., 2012. The British Palaeolithic: hominin societies at the edge of the Pleistocene world. London; New York: Routledge.

Phillips, J.D., 2004. *Manx.* Lincom Europa.

Pickin, J., Worthington, T., 1989. Prehistoric mining Hammers from Bradda Head, Isle of Man. *Bulletin of the Peak District Mines Historical Society* 10, 274–275.

Pinhasi, R., Thomas, M.G., Hofreiter, M., Currat, M., Burger, J., 2012. The genetic history of Europeans. *Trends in Genetics* 28, 496–505.

Pluciennik, M., 1996. Genetics, archaeology and the wider world. *Antiquity* 70, 13-.

Pluciennik, M., 2006. Clash of cultures? Archaeology and genetics. *Documenta Praehistorica* XXXIII, 39–49.

Polzin, T., Daneshmand, S.V., 2003. On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* 31, 12–20.

Poznik, G.D., Henn, B.M., Yee, M.-C., Sliwerska, E., Euskirchen, G.M., Lin, A.A., Snyder, M., Quintana-Murci, L., Kidd, J.M., Underhill, P.A., Bustamante, C.D., 2013. Sequencing Y Chromosomes Resolves Discrepancy in Time to Common Ancestor of Males Versus Females. *Science* 341, 562–565.

Ptolemy 100CE, 1932. Geography of Claudius Ptolemy. The New York public library, New York.

Qamar, R., Ayub, Q., Mohyuddin, A., Helgason, A., Mazhar, K., Mansoor, A., Zerjal, T., Tyler-Smith, C., Mehdi, S.Q., 2002. Y-chromosomal DNA variation in Pakistan. *American Journal of Human Genetics* 70, 1107–1124.

Quilliam, L., 1989. *Surnames of the Manks*. Peel, Isle of Man: Cashtal Books.

Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., Poznik, G.D., Gudmundsdottir, V., Yadav, R., Malaspinas, A.-S., White, S.S., Allentoft, M.E., Cornejo, O.E., Tambets, K., Eriksson, A., Heintzman, P.D., Karmin, M., Korneliussen, T.S., Meltzer, D.J., Pierre, T.L., Stenderup, J., Saag, L., Warmuth, V.M., Lopes, M.C., Malhi, R.S., Brunak, S., Sicheritz-Ponten, T., Barnes, I., Collins, M., Orlando, L., Balloux, F., Manica, A., Gupta, R., Metspalu, M., Bustamante, C.D., Jakobsson, M., Nielsen, R., Willerslev, E., 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* 506, 225–229.

Rasteiro, R., Chikhi, L., 2013. Female and male perspectives on the neolithic transition in europe: clues from ancient and modern genetic data. *PLoS ONE* 8, e60944.

Renfrew, C., 2000a. 'Introduction' , in: Renfrew, C., Boyle, K. (Eds.), *Archaeogenetics: DNA and the Population Prehistory of Europe.* Cambridge: McDonald Institute for Archaeological Research, pp. 3–12.

Renfrew, C., 2000b. At the Edge of Knowability: Towards a Prehistory of Languages. *Cambridge Archaeological Journal* 10 (1), 7-34.

Renfrew, C., 2010. Archaeogenetics--towards a "new synthesis"? *Current Biology* 20, R162–165.

Renfrew, C., Boyle, K. (Eds.), 2000. *Archaeogenetics: DNA and the Population Prehistory of Europe.* Cambridge; Oakville, CT: McDonald Institute for Archaeological Research,

Richards, J.D., 2009. 'The Scandinavian presence.', in: Hunter, J., Ralston, I. (Eds.), *The Archaeology of Britain: An Introduction from Earliest Times to the Twenty-First Century.*

London; New York: Routledge.

Richards, Julian D., 2007. *Viking Age England*, 2nd ed. Gloucestershire, UK: Tempus Publishing.

Richards, M., Côrte-Real, H., Forster, P., Macaulay, V., Wilkinson-Herbots, H., Demaine, A., Papiha, S., Hedges, R., Bandelt, H.J., Sykes, B., 1996. Paleolithic and neolithic lineages in the European mitochondrial gene pool. *American Journal of Human Genetics* 59, 185–203.

Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., Villems, R., Thomas, M., Rychkov, S., Rychkov, O., Rychkov, Y., Golge, M., Dimitrov, D., Hill, E., Bradley, D., Romano, V., Cali, F., Vona, G., Demaine, A., Papiha, S., Triantaphyllidis, C., Stefanescu, G., Hatina, J., Belledi, M., Di Rienzo, A., Oppenheim, A., N? rby, S., Al-Zaheri, N., Santachiara-Benerecetti, S., Scozzari, R., Torroni, A., Bandelt, H.-J., 2000. Tracing European Founder Lineages in the Near Eastern mtDNA Pool. *American Journal of Human Genetic* 67, 1251–1276.

