The Peoples of Britain: Population genetics, archaeology and linguistics

Ellen C. Røyrvik
Hertford College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Michaelmas 2012

Abstract

The history of peoples has always evoked a great deal of both academic and popular interest, and the peoples of Britain, with its island position and semi-mythic serial invasions, have evoked as much as any. As most of the period during which Britain has been inhabited by modern humans lies in prehistory, archaeology has long been the best method for elucidating the past. In recent years, however, genetics has come to complement the reconstruction of peoples' pasts, with its ability to trace lineal human biology instead of transferrable human culture.

The purpose of this thesis is to assess population genetics systems of Britain against the backdrop of archaeologically determined history, informed for later periods by linguistics, and attempt to ascertain any marked congruities or incongruities between this history and modern genetic data. The genetic datasets included in this work are the People of the British Isles Project collection, and some ancillary cohorts from surrounding countries. The genetic systems assessed include mitochondrial DNA, classical marker genes, lactase, pigmentation genes and some phenotypes, and finally a suite of candidate genes for determining normal facial variation. In a self-contained section, the principle of relating population genetic data to population histories is illustrated by a study focusing on Central Asia (a larger area), but using fewer genetic markers.

The chosen markers systems overall reveal modest amounts of genetic differentiation among different groups in Britain, but consistently highlight Wales and Orkney especially as relatively distanced from the rest of Britain. This is in keeping with the historically quite isolated state of the former, and the comparatively recent heavy influx of Norse Vikings in the latter. Further details are observable from subsets of this study: all are discussed in the context of archaeological and linguistic evidence.

These findings provide support and foundation for a forthcoming study from the People of the British Isles Project, using a genome-wide SNP approach rather than selected markers, which will likely increase the nuance of this initial picture and contribute further to answering specific questions regarding Britain's past.

Acknowledgements

This work was carried out in the Cancer and Immunogenetics Laboratory at the Department of Oncology (previously the Department of Clinical Pharmacology), University of Oxford, and my stay in Oxford was economically supported by Overrettssakfører Axel Brynies, Helga Brynies og Adelheid Jordans legat.

I wish to thank my scientific supervisor, Walter Bodmer, for the opportunity to work on such a unique project as the People of the British Isles endeavour, as well as all those who have worked on and volunteered for the project. I am also deeply grateful to my supervisor in archaeology, Barry Cunliffe, for much encouragement and sharing his panoramic knowledge of the European archaeology.

Many thanks are due to Michael Benskin, without whom I would never have been given the opportunity of coming to Oxford, and to my parents, who have been unfailingly supportive.

Finally, this work is dedicated to Iain, for being there for me every step of the way.

Ellen C. Røyrvik

Contents

Chapter 1

Introduction

A notable feature of human existence is the concern we have for the past, especially the specific history of our species. This need to establish a 'historical' position has long been expressed by the ubiquity of origin myths and genealogies in human culture, and latterly by the intense interest in scientific, or broadly scientific, means of elucidating and interpreting the past. A question such as 'Who are the people of Britain?' would have been answered in many different ways throughout the last 13,000 or so years of unbroken human occupation [1]. What the earliest answers would have been we have no way of knowing, but among later theories have been, at least, descent from Celtic[1] and Germanic first peoples, likewise from a Biblical Adam and Eve, founding by a Roman consul descendant of Aeneas, and settlements by Scythians, or, less fancifully, Gauls, Germans and Iberians [2, 3, 4]. Current origin stories are better regarded if they have some basis in verifiable facts. Facts, however, are never as abundant as would be desirable, nor are they always amenable to telling the kind of compelling story that origin legends do. A less narratively satisfactory, but hopefully more accurate, account of the peopling and peoples of Britain through the ages is the overarching aim of this thesis.

At its core, the following work deals with human history, defined as things that have taken place in the past, rather than written history. Such a topic is, naturally,

---

[1] As we may assume existed.

so broad as to be virtually meaningless. Even limiting the geographic range covered to Britain and areas of Western Europe does not significantly diminish the scope in practical terms. Both the questions posed and the conclusions drawn are therefore of a comparatively coarse nature. The following chapter will detail the goals of this research, outline the major and minor approaches that will be used to reach these goals, and give a brief justification for the utility of these approaches in the given setting. Finally, I will give a short overview of the corpora of data that will be drawn upon over the course of the study.

1.1 Research questions

The population history of an area, in its fullest sense, can and must be approached from a number of different angles. The questions of who has lived where, and whence and when they came are only meaningful if we are able to differentiate between human groups in both geographical and chronological dimensions. This kind of discernment can be achieved by examining facets of human existence in which differences accumulate to allow a separation of different groups. The aim of this thesis is to examine three broad fields of evidence that are thought to be informative on the subject of demographic history, and thereby to recover a portion of the processes that shaped the current population of Britain. Specifically, I will consider detailed population genetic data with archaeological, and in later periods, linguistic information, with the aim of identifying consistent themes regarding the processes of migration, cultural diffusion, and autochthonous development during the history of Britain.

In addition to considering Britain alone, I plan to place it, if possible, in the genetic context of western coastal Europe, emerging patterns for which can be compared and contrasted with the patterns seen in the archaeological record and historical sources.

In order to illustrate the principles brought to bear on the question of British population history, Chapter 3 presents an analogical study covering a much larger geographical area, Eurasia, but using many fewer genetic markers.

1.1.1 Population history of Britain

When focusing on Britain in comparative isolation, there are several questions that need to be addressed. The most important of these is, is there any consistent genetic structuring in Britain, and if so, along which elements of the island's geography does it run? Second, do any of the genetic signatures observed bear the traces of specific demographic events? Third, if structure and single-event signatures are observed in genetic data, can they be sensibly related to lengthy cultural trajectories or periods of rapid cultural or linguistic change that are known or inferred from non-genetic disciplines?

1.1.2 Britain in a Western European context

Human genetic variation is generally seen as a continuum [5], and so limiting the scope of research to a single area, however large, will always involve imposing somewhat arbitrary cut-offs. The rationale for including only coastal regions of Western Europe for the purpose of contextualizing the genetic landscape of Britain is the assumption that any large-scale population movements into or from Britain will have been concentrated on the nearest accessible landmasses. We wish to determine the general level of genetic similarity, or conversely, dissimilarity, of Britons as compared to some neighbouring countries for which data is available. As for Britain above, the genetic configuration will then be examined for agreements or incongruities with the landscape and chronology as described by material culture, with reference to written history where this is appropriate.

1.2 Methodological approaches

Omnia mutantur, nihil interit (Everything changes, nothing is lost)

- Ovid, *Metamorphoses*

Combining the powers of biological, cultural and linguistic studies is nothing new.

James Cowles Prichard used this approach in the early 19th century (though drawing on biological morphology and ethnology, rather than the undiscovered field of genetics and the embryonic one of archaeology), in order to demonstrate the unity of mankind through the evolution of human traits [6]. More recently, L.L. Cavalli-Sforza, Colin Renfrew and colleagues in particular have advocated using cultural data to supplement genetic data, or vice versa, in attempting to recount the history of human populations [7, 8, 9]. The following sections provide a brief overview of the kinds of evidence that can be gained from the three main fields that will be drawn on in the course of this work. These three, one primary and two ancillary in this work, differ not only qualitatively (in their raw data, most of their methodological treatments of the data and in conclusions that can be drawn from them), but also quantitatively in the time depth to which they can penetrate (see below, and Figure 1.1). They do, however, subscribe to certain concepts that are common in general terms, if disparate in detail. These are the notions of classification and transformation. The parallels work best for biology and linguistics, whose theoretical developments have been shaped by a sequence of reciprocal influences. The idea of descent with modification, formulated by Darwin in The Origin of Species to explain the existence of separate but related species [10], was illustrated by reference to dialectology and the work done by philologists on the relatedness of different languages, for example William Jones's 1786 model for Sanskrit, Latin, Greek, Persian, Gothic and Celtic being descendants of a single parent language, Indo-European [11]. Jones, in turn, had been influenced by ideas of Linnean taxonomy in his approach to language classification [12]. Taxonomy and evolution were also mainstays of early efforts in archaeology, efforts that were similarly influenced by Darwinian and other evolutionary thinking. *Typology* was one of the first methods used, and it was dependent on ideas of continuity and continuous, if sometimes very slow, transformation; being the ordering of artefacts in chronological or developmental sequences. While typological sequences are not the be all and end all of the archaeological *métier*, they still provide evidence for continuity of tradition, often within a community, and the possibility of relative dating. It is reasonable to contend, on the strength of the

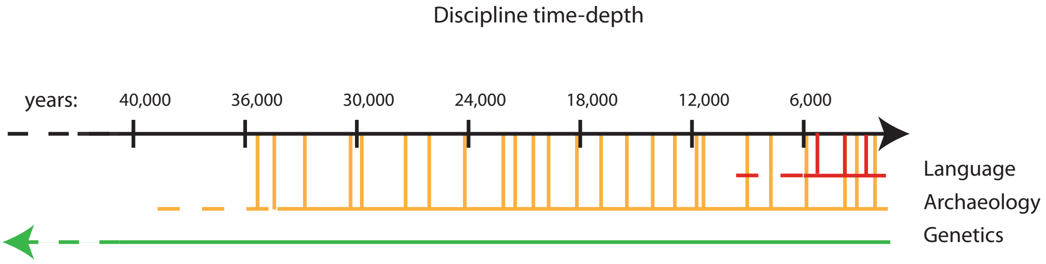


Figure 1.1: **Schematic view of relative time depths of genetics, archaeology and linguistics.** Vertical coloured lines represent absolutely dated anchor points.

above, that genetics, linguistics and archaeology have at least one broadly similar characteristic: ‘objects’ of study have mutable lines of descent, and the changes incurred during descent from an ancestral state can either be reconstructed or directly observed.

To return to the matter of chronology, many, perhaps even most, linguists agree that the limit to which diachronic linguistics can have meaning is of the order of 10,000 years [13]. This reflects a best-case scenario, with early attestations of previous language stages. After a passage of time of this length, any genetic – relatedness by descent – relationships between the first descendants of a parent language are assumed to have degraded to such an extent that they are simply untraceable. Archaeology, however, can reach back to the dawn of human culture, though it is often confined to anatomically modern humans. For Britain seen in isolation, the current phase of continuous human occupation has lasted ca. 13,000 years [1], and so provides something very close to a cut-off point for the scope of archaeological examination for the purpose of this thesis. Genetics, by virtue of the single genesis of life, can theoretically be informative on the scale of billions of years, though naturally population genetics cannot usefully go far beyond species genesis [14].[2]

While time depths vary, the type of dating that can be achieved is also quite different. Archaeology is the only discipline of the above that can offer direct, absolute and often exceedingly accurate dating of the material studied, in the form of dendrochronology, and radiocarbon and thermoluminescence techniques, among others. Relative dating is still common, for example in the form of classical typological sequences, but absolute

---

[2] Although some polymorphisms in human populations also exist in chimpanzees and gorillas.

anchors provided by the aforementioned methods can prevent a large chronological framework being constructed on mistaken assumptions. Genetics and linguistics have, due to their largely superficial similarities of change and descent, been subjected to comparable or even directly transferred methods of dating. The challenge is that neither geneticists nor linguists have stable, physical entities to date. Following McMahon & McMahon [15], such systems must conform to the following requirements to be datable:

1. The system must be mutable, and change must be inherited
2. One must understand the mechanisms of change in this system
3. The rate of change must be fairly constant or normally distributed[3] about a constant, and if biased, the bias must be known and understood
4. Changes must be random, or their directionality known, understood, or predictable
5. The rate of change must be low over the timescale of interest

Few of these requirements are completely fulfilled in linguistics, and they are not even entirely met by genetics. Techniques based on the molecular clock hypothesis have met with severe criticism, though it is considered generally useful if not infallible [16], and glottochronology – based on the presumed existence of a universal and constant rate of lexical replacement – is dismissed by most serious linguists [17]. Mainstream linguistic dating is largely relative, and conducted on the basis of sound changes, the sequences of which cannot on phonological grounds be inverted and yield the same results. Occasional absolute anchor points for literate societies can be provided by texts. Dating in genetics is difficult matter, and its utility is at times questionable – one can possibly date a mutation, but how does this reflect on the population in which it arose? Furthermore, the absolute date estimates are subject to very large standard errors, see e.g. [18], and there are no equivalents of anchor points, unless ancient DNA fortuitously should provide *a terminus ante quem* for the emergence of a variant.

---

[3] There are, however, ways of coping with non-normal distributions.

We are interested in most aspects of human variation along two axes; a temporal one and a spatial one. Along both of these axes, we hope to trace continua, divergences, convergences and discontinuities of the phenomena under consideration. For all fields, we are working with subsets, and descendants of subsets, of past corpora of data. It is more appropriate, here, to paraphrase Ovid, and say that everything changes, and many things may be lost. One can only examine genetic variants belonging to members of past societies who succeeded, generation after generation, in breeding.[4] In linguistics the situation is similar - only languages and language families surviving into comparatively recent times can be considered. Languages about which we know little or nothing may have survived in Northern Europe until as late as 1,000 years ago (see for instance Jackson or Schrijver [19, 20]). Unfortunately, this ignorance does not make the languages themselves less relevant to the present research. Archaeology is particularly aware of such problems, as it is obvious that the vast majority of the total material culture of any given period has disappeared, often without trace (see Chapter 1.2.2), which naturally introduces a large bias.

Crucially, demographic processes - long- or short-range migration, endogamy versus exogamy, etc. - affect the patterns seen in all three disciplines. A human, as a vector, carries with it units of information; alleles, haplotypes, technological skills, religious convictions, words and syntactical patterns, all of which can develop and recombine in new constellations (see Figure 1.2). Some of the links that can be drawn between the disparate data types are obvious, some are obscure, and many can be seen as tenuous. The web of intersecting strands of evidence will undoubtedly be a tangled one. However, through meticulous and critical assessment of all of these strands, it should be possible to reconstruct, in a rigorous fashion, important aspects of demographic history.

Finally, I wish to draw attention to the fact that complex human behaviour is not deterministic, and though similar circumstances will often lead to similar outcomes, the number of variables involved is staggering and many of them, such as communal

---

[4] Specimens of ancient DNA can be analyzed, but this approach is subject to many methodological complications, currently the foremost among them being the issues of sensible and sufficient sampling.

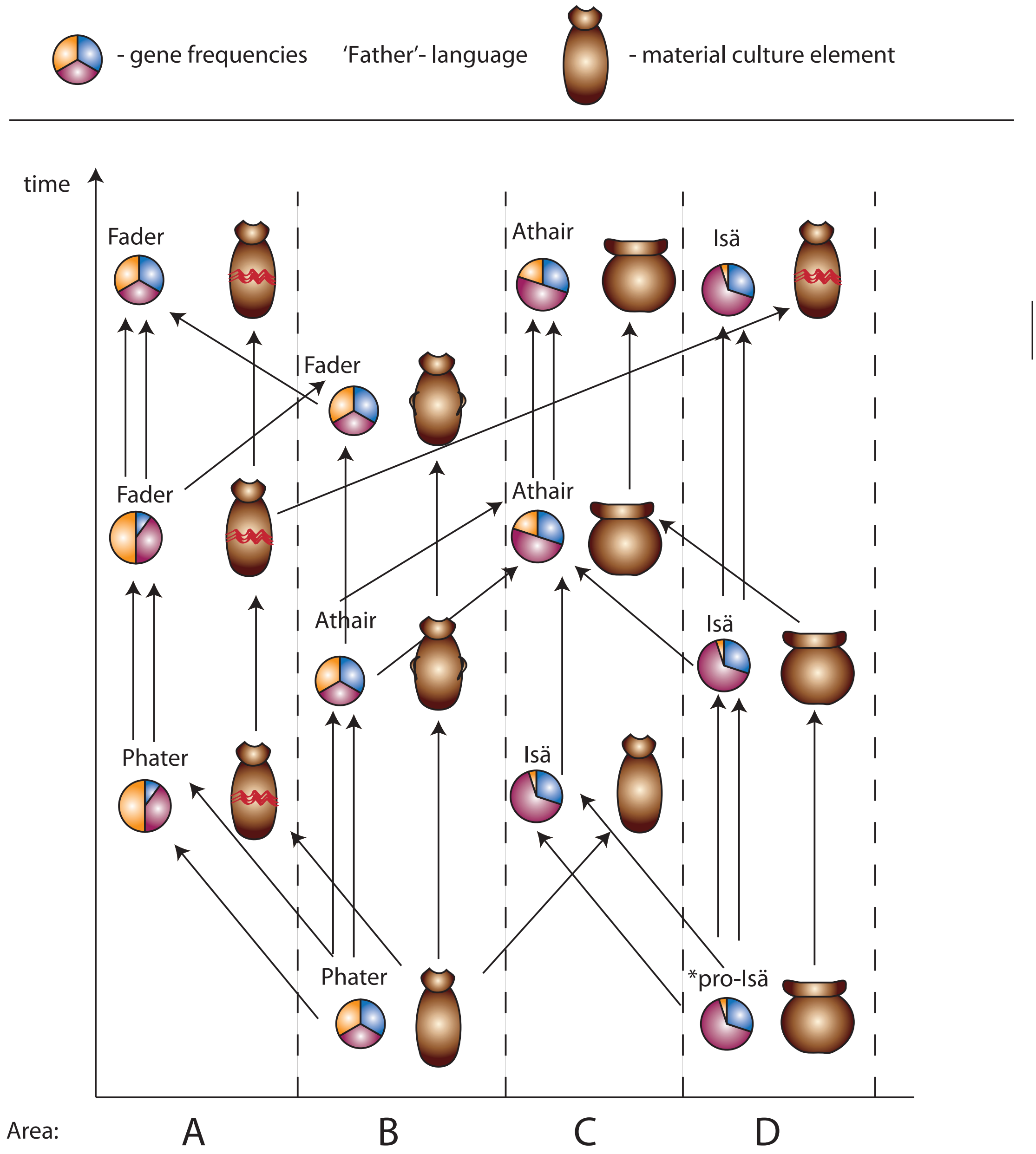


Figure 1.2: **Schematic overview of cultural, genetic and linguistic descent within a larger region** This very simple network of descent (using pots, pie charts of gene or allele frequency variation, and European words for father) illustrates how different coteries of elements may split and recombine in different areas. The chart does not include analogous processes of diffusion/borrowing/migration within each category (material culture, language, genetics), which equally occur.

attitudes, are unrecoverable to researchers far removed from the subjects in time. Popper's refutation of historicism, that "there can be no prediction of the course of human history by scientific or any other rational method" [21], is equally relevant to attempts at creating universal retrodictive models. Thus, a strictly nomothetic attitude to data analysis and interpretation is unlikely to be suitable in all, or even many, situations. The epistemological issues that this uncertainty raises are difficult to resolve, and it seems tacitly agreed that this being the case, heuristic and probabilistic measures must be employed without undue prejudice.

1.2.1 Population genetics

Population genetics can be used to answer, at least in part, many questions of a historical or pre-historical nature. The genetic patterns found in current human populations can contribute to the understanding of past population sizes, migrations and biological exchanges between different groups of people. Issues that population genetics has been brought to bear upon range from large scale demographic processes to single, local events. Examples include the peopling of the planet, over tens of thousands of years, to tracing Phoenician colonies or the creation of the Jewish-like community of the Lemba in southern Africa [22, 23, 24, 25, 26, 27]. The combinations of demographic (e.g. migration or isolation) and biological (e.g. selection) phenomena can create or erase genetic structuring within and between populations. Different parts of the human genome will be affected in different ways by these processes – demographic/chance events average over the genome, while selection is directed to specific loci – and extensive analysis of many different genomic markers is needed to create a comprehensive estimate of how any given population evolved. For example, microgeographic Y chromosomal patterns can be established rapidly [28], and likewise decay in a short period of time. Mitochondrial DNA haplogroup patterning, however, is not in general highly structured and does not appear to be as informative as the Y chromosome, in most cases, on anything less than a continental, or even global, scale. While patterns from genetic data can indicate the probability of occurrence or non-occurrence of population expansion, of limited flow between populations versus isolation or large-scale migration, or of sex specific movements of people, any one kind of pattern can be generated by more than one process [29]. Genetic gradients, for example, have often been interpreted as the result of migration(s), but may equally well have been created by isolation by distance or a combination of both [30, 31, 32].

The genetic variations used to determine the aforementioned patterns are generally neutral as regards outward, physical variation. There are, however, exceptions. Occasionally, molecular genetic variations are mirrored by easily detectable phenotypic variation, such as lactase persistence, which is a selected-for trait in pastoral societies, permitting the digestion of lactose into adulthood [33, 34]. The most obvious examples of such differences, however, is the diversity of human pigmentation, which, disregarding small transient changes caused by tanning or bleaching, is genetically determined, and the molecular variants governing colouring are increasingly well understood. For north-western Europe in particular, this is very fortunate, as this is the area of the (pre-colonial) world with the highest amount of intraregional colouring variation, with blue and green eyes; red, blond and light brown hair being common, where they are uniformly dark brown and black, respectively, in most of the world [35]. The postulated kind of sexual selection that could have caused this chromatic explosion can suggest theories about the population dynamics of the prehistoric societies that promoted the reproductive competition [35]. However, hair and eye colour variation may be hitchiking on the back of pale skin, as several of the genetic variants associated with red hair are also associated with pale skin, and the selective pressures for pale skin in northern climes are quite clear (see Chapter 3). At this juncture it becomes possible, with caution, to attempt a rapprochement between genetic data and directly observed phenotype frequencies, and to test hypotheses suggested by genetics with archaeological data.

In short, the population genetics approach is valid because it can determine how close different groups of people are biologically, and because population processes often leave genetic signatures, traces of their occurrence, in succeeding generations of a population. Its success, however, is dependent on high-resolution data, particularly when the area under investigation is comparatively restricted. In the words of McEvoy *et al.*, this is "[t]he remarkable ability of dense data, in terms of both genome and population coverage, to dissect a range of events from selection to migration [...] across very limited geographic areas"[36].

1.2.2 Archaeology

It is perhaps redundant to specify why the discipline of archaeology is informative on the subject of population history, in that the field has no reference outside the bounds of human past. It may nevertheless be of some use to enter upon *how* archaeology, the study of material culture and remains left by human activity, can be used to construct demographic scenarios from evidence such as potsherds and earthworks. A number of paradigm shifts have occurred in the two hundred years or so that archaeology has been an academic discipline, as separate from the amateur antiquarianism of earlier times. The arguably most influential direction has been that of culture-historical archaeology, which, having its heyday in the western world from the late 19th century through to the 1960s [37], is still the type of approach that is popularly recognized. Culture-historical archaeology sees cultural change as specific and unpredictable, rather than general and deterministic. The basal unit of study is a *culture*, rather than a typological sequence, and the term is itself of fairly recent, anthropological origin [38, 39]. In an archaeological setting, the classical definition was given by V. Gordon Childe:

> "We find certain types of remains – pots, implements, ornaments, burial rites, house forms constantly recurring together. Such a complex of regularly associated traits we shall term a cultural group or just a culture. We assume that such a complex is the material expression of what to-day would be called 'a people'. *Only where the complex in question is regularly and exclusively associated with skeletal remains of specific physical type would we venture to replace 'people' by the term 'race'.*"[5][40]

---

[5] The italics are mine, and highlight a sentence that shows a belief in the possibility of equating a biological type with a cultural type, though Childe himself was cautious in this regard. In the pre-genetic and ancient DNA era, 'physical type' would have been the deepest reachable biological level, and a fairly coarse one at that.

To the list of remains above may now be added a group known collectively as ecofacts, natural remains unshaped by direct human activity, but nevertheless informative of past human behaviour, for example grain or animal bones. Artefacts, ecofacts and features such as house shapes or field systems, can form a cultural unit. Cultural units are naturally not static, and different components of 'material expression' are prone to change over time and with geographical spread. Patterns of directional, morphological changes in and geographical distribution of these components, whether they be dagger styles, brooches, burial practices, etc., can be synthesized to reveal the dynamics of participating societies. Trade can be revealed by ascertaining provenance of raw materials, some of which travelled considerable distances even in prehistory, or by the physical distribution of stylistic types of one or a few kinds of artefacts. When an entire cultural unit appears to be spreading, sizeable migrations or a phenomenon known as élite dominance are often invoked as an explanation (see Chapter 1.2.4 for the place of migration as an explanatory paradigm). A problem with this kind of interpretation is that the subset of a population's material culture that survives as archaeology is most often vanishingly small, in particular for non-urban societies, and it is difficult to know whether the picture that emerges is representative of the actual past situation, or of the population as a whole as opposed to a subset. As an illustration, the materials most likely to survive in standard preservation conditions are stone, pottery, glass, and metals that are minimally prone to corrosion (see e.g. Renfrew & Bahn [41]). Conversely, materials that would have in general be most plentiful, such as wood, plant fibres, bone and skin, have very poor survival in most conditions. There is therefore a preservation bias in favour of more rare raw materials, and by extension, luxury items. Just as textual evidence tends to deal with the rich and powerful, so too can the archaeological record be dominated by items or monuments appertaining to wealthy minorities rather than to the average inhabitant of a given place and time.

In terms of utility for population studies, one of the most prominent benefits of using archaeology is the aforementioned solid spatial and chronological anchorage it can provide for assemblages of material culture, this kind of universal framework is difficult if not impossible to construct for genetic or linguistic data.

1.2.3 Languages

Regarded in a truly rigorous fashion, linguistics is only informative on the subject of languages themselves, not the people who spoke them. It has been quite common to consider linguistic affiliation to be a proxy for genetic line of descent or vice versa, from Darwin's statement that"[i]f we possessed a perfect pedigree of mankind, a genealogical arrangement of the races of man would afford the best classification of the various languages now spoken throughout the world" [10] to Cavalli-Sforza and colleagues' contention that "languages offer a powerful ethnic guidebook, which is essentially complete"[22]. This is unfortunately very far from the case, and language shift – a population changing vernaculars – is a very common phenomenon [42, 43]. This fact severely limits a one-to-one language-people equation, and certainly constrains it to a comparatively short time window. The utility of linguistics is fortunately by no means linked only to its supposed pre-eminent position in determining ethnic identity [42]. In fact, historical linguistics and contact linguistics are most useful in indicating contacts across language barriers, which can usually be fitted into a relative chronological framework, and occasionally even be accurately dated [44]. Linguistic influences and the spread of innovations can give a remarkably detailed picture of the interactions between different language groups in the past. The presence or absence of loanwords, what kind of lexical items are borrowed, the adoption of foreign structural features, and the direction of these influences – all of the above create a probabilistic picture of social, martial and mercantile relationships between speakers of different languages.

Historical linguistics, by the reconstruction of earlier language stages, is thought to throw light on many of the aspects of speakers'[6] environmental surroundings, and religious and social organization [44]. The conclusions drawn from this kind of approach must be treated with caution. For example, cognates in descendant languages may

---

[6] These speakers, of course, may be genetically distant from, or even entirely unrelated to, the majority of modern descendant speakers.

have undergone parallel evolution, giving the impression that a certain concept or object is of greater antiquity than is actually the case. A classical example is the standard Indo-European word for 'king', taking the form of *rex* in Latin, *ríg* in Old Irish, and *raja* in Vedic Sanskrit. It is accepted that Indo-European split up significantly prior to the development of a concept of kingship as such (a male sovereign ruler of an independent state, whose position is either purely hereditary, or hereditary under certain legal conditions [45]), and that an older word corresponding simply to an informal chief, leader, etc., probably developed independently in similar social conditions to refer to a paramount position in a larger, more formally defined community. Furthermore, early loanwords, e.g. for new technologies, can also give a false impression of less recent unity. For example, the wheel is frequently cited as being known to proto-Indo-Europeans, due to the widespread descendants of Indo-European $^{*}k^{w}ek^{w}los$ and **roto-*, both meaning wheel. However, the earliest attested daughter branch of Proto-Indo-European, Anatolian Hittite, has neither of these words [46], making it more likely that words for wheel were borrowed into Indo-European languages, but at a stage before many of their diagnostic phonological changes had occurred. Nevertheless, reconstructed impressions of natural, cultural and social environments can potentially be useful in linking, however ephemerally, a language or language group with a material culture.

One further linguistic aid to historical demography comes in the form of toponymy. Toponyms are known to be in general conservative [17], in particular names for natural features – hydronyms or river names being among the most stable. Because names for settlements and especially landscape features get transferred from an analytic (interpretable) linguistic category to a fixed, onomastic category, they can quite easily persist in the setting of a different language, e.g. after a language shift affects an area. Thus, toponyms can a) provide evidence of earlier languages spoken in an area, b) give an indication of the extent of settlement by an incoming linguistic group, and c) provide something resembling a time-frame for settlements through the periodic productivity of different name elements.