Rienzo, A.D., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M., Freimer, N.B., 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences* 91, 3166–3170.

Roberts, B., 2008. Creating traditions and shaping technologies: understanding the earliest metal objects and metal production in Western Europe. *World Archaeology* 40, 354–372.

Rosser, Z.H., Zerjal, T., Hurles, M.E., Adojaan, M., Alavantic, D., Amorim, A., Amos, W., Armenteros, M., Arroyo, E., Barbujani, G., Beckman, G., Beckman, L., Bertranpetit, J., Bosch, E., Bradley, D.G., Brede, G., Cooper, G., Côrte-Real, H.B., de Knijff, P., Decorte, R., Dubrova, Y.E., Evgrafov, O., Gilissen, A., Glisic, S., Gölge, M., Hill, E.W., Jeziorowska, A., Kalaydjieva, L., Kayser, M., Kivisild, T., Kravchenko, S.A., Krumina, A., Kucinskas, V., Lavinha, J., Livshits, L.A., Malaspina, P., Maria, S., McElreavey, K., Meitinger, T.A., Mikelsaar, A.V., Mitchell, R.J., Nafa, K., Nicholson, J., Nørby, S., Pandya, A., Parik, J., Patsalis, P.C., Pereira, L., Peterlin, B., Pielberg, G., Prata, M.J., Previderé, C., Roewer, L., Rootsi, S., Rubinsztein, D.C., Saillard, J., Santos, F.R., Stefanescu, G., Sykes, B.C., Tolun, A., Villems, R., Tyler-Smith, C., Jobling, M.A., 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *American Journal of Human Genetics* 67, 1526–1543.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N., 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350–1354.

Saint Patrick 470CE, in (Trans.) De Paor L., 1993. *St. Patrick's World: the Christian culture of Ireland's apostolic age*. Notre Dame;Ind: University of Notre Dame Press.

Seielstad, M.T., Minch, E., Cavalli-Sforza, L.L., 1998. Genetic evidence for a higher female migration rate in humans. *Nature Genetic* 20, 278–280. doi:10.1038/3088

Semino, O., Passarino, G., Oefner, P.J., Lin, A.A., Arbuzova, S., Beckman, L.E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., Marcikiae, M., Mika, A., Mika, B., Primorac, D., Santachiara-Benerecetti, A.S., Cavalli-Sforza, L.L., Underhill, P.A., 2000. The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* 290, 1155–1159.

Shennan, S., 2008. 'Archaeology, genes and demographic histories'., in: Matsumura, S., Peter Forster, Colin Renfrew (Eds.), *Simulations, Genetics and Human Prehistory.* Cambridge: McDonald Institute for Archaeological Research, pp. 173–176.

Sheridan, A., 2003. '2. French connections I: Spreading the marmites thinly.', in: Armit, I., Murphy, E., Nelis, E., Simpson, D. (Eds.), *Neolithic Settlement in Ireland and Western Britain.*

*Conference papers of The Prehistoric Society, Queen's University Belfast, April 2001*. Oxford: Oxbow books, pp. 3 – 17.

Simoni, L., Calafell, F., Pettener, D., Bertranpetit, J., Barbujani, G., 2000. Geographic patterns of mtDNA diversity in Europe. *American Journal of Human Genetics* 66, 262–278.

Sims-Williams, P., 1998. Genetics, linguistics, and prehistory: thinking big and thinking straight. *Antiquity* 72 (277), 505-527

Sivia, D.S., Skilling, J., 2006. *Data analysis a Bayesian tutorial*. Oxford; New York: Oxford University Press.

Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M.T.P., Götherström, A., Jakobsson, M., 2012. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* 336, 466–469.

Smyth, A.P., 1977. *Scandinavian kings in the British Isles, 850-880*. Oxford: Oxford University Press.

Soares et al, 2009. Archaeogenetics Research Group - University of Huddersfield [WWW Document]. URL http://www.hud.ac.uk/research/researchcentres/targ/ (accessed 12.3.14).

Soares, P., Achilli, A., Semino, O., Davies, W., Macaulay, V., Bandelt, H.-J., Torroni, A., Richards, M.B., 2010. The archaeogenetics of Europe. *Current Biology* 20, R174–183.

Statistisk Sentralbyra, 2014. StatBank Norway Available at https://www.ssb.no/statistikkbanken/selectvarval/saveselections.asp (accessed 7.26.14).

Steel, J., 2008. 'Population structure and diversity indices.', in: Shuichi Matsumura, Forster, P., Renfrew, C. (Eds.), *Simulations, Genetics and Human Prehistory*. Cambridge: McDonald Institute for Archaeological Research, pp. 173–176.