1.2.4 Migration as an explanatory concept

> “[A]sk the archaeologists what interesting migrations they’ve turned up and [...] see if we can make the linguistic data fi”
>
> -L. Trask, p.357 [47]

A period from the fourth to sixth centuries in Europe is known as the Migration Period. Of the possibility of tracing these migrations at the current time “on the basis of language or material culture, or to identify biological populations named by ancient authors in a period when people changed language and ethnic allegiances within their lifetime”, Chapman & Hamerow contend that this“is a task of immense and absorbing complexity” [48]. One might also add that the task would not necessarily be facilitated by the application of population genetics, as a case study on putative Cimbri tribal descendants in Denmark and Italy shows [49].

Migrations have, in both archaeology and linguistics, frequently been employed to explain changes in the archaeological record or the spread of language families. Archaeology especially was accustomed to rely on migrations to account for the spread of an innovation, and this reliance accounts for the still current tendency to think of most cultural change as the result of an influx of a distinct “people” [37]. This method found evidence for massive invasions of, for example, ‘Celts’, Corded Ware ‘people’ and Beaker ‘people’ in Europe. Indeed, so prevalent was it to consider migrations or invasions the primary vehicle for change that the reaction was nearly three decades (ca. 1960-1985) of the idea being anathema [50]. Following Rouse, migrations should currently be admitted as having occurred, even frequently, but that reasonable causative factors, e.g. greater power and prestige for migrants in the new homeland, and alternative hypotheses must be postulated [51, 52]. As regards the spread of language families, migration hypotheses still enjoy great popularity (see [52] and references therein), largely because it is assumed that for a single family to rise to pre-eminence over a certain area requires a single, directed process. Simulations of language shift, however, indicate that patterns of language families similar to that observed need not hitchhike with any great, unified cultural or biological wave of change [53].

Migrations linked to archaeological change or language shift were very attractive to population geneticists in that it furnished them with clear-cut hypotheses to test, generally working on the assumption that migrants and natives will be genetically distinguishable, and also because migration/admixture are among the easier events to detect. This approach can sometimes yield interesting and valid results, for example regarding the settlement of Iceland primarily from Norway and Norse-controlled regions of the British Isles [54, 55, 56], but it likewise leads to the temptation to link any kind of genetic patterning to known episodes of migration – or at least, migrations rooted in archaeological literature. While it is undoubtedly attractive to unite what seem to be congruent episodes, it is sometimes done to the detriment of unknown demographic events, of which there are doubtless many, and also occasionally lends weight to events which may originally have a dubious archaeological basis.

Migration, as an explanatory concept, must not be invoked simply because it comes to mind readily. Population movements certainly do occur and can have a major impact on the genetics, language and culture of the receiving population, depending on the ratio of immigrants to native, the *a priori* differences between natives and immigrants, the distribution of immigrants, and their relative prestige. They should nevertheless not be allowed to take an unquestioning precedence over other and autochthonous processes.

1.3 Datasets

The datasets used in this work are primarily collections of DNA samples, from the analysis of which geographic patterns in genetic variation are to be recovered. This section is meant to give only a cursory introduction into the types of data that will be used and the kinds of publications consulted, as an aid to understanding the scope of the project. See Chapter 5 for a more detailed description of the primary collections.

1.3.1 Primary

These are the datasets for which the physical samples (DNA or blood for DNA extraction) are either collected by or provided to the PoBI project, and which constitute the samples analysed in the course of this work.

1.3.1.1 The People of the British Isles Project (PoBI)

The People of the British Isles project has collected a large number of DNA samples from all areas of Great Britain (see Chapter 5). The samples are carefully selected to give the best chance of being representative of long-term regional ancestry. This selection process allows the genetic patterns found in these samples to be projected back in time with reasonable confidence.

1.3.1.2 PoBI ancillary collections

These are collections from Western European coastal countries, primarily to be used to put the PoBI collection in its continental or insular context. They consist of samples from Norway (collected by the author), Brittany (Western France), Holland, and Northern Ireland. Whenever possible, the samples have been subject to similar collection criteria to the British samples.

1.3.2 Secondary - Available collections

Vast, publically available depositories of data will be used to provide a background to the study. These collections may allow less restricted grounds for comparison than syntheses of the data they contain, e.g. the possibility to pool or contrast raw data rather than qualitative conclusions.

1.3.2.1 Genetic and morphological databases

After technological revolutions allowed large amounts of DNA-based data to be produced, online databases covering the known extent of human genetic diversity have been constructed, maintained and kept up to date. These repositories contain not only sequence variation data, but also related frequency and geographical information that is invaluable for the rational planning of population genetic studies, as well as for the purposes of comparison or corroboration of studies already undertaken. In addition to the modern genetic databases, anthropometric data collections are available for some present and, more importantly, past populations, and judiciously treated, such morphological data can be very useful as a complementary approach to human demographic history [57].

1.3.2.2 Archaeological descriptions, databases, and catalogues

Central, curated databases like those that exist for genetic data are lacking for archaeological data. Such online resources as exist are usually limited to a single, often small, area or a single project (e.g. the Portable Antiquities Scheme Database or Wessex Archaeology's Artefacts from the Sea Database), and concentrated, comprehensive efforts are rare. Reliance must therefore be placed on published accounts which provide reviews of artefacts, ecofact and site type distributions, and how, where and when these distributions coincide. Potential pitfalls may arise from the fact that such accounts tend to have a focus on interpretations rather than raw data, thereby limiting their utility in comparative studies.

1.3.2.3 Linguistic corpora for Western Europe

These corpora are divided into two sections, both of which will be used to elucidate previous linguistic landscapes of Western Europe. First, sources contemporary or near-contemporary to the era of interest will be considered. These sources are, for most of the area, particularly the Northern section, very patchy, and often non-native for a given language (e.g. place names or personal names for Western Europe given in Latin or Greek texts). Naturally, sources are completely lacking for prehistoric periods. Secondary sources, consisting of reconstructed language scenarios and assessments of historic and prehistoric language contacts and substratum effects will also be consulted.

Third, toponymic (and localized ethnonymic) sources for different times will be consulted, for information on likely diachronic linguistic affiliations. Prime among these sources is Ptolemy's *Geography*, as containing the earliest comprehensive description of Britain and its surroundings, e.g. from de Hoz *et al* [58].

Chapter 2

Human genetics and population structure

2.1 Genetic differences

Genetic differences are, simply, differences in the sequence between two sections of DNA that are homologous, or descended from a common ancestor. In human terms, genetic differences can mean two things on an individual basis: 1) the differences in sequence between each half of a chromosome pair within an individual,[1] and 2) the differences between homologous sequences in two different individuals. These differences can take three basic forms[2] (see Figure 2.1); single nucleotide variation (or polymorphisms, SNPs), repeat number variation and copy number variation – insertions and deletions of differing sizes are included here. SNPs are the difference of a single nucleotide at a given locus, and the term is commonly used even for variants present at sub-polymorphic frequencies (e.g. <1%). Repetitive DNA, such as mini- and microsatellites, have discrete sequence units of up to a hundred base pairs, and the variation consists of the number of units present at a locus [59]. Different numbers of copies of a given stretch of DNA, as well as insertions and deletions of different sizes are features of the genome; but automated assessment of these variation categories is less advanced.

---

[1] Being diploid organisms, humans have two copies of each autosome.

[2] This is a simpification for purposes of clarity.

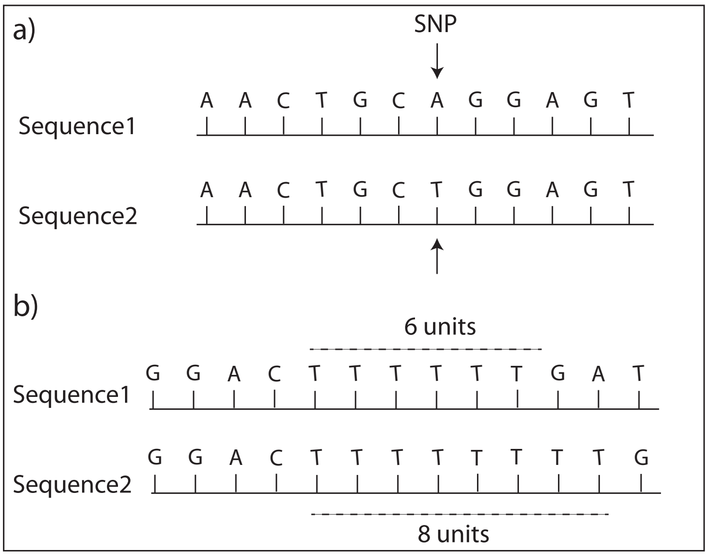


Figure 2.1: Schematic of two types of genetic differences, a) a SNP, and b) a microsatellite.

While microsatellites found early favour in DNA fingerprinting, being easy to analyse and their combinations highly individual, much later focus has been on SNPs. There are over 46 million single nucleotide variants known in the three billion base pair human genome, and every single variant compatible with life probably exists in the current population of seven billion [60, 61, 62, 16].

In addition to variation at single sites, the linear nature of chromosomes means that the variants of a chromosome are not dissociated from other variants on the same chromosome – they are linked, and the combined variation over a stretch of DNA is known as a haplotype [59]. This linkage is usually disrupted every generation at gametogenesis, with an average of 36 crossover events per genome per generation [63]. Linkage disequilibrium – the correlation between alleles – is broken down by crossover events, and the patterns of linkage disequilibrium is of great use in interpreting demographic history, as they provide much more information than that which may be garnered from separate, unlinked loci [64, 65].

All these differences arise from mutations of some sort or chromosomal crossover.

2.2 Population structure

Populations can be defined in several ways and at several levels. At the highest level, 'population' can include all the members of a species. At lower levels it often refers to a group of same-species individuals, defined by the territory they inhabit, on a scale from a continent to a puddle. For humans, the smallest relevant population definition is that of a habitually interbreeding group. Practically, however, such groups are difficult to determine, and heuristic definitions include ethnic or geographically determined groups. If a larger scale population definition is chosen, say that of a modern nation state, one can also speak of subpopulations within it.

Population structure arises when the kind of genetic variation detailed in Chapter 2.1 is unequally distributed between the subpopulations. Such differences between groups of people can be created in several ways, which can be divided into two main categories: demographic events and selection. Selection, whether positive (adaptive) or negative (purifying), can act to differentiate populations if the selection pressure is different, or assimilate populations if the selection pressures are similar. In either case, selection will affect only the locus targeted by selective pressure (and closely linked sequences).

Demographic history events include genetic drift, migration, and mating patterns. Drift is the fluctuation in allele frequencies caused by chance differences in allelic success for any given generation. The effect of drift is related to the size of the population, and the average time before an allele reaches fixation (or extinction) is a function of its frequency at the point of observation and the effective population size. The smaller the population size, the more extreme an effect drift will have. In general, drift will tend to differentiate subpopulations. All subpopulations will be affected by drift, according to the size of the breeding populations, but as the process in random, there is no directed effect – allele A might drift to greater frequency in population X, while allele B might increase in population Y.

Migration, or gene flow, is the exchange of individuals between populations, and has the effect of assimilating their average genetic profiles. The strength of this effect

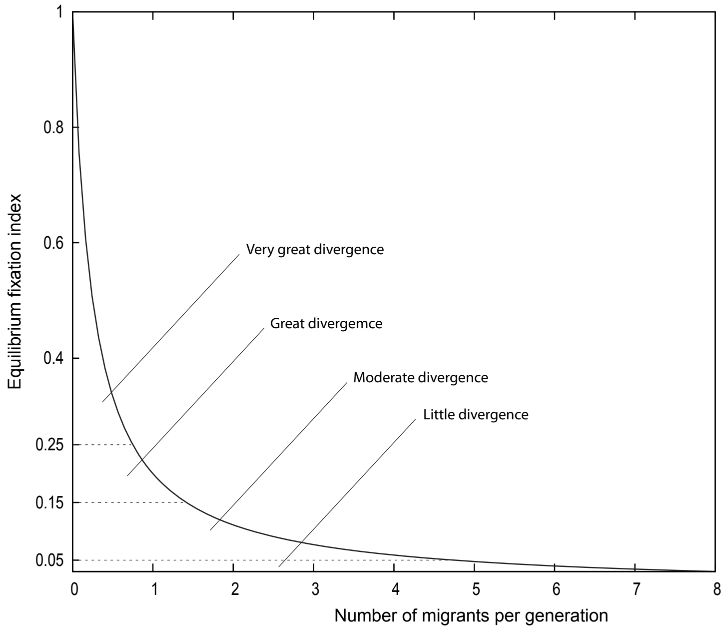


Figure 2.2: **Effect of migration on $\mathbf{F}_{ST}$** Schematic illustrating that with increasing migration, the divergence between subpopulations decreases. After Hartl 2000 [66].

is substantial: “It is remarkable how little migration is required to prevent signficant genetic divergence due to remote inbreeding within the subpopulations”, (see Figure 2.2) [66]. The relative amounts of migration and the amount and directionality of drift that have occurred in the history of a population are unique to that population.

The counteracting processes that create and destroy population structure are important to keep in mind when one attempts to interpret the observed structure in any given population. Most normal variation is assumed to be neutral or near-neutral with respect to selection, so selection is only invoked when the average variation at a locus is unusually high (for differential adaptive selection) or unusually low (for similar adaptive selection or purifying selection) between subpopulations [67].

Founder effects, forming a sub-category of drift, are also important in the development of a population’s genetic character. A newly founded population group will contain a subset of the genetic variation of its parent group, and the chance sampling of those founders will often lead to the new group having a different average genetic profile from the group of origin [68]. This effect is most prevalent with relation to newly arisen variants.

The most basic way of determining population structure is by uncovering differences in allele frequencies at single loci between populations, and more advanced methods are predicated on combining data from single locus allelic variation in terms of haplotypes, genotype, etc. [69].

2.3 Genetics as a tool in historical demography

After describing genetic differentiation between populations, an explanation is sought for this differentiation – a diachronic perspective is thus an inherent part of the field. Ever since the first genetic polymorphisms in humans were discovered, their existence has been used to study population differences [70]. The relationship between genetic structure and population history is oddly reciprocal in terms of explanatory use. Known population history is used to explain observed genetic structure, and observed genetic structure is used to inform population history. Ideally, a combination of population history as determined by palaeontology, archaeology and historical literature – in diminishing time scales – and population genetics can refine the narrative, each contributing aspects the other lacks.

Described below are some of the strategies that have been employed with this goal in mind, from the earliest indirect approaches to mass sequencing.

2.3.1 Classical markers

Classical markers mostly refer to genetic variants that were assessed indirectly, via their protein products – these may reflect a host of different molecular variations. In humans, the first of the markers to be used was the ABO blood group system, whose variant frequencies were seen to vary geographically during the First World War [71]. Other blood groups and serum proteins soon followed (see e.g. Cavalli-Sforza *et al.* [22]), and until about the 1980s classical markers remained the only means of assessing discrete genetic variety.

2.3.2 Uniparental markers

Uniparental markers are those sections of DNA that are passed on to offspring solely through one sex, i.e. the Y chromosome which is only present in and passed down by males, and the mitochondrial chromosome which is present in both sexes but only persists in the female line. Both systems are tractable for the study of populations in that they are inherited as a block, because they are not subject to recombination.[3] This feature is important for their use in phylogeography, which attempts to explain the geographic distribution of branches of a gene tree, or *haplogroups* as the Y and mitochondrial groups are called [72]. Another feature of these markers is their small effective population size, which is only a quarter of that for autosomes. This provides a double-edged sword in terms of patterning observed and its historical-demographical meaning, in that patterning is easy to establish by drift, but is by the same token easily overwritten by the same mechanism. Drift can, especially in the case of males, be accelerated by social practice, e.g. a society with one clan of dominant males could rapidly come to have only a single Y haplogroup present. This type of situation highlights another difficulty with the use of uniparental markers, that is, that they may in some cases present a very skewed version of the demographic history of the population in which they are found.

Even as great advances have been and are being made in both the technological advances needed to genotype or sequence large amounts of autosomal DNA and the statistical techniques to analyse them, the uniparental markers remain useful addenda. They not only constitute a large part of available population genetic data, but for the perspective of gender-specific history they are invaluable.

---

[3] The Y chromosome has a very small portion, 5%, that is susceptible to recombination with the X chromosome [28], so the portion used in population studies is often termed NRY, for non-recombining Y.

2.3.3 Autosomal markers

Autosomal markers are those present in chromosomes 1-22, of which both sexes have two copies. These represent the overwhelming majority of genetic information.

2.3.3.1 Ancestry informative markers

This approach attempts to identify genetic variants that can differentiate between different groups of people. However, a different suite of genetic variants is needed to achieve this discriminatory effect dependent on the populations in question [73]. Very large variant panels may be required for closely related populations, and with the lessening costs of mass genotyping (and sequencing) this approach will become obsolete.

2.3.3.2 High-throughput genotyping

Very high throughput SNP typing has been available for some few years, providing a wealth of SNP data that can be used – in a largely unbiased way – for population genetic studies. Whole genome sequencing is also approaching a cost and efficiency threshold to allow it to be widely used at the population level [74].

2.4 Phenotypes and genotypes

2.4.1 Pigmentation

External pigmentation in humans is the most visually striking example of phenotype variation in humans. Skin colour varies on a global scale, from very pale to very dark, largely correlating with distance from the equator (excluding populations that are the result of recent intercontinental migrations, and see Chapter 3). Hair and eye colour deviation from the standard black and brown, respectively, is almost wholly limited to those of European descent, particularly those from northwest Europe [35]. These phenotypes are, with the exception of modification by bleaching, tanning and freckling, genetically determined. Basal pigmentation depends on the type of melanin produced, yellow-red pheomelanin or brown-black eumelanin, and the size, number and distribution on melanosomes [75].

Both types of melanin are produced by people of all phenotypes, barring those with severe albinism. However, individuals with pale skin and hair and eye colors other than black and brown, respectively, have less eumelanin. In hair and skin, melanin is passed in organelles called melanosomes from melanocytes to keratinocytes, while in the iris the melanosomes are retained in the melanocyte [76]. Over a hundred genes have been implicated in melanogenesis and transport, and others have been linked to colour variation in genome-wide association studies [77, 78, 79, 80, 81].

Melanin is produced via several pathways, all starting with tyrosine (see Figure 2.3). The initial step from tyrosine to dopaquinone is catalyzed by tyrosinase (TYR) for both the production of pheomelanin and eumelanin. After this, the two types diverge, with cysteine being incorporated for pheomelanin, and for eumelanin the steps from initial tyrosine conversion to the final pigment polymer are catalyzed and enhanced by a number of proteins including TYRP1, DCT, P (encoded by *OCA2*), MATP (encoded by *SLC45A2*) and SLC24A5. Elimination of these latter proteins leads to milder forms of albinism, and inefficient variants of them contribute to lighter pigmentation phenotypes [82].

Receptors on the surface of melanogenic cells and their ligands also affect the ratio of pheo- and eumelannin produced. The first gene connected to normal pigmentation variation, *MC1R*, encodes a receptor for MSH (melanogenic stimulating hormone), and functionally impaired variants of this gene cause an overall paler phenotype [83, 84, 82]. *MC1R* is one of best studied pigmentation genes due to its harbouring the strongest variants causative of red hair, and the link with melanoma that the red hair/pale skin phenotype shows [84]. ASIP, a competitive inhibitor of MSH, is likewise associated with normal pigmentation variation. In this case, ASIP binding to MC1R causes a reduction of eumelanin production and a concomitant rise in pheomelanin production, and a variant found in ASIP contributes to darker phenotypes [85].

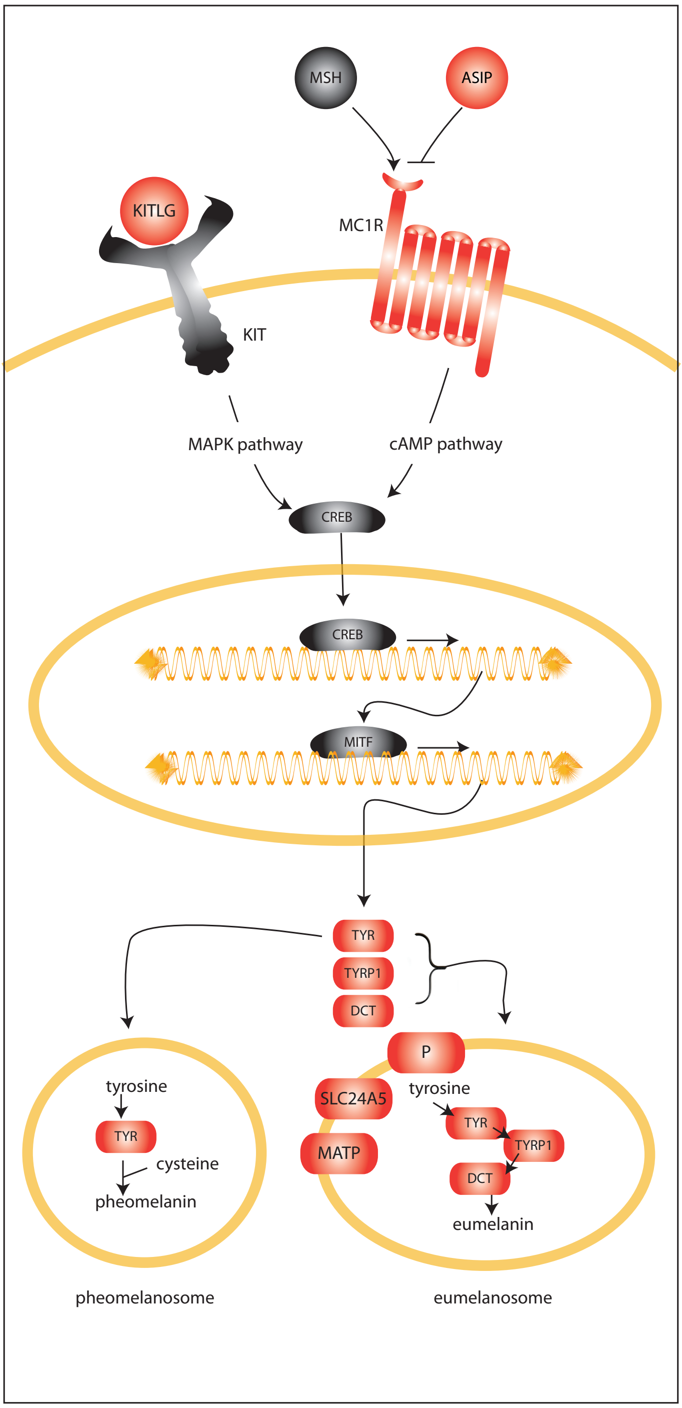


Figure 2.3: **Schematic view of melanogenesis** Proteins with genetic variants known to be associated with normal variation are in red.

While for some genes implicated in pigmentation and melanin synthesis, such as *MC1R*, the causative variants of a change in phenotype are missense mutations, many others – perhaps the majority – are likely to be changes in regulatory sequences. This is the case with the SNP that accounts for the near-Mendelian inheritance patterns seen in blue eyes, where a regulatory SNP for *OCA2* in a *HERC2* intron is nearly perfectly associated with light versus dark eyes [86].

For many genes found to be associated in pigmentation GWASes, the mechanism by which the associated variant or linked locus exerts an effect is unknown. As with disease GWASes, many of these associations may well be spurious, and their associated odds ratios are often very small [87]. So far, over a dozen loci have been strongly implicated in such studies [78, 79, 88, 76].

With few exceptions, most notably the strong red-hair causing variants of *MC1R* and the 'blue eyes' regulatory variant, predicting hair and eye phenotypes on the basis of genotype has been largely unsuccessful, see e.g. Branicki *et al.* 2011 [89].

2.4.2 Genetics of the face

A man finds room in a few square inches of his face for the traits of all his ancestors

- Ralph Waldo Emerson

Facial features are another trait in which there is great variation in humans. On a global scale, the average facial features are very different. Even on smaller scales, such as Europe, certain distinctive combinations can be observed, though in the latter case, extremes are more telling than averages. It is surprisingly often possible, given sufficient background experience, to make a positive regional identification of an individual's European background, if not a negative one.[4] Facial morphology is almost wholly determined by genetics. This is intimated by strong familial resemblances, and

---

[4] A negative identification here implies that one does not feel able to determine regional origin based on facial features and coloring, not that one discounts any specific area.

confirmed by twin studies. Monozygotic twins, even when raised separately, have virtually identical features [90]. Given this, it is reasonable to assume that, in the presence of geographical population structure, there will be a concordant, subtle geographic differentiation of facial features even on sub-continental scales. The genetic basis for normal variation in human facial features, however, is poorly known.

One causative SNP in *FGFR1* is securely linked with changes in cephalic index, and different SNPs in *GHR* are associated with mandibular height in Koreans and Han Chinese [91, 92]. Sonic hedgehog (*SHH*) dosage contributes to upper jaw variation as well, and regions flanking *ENPP1* are associated with face breadth and lower face height [93, 94]. Several recent papers also find associations between *PAX3, PRDM16, TP63, C5orf50, and COL17A1* and 'landmark data' on facial shape [95, 96]. Beyond such indications, which are very weak and leave most of the near-complete heritability unexplained, specifics are unknown.

Facial shape is primarily moulded on the underlying forebrain and skull shape, which starts forming early in development [97]. The vertebrate skull is generated from ectoderm-derived neural crest cell mesenchyme, as opposed to most bone, which is derived from mesodermal mesenchyme [98]. The skull bones and cartilage grow from five tissue buds, or *processes*, that surround the oral cavity, the frontonasal process, the mandibular process, and paired lateral nasal, medionasal and maxillary processes (see Figure 2.4) [98, 99]. Complex spatio-temporal interaction via molecular signals between the cranial mesenchyme and the overlying epithelium then determines the fate of the skull [100]. Many of the genes involved are growth factors, their receptors and homeobox-containing transcription factors (see Chapter 5) [99]. As such, their expression is not limited to the embryonic facial region, and the spatio-temporal precision of expression is crucial [100]. Lack of expression, in the case of *HOX* genes, is also crucial, as skull bones and forebrain fail to develop if they are active [101]. Cell adhesion molecules, including cadherins and integrins, are also critical [98]. The creation of facial skeletal muscles is poorly understood, but in common with the skull bones, it is a system largely separate from that of the trunk skeletal muscles, and also underlies

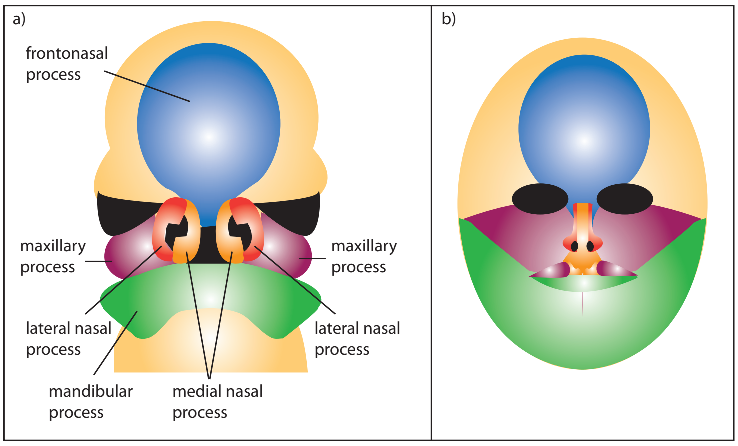


Figure 2.4: **Facial processes** a) Process location in early development, b) Contribution of each process to developed face.

cranial neural crest cell control [100].

Chapter 3

Genetic patterning in Central Eurasia: population history and pigmentation

3.1 Introduction

Western and Central Asia, occupying a vast and pivotal area of Eurasia, have a unique place in the continent's population history. Many migration waves, from prehistoric to early modern times, have started in, ended in or passed through it, possibly changing biometrical phenotypes, language, and customs [102, 103]. However, the genetic profiling of its human populations has often lagged behind that of Europe and East Asia. The two central regions are interesting from two aspects. Western Asia, perhaps Central Asia as well, is considered to be the source area, after Africa, for population expansion into Europe and East Asia, so it is important for putting the latter areas in context. It is also, in its own right, important for the elucidation of global human genetic variation. This chapter examines several markers that have large known allele or haplogroup frequency differences between Europe and East Asia in a large set of populations from across Eurasia, with a particular focus on Central and Western Asia. The markers examined are five functional coding SNPs in *MC1R*, associated with fair complexion and red hair, the 32 base pair deletion in CCR5 and the mitochondrial hypervariable region 1. We also relate these data to Y chromosomal data from the same populations [104].

CCR5, a member of the seven-pass transmembrane G-protein coupled receptor family, is a chemokine receptor notable for being a co-receptor for viruses, including, in particular, HIV [105, 106]. A *CCR5* variant containing a 32 base pair deletion coding for a non-functional receptor has reached surprisingly high frequencies in some populations, possibly due to a role in limiting viral or other infections [107, 108]. This chapter extends the data on the striking geographic distribution of *CCR5del32* [109].

MC1R was the first gene whose variation was shown to have a demonstrable role in human pigmentation variation [83], having variants that cause red hair, pale skin and poor tanning ability in northwestern Europeans. Encoded by a single exon, the coding region is small and has unusually high nucleotide diversity [110]. The *MC1R* variants were investigated as Eurasia hosts some of the most extreme differences in pigmentation in the world, particularly with respect to hair colour. Diversity in external pigmentation is one of the most visually striking differences between human ethnic groups. Basal external pigmentation is largely genetically determined, and there is substantial geographical patterning in the frequencies of alleles of genes involved particularly in melanin production and distribution [111]. As pigmentation variation largely correlates with distance from the equator and UVR intensity, it is most obviously related to natural selection associated with climatic differences [112, 113]. Inevitably, variation in the distribution of pigmentation gene variants may also be influenced by the population histories of the areas in question.

3.2 Materials

Samples novel to this study were collected and treated as described in Wells *et al.* [104] and Chapter 5. Additional allele and genotype data from the literature were collected for *MC1R* variants and *CCR5del32*. In all cases, samples were from unrelated individuals not subject to any known relevant medical conditions (e.g. melanoma or AIDS). See Appendices B and C, respectively, for a synopsis of the samples genotyped for *MC1R* and *CCR5*, and Figures 3.1 and 3.2 for maps of sampling locations. A subset of the 746 samples described in Wells *et al.* [104] was typed for mitochondrial hypervariable region 1 (HVR1). Some of these samples have also been typed for mitochondrial DNA (mtDNA) in Quintana-Murci *et al.* and Comas *et al.* [114, 115]. In total, genetic information on 58,855 individuals was collected. The geographic locations for these samples span all of Eurasia, including areas that are often poorly covered, such as Central Asia and southern Siberia.

When the unique sample identifiers are not given in the results, individual sample groups were pooled. This was done, primarily, on the basis of population identifiers, these being mostly national or ethnic names. For example, all Uzbek groups would be pooled together, as were all Koreans, etc. In cases where population identifiers would group samples from different geographic regions, these were kept separate. e.g. Iranian2-4 (sampled in Iran) were pooled, but Iranian1 (sampled in Samarkand) was not included. Populations from the same nation state whose designation was regional, as opposed to explicitly ethnic/cultural/linguistic, were occasionally pooled, such as Georgian-speaking Kazbegi and Adjar pooled under Georgian. HapMap samples of ethnic Chinese and Indians were included, although they were collected in Denver and Houston (both USA), respectively.

The Turkish1 samples were collected in West Germany but all subjects were born in Turkey [116]. Several smaller groups were included that do not easily fit into any major regional identity in the areas where they were sampled. These are Kurds, sampled in Turkmenistan, and Russians, Arabs, Jews and some Iranians sampled in Uzbekistan. The Kurdish people are widespread across West Asia, and those living in Turkmenistan represent the eastern limit of their main distribution. The Romani (gypsies) are likely to be of Indian extraction originally, from ca. 1000 AD, as suggested by multiple lines of evidence (see Pamjav *et al.* and references therein [117]). The Uzbek Jews, from Bukhara, have had a presence in the area from around 500 BC, but it is thought that

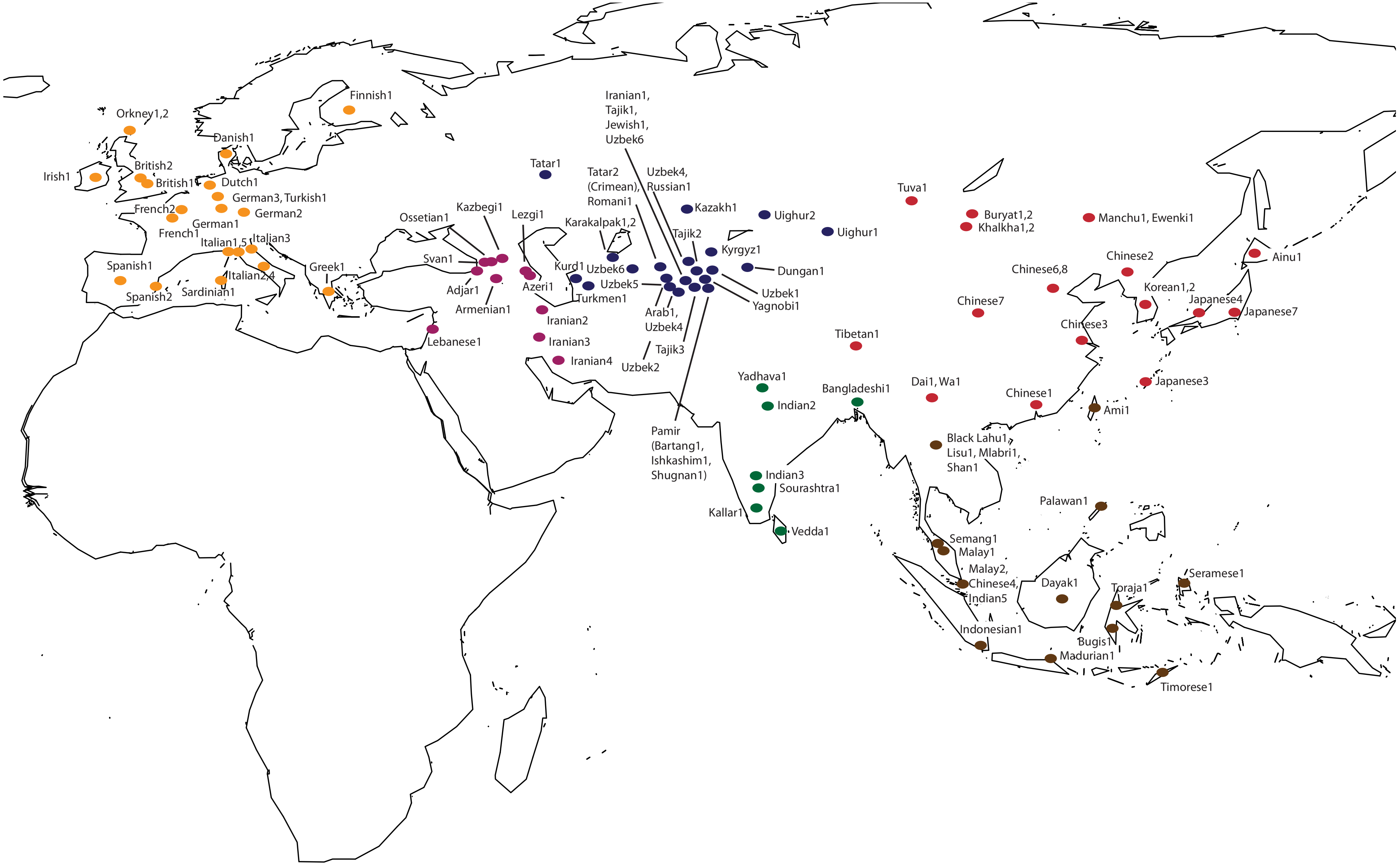


Figure 3.1: **Sampling locations for individuals genotyped for *MC1R* variants** Locations are colored according to region, with Europe - orange, West Asia - purple, Central Asia - blue, South Asia - green, East Asia - red, and Southeast Asia - brown.

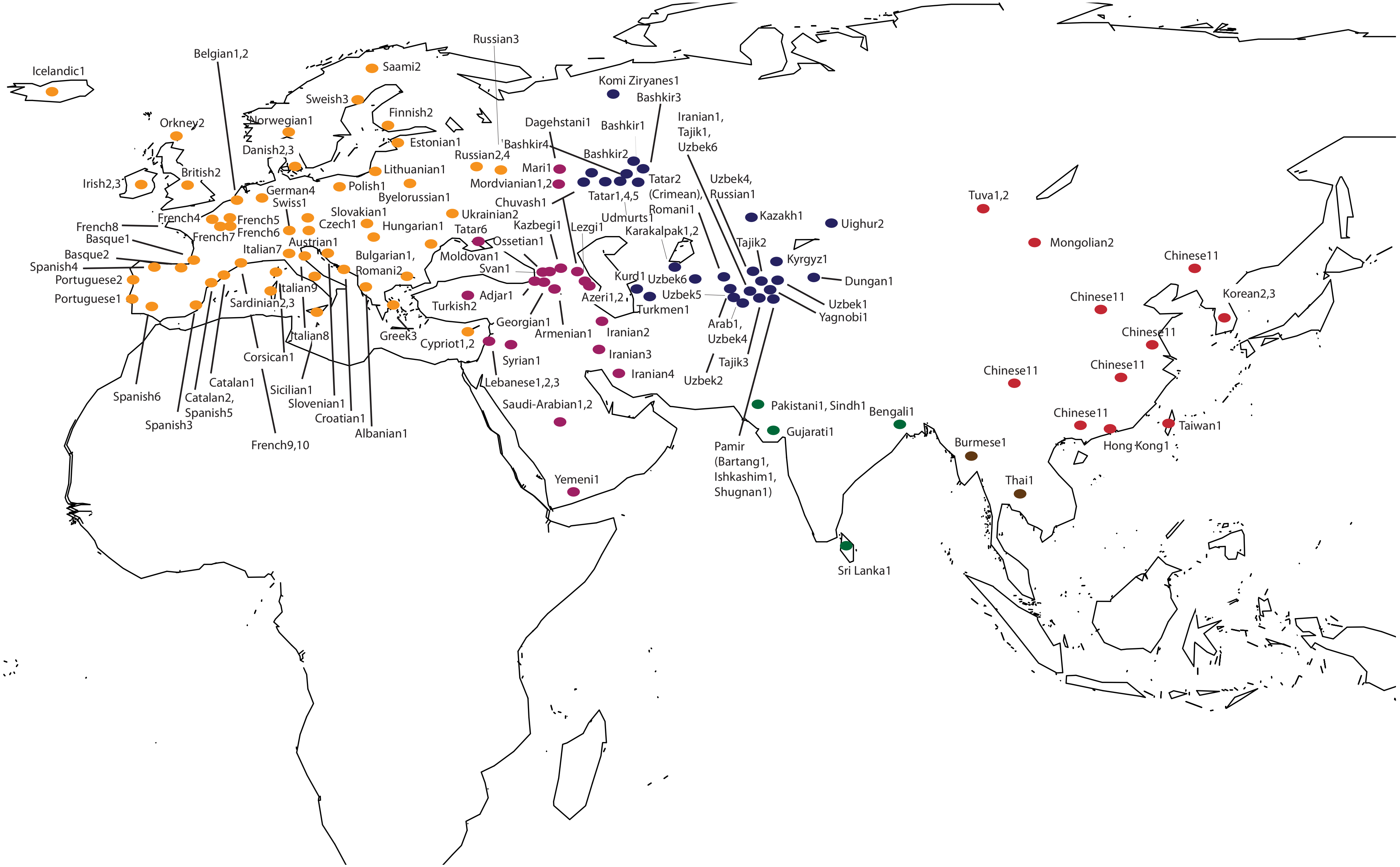


Figure 3.2: **Sampling locations for individuals genotyped for the *CCR5* 32 bp deletion** Locations are colored according to region, with Europe - orange, West Asia - purple, Central Asia - blue, South Asia - green, East Asia - red, and Southeast Asia - brown.

the community is descended from several different influx events [118]. Their language, until recently, was a version of Tajik with a strong Hebrew influence. The Arabs in what is currently Uzbekistan date to around the mid-7th century AD, and the community keeps a distinct identity. Some even used the Arabic language until the 20th century [102]. The Iranians are said to have arrived some time before 1785. Finally, the Russians present in Uzbekistan are descended from ethnic Russians arriving in the country throughout the 20th century, and are not known to have intermarried with other ethnic groups in the region. Though also influenced by the socially dictated degree of endogamy/exogamy, the above groups are likely to be more similar to their geographic neighbours the longer they have been present in their current sampling location.

3.3 Methods

3.3.1 Genotyping

[This was performed by previous members of the Cancer and Immunogenetics Laboratory.]

Genotyping for *MC1R* variants (V60L – rs1805005, V92M – rs2228479, R151C – rs1805007, R160W – rs1805008, and R163Q – rs885479) and the *CCR5* 32bp deletion was carried out in 10 μl PCR reactions containing 10 pmol of each primer, 30 ng DNA, 0.2 units AmpliTaq Gold (PerkinElmer, Vaterstetten, Germany), 175 μM of each dNTP, 1X AmpliTaq Gold PCR buffer, and 2.3 mM MgCl2. PCR products were resolved by gel electrophoresis or alkaline-mediated differential interaction (AMDI) [119]. HVR1 was sequenced between nucleotide positions 16001 and 16571 of the revised Cambridge Reference Sequence [120]. The sequences were edited and aligned using Sequencher software (Gene Codes Corporation. Ann Arbor. Michigan). To maximise the number of sequences available for analysis the sequences used were restricted to a 376 base pair section from position 16039. Haplogroups were determined using MacClade and assigned according to published data [121, 122, 123, 55, 124, 125, 115, 126].

3.3.2 Statistics

F_{ST} for all variants was calculated using Arlequin 3.11 [127]. Regression analyses for the *MC1R* variants, mitochondrial heterozygosity, and the *CCR5* deletion were performed with default linear models, and Pearsons correlation coefficient for mitochondrial and Y heterozygosity was calculating using R v.2.9.2. Principal components analysis was performed in R v.2.9.2 using default settings. Partial Mantel tests were performed using the R package *vegan*, using a default of 999 iterations. For a Mantel test examining residual correlations between MC1R variant F_{ST}s and UV exposure after controlling for geographical distance, exposure was calculated by taking the average of daily erythemal UV measurements between August 1996 and August 2003, with raw data files downloaded from the NASA TOMS (Total Ozone Mapping Spectrometer) website, *http://ozoneaq.gsfc.nasa.gov.* For a Mantel test examining residual correlations between MC1R variants/mitochondrial DNA F_{ST}s and language families, a heuristic distance measure was employed, were dialectal variants were given a distance score of 0.5, members of same sub-family were scored as 2, members of same family as 3, and unrelated languages as 1000. For both partial Mantel tests, great circle geographic distances were employed, calculated using the Haversine formula.

3.3.3 Mapping

For the CCR5del32 data, population samples were given latitude and longitude values. For samples where cities were cited as the sampling location, the position of the city was used. For other samples, the centre of the most detailed geographic position available was chosen, i.e. the centre of a municipality or country (see Appendix C for sampling locations). The longitudinal and latitudinal positions of sampling locations were determined using iTouchMap on GoogleMaps, available from http://itouchmap.com/latlong.html. To construct a frequency surface, a quadratic polynomial regression was performed on the data, using the loess function in R with a scaling parameter alpha = 0.5, chosen to retain some local detail without displaying large fluctuations due to data variability

[128]. The vector map was provided by the Bodleian Map Room.

3.4 Results

3.4.1 CCR5del32

Frequencies of this variant are presented in Table 3.1. The deletion reaches its highest frequencies in the circum-Baltic area, with the Byelorussian sample having a frequency of 16.25%, and its lowest in East and Southeast Asia, with no carriers in the Mongolian, Taiwanese and Burmese samples, and only 16 among 3702 Chinese (0.43%). The frequency of the deletion decreases approximately radially from the eastern Baltic (see Figure 3.3). F_{ST} shows significant differences between Northwest Eurasians, and populations from East Asia, and populations from Southern Europe (see Appendix D).

Table 3.1: ***CCR5del32*** **counts and frequencies**

Population	**N(chromosomes)**	**N(CCR5del32)**	**Frequency**
Albanian	146	12	8.2
Austrian	72	6	8.3
Basque	458	20	4.4
Belgian	2028	204	10.1
British	1512	174	11.5
Byelorussian	160	26	16.3
Catalan	302	22	7.3
Corsican	208	2	1.0
Croatian	606	43	7.1
Cypriot	2172	65	3.0
Czech	322	33	10.2
Danish	726	85	11.7
Estonian	316	42	13.3
Finnish	586	66	11.3
French	3448	365	10.6

Continued on the next page

– continued from the previous page

Population	**N(chromosomes)**	**N(CCR5del32)**	**Frequency**
German	614	66	10.7
Greek	732	29	4.0
Hungarian	198	17	8.6
Icelandic	204	30	14.7
Irish	150	11	7.3
Italian	2560	144	5.6
Jewish	1006	98	9.7
Lithuanian	566	65	11.5
Moldovan	112	9	8.0
Norwegian	200	21	10.5
Orcadian	144	13	9.0
Polish	60	8	13.3
Portuguese	450	26	5.8
Romani	144	3	2.1
Saami	240	20	8.3
Sardinian	728	25	3.4
Sicilian	1802	81	4.5
Slovakian	60	8	13.3
Slovenian	220	17	7.7
Spanish	1424	108	7.6
Swedish	670	94	14.0
Swiss	128	11	8.6
Ukrainian	514	51	9.9
Adjar	70	1	1.4
Armenian	154	3	1.9
Azeri	174	9	5.2
Bulgarian	58	3	5.2
Daghestani	220	14	6.4
Georgian	100	0	0.0
Iranian(W)	1098	71	6.5
Kazbegi	94	9	9.6
Kurdish	74	2	2.7
Lebanese	620	14	2.3
Lezgi	84	4	4.8
Mordvinian	274	34	12.4
Ossetian	50	1	2.0
Russian	708	93	13.1
Arabian	750	10	1.3
Svan	60	1	1.7
Syrian	212	3	1.4
Tatar	712	70	9.8
Turkish	288	18	6.3

Continued on the next page

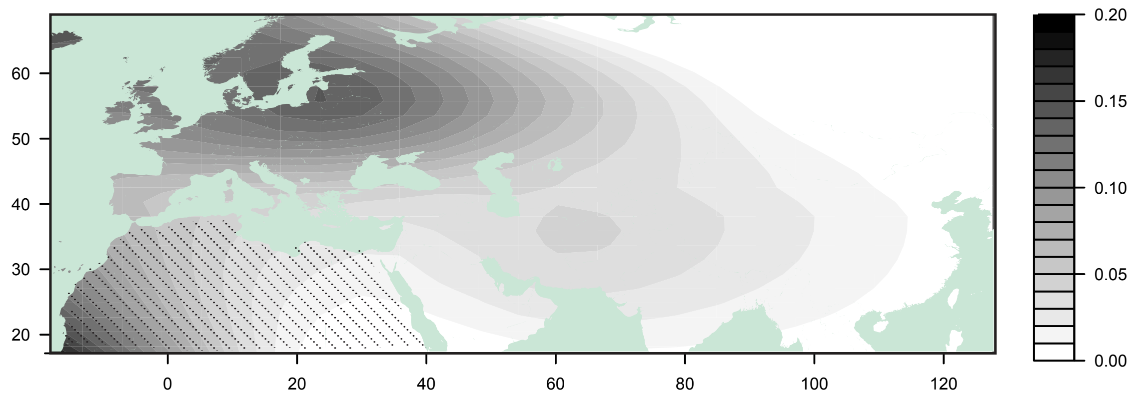


Figure 3.3: **Frequency distribution of *CCR5del32*** The darker color indicates higher frequencies, according to the bar on the right, and the contour lines separate zones of different frequencies. The construction of the frequency surface was based on a quadratic polynomial regression. Note that no samples were sourced from Africa.

– continued from the previous page

Population	**N(chromosomes)**	**N(CCR5del32)**	**Frequency**
Udmurts	104	12	11.5
South Asian	382	7	1.8
Arab	106	1	0.9
Pamir	200	7	3.5
Bashkir	410	15	3.7
Chuvash	158	8	5.1
Dungan	126	3	2.4
Iranian(E)	192	10	5.2
Karakalpak	190	10	5.3
Kazakh	392	15	3.8
Komi Ziryanes	100	6	6.0
Kyrgyz	134	4	3.0
Mari	94	9	9.6
Russian(E)	186	14	7.5
Turkmen	132	5	3.8
Tajik	210	11	5.2
Tuva	302	7	2.3
Uighur	302	9	3.0
Uzbek	1110	59	5.3
Chinese	3702	16	0.4
Korean	202	2	1.0
Mongolian	118	0	0.0
Taiwanese	166	0	0.0
Burmese	134	0	0.0
Thai	202	1	0.5

3.4.2 mtDNA

The haplogroup frequencies observed (see Table 3.2) are in line with what has been previously observed for the study region [114, 115, 129]. Populations of known higher Turkic input (Kyrgyz, Kazakhs, Karakalpaks) have frequency profiles displaying more typically East Asian haplogroups (A, B, C, D, M*), populations from the Caucasus principally have haplogroups associated with western Eurasia, and sedentary Uzbek and Tajik profiles are intermediate (see Figure 3.4). The Iranians from Uzbekistan likewise have an intermediate profile, similar to their Tajik and Uzbek neighbours, and different from those living in Iran. F_{ST} values are low (0-0.073), but a broad differentiation between eastern and western parts of Eurasia may be discerned (see Appendix E). In terms of variability, the Central Asian populations have, with few exceptions, the highest heterozygosity measures (see Table 3.3). This is a similar, but stronger, pattern to what was seen in the Y chromosome data from the same samples for which the heterozygosity is more variable. The correlation between the two heterozygosity measures is 0.3, but statistically it is not significant.

3.4.3 *MC1R*

Variant frequency data for *MC1R* is presented in Table 3.4, and graphically in Appendix F. The five *MC1R* variants have differing geographic distributions (see Figure 3.6 for a schematic overview). Broadly speaking, they fall into two groups: those that reach their highest frequencies in West Eurasia (V60L, R151C, and R160W), and those that have their frequency peak in East Eurasia (V92M and R163Q). For V60L, the basal trend is of ‘high’ frequencies (>7%) west of the Caspian Sea, and ‘low’ frequencies (<5%) east of it. V92M has the most even distribution, for most of the continent, but the higher levels follow an arc from Southeast Asia to Northern Europe. R151C and R160W both reach their highest frequencies (up to 12-13%) in Northwest Europe, with levels decreasing in all directions, though R160W has more of an easterly spread showing higher frequencies in several population samples from the Caucasus than in Southern

Table 3.2: **Mitochondrial haplogroup frequencies** The second column, 'N' gives the number of samples typed for each population group.

Population	**N**	**A**	**B**	**C**	**D**	**F**	**G**	**H/R**	**I**	**J**	**JT**	**K**	**L1/2**	**L3**	**M**	**N**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
Adjar	10	0.0	10.0	0.0	10.0	0.0	0.0	10.0	0.0	0.0	0.0	10.0	0.0	20.0	0.0	0.0	30.0	10.0	0.0	0.0	0.0	0.0	0.0
Arab	17	5.9	0.0	0.0	0.0	11.8	5.9	23.5	0.0	11.8	0.0	0.0	0.0	5.9	0.0	0.0	11.8	17.6	0.0	0.0	0.0	5.9	0.0
Armenian	27	0.0	3.7	0.0	0.0	7.4	0.0	0.0	0.0	7.4	0.0	7.4	0.0	14.8	0.0	0.0	14.8	37.0	0.0	7.4	0.0	0.0	0.0
Azeri	24	0.0	0.0	0.0	4.2	0.0	0.0	58.3	0.0	4.2	0.0	8.3	0.0	0.0	0.0	0.0	8.3	8.3	0.0	8.3	0.0	0.0	0.0
British	57	0.0	0.0	0.0	0.0	1.8	0.0	36.8	3.5	7.0	0.0	3.5	0.0	8.8	0.0	1.8	10.5	17.5	8.8	0.0	0.0	0.0	0.0
Dungan	40	12.0	5.0	5.0	7.5	2.5	0.0	22.5	0.0	0.0	0.0	0.0	0.0	2.5	12.5	5.0	10.0	7.5	0.0	5.0	0.0	0.0	2.5
Iranian(E)	29	3.4	0.0	3.4	6.9	6.9	3.4	17.2	6.9	3.4	3.4	0.0	3.4	17.2	0.0	3.4	6.9	13.8	0.0	0.0	0.0	0.0	0.0
Iranian(W)	13	0.0	0.0	0.0	0.0	0.0	0.0	30.8	0.0	7.7	7.7	0.0	7.7	15.4	0.0	0.0	15.4	15.4	0.0	0.0	0.0	0.0	0.0
Karakalpak	20	0.0	5.0	10.0	10.0	10.0	10.0	10.0	0.0	5.0	0.0	0.0	0.0	0.0	10.0	0.0	0.0	25.0	5.0	0.0	0.0	0.0	0.0
Kazakh	61	4.9	4.9	4.9	3.3	3.3	9.8	19.7	0.0	1.6	0.0	1.6	0.0	9.8	4.0	8.2	6.6	9.8	1.6	1.6	0.0	1.6	1.6
Korean	38	0.0	5.3	2.6	36.8	7.9	2.6	7.9	2.6	0.0	0.0	0.0	0.0	10.5	10.5	5.3	0.0	5.3	0.0	0.0	0.0	2.6	0.0
Kurd	14	0.0	7.1	0.0	0.0	0.0	0.0	21.4	0.0	7.1	0.0	7.1	0.0	14.3	0.0	0.0	7.1	35.7	0.0	0.0	0.0	0.0	0.0
Kyrgyz	33	3.0	6.1	3.0	27.3	6.1	9.1	9.1	3.0	0.0	0.0	0.0	0.0	6.1	3.0	3.0	3.0	12.1	0.0	3.0	0.0	0.0	3.0
Lebanese	24	0.0	0.0	0.0	4.2	0.0	0.0	37.5	4.2	12.5	0.0	8.3	0.0	8.3	0.0	0.0	12.5	0.0	8.3	4.2	0.0	0.0	0.0
Lezgi	20	0.0	5.0	0.0	0.0	0.0	0.0	25.0	0.0	25.0	0.0	0.0	0.0	15.0	0.0	0.0	10.0	15.0	5.0	0.0	0.0	0.0	0.0
Ossetian	11	0.0	9.1	0.0	0.0	9.1	0.0	45.5	9.1	0.0	0.0	0.0	0.0	9.1	0.0	0.0	0.0	18.2	0.0	0.0	0.0	0.0	0.0
Pamir	27	0.0	0.0	0.0	0.0	0.0	0.0	25.9	0.0	18.5	0.0	7.4	3.7	11.1	3.7	0.0	3.7	22.2	0.0	3.7	0.0	0.0	0.0
Romani	8	0.0	0.0	0.0	12.5	25.0	12.5	25.0	0.0	0.0	0.0	0.0	0.0	12.5	0.0	0.0	12.5	0.0	0.0	0.0	0.0	0.0	0.0
Russian(E)	15	0.0	0.0	0.0	0.0	0.0	0.0	53.3	6.7	13.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	13.3	13.3	0.0	0.0	0.0	0.0
Svan	9	0.0	11.1	11.1	11.1	0.0	0.0	44.4	0.0	0.0	0.0	11.1	0.0	0.0	0.0	0.0	11.1	0.0	0.0	0.0	0.0	0.0	0.0
Tajik	25	4.0	0.0	4.0	4.0	0.0	0.0	16.0	4.0	12.0	0.0	4.0	4.0	8.0	8.0	0.0	16.0	8.0	0.0	8.0	0.0	0.0	0.0
Tatar	23	4.3	0.0	0.0	8.7	0.0	0.0	30.4	4.3	13.0	0.0	0.0	0.0	4.3	0.0	0.0	17.4	13.0	0.0	4.3	0.0	0.0	0.0
Turkmen	22	0.0	4.5	4.5	0.0	0.0	0.0	36.4	0.0	9.1	0.0	0.0	4.5	22.7	0.0	0.0	4.5	9.1	4.5	0.0	0.0	0.0	0.0
Uzbek	179	3.4	2.2	5.0	9.5	3.4	3.9	22.3	0.0	6.7	0.6	2.2	0.6	10.1	4.5	0.6	1.7	19.0	1.1	1.7	0.0	0.0	1.7

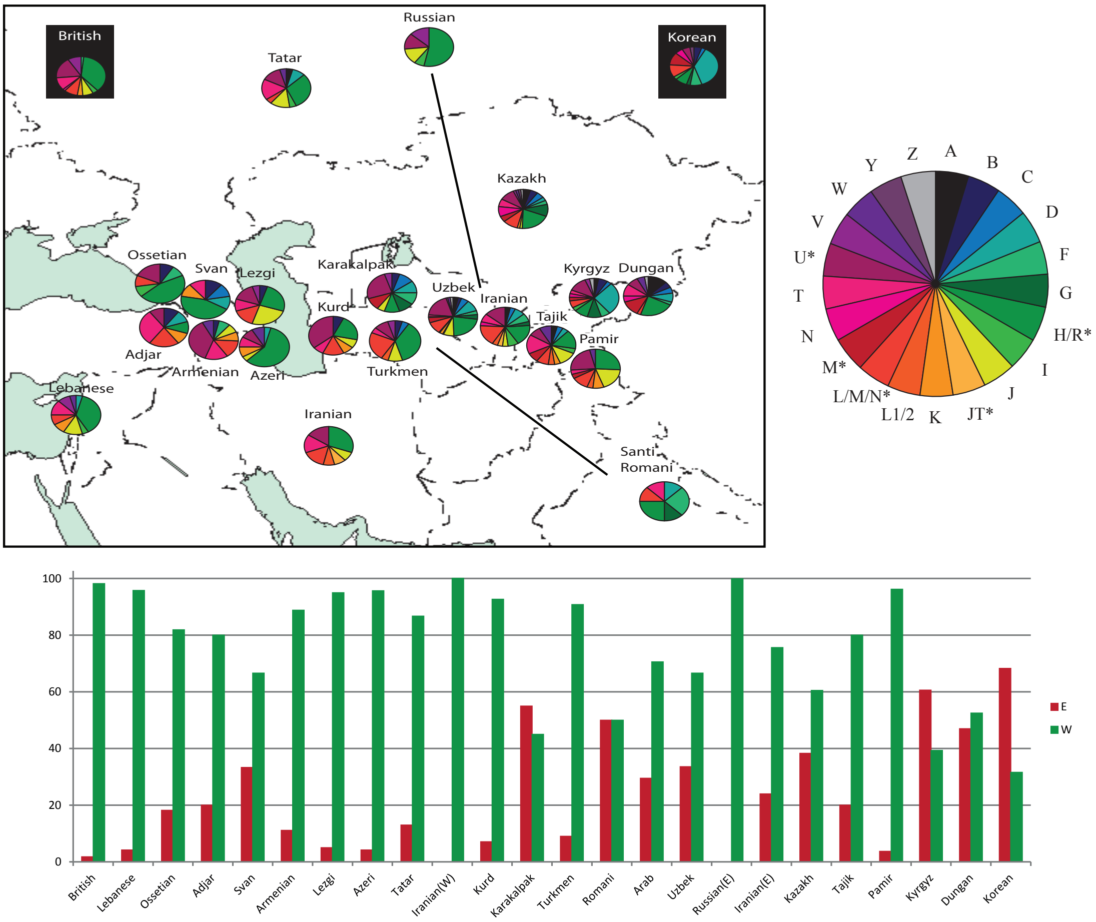


Figure 3.4: **Mitochondrial haplogroup profiles in West and Central Asia** The wheel to the right gives the colour-haplogroup correspondences. The bar graph below gives the frequency per population of combined haplogroups, where green bars comprise haplogroups which are typically 'West Eurasian', and red bars the typically East Asian haplogroups.