Steinsland, G., 2011. *Ideology and Power in the Viking and Middle Ages: Scandinavia, Iceland, Ireland, Orkney and the Faeroes*. Leiden: Brill publishing.

Stoneking, M., Krause, J., 2011. Learning about human population history from ancient and modern genomes. *Nature Review Genetics* 12, 603–614. doi:10.1038/nrg3029

Sunnaker, M., Busetto, A.G., Numminen, E., Corander, J., Foll, M., Dessimoz, C., 2013. Approximate Bayesian Computation. *PLoS Computational Biology* 9.

Sykes, B., 2007. *Blood of the Isles*, London: Corgi.

Takahata, N., Satta, Y., 1997. Evolution of the primate lineage leading to modern humans: Phylogenetic and demographic inferences from DNA sequences. *Proceedings of the National Academy of Sciences USA* 94, 4811–4815.

Talbot, T., 1924. *Index Page -The Manorial Roll of the Isle of Man, 1511-1515.* Available at http://www.isle-of-man.com/manxnotebook/fulltext/manroll/ (accessed 6.7.14).

Taylor, J.J., 1994. The Oliver Davis Lecture: The First Golden Age of Europe Was in Ireland and Britain (Circa 2400-1400 B.C.). *Ulster Journal of Archaeology, Third Series* 57, 37–60.

Tetzlaff, S., Brandstätter, A., Wegener, R., Parson, W., Weirich, V., 2007. Mitochondrial DNA population data of HVS-I and HVS-II sequences from a northeast German sample. *Forensic Science International* 172, 218–224.

Thomas, J., 2006. Gene-flows and social processes: The potential of genetics and archaeology. Documenta Praehistorica XXXIII.

Thomas, J., 2008. 'The Mesolithic-Neolithic Transistion in Britain.', in: Pollard, J. (Ed.), *Prehistoric Britain*. Malden, Mass: Blackwell, pp. 58–89.

Thomas, M.G., Kivisild, T., Chikhi, L., Burger, J., 2013. 18 'Europe and western Asia: genetics and population history.', in: *The Encyclopedia of Global Human Migration*. Blackwell Publishing Ltd, Chapter 8

Thomas, M.G., Stumpf, M.P.., Harke, H., 2008. Integration versus apartheid in post-Roman Britain: a response to Pattison. *Proceedings of the Royal Society B: Biological Sciences* 275, 2419–2421.

Thomas, M.G., Stumpf, M.P.H., Härke, H., 2006. Evidence for an apartheid-like social structure in early Anglo-Saxon England. *Proceedings of the Royal Academy Society B: Biological Sciences* 273, 2651–2657.

Tillmar, A.O., Coble, M.D., Wallerström, T., Holmlund, G., 2010. Homogeneity in mitochondrial DNA control region sequences in Swedish subpopulations. *International Journal of Legal medicine* 124, 91–98.

Tishkoff, S.A., Verrelli, B.C., 2003. Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annual Review Genomics and Human Genetics* 4, 293–340.

Tresset, A., 2003. 'French Connection II: of cows and men.', in: Armit, I., Nelis, E., Simpson, D. (Eds.), *Neolithic Settlement in Ireland and Western Britain*. Oxford: Oxbow, pp. 18–30.

Underhill, P.A., Passarino, G., Lin, A.A., Shen, P., Mirazón Lahr, M., Foley, R.A., Oefner, P.J., Cavalli-Sforza, L.L., 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Annals of Human Genetics* 65, 43–62.

Valante, M.A., 2008. *The Vikings in Ireland : settlement, trade and urbanization*. Four Courts, Dublin.

Van Geel, B., Buurman, J., Waterbolk, H.T., 1996. Archaeological and palaeoecological indications of an abrupt climate change in The Netherlands, and evidence for climatological teleconnections around 2650 BP. *Journal of Quaternary Science* 11, 451–460.

Van Oven, M., Kayser, M., 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation* 30, E386–E394.

Verdugo, M.P., 2012 The Isle of Man: Investigating Viking ancestry in Y-chromosomal lineages. *BSc Biological Sciences (Genetics) third-year dissertation*, University of Leicester, UK.

Waddell, J., 1998. *The prehistoric archaeology of Ireland.* Galway University Press, Galway.

Wakeley, J., 1999. Nonequilibrium migration in human history. *Genetics* 153, 1863–1871.

Wang, C.-C., Gilbert, M.T.P., Jin, L., Li, H., 2014. Evaluating the Y chromosomal timescale in human demographic and lineage dating. *Investigative Genetics* 5, 12.

Weale, M.E., Weiss, D.A., Jager, R.F., Bradman, N., Thomas, M.G., 2002. Y chromosome evidence for Anglo-Saxon mass migration. *Molecular Biology and Evolution* 19, 1008–1021.