Table 3.3: **Mitochondrial heterozygosity** Populations ranked by increasing heterozygosity, defined as one minus the sum of squares of all the allele frequencies.

Population	**Heterozygosity**
Azeri	0.63
Russian	0.66
Ossetian	0.73
Svan	0.74
Kurd	0.79
Turkmen	0.79
Armenian	0.80
British	0.80
Lebanese	0.80
Romani	0.81
Lezgi	0.82
Iranian(W)	0.82
Korean	0.82
Adjar	0.82
Pamir	0.83
Tatar	0.83
Arab	0.86
Karakalpak	0.87
Kyrgyz	0.88
Uzbek	0.88
Dungan	0.89
Iranian(E)	0.89
Tajik	0.90
Kazakh	0.91

Europeans. R163Q is most frequent in East Asia, and is the only *MC1R* variant under consideration to consistently reach frequencies of more than 50%. The Finns and Saami have, for Europe, uncharacteristically high R163Q levels (30% and 15%, respectively) which may be pertinent in elucidating the origins of these populations. For all *MC1R* variants, the combination of single frequencies show South Asians to be significantly different from almost all of the other populations we have studied (by Fisher's exact test, see Figure 3.5), in having a low frequency for all the minor alleles we have typed for, in contrast to the other populations that have relatively high frequencies for at least one of these alleles. This is consistent with these southern populations not having been subject to pigmentation-reducing selection due to the higher levels of UV-radiation on the subcontinent. F_{ST} measures reveal geographical distinctions along a longitudinal axis, excluding South Asia (see Figure 3.7, Appendix G and Table 3.5). R151C and R160W distinguish Northern Europe from the rest of the continent. V60L shows a tendency to separate Europe and West Asia from Central and East Asia. R163Q, for which F_{ST} values between populations reach a maximum of 0.94, divides the continent into blocks approximately consisting of Europe and West Asia, Central Asia, and East and Southeast Asia. Central Asian populations that stand out are the Karakalpaks, Kyrgyz and Kazakhs of northern Central Asia, and the Mandarin-speaking Dungans from the border regions with China, all of which show a greater affinity with the East Asian block by more similar allele frequencies and lower F_{ST} values. The last variant examined, V92M, separates Southeast Asian and southern Chinese groups (Dai and Wa) from other populations. A principal components analysis of the MC1R variants confirms the above (see Figure 3.8, with PC1 accounting for 0.91 of the variance) separating East Asia from West Eurasia and South Asia, with Central Asia in a medial position. PC2 (0.04 of the variance) further separates Northern Europe from the rest of West Eurasia. PC1 largely mimics an east-to-west axis, probably principally due to the influence of R163Q, while PC2, splitting the north of West Eurasia from the south, is probably influenced by the higher R151C and R160W levels in the north (see below).

One way to look at the variation in the frequencies of the *MC1R* variants is to con-

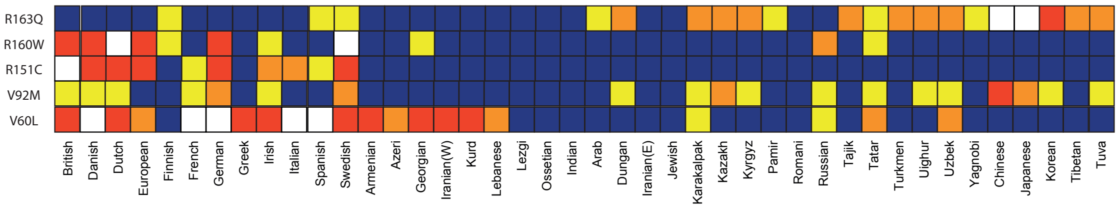


Figure 3.5: **South Asian *MC1R* differentiation** Heatmap of p-values for pairwise Fisher's exact tests for *MC1R* variant frequencies between South Asians and other groups. South Asian show consistent differences from all populations with noticeable frequencies of the given variant. Blue is above the Bonferroni-corrected significance threshold, yellow $< 7.9e-5$, orange $< 1e-7$, red $< 1e-10$, and white $< 1e-15$.

sider the relationship of their frequencies with latitude and longitude. Within Europe R151C and R160W frequencies are good predictors of latitude, with r^2 values of 0.72 for R151C and 0.88 for R160W, both being positive correlations (see Figure 3.9, where latitude is regressed on frequency). As these two variants are the strongest inducers of pale skin, and UV-radiation decreases with increased latitude, this is not unexpected [84]. Furthermore, the combined information provided by the frequency data for each of the *MC1R* variants in a population sample is a reasonable predictor of that samples longitudinal origin within the continent, with a multiple regression $r^2 = 0.69$, where a linear combination of the frequencies is used as the predictor variable in the model and longitude is the response variable.

A partial Mantel test was performed to see if there was any residual correlation between *MC1R* variant frequencies and erythemal UV exposure, after geographical distance was controlled for (see Figure 3.10 for scatterplots of UV exposure by variant frequencies). No significant such correlations were observed for V60L, V92M or R163Q, but moderate and signficant correlations were found for R151C (Mantel r=0.47, p=0.001) and R160W (Mantel r=0.27, p=0.001).

Table 3.4: ***MC1R* variant frequencies** Count and frequency information for V60L, V92M, R151C, R160W, and R163Q in populations from Appendix B.

Population	**V60L**		**V92M**		**R151C**		**R160W**		**R163Q**	
	N=	**%**	**N=**	**%**	**N=**	**%**	**N=**	**%y**	**N=**	**%**
Arab(E)	6/106	5.6	3/106	2.8	1/106	0.9	1/106	0.9	19/106	17.9
Armenian	18/154	11.7	3/154	1.9	4/154	2.6	5/154	3.3	5/154	3.3
Asian	NA	NA	28/120	23	0/120	0	NA	NA	84/120	70
Azeri	11/94	11.7	4/94	4.3	0/94	0	5/94	5.3	9/94	9.6
British	25/166	15.1	19/166	11.5	181/1970	9.2	141/1958	7.2	6/166	3.6
Buryat	NA	NA	10/100	10	0/100	0	NA	NA	259/386	67.1
Chinese	2/546	0.4	238/964	24.7	2/470	0.4	2/460	0.4	1054/1644	64.1
Danish	91/756	12	53/756	7	68/756	9	60/756	7.9	30/756	4
Dungan	3/126	2.4	14/126	11.1	3/126	2.4	1/126	0.8	61/126	48.4
Dutch	63/770	8.2	54/770	7	39/770	5	81/770	10.5	42/770	5.5
European	17/222	7.7	15/222	6.8	27/222	12.2	27/218	12.4	23/222	10.4
Ewenki	NA	NA	20/90	22.2	0/90	0	NA	NA	63/90	70
Finnish	1/30	3.3	2/30	6.7	3/30	10	4/30	13.3	9/30	30
French	62/554	11	44/554	7.9	21/554	3.9	17/554	3.1	26/750	3.5
Georgian	22/224	9.9	6/224	2.7	3/224	1.3	13/224	5.8	10/224	4.5
German	228/2076	11	208/2076	10	125/2076	6	125/2076	6	151/2658	5.7
Greek	35/310	11.3	13/310	4.2	6/310	1.9	3/310	1	2/310	0.6
Inuit	0/40	0	1/40	2.5	0/40	0	0/40	0	37/40	92.5
Iranian(E)	8/180	4.4	6/180	3.3	2/180	1.1	7/180	3.9	14/189	7.8
Iranian(W)	19/170	11.2	5/170	2.9	3/170	1.8	5/170	2.9	4/170	2.4
Irish	16/122	13.1	15/122	12.3	11/122	9	8/122	6.6	6/122	4.9
Italian	212/1194	17.8	39/1192	3.3	45/1194	3.8	24/1194	1.2	24/1192	2
Japanese	1/200	0.5	81/746	10.9	0/188	0	3/218	1.4	1192/1528	81.8
Jewish	3/20	15	1/20	5	0/20	0	0/20	0	1/20	4
Karakalpak	10/212	4.7	16/212	7.6	5/212	2.3	1/212	0.5	107/212	50.5
Kazakh	8/292	2.7	32/292	11	3/292	1	6/292	2.1	131/292	44.9
Korean	1/102	1	32/324	9.9	0/102	0	1/102	1	191/250	76.4
Kurd	14/74	18.9	3/74	4.1	0/74	0	2/74	2.7	4/74	5.4
Kyrgyz	2/134	1.5	15/134	11.2	3/134	2.2	1/134	0.8	79/134	59
Lebanese	11/100	11	0/100	0	0/100	0	2/100	2	1/100	1
Lezgi	2/84	2.4	9/84	10.7	1/84	1.2	2/84	2.4	4/84	4.8
Mongolian	NA	NA	10/80	12.5	0/80	0	NA	NA	399/556	71.8
Orcadian	8/144	5.6	10/144	6.9	40/340	11.8	30/338	8.9	9/44	6.3
Ossetian	1/50	2	4/50	8	2/50	4	3/50	6	3/50	6
Pamir	9/200	4.5	14/200	7	4/200	2	4/200	2	27/200	13.5
Romani	0/50	0	1/50	2	0/50	0	0/50	0	8/50	16
Russian(E)	12/186	6.5	21/186	11.3	6/186	3.2	16/186	8.6	17/186	9.1
Saami	0/26	0	6/26	23.1	3/26	11.5	0/26	0	4/26	15.4
Southeast Asian	NA	NA	543/1316	41.3	0/1140	0	NA	NA	701/1526	45.9
Spanish	217/1464	14.8	63/1464	4.3	37/1464	2.5	13/1464	0.9	28/1464	1.9
Swedish	100/1378	7.3	102/1378	7.4	108/1378	7.8	139/1378	10.1	127/1378	9.2
Tajik	9/242	3.7	12/242	5	7/242	2.9	3/242	1.2	49/242	20.3
Tatar	13/166	7.8	16/166	9.6	2/166	1.2	10/166	6	30/166	18.1
Tibetan	0/40	0	4/40	10	NA	NA	NA	NA	34/40	85
Turkish	NA	NA	NA	NA	NA	NA	NA	NA	30/400	7.5
Turkmen	4/132	3	10/132	7.6	1/132	0.8	0/132	0	40/132	30.3
Tuva	0/102	0	13/102	12.8	0/102	0	102	0	63/102	61.8
Uighur	3/182	1.7	21/182	11.5	0/112	0	0/112	0	69/182	37.9
Uzbek	47/972	4.8	71/972	7.3	13/972	1.3	19/972	2	250/972	25.7
Yagnobi	3/62	4.8	2/62	3.2	0/62	0	1/62	1.6	18/62	29
South Asian	3/762	0.4	7/1006	0.7	2/840	0.2	23/764	0.4	72/1458	4.9

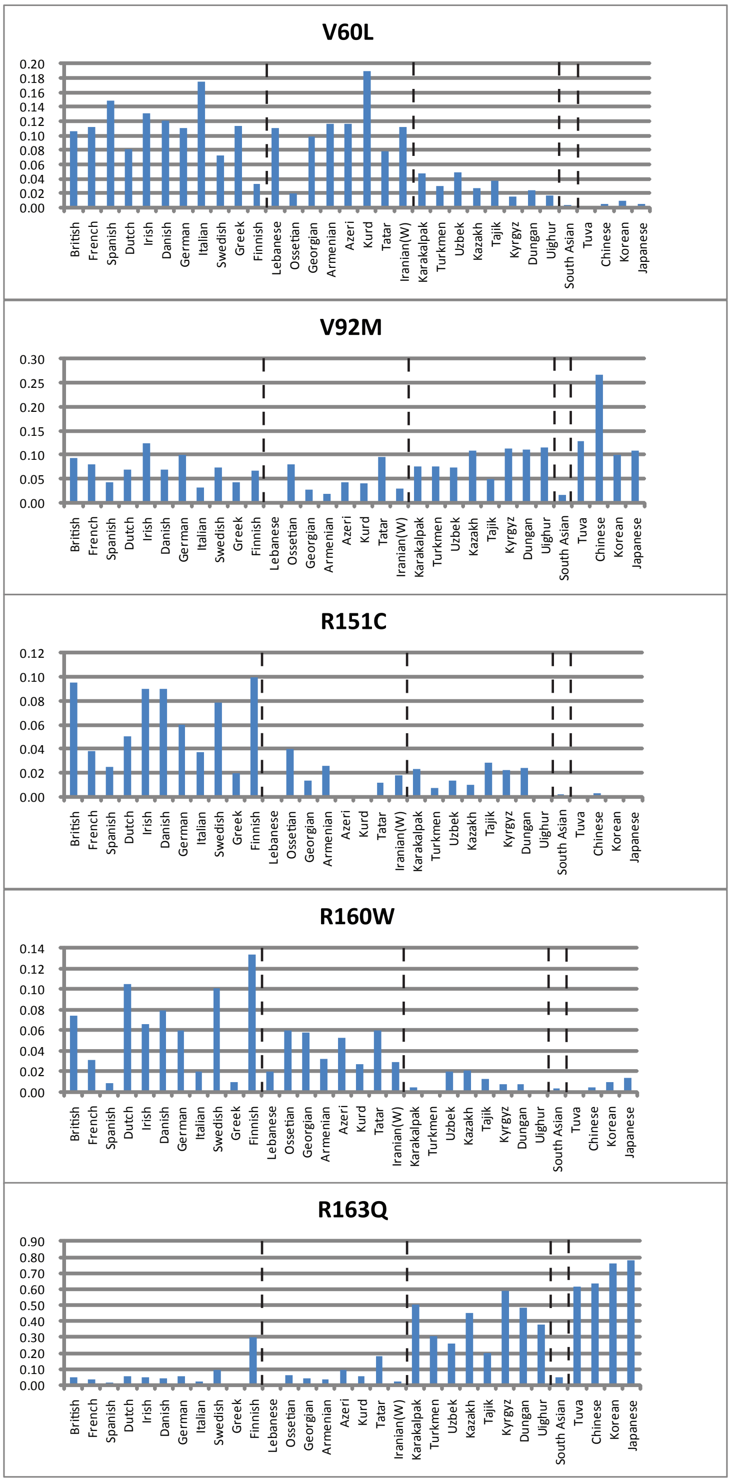


Figure 3.6: ***MC1R* variant frequencies by longitude** Frequencies of the five *MC1R* variants for representative populations, plotted on a west-to-east (left to right) axis. Vertical dashed lines delineate geographical regions, from left to right: Europe, West Asia, Central Asia, South Asia and Eastern Asia.

Table 3.5: **Regional F_{ST} values for five MC1R SNPs** Numbers in bold are considered significant.

	Northern Europe	Southern Europe	West Asia	South Asia	Central Asia	East Asia	Southeast Asia
V60L							
Northern Europe	0.00						
Southern Europe	**0.01**	0.00					
West Asia	0.00	**0.01**	0.00				
South Asia	**0.10**	**0.13**	**0.08**	0.00			
Central Asia	**0.02**	**0.08**	**0.02**	**0.02**	0.00		
East Asia	**0.09**	**0.12**	**0.08**	0.00	**0.02**	0.00	
V92M							
Northern Europe	0.00						
Southern Europe	**0.01**	0.00					
West Asia	**0.01**	0.00	0.00				
South Asia	**0.03**	**0.01**	**0.00**	0.00			
Central Asia	0.00	**0.01**	**0.02**	0.03	0.00		
East Asia	**0.07**	**0.12**	**0.12**	**0.15**	**0.07**	0.00	
Southeast Asia	**0.22**	**0.36**	**0.37**	**0.45**	**0.24**	**0.03**	0.00
R151C							
Northern Europe	0.00						
Southern Europe	**0.03**	0.00					
West Asia	**0.05**	**0.01**	0.00				
South Asia	**0.06**	**0.02**	**0.01**	0.00			
Central Asia	**0.07**	**0.01**	0.00	**0.01**	0.00		
East Asia	**0.06**	**0.02**	0.00	0.00	**0.00**	0.00	
R160W							
Northern Europe	0.00						
Southern Europe	**0.03**	0.00					
West Asia	**0.01**	**0.01**	0.00				
South Asia	**0.04**	**0.01**	**0.03**	0.00			
Central Asia	**0.04**	0.00	**0.01**	**0.01**	0.00		
East Asia	**0.04**	**0.00**	**0.02**	0.00	**0.00**	0.00	
R163Q							
Northern Europe	0.00						
Southern Europe	**0.02**	0.00					
West Asia	0.00	**0.02**	0.00				
South Asia	0.00	**0.19**	0.00	0.00			
Central Asia	**0.17**	**0.21**	**0.17**	**0.17**	0.00		
East Asia	**0.55**	**0.58**	**0.55**	**0.56**	**0.25**	0.00	
Southeast Asia	**0.53**	**0.61**	**0.54**	**0.54**	**0.10**	**0.06**	0.00

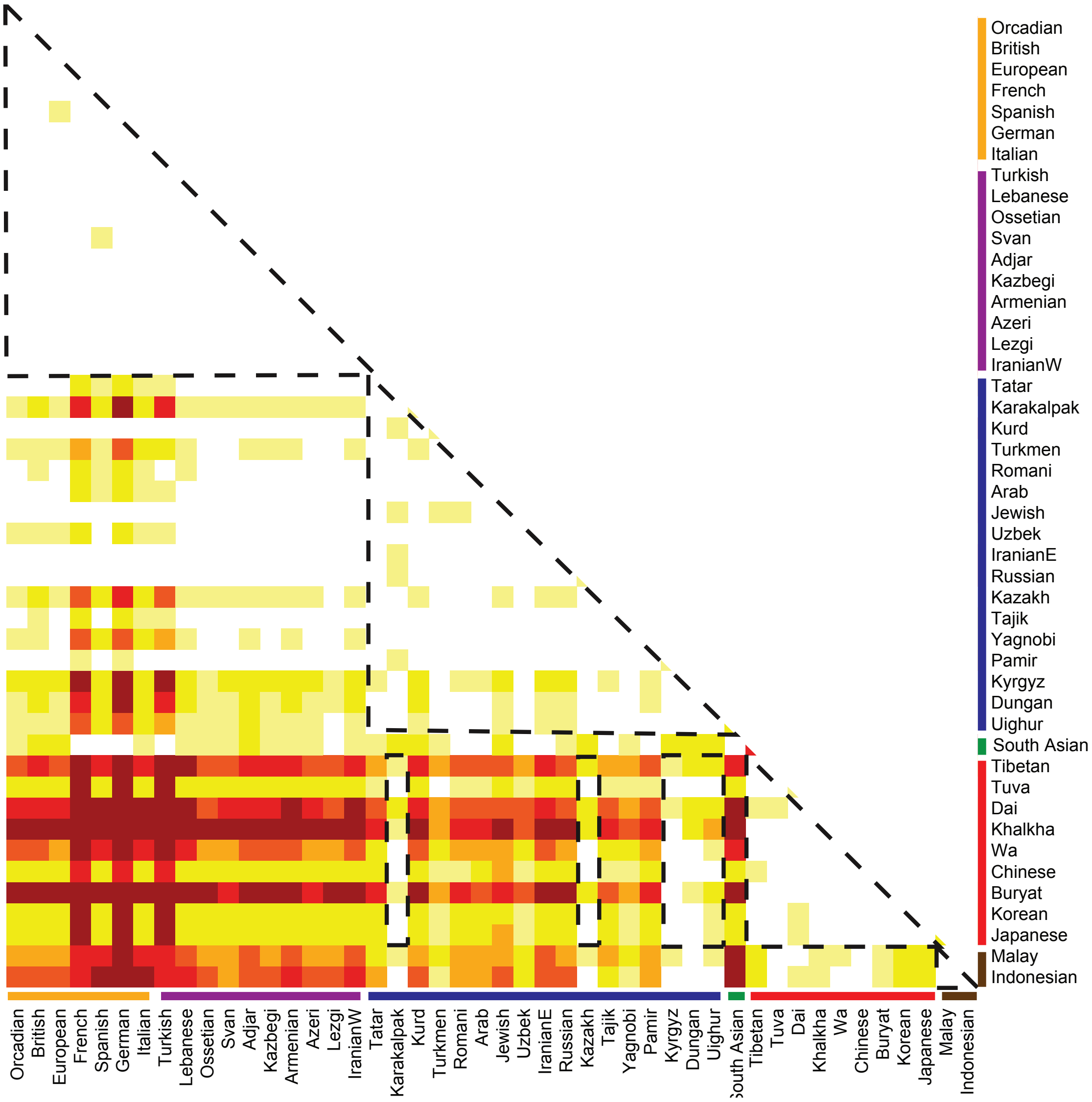


Figure 3.7: **Graphical representation of *MC1R* variant $\mathbf{F}_{ST}$ values** Averages of separately obtained values for each variant, where white <0.05, pale yellow=0.05-0.1, yellow =0.1-0.2, pale orange=0.2-0.3, orange=0.3-0.4, red=0.4-0.5, dark red >0.5. Blocks of similar populations are indicated with dashed triangles along the diagonal. Dashed boxes, as an extension of the East Asian triangle, highlight the greater similarity of the Karakalpak, Kazakh, Kyrgyz, Dungan and Uighur groups to the East Asian populations than that shown by other Central Asian populations. Population names are colour-coded and arranged by continental region, see above and Figure 3.1.

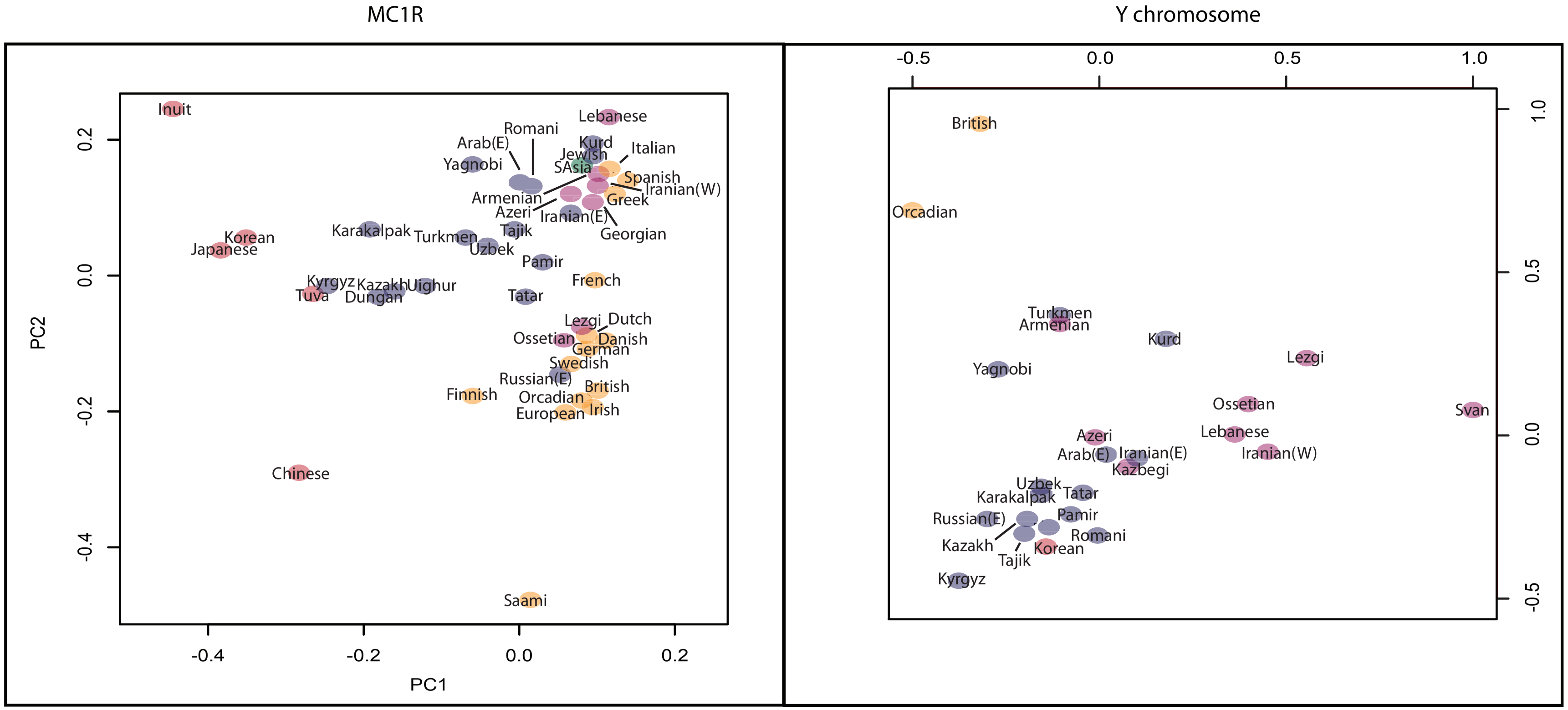


Figure 3.8: **Principal components plots** Left: MC1R variants
, where PC1 accounts for 91% of the variance, and PC2 4%. Right: Y chromosome data (from Wells et al. 2001 [104]), where PC1 accounts for 23% of the variance and PC2 19%. Populations are colour-coded; Europe – orange, West Asia – purple, Central Asia – blue, South Asia – green, and East Asia - red.

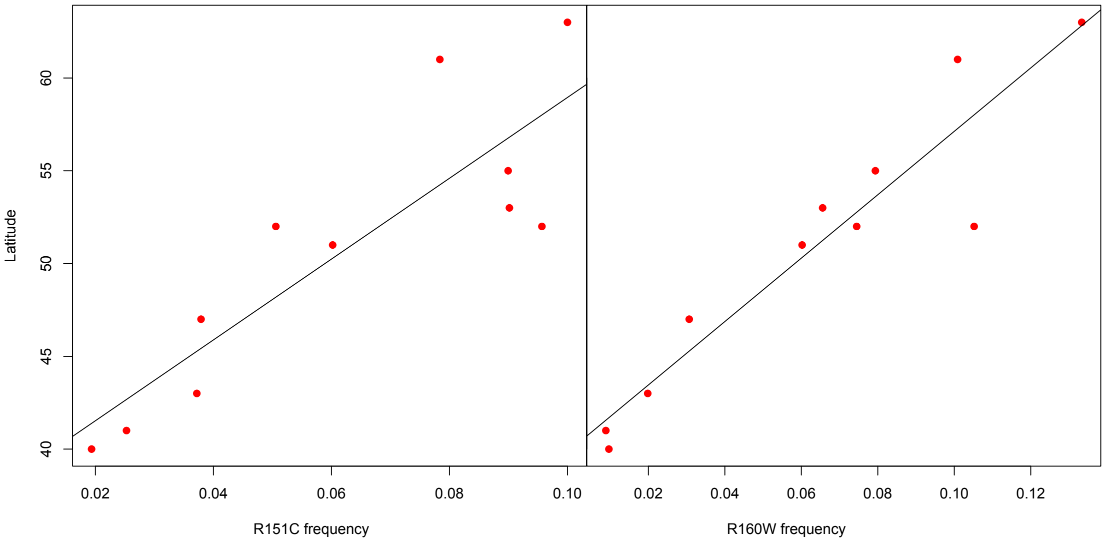


Figure 3.9: **Scatterplots of the two ‘red hair’ variant frequencies** R151C and R160W, by latitude, with the fitted regression line.

3.5 Discussion

3.5.1 Population history

Central Eurasia holds a wealth of cultural and linguistic diversity, as well as being at a genetic crossroads. Central Asia’s earliest modern human inhabitants are deemed to have been Caucasoid by physical anthropologists, though their methodology has occasionally met with scepticism (see e.g. Frachetti [130]), and all its earliest languages on record (from the 6th century BC) belong to the Iranian branch of Indo-European. Nomadic pastoralists of an eastern Asian type (“Mongoloid”) appear in what is now Kyrgyzistan in the 5th-8th centuries AD, and this physical type is claimed to become dominant from the 13th century [102]. There are many historically attested waves of Turkic migration from the east into Central Asia, and across it to Western Asia, starting from around the 5th century AD, however, the number of individuals involved is difficult to estimate [103]. There are still remnants of people speaking the once-widespread East Iranian languages (e.g. the Pamirs), but most inhabitants now speak Turkic languages. Two of the largest ethnic groups in Central Asia, Uzbeks and Kazakhs, are recent (late

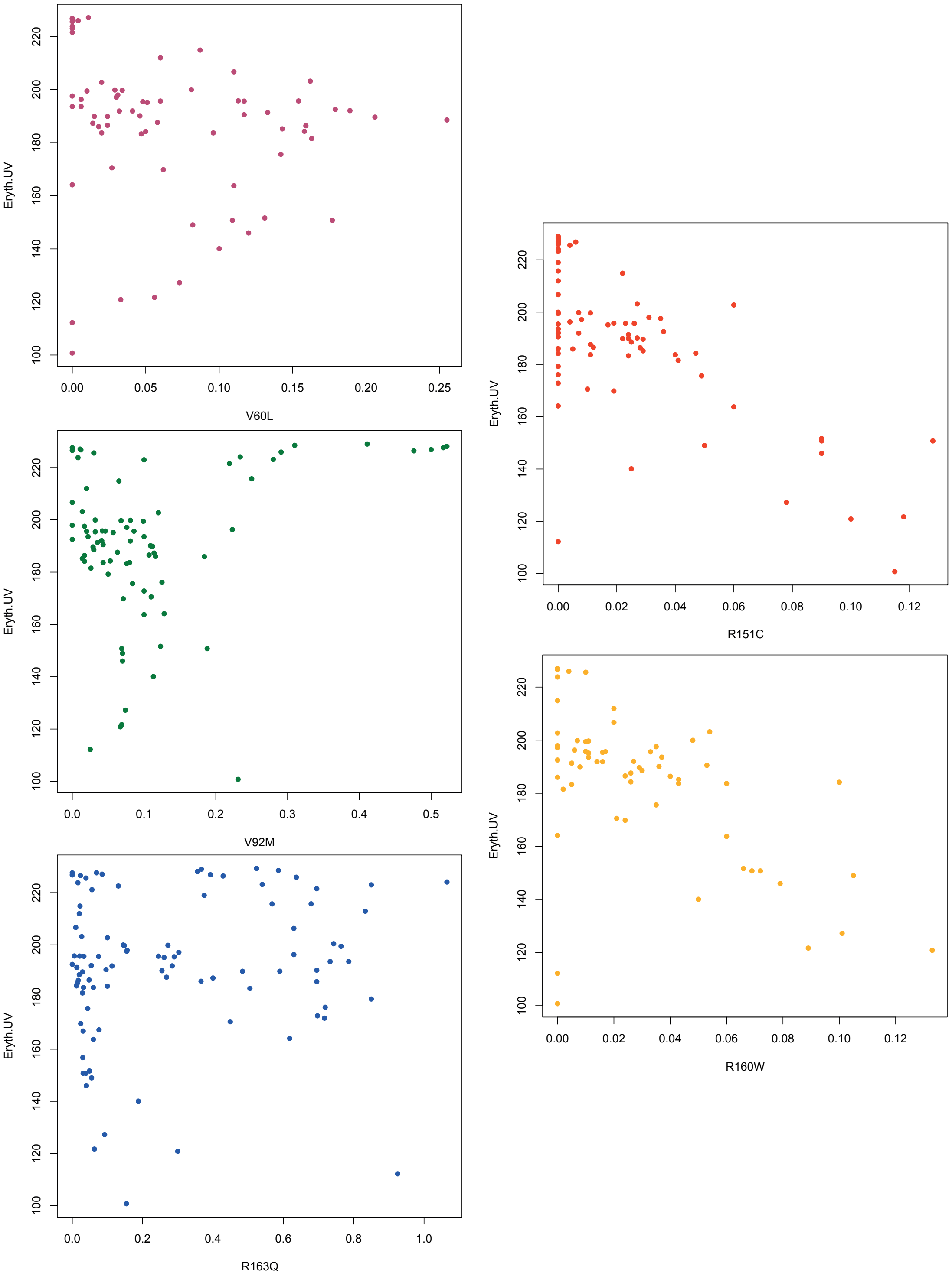


Figure 3.10: ***MC1R* variant frequencies by erythemal UV**

15th century) agglomerations under new names of former groups living in their current areas.

The above illustrates a very small portion of the genetic complexity one may expect to find in the region, though the general stability of the area's population as a whole must be considered. Two main theories have been put forth as to the origins of Central Asias genetic character, which displays both typically European/West Asian and East Asian patterns. The first is that Central Asia has been a primary source of genetic variation for both Europe and East Asia, and the second that most currently observed genetic variants arrived with separate streams of migrants from Europe and East Asia, largely swamping the sparse extant population of the arid region. Different studies find cases for either hypothesis, mostly based on mitochondrial and Y chromosomal DNA markers [104, 114, 129, 131, 132, 133]. Mitochondrial DNA data tends to support a situation where Central Asia has been a contact zone for East and West Eurasian peoples, possibly outnumbering an extant and likely sparse population in the region that is, and was, primarily desert [114, 115, 134]. Arguments for this position include the observations that there is little haplogroup overlap between East and West Eurasia (a smoother frequency gradient might be expected if they had all diffused from Central Asia), and that there is no great divergence of HVRI sequences between Central Asia and its neighbours to the east and west - that most haplotypes found in Central Asia are also found in Europe and East Asia - suggesting a more recent arrival [115]. However, it may be argued that frequency differences arising by drift in sparse, separated populations would not be expected to create such smooth gradients. A study of 54,794 SNPs in East and Southeast Asian populations concludes that the main genetic input for East Asian populations comes from Southeast Asia, with only a very small proportion from Central and South Asia [133]. Conversely, Y chromosome data are consistent with Central Asia being a source of pan-Eurasian genetic variation, detecting several emigratory events [104].

For the data presented here, CCR5del32 is the least informative of general population history. It is estimated to be around 5000 years old (3150-7800 years, 95% CI), has been found in Bronze Age skeletal remains [135, 136], and has most likely been subject to strong positive selection pressures [137, 138], though see Bollback *et al.* and Novembre *et al.* for a dissenting opinion [139, 140]. These selective pressures may be more intense in the northern part of the continent [138] as is suggested by the recent expansion of the deletion allele to high frequencies in otherwise quite dissimilar populations (e.g. Belgians and Tatars). However, this pattern could also be due to drift, once the allele was established; as a frequency-dependent process, drift is stronger when frequencies are low.

The tendency observed in mitochondrial DNA for several Central Asian groups of ethnographically presumed high input from Turkic populations to have more East Asian haplogroups is mirrored in *MC1R*. For all *MC1R* SNPs, the F_{ST} values are consistently higher between more 'Iranian' groups (like the Pamirs, Tajiks, and Uzbeks) and Chinese (representing East Asia) than between more Turkic groups (such as the Kazakhs, Karakalpaks, and Kyrgyz) and the Chinese: 0.08-0.10 versus 0.01-0.04. In addition, for the Y chromosome data from the same populations, Uzbeks, Tajiks, Pamirs and Turkmens display a higher percentage of 'western' haplogroups and Kazakhs and Karakalpaks have more haplogroups typical of Mongolia and China [115]. The *MC1R* SNPs further reveal the notable property that, when the geographical structure from each of them is combined, they define geographic subdivisions of Eurasia quite distinctly: Northern Europe, Southern Europe, West Asia, South Asia, Central Asia, East Asia, and Southeast Asia.

Though based on few loci, the data seem reasonably congruent with known historical incursions of Turkic peoples into, and occasionally through, Central Asia, overlaying a more Caucasoid (West Asian) character of prehistoric peoples in the region. Conversely, the general picture of higher genetic diversity in Central Asia compared with the rest of Eurasia may best be interpreted as supporting a position for this region as a source of diversity for the east and west. A further demonstration that certain markers can be indicative of past population movements is seen in the uncharacteristically high frequency of R163Q in Finns and Saami. If the variant can be seen as a marker of

East Asian genetic influence, this eastern input is supported by a wide range of markers throughout the genome, from large SNP arrays as well as mitochondrial DNA, Y chromosome data, HLA haplotypes, and blood groups [141, 142, 143, 33, 144].

3.5.2 Linguistic groups

The language groups that are most prevalent in West and Central Asia are Iranian and Turkic, represented in our samples by languages of both the western and eastern branches of Iranian, and languages from four different branches of the Turkic family (see Figure 3.11 for the relationships between these different languages, and their distribution within West and Central Asia) [102]. The populations speaking these languages were analysed separately for mitochondrial haplogroups and MC1R variants to see if there was any substructure in our genetic data that indicated affinities within a language grouping that was in contrast or in addition to the geographical locations of the samples. A limited amount of such correlation was found.

R151C and R160W, being largely limited to Europe, showed no correlation with language in these mostly Central Asian populations. Partial Mantel tests controlling for geographic distance reveal close to statistically significant correlations between mitochondrial and V60L F_{ST}s and language affiliation (p=0.07), and significant residual correlations were found between language affiliation and V92M/R163Q F_{ST} values. The R163Q correlation in particular is highly significant (Mantel r=0.299, p=0.001 by permuation), where populations speaking Turkic generally have higher levels of R163Q than geographically proximate Iranian-speaking groups. A recent publication, covering some of the same population groups, finds genetic structuring that is explained by linguistic affiliation [145]. However, their study, based on fewer populations and language groups, and a smaller study area, than the current one, is essentially comparing the genetics of ethnic groupings, rather than that of hierarchical linguistic relationships as represented by the speakers of language family subgroups. Nevertheless, all these results intimate that, while geographic distance remains a prime explanatory factor for genetic variation distribution, a genetic signature of population history, as represented by inherited language, may remain.

3.5.3 Pigmentation evolution

Despite the demonstrated function of *MC1R* in integument pigmentation [83, 146, 147, 78], and that skin colour is likely to be under considerable selective pressures [148], tests for selection in the gene have not provided clear-cut answers [110, 111, 78, 149, 150, 151, 152, 153]. Intuitively, as its product is one of the most influential elements in melanin production and distribution [84], *MC1R* is a prime candidate for selection to act upon in producing lighter skin tones, as is the case for other genes that affect pigmentation [111, 154]. Several studies also show that Europeans and East Asians owe their comparably light skin to different variants of pigmentation genes [111, 155]. Some assessments that did not find evidence for positive selection used methods unsuited to the probable selection pressures, both positive and negative. For example, Harding *et al.*'s assessment depends on there having been no selection on *MC1R* in the course of evolution along the human and chimpanzee lineages [152]. Savage *et al.* do, however, find evidence for selection at the locus, at least in Europe [110].

Two main characteristics of *MC1R* variation highlighted in this chapter support the contention: The first is that of much higher levels of nucleotide variation in this very small gene (55 SNPs in 951bp) compared to a variety of other gene classes, which is a strong indication of selection [110]. The second is the geographical distribution of variation. In contrast to most loci, the majority of nucleotide diversity is found in Europe (44 SNPs), not Africa (11 SNPs). Furthermore, the European samples display a non-synonymous/synonymous substitution ratio of 4.5 (northern Europeans, 6), versus an African ratio of 0.37, which indicates adaptive selection in Europe and purifying selection in Africa. The derived alleles of the five *MC1R* SNPs assessed in this study are all associated with a lighter skin tone [84, 156], and V60L, V92M and R163Q exhibit higher than expected differentiation between population groups as measured by

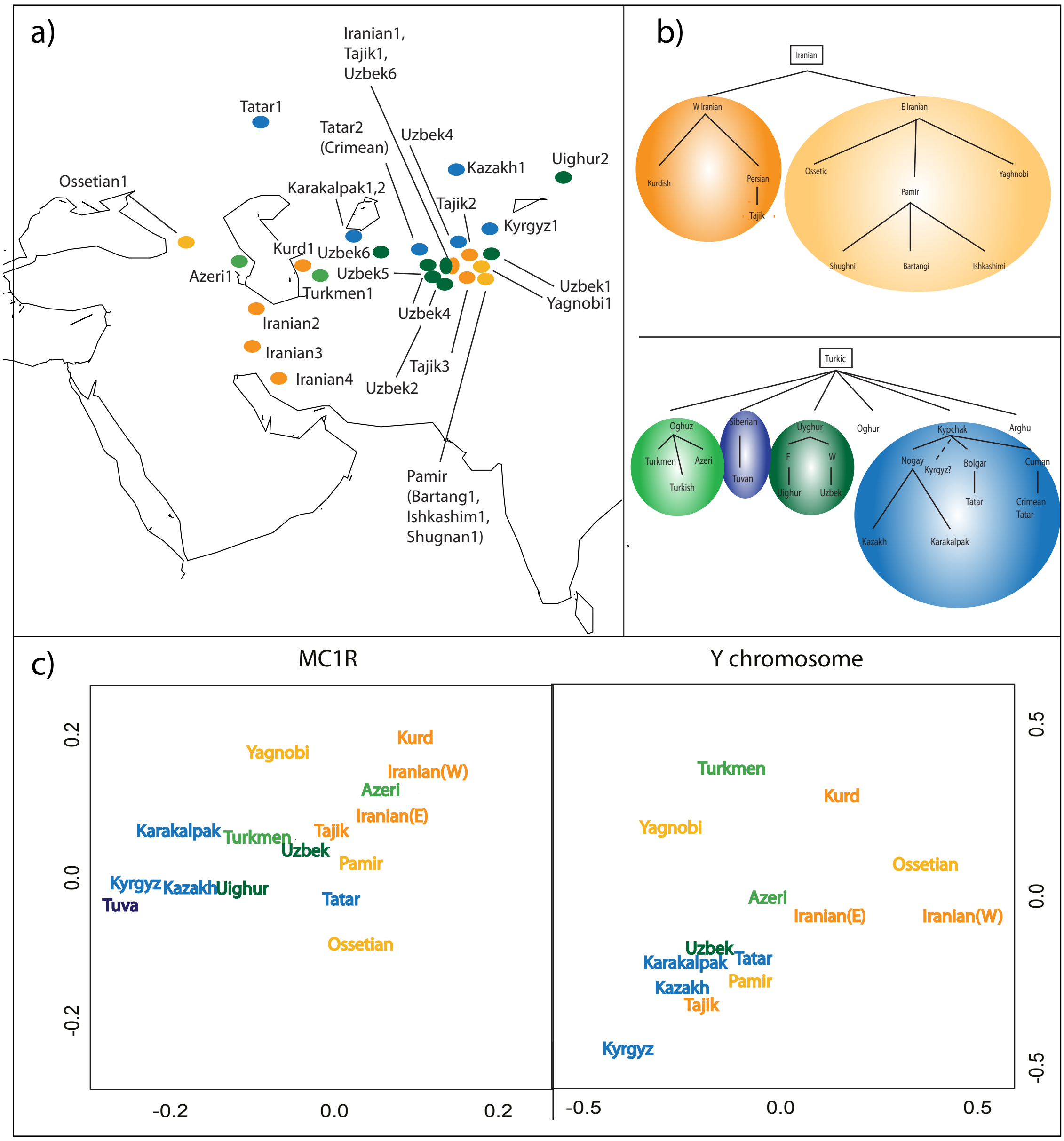


Figure 3.11: **Language families in West and Central Asia – their geographic distributions and linguistic relationships** a) The sampling locations of population samples speaking Iranian and Turkic languages, with each sample colour-coded according to language branch. b) Which languages correspond to the colour codes, and the linguistic relationship between the languages in question (upper panel for the Iranian language family, lower panel for the Turkic language family). c) PC plots of *MC1R* and Y chromosome (from Figure 3.8) with populations colour-coded.

F_{ST} (compared to subcontinental genomic averages from Norton *et al.* based on 11,078 autosomal SNPs [111]).

The distributions of the variant alleles show a high degree of population differentiation, and this type of genetic structuring has been predicted and shown to be a feature of a sizeable proportion of selected variants [157, 158, 159]. A further indication of the importance of *MC1R* variants as targets of selection is the discovery that Neanderthals carried a variant, different from any known in modern humans, which is predicted to have been functional and cause pale skin [160]. These results strongly support the positive selection of *MC1R* variants, rather than simply relaxation of purifying selection in human groups that left high-UV zones. The positive correlation between strong lighter skin variants with latitude and decreasing UV exposure is likewise a strong indication of this. The lack of correlation between latitude/UV exposure between the 'weaker' variants (V60L, V92M and R163Q) may be due decreased selective pressure – most east Asian/east-central Asian populations live at a latitude comparable to southern Europe, so the selective pressure might be weaker and a smaller depigmentation effect might be sufficient. The very high frequency in east Asian and Inuit populations also indicate that this variant may have come under selection and spread in the region early on in the latter's human occupation history. Additional evidence is also provided for the convergent evolution of lighter skin within the Eurasian continent, as different variants in a single, small gene, but with similar functions, appear to have been selected in western and eastern populations.

3.6 Conclusion

The high variability of mtDNA types, and congruent Y chromosome data, may be a source of human variation in both the east and the west of Eurasia. However, the role of subsequent migrations is difficult to evaluate. The lack of correlation between language family groups and genetics beyond that determined by geography emphasises the fact that genetic variation is usually stable compared to cultural variation, and that the association between the two is ephemeral. The *CCR5del32* distribution indicates its origin, but has no obvious association with any known epidemic – selection with respect to viral pathogens remains a distinct possibility. The *MC1R* variant data provide another example of convergent selection for pale skin in the east and west. Large scale SNP data, to cover the central part of Eurasia rather than focusing on the eastern and western peripheries, are needed to help clarify these issues and put our data into a wider context.

[The preceding chapter has been submitted for publication.]

Chapter 4

Archaeology, Linguistics and Genetics of Britain in northwest Europe – A background

4.1 Archaeology of Britain and Western Europe

This chapter is intended to give a very brief introduction to material culture trends in western Europe of the past fifteen thousand years, focusing on themes of demographic continuity and mobility. It is structured around the Three Age System (see Figure 4.1), as a familiar chronological framework only. This format should not be seen as adherence to a cultural evolution paradigm, or to the idea that these periods were necessarily separated by pervasive changes in the material culture overall.

I intend to present broad cultural trajectories from British archaeology in its European context, from the beginning of the current period of unbroken human occupation to the early modern era. I will specifically focus on periods of episodes of demonstrated or perceived discontinuity across different categories of archaeological remains, such as are presumed to be indicative of substantial external influence, which may or may not be accompanied by an equally substantial genetic impact.

Britain has been continuously occupied since the Palaeolithic, after the last glacial maximum. At this stage Britain was not an island, but an extension of northern Europe, as it remained for half of the Mesolithic until rising sea levels severed it from the mainland around 6000 BC [161, 162]. Until this severance occurred it is questionable how appropriate it is to separately consider Britain [163], and likewise the British population, which would have been continuous with that of the area that is currently the North Sea.

It is often presupposed, probably justifiably, that as islands, Britain and Ireland have been less susceptible to frequent and large-scale immigration from either surrounding parts of the continent or each other. Thus, the assumption is that Britain experienced only a limited number of migratory events of any size. Following standard archaeological/anthropological theory (see Chapter 1.2.4), a demic introgression is invoked when sufficiently radical discontinuities in the archaeological record are observed, however, the threshold for sufficiency is not easily or uncontroversially defined [51, 164, 37]. The archaeological record of Britain was once interpreted as evidence of successive waves of invaders from the Continent, eliminating or displacing the existing population. When this simplistic paradigm was replaced, the notion of 'peoples' largely disappeared as well, leaving the archaeology more open to different interpretations of the involvement of demic movements in cultural change and stability. In brief, a traditional view of immigration into Britain as evidenced by changes in material culture will be outlined below. The validity of this view will subsequently be assessed.

Neolithic - Starting around 4000 BC, this period sees the introduction not only of the eponymous new stoneworking technologies, but of a new economy with roots in the Near East, pottery, and increased social stratification evinced by the development of monumental architecture [165, 166]. The advent of farming is, in retrospect at least, regarded as causative of a watershed in the human history of Britain, both culturally and demographically.

Bronze Age - The chief innovations seen in this period (from the late third millennium BC) were that of widespread metalworking, the increased prevalence of single (rather than communal) burials, and the extensive distribution of a distinctive ceramic style in much of western Europe, including Britain: the Atlantic Bell Beaker. The wideranging Beaker network would have been dependent on efficient channels of communication, but the demic impact is frequently thought to have been limited to mate exchange and migration among the social élites.

Iron Age - The Iron Age (from ca. 800 BC) was once the preferred era for the arrival of Celtic tribes in Britain, but the evidence is now interpreted as being mainly a spreading of artistic styles, and much less as any fundamental xenochthonous change of population or material culture as a whole.

Roman era - As part of the Roman Empire, from 43 AD, the southern and eastern parts of Britain were quite heavily Romanized, as luxury items and social standing were increasingly identified with Rome. There is even evidence that Vulgar Latin became a, or even the, vernacular in certain areas [167, 168]. In prehistory, the archaeological record of this period would almost certainly have been interpreted as indicative of a sizeable and permanent population influx - it is likely that very few 'Romans' settled in Britain [169]. Those that did would have been drawn from all corners of a far-flung empire, but most particularly from adjacent regions of the European continent.

Anglo-Saxon England - The *Adventus Saxonum*, commencing in the mid-fifth century AD, is known from both archaeology and written records. From the former, there is no doubt that the Germanic invasion of Britain was significant (changes in burial practice, jewelry, weaponry, pottery and household goods), from the latter it is obvious that the impact on the indigenous population was felt [170, 171, 172]. The fact that the immigrant language rapidly became widely and, in parts, exclusively spoken is testament either to the number of settlers, the power they wielded, or both.

Viking Age - Viking incursions, starting in the late eighth century AD, are likewise attested in the archaeological record, contemporary writings, and the linguistic impact of Norse speakers on toponyms, anthroponyms, and Old English dialects. The ratio between permanent settlers, and more or less transitory 'raiders', is uncertain, and the separate evidentiary categories frequently tell contradictory tales.

Norman Conquest - This event of 1066 AD can be seen as a classic example of

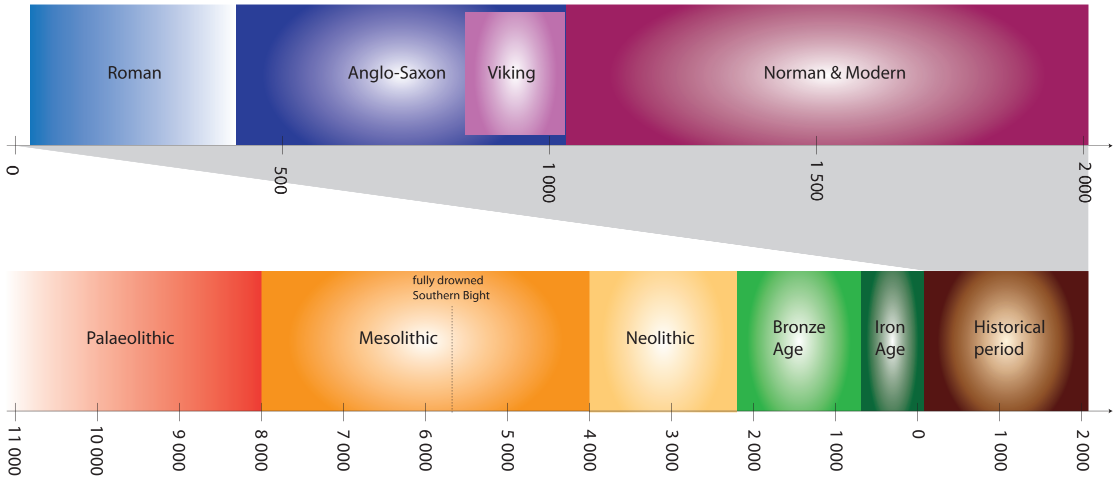


Figure 4.1: **Timeline of human occupation in Britain** Years indicated are BC/AD, the upper section represents the shorter eras of the historical period.

élite dominance - the Norman French seized most of the political power in Britain, and heavily influenced the art, architecture and language of the island. However, it must be noted that this was primarily among the higher, and therefore smaller, echelons of society. Actual settlement in Britain was minimal, with most Normans playing the role of absentee landlords.

4.1.1 Europe

4.1.1.1 Late Glacial and Mesolithic

This discussion of the European Palaeolothic will start with the Late Glacial period, and the cultures expanding from refugia in southern Europe at the end of the Last Glacial Maximum (LGM). Palaeolithic cultures are defined almost wholly by the most durable fraction of material culture, lithics, as well as bone and antler artefacts, and it can be difficult to determine if differences in cultural asssemblages are due to them being created by different people at a different time, or by the same people for a different purpose (e.g. seasonal camps requiring distinct toolkits for different processes) [173].

During the LGM (24,500-17,000 BC [174]), the populations of northern Europe contracted southwards into so-called glacial refugia; in western Europe, primarily the

Franco-Cantabrian refuge in northern Spain/Aquitaine. The main eastern refuge is thought to have been in the Ukraine and the Balkans, and there was likely also a minor refugium in the Italian peninsula [175]. As the ice receded, but probably before a much warmer climate was established [176], the refugial populations expanded northwards from the Franco-Cantabrian refuge, bearing the Magdalenian cultural complex. The Magdalenian characterized much of late Palaeolithic western Europe, in regional but nonetheless related forms.

Palaeolithic northwest European cultures are usually termed 'Magdalenian' in France, 'Hamburgian' in northern parts of Germany/southern Denmark and the Low Countries, and Creswellian in Britain. The exact relationship of these cultures is not a matter of consensus. The Hamburgian is variously described as being a development completely independent of the Magdalenian, and having its origins in the Gravettian of eastern Europe, or as an early northern offshoot of the Magdalenian, or as the result of a fusion between the two above cultural complexes [177, 178, 179]. The 'Creswellian' has been presented as anything from part of a 'Creswello-Hamburgian' entity, dissociated from the Magdalenian and anything else, to a local but identifiable variant of the Magdalenian, or simply 'Final Magdalenian' [177, 180, 181, 182]. With finds of diagnostic 'Hamburgian', 'Magdalenian' or 'Cheddar' points in sites that are classified as belonging to one of the others, the weight of opinion seems to favour the idea that all of the above form part of the Final Magdalenian continuum.

This continuity is not noticeably interrupted in the Final Upper Palaeolithic/Early Mesolithic. Federmesser and Ahrensburgian type lithics are found throughout northwest Europe, though assemblages are not completely homogenous [183, 184, 185, 186]. The 'cultural' shift to the Mesolithic is characterized by an increasing use of microliths.[1] The continued similarities, and similarity of developmental trajectories, of northwest European Early Mesolithic communities is remarkable, and implies comparatively extensive contacts throughout the area [161, 187, 186, 188]. Local variation certainly existed and continued to develop, e.g. differences between northwest and northeast France,

---

[1] Small blades/bladelets, frequently employed in composite-blade tools.

different areas of Britain and even in the comparatively small area of the Netherlands and Belgium [187, 189, 186], but “the similarities across [the North European Plain, including Britain] are as impressive as any differences”, p.358 [188].

The demographic implications of this narrative are obvious. Northwest Europe was repopulated, probably nearly exclusively, by people originating in the Franco-Cantabrian refuge, expanding northwards to reach southern Scandinavia by around 10,000 BC (see Figure 4.2) [190]. Estimations of actual population sizes are not without severe methodological problems, but a high end estimate sees the Glacial Maximum refuge population as approaching 9,000 and more than quadrupling with the Magdalenian expansion [191].