Wei, W., Ayub, Q., Chen, Y., McCarthy, S., Hou, Y., Carbone, I., Xue, Y., Tyler-Smith, C., 2013. A calibrated human Y-chromosomal phylogeny based on resequencing. *Genome Research* 23, 388–395.

Wilkins, J.F., Marlowe, F.W., 2006. Sex-biased migration in humans: what should we expect from genetic data? *BioEssays* 28, 290–300.

Willuweit, S., Roewer, L., 2007. Y chromosome haplotype reference database (YHRD): Update. Forensic Science International: Genetics, *DNA in Forensics* 2006 1, 83–87.

Wilson, D.M., 2008. The Vikings in the Isle of Man. Aarhus University Press, Aarhus.

Wilson, J.F., Weiss, D.A., Richards, M., Thomas, M.G., Bradman, N., Goldstein, D.B., 2001. Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proceedings of the National Academy of Science U.S.A*. 98, 5078–5083.

Winney, B., Boumertit, A., Day, T., Davison, D., Echeta, C., Evseeva, I., Hutnik, K., Leslie, S., Nicodemus, K., Royrvik, E.C., Tonks, S., Yang, X., Cheshire, J., Longley, P., Mateos, P., Groom, A., Relton, C., Bishop, D.T., Black, K., Northwood, E., Parkinson, L., Frayling, T.M., Steele, A., Sampson, J.R., King, T., Dixon, R., Middleton, D., Jennings, B., Bowden, R., Donnelly, P., Bodmer, W., 2012. People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population. *European Journal of Human Genetics* 20, 203–210.

Woodcock, J.J., 1999a. 'The Bronze Age Burials from Cottiers Feild, Bishopscourt Farm, Ballaugh.', in: Davey, P., White, S.D. (Eds.), *Recent Archaeological Research on the Isle of Man*. Oxford: Archaeopress, pp. 99–110.

Woodcock, J.J., 1999b. 'A report on the excavation carried out by B. R. S. Megaw at Staarvey Farm, Kirk German', in: Davey, P., White, S.D. (Eds.), *Recent Archaeological Research on the Isle of Man.* Archaeopress, Oxford, pp. 89–98.

World Population Review, 2014. Sweden Population 2014 - World Population Review Available at http://worldpopulationreview.com/countries/sweden-population/ (accessed 9.10.14).

Wormald, P., 1991. The Ninth Century., in: Campbell, J. (Ed.), *The Anglo-Saxons.* London, England; New York, N.Y., USA: Penguin.

Wright, S., 1931. Evolution in Mendelian Populations. *Genetics* 16, 97.

Wright, S., 1943. Isolation by Distance. *Genetics* 28, 114–138.

Xue, Y., Wang, Q., Long, Q., Ng, B.L., Swerdlow, H., Burton, J., Skuce, C., Taylor, R., Abdellah, Z., Zhao, Y., Asan, MacArthur, D.G., Quail, M.A., Carter, N.P., Yang, H., Tyler-Smith, C., 2009. Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Current Biology* 19, 1453–1457.

Yang, I.S., Lee, H.Y., Yang, W.I., Shin, K.-J., 2013. mtDNAprofiler: A Web Application for the Nomenclature and Comparison of Human Mitochondrial DNA Sequences,. *Journal of Forensic Sciences* 58, 972–980.

Y Chromosome Consortium, 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Research* 12, 339–348.

Yonsei DNA Profiling Group, 2013. mtDNAprofiler [WWW Document]. URL http://mtprofiler.yonsei.ac.kr/index.php?cat=1 (accessed 10.28.14).

Zerjal, T., Xue, Y., Bertorelle, G., Wells, R.S., Bao, W., Zhu, S., Qamar, R., Ayub, Q., Mohyuddin, A., Fu, S., Li, P., Yuldasheva, N., Ruzibakiev, R., Xu, J., Shu, Q., Du, R., Yang, H., Hurles, M.E., Robinson, E., Gerelsaikhan, T., Dashnyam, B., Mehdi, S.Q., Tyler-Smith, C., 2003. The genetic legacy of the Mongols. *American Journal of Human Genetics* 72, 717–721.

Zhang, H., Singh, K.K., 2014. Global Genetic Determinants of Mitochondrial DNA Copy Number. *PLoS ONE* 9, e105242.

Zvelebil, M., 1989. On the transition to farming in Europe, or what was spreading with the Neolithic: a reply to Ammerman. (1989) *Antiquity* 63 (239), 379-383

Zvelebil, M., Zvelebil, K.V., 1988. Agricultural transition and Indo-European dispersals. *Antiquity* 62, 574–583.