4.1.1.2 The Mesolithic-Neolithic Transition

The ‘Neolithisation’ of Europe was a distinctly heterogeneous process. The constituents of the so-called Neolithic package – pottery and plant and animal domesticates – spread, not necessarily concomitantly, to Scandinavia and the British Isles by around 4000 BC [192]. Animal and plant domesticates, and the southeastern pottery tradition, were first introduced to southeast Europe. It is difficult to identify a single source for these first European Neolithic cultures in Greece/the Balkans, but their immediate cultural predecessors were likely situated in Asia Minor [193].

Two main routes trace the westwards expansion of the Neolithic to reach the Atlantic/North Sea Europe; one along the north coast of the Mediterranean, and another up the Danube and associated river valleys and across north-central Europe [192, 194, 195]. The archaeological record appears to show a ‘leapfrogging’ effect in Neolithic site distribution. The sites tend to be geographically disjacent, occupying a restricted set of topographical and soil niches in the northern movement, and usually areas lacking in Mesolithic acitivity in the Mediterranean [196].

The northern arm of the early Neolithic, Linearbandkeramik (LBK), is named after its distinctive fineware, three-quarter spherical pottery with linear decoration, and initially covered north-central Europe, then extending from northern France/Belgum

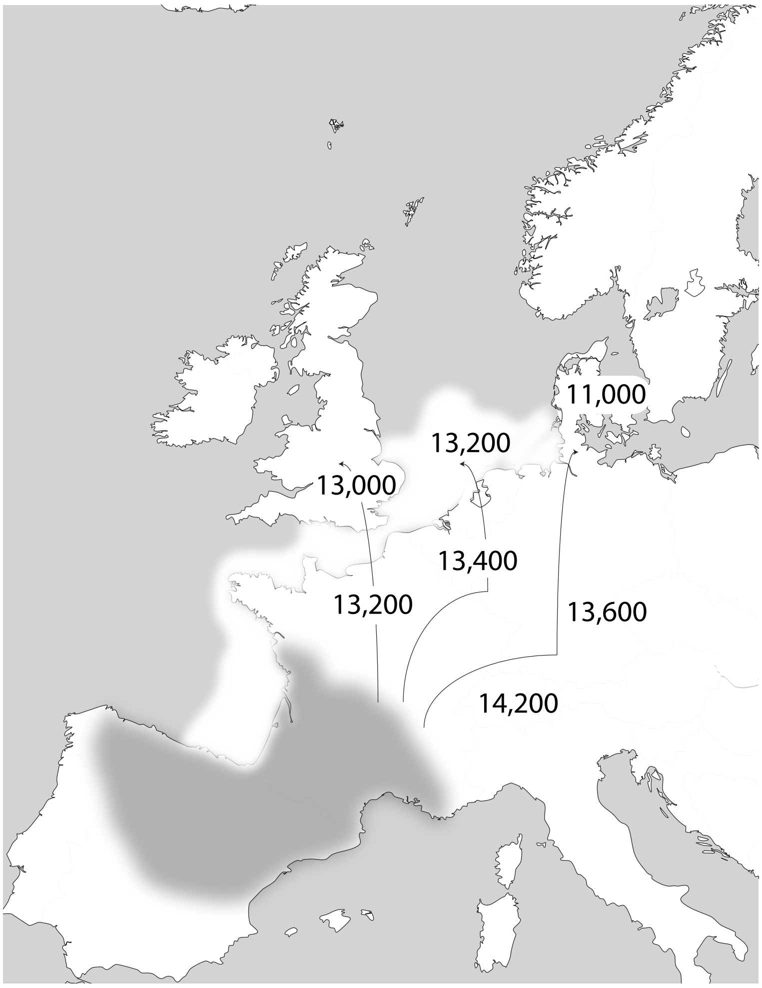


Figure 4.2: **Late Glacial repeopling of western Europe** The grey shaded area is the Franco-Cantabrian refugium, and the dates (BC) indicate the time at which a given area was repopulated. Overlaying modern coastlines in white shading are approximate Late Glacial coastlines of western continental Europe.

nearly to the Black Sea (see Figure 4.3) [194]. In addition to the pottery, characteristic features of LBK are sturdy, substantial longhouses, cereal crops (einkorn and emmer wheat, barley), and domesticated animals (primarily cattle) [194]. LBK emerged in western Hungary/southwest Slovakia, and quickly spread over about two hundred years to occupy most of northern continental Europe by 5500-5300 BC. Its antecedents were Neolithic cultures of the Balkans, which show evidence of external settlement - ultimately from the Levant - followed by extensive, bi-directional cultural assimilation with indigenous cultures in the area of the Iron Gates [193]. The extremely fast expansion of LBK took place through serially settling loess soils in terraced river valleys. The spread has usually been seen interpreted as a classical case of population movement being the most important vector of cultural change, due to its rapidity (averaging about 5-6 kilometers per year for the original expansion, though this encompasses the entire area in which LBK was found and does not take into account its patchy distribution), as well as the uniformity of its cultutral expression. However, this uniformity of settlement type and agro-pastoralism masks variability in other aspects, such as the continuity of local Mesolithic stoneworking traditions in many, though not all, sites in the core LBK area [195, 194]. Furthermore, Robb and Miracle [197] make a case for primarily acculturation of Mesolithic populations for almost the same reasons, citing the fact that population growth in the nascent farming communities would have to be approaching the the bounds of what is theoretically possible in order to sustain the expansion, and there is little or no archaeological evidence to support that such an explosive growth took place. In the western extremity of the LBK culture area, southwestern Germany into northern France and the Low Countries, there were other, largely contemporary Neolithic facies of a different provenance: Limburg and La Hoguette pottery and their associated characteristics appear to stem fron a northern projection of the Cardial culture and local Mesolithic traditions [198, 199], though the opinion that they are the product of LBK influence on local western European foragers is also held [194].

After a period of stasis, a secondary LBK expansion took place, taking the cultural complex to its fullest extent. Subsequently, the ‘Neolithisation’ of northern Europe

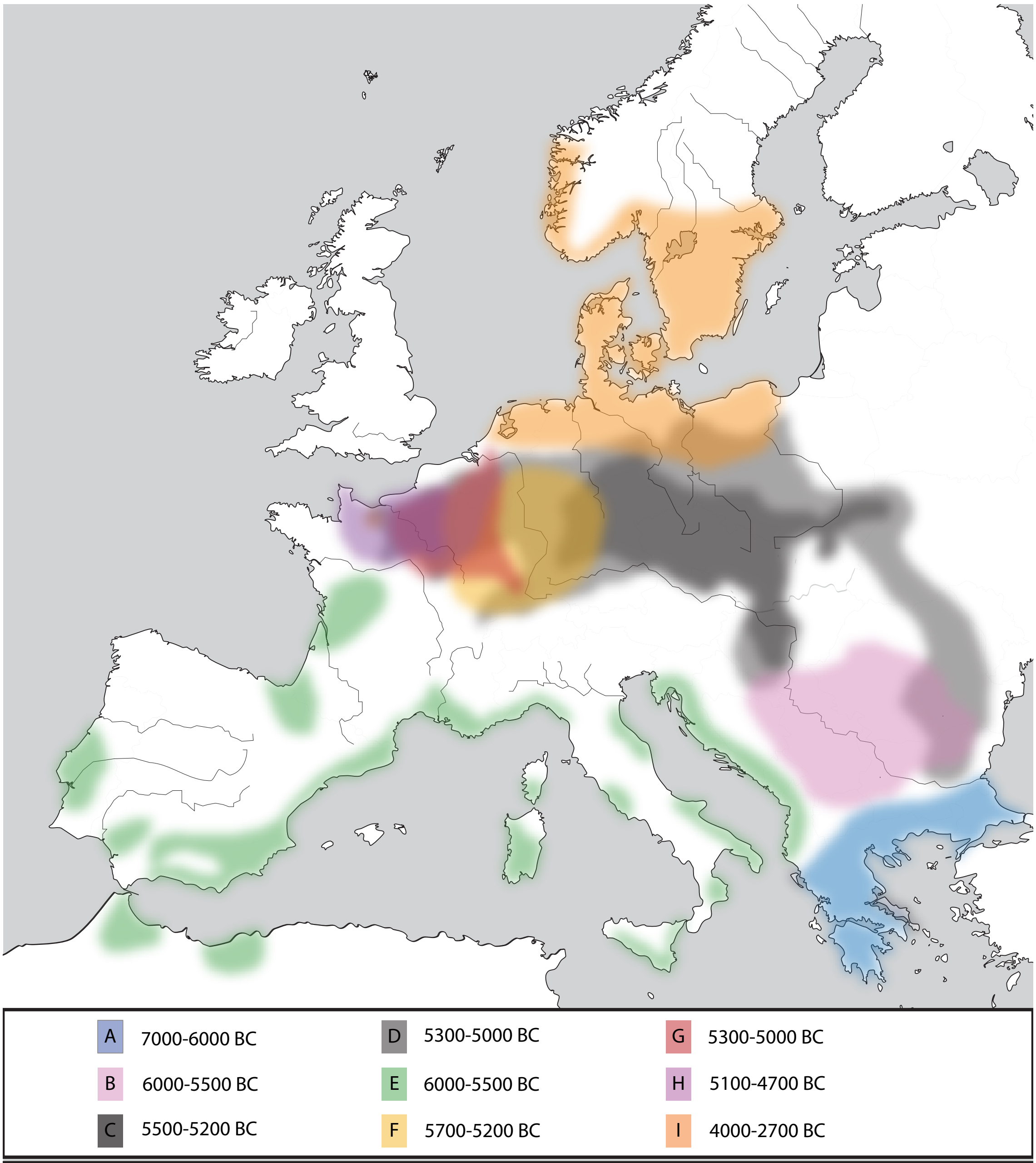


Figure 4.3: **Early Neolithic cultures of Continental and Scandinavian Europe** A - First European Neolithic, B - Balkan Neolithic, C - LBK core area, D - Subsequent LBK expansion, E - Cardial/Impressed Ware, F - La Hoguette pottery distribution, G - Limburg pottery distribution, H - Villeneuve-Saint-Germain, I - Funnel Beaker/Scandinavian Neolithic. For Neolithic Britain, see Chapter 4.1.2.3

stalled at this frontier for nearly a thousand years [194]. There have been many hypotheses advanced regarding why the LBK movement came to such an abrupt halt after its monumentally successful periods of expansion. Most of these cite the likelihood of coastal Europe having a higher Mesolithic population density, due to the rich natural resources provided by a marine environment compared to the wooded inland [192]. This would have had a dual effect on Neolithisation: it would have made it less easy for any colonizing farmers to lay claim to such an area without damaging disputes regarding land access, and the sufficiency of the Mesolithic system of food procurement would have meant that there was no pressing economic impetus to shift to agriculture. The idea that agriculture and pastoralism provided a vastly more productive means of existence than hunting and gathering (with indications of resource management [200]) is still often implicitly invoked, though it has been shown that primitive agriculture - particularly with crops unsuited to alien climates and soil types - likely required more effort in order to sustain a livelihood and led to a general decrease in health and robustness among at least some Neolithic populations [201, 202, 203, 204, 205].

The introduction of the Neolithic in the Low Countries (or Lower Rhine Basin), in contrast to many other areas of Europe, was a long drawn and piecemeal process wherein the Swifterbant cultural group gradually incorporated pottery (ca. 5000 BC), Neolithic flintworking (ca. 4800 BC), livestock (ca. 4500 BC) and finally cereals (ca. 4100 BC) into their repertoire from neighbouring LBK communities [206].

The late Mesolithic Ertebølle culture of northern Germany and Denmark adopted pottery, possibly from pottery-producing Mesolithic peoples of Russia, rather than LBK [195], and also seem to have had some familiarity with domesticated animals, though this was likely a question of acquiring exotic items rather than an actual engagement in husbandry [207]. The Funnel Beaker culture, the first fully Neolithic culture of northern Germany and south Scandinavia with agriculture and animal breeding, was introduced around 4000 BC. The full extent of the Funnel Beaker complex lies from Holland through northern Germany to northern Poland, and from Denmark throughout southern Sweden and along the Norwegian coast (see Figure 4.3). The antecedents of the

Funnel Beaker culture are not entirely clear - it appears to have little in common with the LBK descendants directly to the south, beyond an increased reliance on a farming economy - and some identify the Lengyel-type LBK derived Sarnowo assemblages of central Poland as a more likely candidate, in addition to an indigenous Ertebølle element [192]. Whatever its origins, it spread quickly from its inception, reaching the west coast of Norway at about 3650 BC, from its core region of northern Germany/southern Denmark.

The southern, Mediterranean arm of the Neolithic pincer movement into western Europe is collectively known as the Cardial/Impressed Ware culture, again after its typical pottery which is decorated with patterns pressed into the clay, in many cases using *Cardium* shells as a tool. The origin of this movement may lie directly in the Levant - any south-eastern European involvement in the evolution of Impressa pottery is tenuous [208]. It is characterized by wheat, barley and a predominance of ovicaprids in addition to the pottery. The cultural complex first appears along the southwest Adriatic coast, subsequently extending northwards to about the level of Trieste, where it stops, possibly due to an increased density of Mesolithic peoples in that area [209]. Cardial/Impressed ware is found westwards along the north coast of the Mediterranean, in disjacent regions, on islands such as Sicily, Sardinia and Corsica, in most of coastal Iberia and parts of France (see Figure 4.3). While Impressed Ware and Cardial Ware are frequently combined, there is evidence for a certain amount of geographical and spatial differences between them. Generally, Impressa is earlier and more eastern than Cardial. There are also inidications of repeat Impressed/Cardial influences flowing from Italy into France [210]. The overall mode of Cardial/Impressed spread is, like LBK, very rapid and the components of the cultural package largely concomitant. Again like the LBK phenomenon, the speed and coherence is most frequently interpreted as evidence for the cultural spread being mediated largely by migrating Neolithic peoples, with a minimal impact from Mesolithic inhabitants. In many areas of Neolithic settlement in this region, the evidence for Mesolithic activity is sparse or absent, strengthening the colonization hypothesis, but many coastal sites are thought to be submerged, thereby skewing the overall impression [196, 210]. The colonization scenario is more probable with respect to the Cardial of the Mediterranean than that of Portugal and western France, where the Mesolithic remains overlapping chronologically with the incipient Neolithic are much more apparent [211].

The extent to which the spread of the Neolithic in Europe was due to population movements for the transmission of ideas continues to be a much debated issue, but the transition is almost uniformly seen as a watershed in the history of the continent, both in terms of culture and biological ancestry. This assumption owes much to the modern, post-Colonial-influenced idea that the introduction of crop plants and domesticated animals was a revolution that rapidly and radically transformed people's way of life and gave cultivators and pastoralists a significant advantage over Mesolithic hunter-gatherers [212].

There is no denying that the mainstays of the new subsistence economy originated in the Levant: Most of the domesticates that characterize Neolithic economies, such as wheat, barley, sheep and goats are descended from wild species alien to Europe, but native to the Fertile Crescent [213, 192]. Wild pigs and the now extinct aurochs, the ancestor of modern cattle, were native to Europe as well. However, there is no evidence of true domestication of wild species *in situ* in Europe prior to the Neolithic, though during the Mesolithic pigs were probably managed in some way in certain areas, and modern European pigs are descended from wild European ancestors, and not the Near Eastern stock that was imported in the Neolithic [214]. Aurochs may also have contributed to the gene pool of modern European cattle. The evidence for this is slim, but solid. Mitochondrial DNA, which is the basis for these comparisons, indicates that the vast majority of extant maternal cattle lineages originate in the Near East, but that a few are closer to DNA extracted from European aurochs bones [215, 216]. It is possible that a paternal contribution might be greater: Intuitively, on the basis that hybridisation would simply require instances of male aurochs copulating with a domestic cow, while female aurochs introgression would involve capture and taming of a wild individual, but the evidence for this is inconsistent [217, 218]. European

Neolithic pottery traditions can also be traced back to the Near East, on stylistic grounds [208]. It follows from the above, that for for these technologies, pottery, reaping and sowing in particular, to spread by acculturation alone, a fairly intesive amount of contact must have occurred between pottery-producing farmers and aceramic hunter-gatherers. European Neolithic pottery traditions not only include pots that *look* the same superficially, but have been shown to have been manufactured in the same fashion [196], and growing cereals even on a small scale requires at least an intensification of skills the Mesolithic peoples probably already possessed in preparing soil and assessing growing seasons and ripeness.

Neither a pure acculturation nor a pure colonization model is likely to reflect the historical situation, nor is there any pressing reason to believe, given the different trajectories described above, that the same kind of process was involved in the Neolithic transition throughout all of northern and western Europe. While few archaeologists currently adhere to any such simplistic models, there is a clear divide between those who stress demographic continuity with Mesolithic cultures and those more liable to underscore the role of immigrant farmers [195]. The latter, following the logic if not the relatively totalitarian approach of culture-historical archaeologists, note the comparative cohesion of Neolithic cultural trends like Cardial/Impressa and Linearbandkeramik, and the speed with which they spread. This, it is argued, would necessitate the transmission via fully Neolithic people participating in the demographic increase that is supposed to have followed the introduction of farming. The other camp puts more emphasis on the lack of evidence for such a population growth in the earliest farming communities, the continuing Mesolithic traditions in many areas, and the sudden halt in especially the explosive Linearbandkeramik expansion for no discernable reason [219].

4.1.1.3 Northwest Europe in the Neolithic

The classical Linearbandkeramik culture lasted until around 5000 BC when more distinct regional successors began to appear (see Figure 4.4 for a schematic overview of regional post-LBK cultural trajectories). Most of these are initially very similar to

LBK, and differences are on the level of subtly changing longhouse shapes, pottery decoration and burial traditions [219]. By the mid-fifth millennium BC, significant changes set in in the northwest fringe of continental Europe. The importance of the longhouse had declined [219], and they were no longer found in the Cerny culture of the Paris Basin and associated groups [220], nor are they present in the following Michelsberg-Chasséen continuum, stretching from Rhineland/Belgium to at least the borders of Brittany through the Paris Basin and Normandy [221]. Brittany, largely outside the realm of LBK influence, had no archaeologically visible traces of domestic dwellings prior to this date either [222]. Largely concurrently, one of the most spectacular western European Neolithic innovations was developing: monumental architecture, in the form of tombs. The earliest examples of these come from western France [223], and take two major forms, passage tombs and long mounds. Long mounds tend to appear in areas of marginal LBK influence, and are usually seen as a re-interpretation of the longhouse tradition in places were such structures were not, or were no longer, primary dwellings [224, 225, 165, 226, 163]. Passage tombs - megatlithic structures - are found mostly in regions beyond the LBK sphere, where its influence was minimal at best, and where Mesolithic continuity was stronger. In northwest Europe, Brittany is the primary region where passage tombs developed, as part of an Atlantic process decoupled from post-Danubian tradition [226, 165, 227].

The developmental trajectories of fifth millennium cultures of the region from Brittany to the Low Countries is particularly relevant with respect to the later introduction of the Neolithic into the British Isles, as they provide the most probable departure points from the continent, geographically as well as typlogically (see below). For a large part of this region, the LBK phenomenon is the most evident cultural (sub)stratum, with the LBK-derived Villeneuve-Saint-Germain (VSG) and Blicquy in the Paris Basin and Belgium, respectively. Subsequently, the Cerny culture of northern France, whose sphere of influence reached Armorica, introduced elements of southern European Neolithic and/or local Mesolithic traditions [220, 228, 229].

The Michelsberg complex is usually viewed as a clear descendant of the Rössen culture of the Rhineland (another post-LBK tradition) [230], while some French archaeologists contend that it is equally influenced by cultures of the Paris Basin [231], or even that the latter are primary determinants of it [232]. The Michelsberg heartland lay in the Belgium/Rhineland zone, but also impacts and coexists with the post-Cerny Chasséen of the northern Paris Basin in particular, and provides the final 'Neolithising' influence on the long-lasting Swifterbant complex of the Netherlands [233]. The Chasséen designation is one given to Middle Neolthic cultures of virtually all of France, and this term has with varying degrees of severity been castigated as being anything from "broadly defined" [234], to imperialist and falsely unifying [235]. As such, it is subject to several subdivisions, of which the North Chasséen is relevant to the current discussion. North Chasséen covers the Paris Basin and Normandy, along with Michelsberg, but in Brittany its presence is disputed, and if present, is not strong unless it be considered 'Chasséen-Atlantique' [234, 235, 236, 237, 229].

This brief sketch of the coastal strip of the northwest corner of continental Europe brings us to the late 5th millennium BC, the period when elements of the Neolithic package start making their way to the British Isles.

4.1.1.4 The Bronze Age in western Europe

The most widespread material cultural phenomenon of the western European Late Neolithic/Early Bronze Age is the Beaker complex. The loosely linked assemblage types belonging to it are scattered throughout most of the peninsula of Europe and its associated islands, as well as northern Morocco (see Figure 4.5). Apart from the eponymous beakers, which came in many regional forms, the 'Beaker Culture' was unsurprisingly heterogeneous across this vast area.

Several partially contradictory views exist regarding the origin(s) of the phenomenon. Traditionally, beakers are seen as a development from late Corded Ware.[2] Similarities

---

[2] The Corded Ware culture is a late Neolithic culture of northern and northeastern Europe. The Single Grave culture, part of the Corded Ware culture, occurs in the north and west of the preceding TRB cultural distribution. Funerary characteristics of this group include single burials, often under barrows, accompanied by a Corded Ware vessel and a stone 'battle axe'.

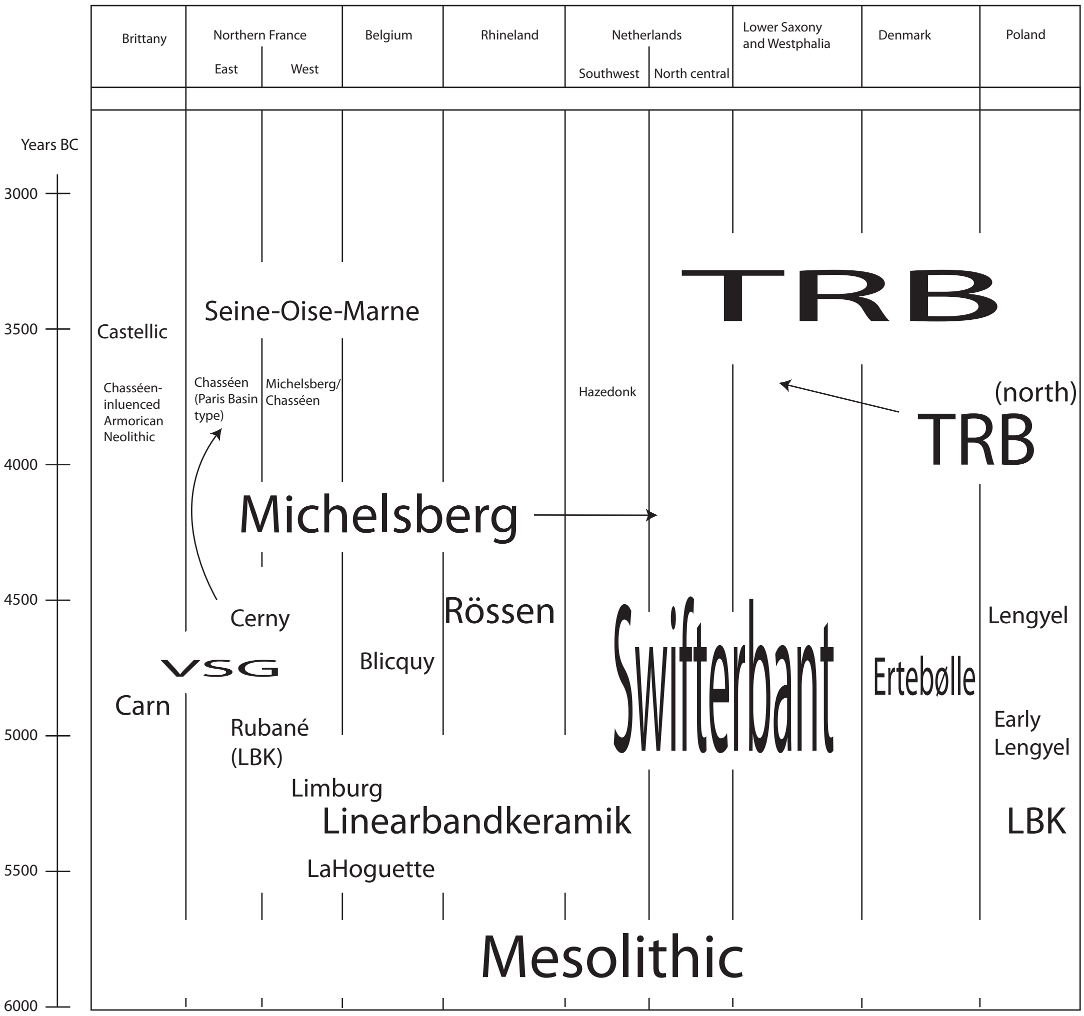


Figure 4.4: **Northwest and north central European Neolithic cultures** LBK - Linearbandkeramik, TRB - Funnel Beaker culture, VSG - Villeneuve-Saint-Germain; following Louwe Kooijmans, Bogucki and Whittle [206, 194, 219].

exist between the late Corded Ware vessels of the Lower Rhine delta and beakers [238]. However, the earliest known beakers have been found in Portugal, and stylistically they could also be a local development on the Tagus estuary, perhaps influenced by northern traditions [239, 240]. To fully account for the observed stylistic and chronological distributions, axes of reciprocal influence have been invoked, and a fusion corridor in northern France/Lower Rhine for Atlantic bell beaker influences to merge with those of the Corded Ware area proposed [241]. These interactions may have been facilitated by the extant Late Neolithic network of distribution for Grand Pressigny flint which, mined in west-central France, is found throughout much of western and central continental Europe [163].

In addition to the beaker pottery itself, stereotypical elements of a Bell Beaker assemblage include copper (earlier flint) daggers and stone archer's wristguards in a funerary setting [242, 243]. These traits, however, are not universal, and the wider contexts of Beaker pottery differ considerably [244, 242]. The burial types are not identical. In Brittany, Atlantic France, Ireland, and northern and western Scotland, beakers are often found in re-used megaliths, while inclusion in single burials is the norm in Britain, central and northern Europe [163, 245, 240].

The coeval introduction of copper metallurgy and Bell Beakers in northwestern Europe has contributed to the interpretation of the spread of Beaker culture being mediated by a caste of élite smiths, a process which is not compatible with the pattern seen in the rest of Europe [246, 242]. The comparative homogeneity that allowed a Beaker culture to be identified must be viewed in conjunction with the local continuity observed in most spheres that form the cultural stratum, which Beaker funerary customs overlaid in a geographically discontinuous network. The gaps between Beaker areas are variously explained as being due to geographic barriers in central Europe, a genuine lacuna for most of France, and poor surveying in the British Isles. In northern and eastern Europe and the Danube basin the distinction is made on the basis of Beaker pottery alone, the removal of which would see any 'Beaker society' recede seamlessly into their neighbours' cultural groups [242].)

Figure 4.5: **Beaker distribution** Areas of Europe and north Africa where Beakers have been found. Inset is an example of a type of Beaker.

Explanatory paradigms for the Beaker phenomenon range from the earliest migrationist and élite dominance models, to religious movements and marriage exchanges [242, 246, 243, 247], but the vehicle(s) for the spread is still unproven [248]. What is certain is that some kind of network, which was not simply mercantile, must have operated in the late Neolithic/Chalcolithic/earliest Bronze Age transitional period.

In areas outside of the Beaker sphere, continuity with the preceding Neolithic appears even stronger, the eponymous metal being transformative for only a small fraction of society [249]. In the Early Bronze Age, settlement sites of northwest Europe are rare, and the buildings seem to be largely ephemeral. Structures, where they have been found, were usually round or subrectangular, but in the northern areas, northern Netherlands and southern Scandinavia, rectangular longhouses occur [250, 249]. Settlement sites in the Paris Basin are virtually non-existent [251], but in south and central Netherlands, northern Belgium, Lower Normandy and Brittany roundhouses are found [252, 249, 253, 254].

Despite the new metal technology, stone tools were still very much the norm, and traditions from the Neolithic carried on into the Bronze Age, mostly using regional sources [250]. In contrast to the preceding period, status burials changed from communal deposition – often fragmentary – to individual inhumation. While flat graves were standard in the Unetice culture of central Europe, north and west Europe, especially coastal areas, saw the growing popularity of burial mounds or tumuli (see Figure 4.6 [250]).[3] The character of these burial mounds varied, with regional and temporal traditions dictating detail such as re-use of the mound for further burials, surrounding ditches or timber posts, grave goods, and construction media (turf, earth, stone, or any combination of the three) [249, 252, 251, 255, 253, 250]. Burial mounds are also found more rarely in southwest Europe, and unrelated burial mound traditions exist in eastern Europe and west Asia [250, 249]. Mounds could contain cists, as in northern Britain and Armorica (see below, Chapter 4.1.2.5), or coffins, as in lowland Britain, the Netherlands and the fabulously preserved ones in Denmark [250, 243, 249]. Cist buri-

---

[3] Often called barrows in England and anglophone literature.

als without covering mounds are also common in the same areas as they occur under mounds [249]. Grave goods ranged from virtually nothing, to grand displays of wealth, even within the same area and period [250]. Two main Early and Middle Bronze Age groups are the rich grave clusters of Wessex and Armorica (see below, Chapter 4.1.2.5), and the urn traditions of northern France, Belgium, the Netherlands, and the British Isles.

The metalwork of northwest Europe is particularly important as a proxy for establishing both societal regionality and interaction. Like pottery, metal is a comparatively plastic medium and allows for a great range of stylistic differences to be combined with the functional forms of an object. Early metalwork includes knives, axes (continuing their extra-functional status from the Neolithic), and some items of personal decoration. As copper and tin resources are not evenly distributed, developing and sustaining bronze-working traditions would for many areas have necessitated prolonged contact and exchange with mining centres of Atlantic and Alpine Europe [249]. This being the case, while regional artefact types developed, the fact that there is much evidence for interregional exchange of artefacts and styles is unsuprising. In the Early Bronze Age and increasingly in the Middle Bronze Age, both actual imports and stylistic influences are found to link the British Isles, western coastal and northern France, the Low Countries and occasionally Scandinavia and Central Europe [256, 250, 257, 249, 163]. Some of the contact may well have been sporadic, but in the case of the Nordic Bronze Age it must have been sustained. Scandinavia has quantities of distinctive metalwork despite its comparatively late adoption of metallurgy, and no known metal deposits that were exploited during the period [250]. The widespread finds of amber, often of Baltic origin, in other areas of Bronze Age Europe have been suggested to represent one of the commodities, many of which are archaeologically invisible, which might be exchanged for metal [250, 249].

Depending on the authority, and very likely the differing implicit yardsticks for 'continuity' that they employ, the later Middle and Late Bronze Age (mid 2nd millennium BC) was a period that saw great change or great continuity in Europe [250, 258]. It is

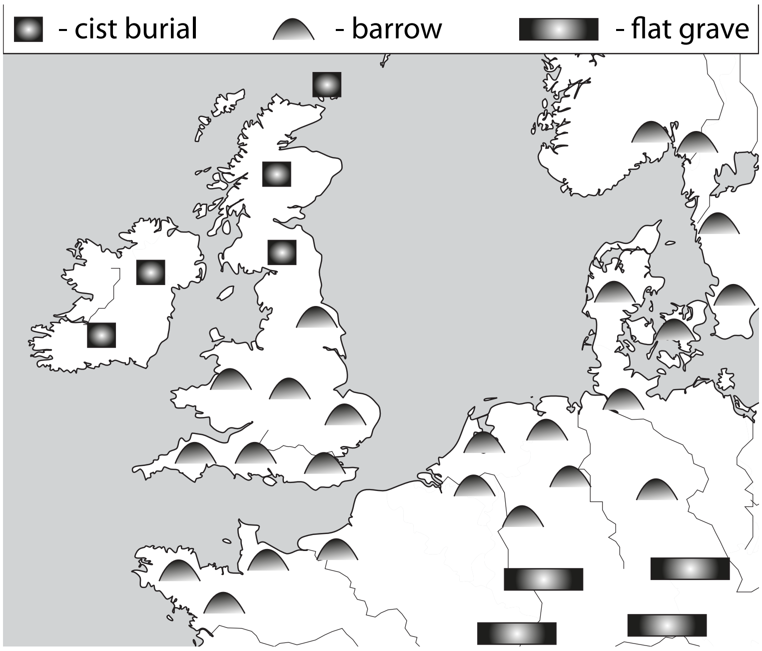


Figure 4.6: **Early Bronze Age grave types**

perhaps correct to say that, while an earlier to later Bronze Age transition is characterized by more pervasive changes than that of the Late Neolithic to Early Bronze Age, most of these changes are gradual and continue local trajectories rather than embodying sweeping external influences.

Several supra-regional culture areas are of relevance to western Europe in the later Bronze Age: 'continental' central Europe, the Nordic area and the coastal 'Atlantic' area (not including Scandinavia). Commonalities in these areas include indirect evidence of population increase (larger settlement sites, more and fuller cemeteries), increasing concern with weaponry and protective equipment, an increase in metal deposition and a decrease in the construction and use of large ritual sites [250, 249, 163]. In addition, most areas saw cremation as the primary mode of burial preparation, though the tradition emerged at different times and in different contexts [249].

Much of central Europe in the Late Bronze Age belonged to a cultural sphere known as Urnfield, though the unity of material expression was not limited to the burial practice of placing cremated remains within a ceramic urn in a pit. Regional variants of the basic paradigm spread to many areas including southern France, western Iberia and southern Belgium, though in the latter case it was incorporated into the local cinerary urn tradition, and most metalwork was Atlantic rather than Urnfield in style [259]. Continental urns were also adopted in the Nordic area (especially in the south), but here they were often buried in earlier barrows, and the metalwork continued to be distinctively autochthonous [260]. Some Urnfield influence, seen in artefact styles, is recognized in the Atlantic zone as well, but it is usually, if not exclusively, seen to be weak [250, 257, 260]. The Atlantic contact zone, developed throughout the earlier Bronze Age was intensified in the later Bronze Age, to the extent that there the term ‘Atlantic Bronze Age’ is occasionally employed, when the links between the northern part of the region expanded to include an emerging southern network [261]. The evidence for the Atlantic area is based mostly on typological affinities of metalwork, with shared sword types, hoard components and feasting equipment (cauldrons and fleshooks) [261, 262]. Occasional non-metallic testimonies to Atlantic communication are found, such as V-notched shields in Iberian scultpture and, in an Irish bog, a leather example.[4]

That there are very distinct subdivisions within the Atlantic zone is not in dispute, nor is the fact that most contact would have been local, and that the indications of direct long-distance communication is very slight [261, 263, 264, 262]. Instead, sustained chains of association along the length of the larger zone is thought to account for the shared artistic expresssions and social concerns evidenced by the artefact types and hoarding practices in liminal (most frequently wet) settings. In addition, shared value systems along the Atlantic fringe are evidenced by commonalities in ritual behaviour in these coastal regions [240].

The comparative stability and interconnectedness of the Bronze Age sees reordering in the 8th century BC, concomitant with, and perhaps partially caused by, the introduction of iron-working.

---

[4] The rivet skeuomorphs in this shield, however, indicate that these too come within the sphere of metallurgy.

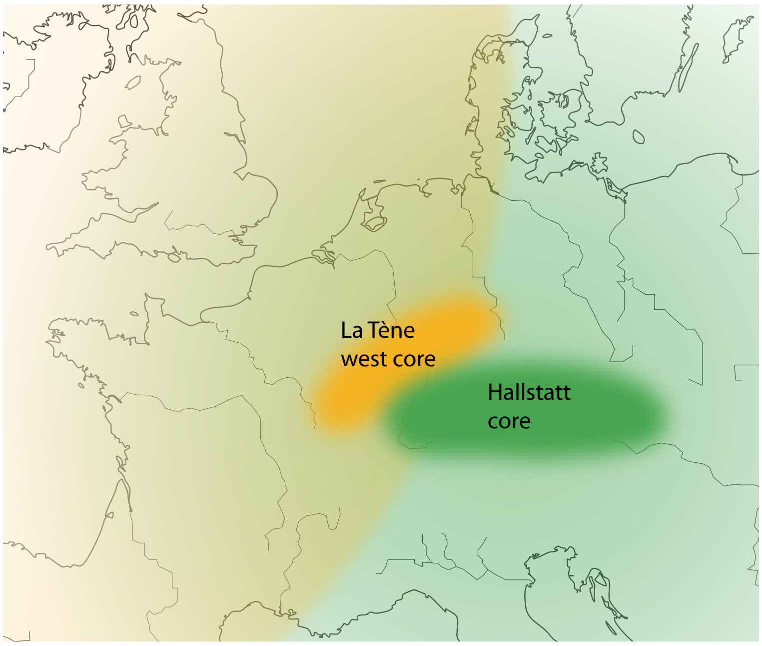


Figure 4.7: **Hallstatt and La Tène** The earlier Hallstatt core and radial influence, overlaid by those of the western La Tène phenomenon.

4.1.1.5 Iron Age in northwestern Europe and Ireland

4.1.1.5.1 Northwestern Continental Europe The great supraregional cultural trend of central and western Europe in the Early Iron Age was named after the Austrian type site of Hallstatt. Hallstatt culture grew out from and succeeded the Late Bronze Age Urnfield culture, and its western sphere of influence eventually covered modern France, northwest and central Iberia, and moved into the British Isles (see Figure 4.7) [265]. This influence, was, at least in the periphery, largely restricted to weapons, both bronze and early iron examples [266, 267].

Around the mid-fifth century BC, on the northern fringe of the core Hallstatt zone, the La Tène culture developed, which came to characterise the latter half of the Iron Age. The La Tène core zone was bipartite; focus here will be given to the western core, centred on the Marne and Moselle valleys. Frequently, La Tène and Celtic are used interchangeably, but the justification for this is not great (see Box 1). La Tène influence is seen, outside the core zones, chiefly in the decorative sphere, to the extent that La Tène is often used to refer to the artistic style alone. In the beginning, the artistic tradtion was heavily influenced by Mediterranean (Greek and Etruscan) items, and later eastern art originating in the Pontic steppes [265]. While the origin of La Tène style is obvious in the areas outside the core zone, in the British Isles, for example, it is usually recognizably different in both repertoire and artefact type from the canonical Continental version [266, 267].

The transition from Bronze Age to Iron Age in temperate Europe is not characterized by any great cultural interruption. By and large, the defining characteristics of the last eight centuries of the first millennium BC were direct continuations of Bronze Age patterns, or piecemeal adaptations and adoptions.

At the beginning of the period, fortified sites placed on natural high ground were the most obvious non-funerary category of landscape evidence. Such enclosures, their British counterparts termed hillforts, were in the main not sites of permanent settlement, but probably served functions related to assembly and/or defense [268]. Hillforts saw a chronologically staggered period of currency, with several periods of construction, decline and new construction/re-occupation, but most finally declined towards the middle of the Iron Age [268, 269, 270].

In the later Iron Age, the first nucleated and proto-urban settlements, known as *oppida*, were seen in temperate Europe, from ca. 200 BC until they subsided at the end of the first century BC [271, 272, 273]. They had large surface areas and, frequently, hilltop locations, often making them – especially in cases of poor excavation – difficult to distinguish from hillforts in Gaul [274]. Where well excavated, however, oppida reveal roughly quadratic street plans, and different quarters for specialist use [271]. In the north part of the region (northernmost France, Lower Rhineland) oppida are rare, and the paucity of oppida corresponds with few Mediterranean goods and overall poorer material culture [275, 276]. This observation has lead to a perceived general poor versus rich, rural versus urban, and even 'German' versus 'Celtic' dichotomy, and such a division corroborates Caesar's observations of Belgica [275, 276, 277].

Box 1. Celtic - Uses and abuses

There are several problems with the term *Celtic.* Though it is used in the current work, it is necessary to distinguish three main usages, and stress that their fields of reference do not necessarily overlap.

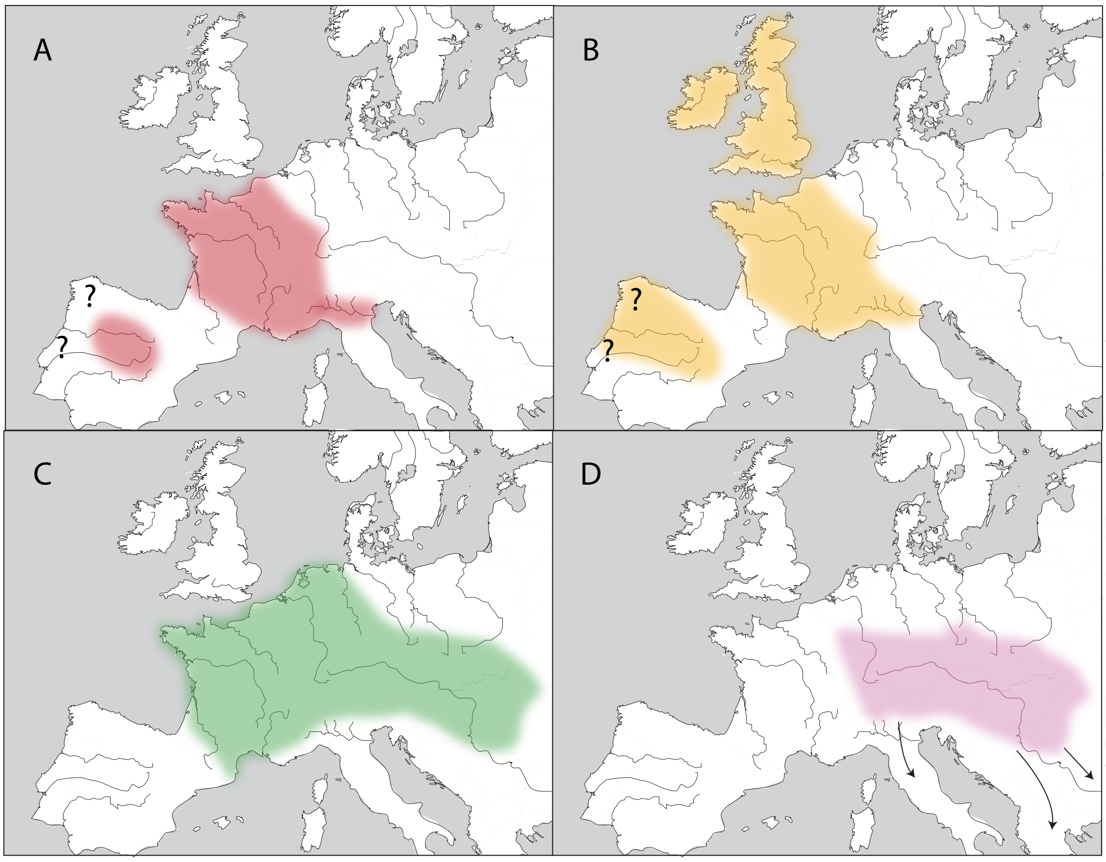
First, the original use of the term, by Classical writers, was to denote an ethnic group comprising certain tribes of western Europe, by and large those of modern-day France and certain areas of central and western Iberia. It is unclear to what extent these peoples would have attached the name to themselves, but Julius Caesar claims that people of Gaul, at least, did so.

Second, ‘Celtic’ was applied by Edward Lhuyd, in part following Buchanan, to a language group consisting of Irish, Scots Gaelic, Welsh, Cornish and Breton, having noticed similarities between them and forms recorded by Classical authors as being in the language of Celtae/Keltoi [3, 278, 279]. Later, Continental languages were discovered from inscriptions and glosses, and were included in the Celtic family.

The third refers to ‘Celtic culture’ or ‘Celtic art’ – essentially the La Tène phenomenon of the latter half of the Iron Age. The use originated in an understandable if unfortunate tautological series of equivalences made in early modern scholarship [280].This usage is particularly troubling when, as has happened, it has been uncritically extended to an idea that Celtic ethnicity has some kind of immutable identity.

The geographic distribution of these three definitions of Celtic can only reasonably be said to properly meet in France, especially in the northeast (see below).

A – Celtic areas by Classical authors, **B** – Areas of Celtic speech in the early historical period, **C** – Areas of marked La Tène influence, **D** – Direction of attested Celtic population movements.



As the Iron Age progressed, enclosed settlement sites increased (though not to the extent seen in Britain, see Chapter 4.1.2.6)[274, 281, 273]. While disparate opinions appear to exist, enclosures seem to be more common in the north and west, less so in central areas [268, 282, 283]. Mostly they are rectilinear, but round ones are known from Normandy [270]. Enclosures in Belgium have similarities with those from the Upper Rhine, and also southeast Britain [274]. Armorican sites are very rare for the Early Iron Age, but later they are oval and rectilinear in shape, and mostly univallate. Ditches and pallisades were common delineators of enclosures across the region [282, 270, 273]. However, unenclosed settlement sites were common everywhere, despite being archaeologically less visible [270].

Rectangular houses were the norm, as they were throughout most of prehistory, but exceptions may be found in the northern and western extremities, such as Normandy and Belgium, where circular wattle and daub houses very similar to those in Britain have been excavated [281, 283].[5] Other structures associated with settlement sites are post-built storage buildings, and storage pits, which may have had some ritual significance (see below and Chapter 4.1.2.6).

Burials constitute quite a heterogenous class of evidence. In many areas, such as parts of Low Countries and northwest France, the urnfield and cremation traditions continued well into the Iron Age [285, 286, 287, 288]. Overall, inhumation became dominant until cremation again gained ground in the last few centuries BC [289, 283]. However, much variability existed [288, 287, 289, 273, 290, 283]. The majority of the population were accompanied by few (if any) and poor grave goods, in simple pits – or sometimes cists in Armorica – whether they underwent cremation, inhumation or even excarnation [267, 290, 273].

The classic high status Hallstatt graves include features such as weapons, exotic Mediterranean goods, and accoutrements of a luxurious life enclosed with the corpse in a wooden chamber under a mound. Outside the core Hallstatt zone, these occur

---

[5] Circular buildings are also found from the 1st century BC in Armorica, again very like British examples in plan, but they are suggested to be ritual structures, much like the rectangular shrines of contemporary Britain [284].

in northeast France, and in the Low Countries where many Halstatt elements are incorporated into the existing cremation ritual [286, 290]. In northwest France, as in Britain (see Chapter 4.1.2.6), human bodies and disarticulated bones are sometimes found within settlement areas, in particular in ditches and disused storage pits.

High status La Tène burials of the latter half Iron Age share several features with those of Hallstatt, but often have a more martial aspect. Inhumation was by now standard and grave goods include weapons and protective gear, drinking paraphernalia, and some imported goods from the south [291, 290, 265]. They also might contain a wheeled vehicle with attendant fittings [290].[6] Such vehicle burials centre on what was later known as Belgian Gaul, from the Lower Seine across to the Rhine, but a few have also been found in Britain (see Chapter 4.1.2.6) [291]. Barrows surrounded by square ditches usually cover these vehicle burials [267, 290].

4.1.1.5.1.1 La Tène, Celtic and population movement For a long time, and by some still, the archaeologically observed expansion of elements of the La Tène culture was equated with ‘the Coming of the Celts’. While this is moderately reasonable for the historically known Celtic incursions into the east and the south, the Celtic language and classically asserted presence of Celts in western Europe overlaps considerably less with La Tène [293]. Iberia was virtually untouched by the La Tène phenomenon, despite having both ‘ethnic’ Celts and Celtic dialects in later prehistory [294, 295, 296]. The British Isles, while not ethnically Celtic in the Classical sense, were most likely linguistically Celtic in prehistory – though not necessarily before the La Tène phenomenon – but again, the islands were only lightly touched by La Tène material culture [266, 267].

Some westward population movement may yet be considered. There is archaeological evidence for a reduction of population compatible with migration in the Champagne region, and suggested burial evidence for immigration from Champagne in the Lower Seine region [265, 297]. Additionally, there may be some justification for equating

---

[6] Both ‘chariot’ and ‘cart’ have been used to describe these vehicles, but they have equal, if opposite, connotations that can be misleading in terms of intended use and prestige [292].

La Tène expansion with that of Gaulish and its speakers, specifically, as opposed to undifferentiated 'Celtic' [298].

4.1.1.5.2 The northern provinces of the Roman Empire In the last century BC, the expanding Roman Empire added its northernmost Continental provinces: the two northern Gallic territories (*Lugdunensis* and *Belgica*, roughly equivalent to the northern half of modern France), and the two Germanys (*Superior*, comprising western Switzerland, Jura, Alsace and modern southwest Germany, and *Inferior*, consisting of Luxembourg, southern Netherlands, parts of Belgium and North Rhine-Westphalia) [299, 300, 169]. The transition to a Roman possession was on average less smooth in the north, compared to southern Gaul and Iberia, which with their Mediterranean coasts had long lain within the Hellenistic and nascent Roman spheres of influence [299].

Roman towns were founded, but in northern Gaul and Germania these were only exceptionally on the same site as earlier native oppida. The oppida may have predisposed the natives to an increasingly urban lifestyle, but they were not direct precursors of towns [299]. Villas, though of distinctly regional characters, quickly became the norm. 'Native' sites are rare, and most common in the north and west peripheries (Armorica and the Low Countries), but even they have items of imported pottery, show evidence of Roman style funerary rites, and provide the occasional inscription [299]. These areas also have fewer, and less Roman-like, towns and villa complexes.

Roman habits seem to have been adopted relatively swiftly. The provinces eventually had their own vineyards and Gaul even became the pre-eminent exporter of Roman-style fine pottery tableware [299]. Early on, local gods are prominent in inscriptions, but later Roman deities predominate. This change may possibly be linked with decreasing use of local Celtic dialects [299]. Classic Roman temples, however, remained rare, and the Romano-Celtic temple layout, also seen in southeast Britain, was the standard religious structure until the medieval church organization superseded it [299, 301].

The division between Gallia and Germania must be considered one of political expediency, and people of both Celtic, Germanic and mixed heritage were found either side of the province borders [299]. As a general rule, within the northern provinces, those further south showed greater degrees of civil Romanization. Northern Belgic Gaul and Germania Inferior have a more Germanic character, while southern Belgic Gaul and Germania Superior are more Gallic, as well as more Roman [299, 276]. The experience of the Germanic frontier has many parallels with that of northern Britain, and will therefore not be discussed here (see Chapter 4.1.2.7).

4.1.1.5.3 Ireland If northwest Continental Europe, especially in the Early Iron Age, lacks extensive settlement evidence and Britain lacks plentiful burial evidence (see below, Chapter 4.1.2.6), Ireland lacks both. There is, overall, very little Iron Age material from Ireland, and what there is tends to be skewed towards the later part of the period, luxury items and élite sites [266]. There are few signs of rupture between the Bronze Age and the Iron Age in terms of innovation, apart from the comparatively marginal introduction of iron. Such domestic structures as existed were probably still round, as indicated by postholes [302], and there was continued use of Late Bronze Age hillforts [266]. Across the board, there was much focus on and reuse of earlier prehistoric sites. Ireland in the Iron Age shows signs of isolation. There are a limited number of Hallstatt C objects (Early Iron Age), which are overwhelmingly of domestic manufacture, and probably find their immediate antecedents in southern England rather than anywhere on the Continent. After Hallstatt C, one must wait until the introduction of La Tène styles in the third century BC for the next evidence of xenochthonous trends. Nevertheless, excluding the unequivocally foreign nature of the original impetus, La Tène objects on Ireland are mostly Irish-made and lack many features typical of Continental La Tène crafts. No single source for La Tène Ireland can be easily found, though there is limited evidence for reciprocal exchange between the northern half of Ireland (to which La Tène is virtually exclusive) and Britain. Pottery, which for much of prehistory is used as one of the foremost cultural diagnostic categories, is absent from the Irish Iron Age except for a few imported examples at large ritual sites

[266].[7] Until the Roman Empire enroached upon its neighbor, Ireland appears to have been largely self-contained – even bisected along a south-southwest/north-northeast axis – and only marginally susceptible to external influences.

4.1.1.6 Medieval northwest Europe

Early Medieval Europe is often called the Migration Period because of the large number of Germanic tribespeople migrating into and across parts of the Roman Empire. Tribes such as the Ostrogoths, Visigoths, Suebi, Burgundians, Franks, Langobards and Vandals settled as far from north-central Europe as Spain and North Africa [299, 303]. The kingdoms these Germans set up became more or less Romanized, as a function of two main variables: The previous exposure of the given tribe to Roman ways, and the depth of Romanization in the settled area [304].

The insidious infiltration of Franks into northern Gaul started in the fourth century, and during the course of the fifth century, the northern part of the province can properly be said to be under Frankish, rather than Roman, control [299]. At such a distance from the epicentre of the Empire, northern Gaul was heavily Romanized only compared to its neighbors in the north and west, and the Franks imposed their ethnonym and certain habits on the Gallo-Romans [301]. However, they also – eventually, at least – adopted the native Latin[8] vernacular, their church and much civic organization [299]. The fringes of Gaul and much of Germania lost many signs of *Romanitas* in the interrelated decay of the Western Roman Empire and the German expansion [299].

The experiences of Germanic movement on the proximate Continent provide several interesting parallels and comparanda for the largely contemporary Anglo-Saxon invasion of Britain (see Chapters 4.1.2.8 and 4.2). Unlike Britain, the later Roman Empire settled a number of largely Germanic people within its Gaulish boundaries. These *laeti* and *foederati* formed something akin to internal buffers against other Germans without a vested interest in the stability of the Empire [305].

---

[7] Due to their mention as ‘royal compounds’, sites such as Navan Fort and Tara have attracted much more archaeological attention than equivalent sites without a written pedigree.

[8] Perhaps partially Celtic.

4.1.1.7 Later Medieval Europe

The ninth and tenth centuries saw a new set of power bases arise in Europe, that of the Scandinavian areas, or the 'Vikings'. These Norwegians, Swedes and Danes eventually created a wide maritime and riverine network to support trading, raiding and eventually settlement activities [306]. The Swedish Vikings' sphere of operations was primarily the Baltic and Russian rivers, while the Danes and Norwegians focused on the North Atlantic and the North Sea [306]. Within the latter areas, though there was much interaction, the Danes were more active along the North Sea coast, and the Norwegians among the North Atlantic islands. In addition to various settlements in Britain (see Chapter 4.1.2.9), there is evidence of a certain level of Scandinavian immigration into two areas of western Europe: in northern coastal France, and Ireland.

The northern French region, which became the core of the duchy of Normandy, was ceded to new Scandinavian rulers partially in an attempt to ensure a secession of raiding up the Seine, in 911 [307]. There is little by way of archaeological evidence for settlement, being mostly limited to finds of Scandinavian weaponry [307]. Written records and linguistic repercussions of the settlement period form most of the proofs [308]. While the dukes of Normandy descend from the Norse chieftain Rolv (Rollo), the majority of Scandinavian settlers were probably Danish, with a predominance of males [307, 309]. In addition to Scandinavians directly from their homelands, there is evidence – both in weaponry and toponyms – of an influx from earlier established Viking enclaves in the British Isles; Hiberno-Norse and Anglo-Danish [307]. The Scandinavian adstrata to the Franco-Gallic population is also witnessed by place-names, both purely Norse and hybrid, such as Scandinavian personal names combined with *-ville*, and the large number of loanwords from Norse in various local dialects [308]. The Norse language, however, probably did not survive for much more than a century, if that, in Normandy, and culturally speaking, the region was quickly re-integrated into a Francophone, Christian polity with a population not dominated by Scandinavian descent [307].

The Scandinavian settlements in Ireland were of a rather different nature. First, they were dominated by Norwegians and a recognized hybrid Hiberno-Norse group [310, 311]. Second, they were not rural, but based around emporia in Dublin especially, Wexford, Waterford, Cork, and Limerick, in a country that had no pre-existing urban centres [312].[9] Archaeological evidence is limited as well, but Viking Dublin – like Viking York – had some well-preserved sections [313]. In another contrast to Viking settlement in Normandy and Britain, very few toponyms are Norse [314]. Waterford and Wexford are among the few exceptions [314], and the difference is perhaps due to the overwhelmingly urban and mercantile nature of Scandinavian settlement in Ireland.

4.1.2 Britain

4.1.2.1 Highland and Lowland Britain in the archaeological record

The separation of Britain into two broad zones of different physical geography, the highland and lowland zones, is one which serves as a useful shorthand throughout much of British pre-history and history. The term was first used in this context by Fox in his *The Personality of Britain* [315], and defines Lowland Britain as the section southeast of a line drawn between the mouths of the Exe in Devon and the Tees in County Durham, and Highland Britain as the regions to the north and west of this line – essentially, the Cornwall/Devon peninsula, Wales and Scotland (see Figure 4.8). This is the definition which will be used whenever the term is referred to in the following text, however, I do not intend it to reflect adherence to the kind of geographic determinism which was the original reason for its introduction.

In addition to this dichotomy, an east-west axis of dissimilarity may also be discerned. The principal difference between this scheme and the Highland/Lowland one is the separation of eastern and western parts of Scotland. Ireland, while physically separate from Britain, will frequently be touched upon under this heading, as it is often justified to treat its archaeological trajectory in common with highland Britain, perhaps, indeed, more so than to consider Britain as a whole.

---

[9] Some proto-urban functions were filled by monastic centres, which had artisanal workshops, but occasional tribal markets dealt with what little trade was present in before the Viking Age [312].

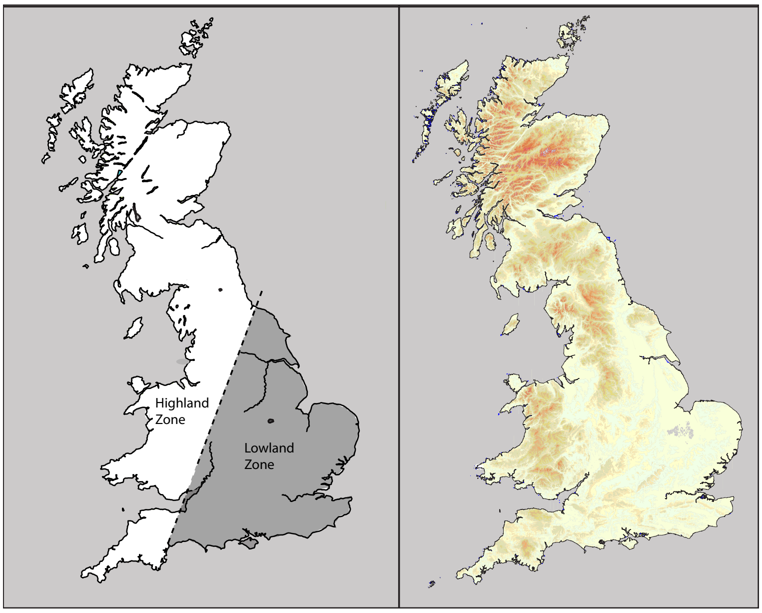


Figure 4.8: **Highland-Lowland Zone Division in Britain** The left panel defines the conceptual zones reflected in much of the archaeological record, and the right panel shows the topographical reality of Britain.

4.1.2.2 Late Glacial and Mesolithic

It is accepted that, since the last glacial maximum, Britain has been subject to at least one episode of human recolonization, and probably two. The first occurred around the time of climatic amelioration (ca. 13,000 BC), by a group carrying a regional version of the Magdalenian lithics industry, often termed Creswellian [182]. The Palaeolithic industries of Britain bear great resemblances to those of the proximate continent [183, 316, 182], a not unexpected development considering that the landmass at the time was simply an extension of the North European Plain, and the people probably highly mobile [317]. Later Upper Palaeolithic assemblages in Britain show strong affinities to the Federmesser group of northwestern Europe, dated roughly to the Allerød interstadial (12,000-11,000 BC) that preceded the Younger Dryas, as well as the Ahrensburgian of northern Germany/southern Scandinavia [318, 183, 319]. Firmly dated Palaeolithic sites have only been found in England and Wales, but there is some typological evidence for Palaeolithic populations as far north as southern Scotland [320].

Following this period of occupation there was an apparent hiatus in human activity during the Younger Dryas, a climatic cooling lasting roughly a thousand years, with an uncertain onset around 11,000 BC and ending around 9,800 BC [316, 321]. This postulated break in occupation, for which there was once only quite tenuous evidence, has been strengthened by more exact dating. Currently dated artefacts have estimation profiles that cover much of the Late Upper Palaeolithic, but the Younger Dryas period is empty [T. Higham, pers.comm.]. While the absence of occupation evidence does not necessarily entail the complete depeopling of Britain, it is probably indicative of at least a population contraction. In the final Palaeolithic there seems to have been cultural continuity in both spatial (with northwest Europe) and temporal (across the somewhat arbitrary boundary with the Mesolithic) dimensions, a situation incompatible with neither population continuity nor discontinuity *in situ* since the first post-LGM repopulation.

A problem in assessing the Palaeolithic (and Mesolithic) occupation of Britain lies in the rising sea levels after the Last Glacial Maximum flooding much land that was likely inhabited along the coast. Palaeolithic findspots in Britain tend not to be very far inland [180], and it may be that the majority of the population sites that existed are now submerged. The distribution of early Mesolithic sites is not dissimilar to that of Palaeolithic sites, and are limited to England and Wales. The earliest dated occupations in Scotland and Ireland are essentially contemporary, from the early to mid-eighth millennium BC [322, 323].

Early Mesolithic British sites, belonging to the same developmental sequence as the rest of northwest Europe, frequently show very strong similarities to sites in northeast France, the Netherlands, Belgium, and northwest Germany, and the axes of connectivity with current continental Europe are suggested to have run along the southern North Sea Basin [187], then dry or marshy land (see Figure 4.9). Subsequent to the drowning of the land between eastern England and the northwest of the continent, Britain entered a period of (perhaps total) isolation from the surrounding countries, and in the Late Mesolithic more distinct regional entities had emerged on both sides of the Channel [324, 325, 326, 199].

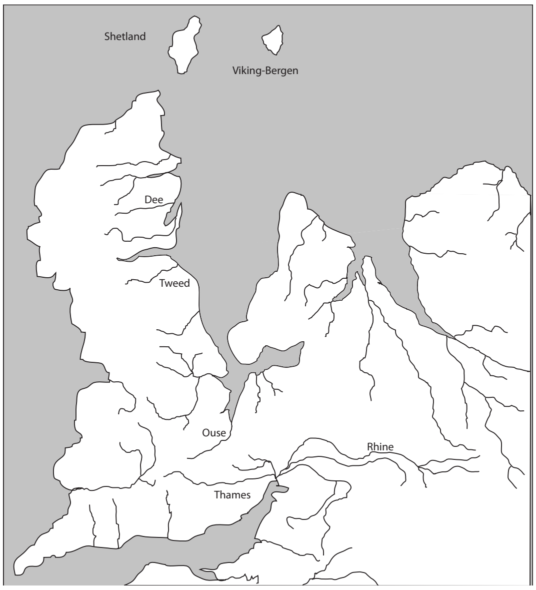


Figure 4.9: **'Doggerland'** The pre-transgression North Sea Basin.

4.1.2.3 The Mesolithic-Neolithic Transition

As with many territories in Europe, the arrival of the Neolithic in Britain (and Ireland, with which Britain shall be discussed) was a swift affair, such as to appear virtually instantaneous at the level of resolution provided by archaeological dating. This occurred, like the Neolithisation of Scandinavia, around 4000 BC [327, 328], with the first examples of objects such as pottery and domesticates. However, dating schemes have indicated that the development of the classical suite of insular Neolithic culture, with monumental architecture and predominant sedentary agriculture, may have taken as long as two hundred to a thousand years [163, 329, 330]. The origins of the British and Irish Neolithic are disputed, with perhaps the only element of consensus being that it is unlikely to have been the result of a single cultural impulse or migration event. Cultural analogs for most novel technologies in the Early Neolithic of the British Isles are found along the Altantic/North Sea coast, but not in the same place (see Figure 4.10). Julian Thomas summarizes the situation, and the potential confusion, succinctly [327]:

"Passage tombs like that at Achnacreebeag have affinities with those in Brittany, yet

> the trapezoidal long cairns of the Cotswold-Severn region are more easily paralleled in Normandy. Grimston [Carinated] bowls are related to Dutch Hazendonk or Belgian Michelsberg pottery, but the globular pots of the southwest of England are more redolent of the Chasséen of western France. The earthen long mounds of eastern Britain, and certainly the linear timber structures that they contain, find close paralleles in southern Scandinavia[...]. Individual causewayed enclosures in Britain suggest affinities with those of western France, the Paris Basin or Scandinavia [...]."

Furthermore, the aspect of the Neolithic appears quite different in different regions: In Ireland and the north and west of Britain the evidence is quite distinct from that of lowland Britain. In the latter, most hallmarks of a Neolithic lifestyle are present, including pottery, domestic cereals and animals, ground stone axes and the first indications of a typical western European feature: tombs and shrines of a monumental character [329]. Lacking, however, are strong indications of significantly increased forest clearance, fields, permanent settlements or any particular reliance on domestic plants and animals as opposed to wild ones [163, 331, 327, 332]. This situation contrasts with that of Ireland, where there is early and comparatively abundant evidence for a drastic shift to houses of some permanency, and in subsistence to domesticated resources - the island's lack of native land mammals larger than the pig could be a contributory factor [163, 327]. The northern part of Britain presents yet another facet of the Early Neolithic. The islands and coast of western Scotland bear significant similarities to Ireland, and cereals are consistently found throughout all of Scotland. However, as in England, permanent domestic structures are rare [333]. The situation in Wales is largely analogous to that of Scotland, with houses occurring more often than in England but more rarely than in Ireland [163].

Taking the evidence at face value, arguments have been made for a comparatively modest initial impact of the Neolithic on the lives of people in southern Britain, and that domesticates, pottery and early monument types lay lightly on a society that still relied heavily on wild resources and was not fully sedentary [329, 331]. Others - frequently those whose primary area of research lies outside England - attribute the Scotland/Wales versus England discrepancies to taphonomic concerns (e.g. differential preservation) and less excavation in the latter, but this is difficult to credit in light

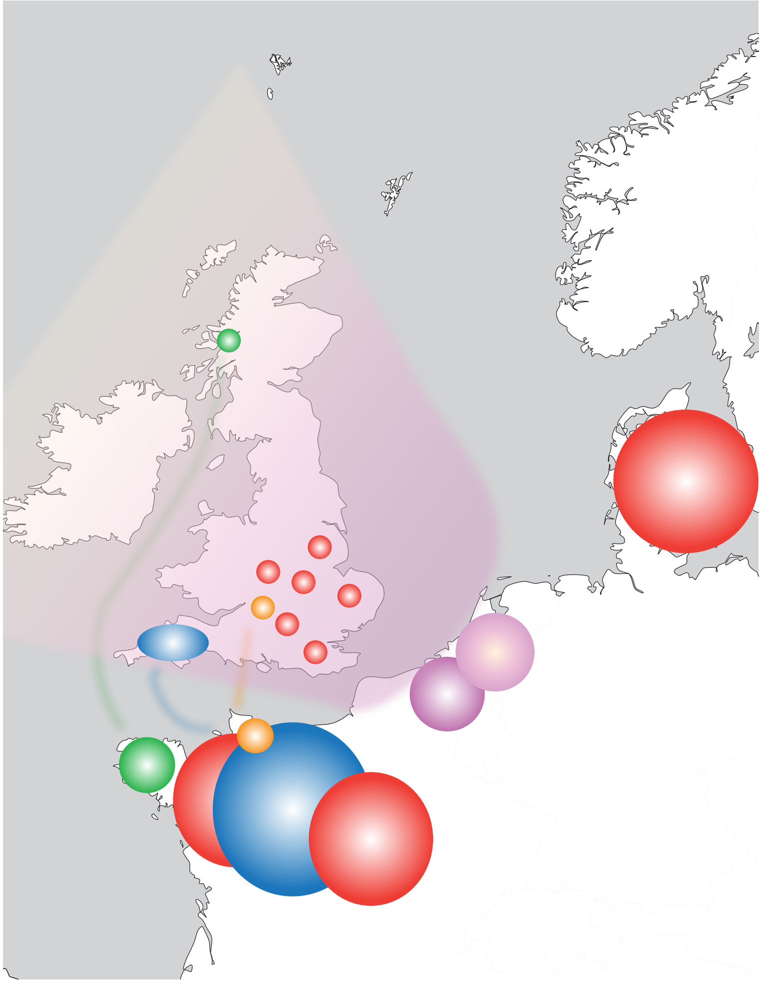


Figure 4.10: **Continental Neolithic influences in Britain** Yellow - trapezoidal long cairns in the Cotswold-Severn region and Normandy; green - passage tombs in Brittany and Achnacreebeag; blue - Chasséen type pottery in western France and southwest England; red - causewayed enclosures in England showing different Continental likenesses; pink/purple - Carinated bowls in the Low countries and throughout the British Isles. Color shading indicates the approximate area covered by the material culture category in question.

of very good preservation in the same area for other time periods and the fact that few places are as archaeologically well canvassed as the English lowlands [327]. The differences appear to reflect a real situation [163].

Certain features unite the shift to the Neoltihic in the British Isles, such as the Carinated Bowl tradition and ground stone axes of a not strictly utilitarian nature, but the heterogeneity of continental links related to innovations makes interpretation difficult. It seems clear that 'the' British Neolithic - though it is probably unproductive to view it in a unified fashion - was an amalgam of native substrata and several different continental sources. The next question to be addressed is, how were these traditions (not necessarily as physical entities) and domesticates (by necessity in physical form) introduced? A simple attempt at an answer is: by sea. As a clarification it is unhelpful, but it raises an important point. Any contact between farming communities of the continent and Mesolithic communities in Britain would have had to have been by boat. No direct evidence of these craft exists, but the most likely candidate for seagoing vessels at this time are light, wood-framed boats with hide covering [240].[10] Writing at a period when Childe and Piggott's visions of Neolithic pioneers overwhelming the British Isles were still current [335, 336], Humphrey Case made the point that any large scale immigration would be logistically extremely difficult and therefore implausible, though he did not argue against colonization being the vector for change [337]. All domesticates found in the Early Neolithic in Britain and Ireland derive from non-native predecessors, and therefore must have been transported by boat. Case further argues that it would not have been possible to feed or water the animals during the voyage, and that keeping seed corn dry would have been a substantial challenge, and concludes from this that actual episodes of importation were likely to have been small and infrequent.

Moving on to the question of by whom the innovations were introduced, there is again a dichotomy of opinion, which by and large reflects the aforementioned split between archaelogists who work in England versus those who work in other areas of the

---

[10] Logboats, for which archaeological evidence is known from the sixth millennium BC, are also a possibility. They have often been assumed to be unsuitable for seafaring, but ethnographic research indicates they could be used for such a purpose [334].

British Isles. The latter cite the paucity of evidence for any contact between the British Isles and the continent in the Late Mesolithic as proof that Insular Mesolithic people were introverted and would not have imported a foreign mode of existence without external influence [328, 338], while the former contend that the timing of the event and different continental origins argues for significant native agency being involved in the shift: They can identify no reason for disparate Neolithic groups along the Atlantic and North Sea coasts of Europe, virtually simultaneously and presumably not consciously coordinated, to dispatch colonists to the British Isles [331]. Bidirectional maritime connections between the islands and the continent being equally and multilaterally responsible for the introduction of the Neolithic has also been posited [339].

Alison Sheridan, of the camp promoting sizeable migrations of continental origin, envisions three 'Neolithic strands' of subtly different timing: i) a Breton strand moving up along the Atlantic/Irish Sea coasts into Wales, western Scotland and northern Ireland, around 4200-3900 BC, ii) a Normandy-southwest England link, 4000-3700 BC, including ceramics and funerary monuments; and iii) the Carinated Bowl tradition, which she envisions spreading from Picardie/Pas-de-Calais around 3900-3700 BC across the British Isles, to Ireland via Britain [338, 328]. Sheridan particularly stresses the uniformity and competence with which the pottery (especially Carinated pieces) were made in Britain, and sees this as proof of immigration as opposed to acculturation, a view which is echoed by others working in northern and western Britain and Ireland [338].

A similar reasoning lies behind the supposition (*contra* Case [337]) that domesticates were necessarily accompanied by substantial groups of immigrants familiar with husbandry and agriculture – that the specialist nature of the tasks required a community raised to their performance. Osteomorphological and assemblage evidence suggests, at least, that domestic animals were introduced to Britain and thence to Ireland, from the region of the Paris Basin, with the possibility that a 'southern' Atlantic stock, ultimately of Cardial origin, also played a part [340, 338].

Another element which may support a rapid upheaval of established tradition is an apparent abandonment of marine foods at the onset of the Mesolithic [341, 342].

However, the relevant samples postdate the very early Neolithic, and there are other interpretations possible than simply the impact of introduced domesticates [330, 343].

Collard *et al.*, using the (radiocarbon) dates-as-data approach, identify a "population boom" in the Neolithic that they assert is due to two events of immigration by Continental farmers [344]. They identify these as two of the strands in Sheridan's model, i.e. as strands i) and ii) above, but prefer a reversal of their chronological order. However, their conclusions rely heavily on the validity of both the dates-as-data approach (critiqued in Bronk-Ramsey [345]) and their models of differences in 'learned Neolithic' versus 'born-Neolithic' population growth. The latter seems poorly defined. Another recent, model-based approach covering Europe, concludes that, while "migration and cultural diffusion by trade may explain the western European transition equally well", where local adopters are more important to the establisment of agriculture than are migrants [346].

Recent re-sampling and reassessment schemes of radiocarbon dates in southern England, with the view to creating a less coarse chronology, have shown that the Neolithic package was not introduced as immediately as previously thought, demonstrating that monumental architecture was not among the earliest features. Long barrows (mounds) do not appear until two hundred years after the earliest Neolithic signs in Britain, which produces both a temporal and geographic disconnect with their supposed longhouse antecedents on the continent [330, 347]. Causewayed enclosures, large earthworks with continental parallels, come in yet later. There are also suggestions that the passage tombs around the Irish Sea – though obviously derived from the Atlantic tradition – mostly post-date the 4000 BC timing for the Neolithic onset by several hundred years [348].

The evidence for continuity of site use, often seen as continuity of population, is again different, or at least differently interpreted, in the south and east of Britain and in the north and west and Ireland. Continuity is perceived to be strong in the south and east, with Mesolithic layers being directly overlaid with Neolithic culture in a non-intrusive manner, and weaker elsewhere [330, 349, 338].

The 'acculturationist' opinions, current among archaeologists of the south and east, sees the somewhat piecemeal introduction of Neolithic elements as a reflection of less than wholesale immigration. For example, while not denying strong continental links, a mass of incomers would not be expected to bring with them only a definite subset of their material culture, as is the case with the Michelsberg-Chasséen links. This view holds that the evidence is best explained by a quite small, but non-negligible, level of migration, with the active participation of the indigenous people in the establishment of a new way of life.

We are left with a picture of fragmented units across the geographic area of the British Isles and the temporal unit of the earliest Neolithic (ca. 4000-3700 BC). Many innovations are shared across the Isles, such as cereals, domesticated animals, polished stone axes and leaf-shaped arrowheads, pottery and monumental architecture. However, their expression was not always uniform, either in space (e.g. different monument types in different parts of the country) or in time (with several centuries passing between the introduction of different elements).

It may well be that the Highland Zone and Ireland were subject to proportionally more continental immigrants than the Lowland Zone, but it is difficult to see the introduction of the Neolithic to the British Isles as a holistic, if 'three-stranded', process in which the local Mesolithic inhabitants were completely passive and wholly obliterated, culturally if not physically.

4.1.2.4 Neolithic

After the earliest Neolithic, regional traditions developed further, and contacts within the British Isles, and between them and the continent, were reduced [329, 350]. New monument types developed, such as the bank-defined cursus, emerging in Scotland from timber predecessors contemporary with causewayed enclosures, and spreading south [163]. Loosely related pottery styles known collectively as Peterborough Ware were current, and the prestigious production and trade of axes continued [329]. The passage tomb tradition continued likewise, and was subject to a process of competitive emulation in eastern Ireland, northern Scotland and Orkney, giving rise to such spectacular examples as Knowth and Newgrange in the Boyne Valley and Maeshowe in Orkney [351]. Out of this passage tomb environment, with its distinctive abstract art, came the Grooved Ware tradition. This ceramic style probably originated in Orkney, in a funerary context, but was also used domestically in the islands' permanent, stone-built villages [163, 350]. Indicating increased connectivity within the British Isles, Grooved Ware is later adopted in mainland Britain and Ireland in the early third millennium BC [352, 350]. The changing contexts in which the pottery is found makes it likely that its significance is re-interpreted in different regions from everyday domestic in Orkney to ritualising domestic settings in the house-poor south of England [350]. In mainland Britain, Grooved Ware is usually associated with ceremonial centres, such as the later Neolithic henges, and the currency of these monuments also saw the most extensive distribution of ground stone axes [163].

The Neolithic in Britain was introduced via centuries of contact, possibly intermittent, with the proximate continent. After a period of comparative isolation, Britain re-emerges into the wider geographical scene at the end of the period, at the beginning of the Bronze Age.

4.1.2.5 Bronze Age

The onset of the Bronze Age, to a much larger extent than the Neolithic, was not an 'event' with particularly wide repercussions for material culture in Britain as a whole. Any chronological boundary in archaeological records is to some extent arbitrary, and the Neolithic-Bronze Age boundary particularly so. Any time in the five hundred year period of 2700-2200 BC are common choices, and these dates are almost wholly related to the introduction of metal objects and metallurgy.

Overall, the distinctions between late Neolithic and Early Bronze Age in Britain (and Ireland) are difficult to draw. Domestic architecture saw no revolution. House remains speak of comparatively ephemeral, smallish, round structures across most of the British Isles, with exceptions in upland areas where foundations were occasionally stone-built [243, 163]. Economic practices likewise continued much as before, but with some indications of intensification of agriculture, and flint- and stoneworking practices continued along trajectories from the Neolithic [243].

The most prominent Neolithic burial tradition, placement in monuments or tombs, with the remains of other individuals, did not continue into the Bronze Age, when individual burials became the norm for 'important' members of society.[11] Single burials may have been common throughout the Neolithic, but this is disputed [219, 163]. Within the scope of a tradition for individual burial, the form varied widely in the British Isles. In western Ireland, and western and northern Scotland one finds tendencies to continue megalithic traditions, albeit on a smaller scale than before, and in eastern parts of Ireland and Scotland single inhumations in stone cists are common. In England and Wales, mounds or 'barrows'[12] are the defining Early Bronze Age funerary type, but these are also found in Ireland and Scotland, though less frequently [243, 163].

In terms of pottery styles, Grooved Ware use continued into the Early Bronze Age and so probably did descendants of Peterborough Ware [243]. The Beaker tradition (see Chapter 4.1.1.4) was also widespread in Britain. Here it corresponds to the first evidence of metal use, in conjunction with the copper knives frequently found in 'Beaker style' burials. Other common elements of these burials are barbed and tanged arrowheads, stone wristguards and the eponymous beaker [163]. The pottery in Britain occurs in domestic contexts as well, and so was not a purely ritual or funerary style. Later Beakers are frequently associated with another, larger ceramic type of indigenous origin, the Food Vessel, particularly in northern Britain [243].

Although Britain adopted metallurgy comparatively late and has no Chalcolithic (Copper Age) to speak of, the island was at the forefront in the western European Bronze Age thanks to its deposits of tin (in Cornwall) for making tin bronze [163]. The new technological introduction, as with the introduction of farming, was once seen to be

---

[11] Neither Neolithic monumental tombs nor the rich/labour-intensive graves of the Bronze Age would have been the final resting places of the majority of the population, though the precise nature of how the favoured few acquired their status is unknown [243].

[12] The word is derived from the Old English '*beorg*' meaning 'rounded hill' which was frequently applied to these man-made mounds [353].

mediated by a 'Beaker People' [247]. Subsequently, this was amended to an élite class of peripatetic, Beaker-bearing smiths. Latterly, metallurgy is mostly interpreted as a skill adopted from continental metalworkers, but dissociated from much of their other cultural context [354]. The de-emphasis on migration-borne metallurgy owes much to the aforementioned amount of material culture continuity, as well as the very different facies of the wider contexts in which Beaker assemblages are found in different countries [244].

The extremely large area that the Beaker culture spanned, rather than covered, has given rise to the idea that there was extensive mobility of people across Europe at this period. In order to get a measure of actual migrants, analyses of stable isotopes in bone and tooth enamel have been employed. The theory is straightforward: The ^{87}Sr/^{86}Sr and ^{18}O/^{16}O ratios vary with geology and microclimate, and comparatively detailed maps can be drawn up for local values. During bone and enamel formation these elements will be incoporated, reflecting the geological composition and water of the area in which an individual lived [355, 356]. Enamel isotope values will be those of early life, bone those of the last few years of life. Using this technique, a surprisingly high percentage of 'migrants' were found in Beaker cemeteries of Central Europe (63% of 81 individuals), and three of five Beaker interments near Stonehenge contained 'non-locals' [357, 358]. Criticisims, however, may be levelled at the interpretation of isotope data. Often, isotope ratio maps are insufficiently detailed, many other elements than simple geology may affect strontium ratios, and such sources of error are infrequently accounted for [356]. It is likely, therefore, that the number of migrants claimed in isotope studies in general, and not just Early Bronze Age, are overstated, and perhaps substantially so [356].

In the later phase of the Early Bronze Age/Middle Bronze Age increasingly regional traditions of pottery developed in the British Isles, which are thought to reflect, to some extent at least, regionality in other cultural spheres. While Food Vessels and Collared Urns are traditions that occur throughout the British Isles, in slightly differing contexts, later urns separate them into three areas; Scotland/northern England and

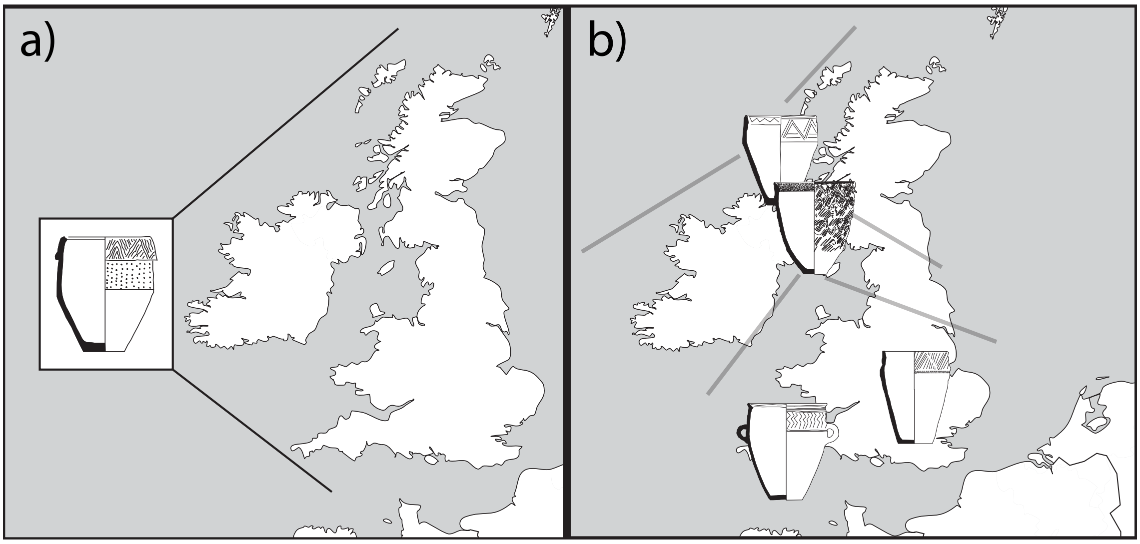


Figure 4.11: **Early and Middle Bronze Age cinerary urns** a) Collared Urns, found across the British Isles, though in Ireland confined to the east. b) Cordoned Urns in Ireland and Scotland, Encrusted Urns, found in northern Engand as well; Trevisker ware found in southwest England, and Biconical Urns in lowland England. Urns redrawn after the following: [360, 361, 359, 243].

Ireland, southwest England, and lowland England (see Figure 4.11) [359, 243]. As the period progressed, an increasing trend towards cremation, rather than inhumation, is seen [243]. Often the urns above are found in conjunction with cremated remains, both in flat graves, reused megaliths, and in the final period of barrow building [243, 255].

In terms of Britain's links with Ireland and continental Europe at this time, the evidence is ample. Primarily, pottery and metalwork provide this proof, but connections can also be seen in structural remains. Britain and Ireland, as noted above, share pottery traditions of Food Vessels and Collared Urns[13], and Ireland and northern Britain share later regional urn types. Metalworking traditions likewise show connections between the islands. Indeed, some of the earliest bronzeworking traditions, as well as gold objects such as lunulae, likely entered Britain from Ireland [256, 163]. Many style trajectories in metalworking continued in parallel between the islands. Burial traditions also tend to associate parts of Ireland and western Scotland, such as Beaker burials reusing megalithic tombs and cists. Housing norms, too, followed similar patterns of

---

[13] Originating in southwest Britain.

increasingly substantial structures by the Middle Bronze Age.

Continental connections are no less obvious in the first half of the Bronze Age. O'Connor even states that in the latter part of the period, south and central England have more in common with the proximate continent than with the other parts of the British Isles [257], and Marcigny speaks of genuinely cross-Channel communities including the Low Countries, eastern and southern England, and northwest France based on shared ceramics, lithics, house types and deposition practices [254, 362]. There have been suggestions that some communities in Belgium and the Netherlands originated in southern Britain, such are the detailed similarities of their recoverable material culture (urns, house shapes, barrow types, etc.) [363, 249], but this has been refuted [250, 252, 364, 253, 365]. Similarly, Burgess rates the resemblances between Middle Bronze Age southeast England and Picardy as such that "there had to be movement of populations" [366]. To mention but a few specifics of portable artefacts: Norman palstaves are found in southern coastal England, Norman ceramics in southern England and North Sea coast, southern English biconical urns and palstaves in Picardy and the Low Countries, Irish lunulae in Britain and a few examples on continent, Armorican *vases à anses* in southern England, tin and bronze from the southwest of both England and Ireland, respectively, in in western Scotland, and a triangle of exchanged local axe types between southwest England, Armorica, and Ireland (see Figure 4.12) [257, 363, 366, 249, 163, 254]. A special case of undisputed and intense contact occurs between the cultures that produced the 'rich graves' of Armorica and Wessex. These barrow graves, which define the Wessex 'culture', are particularly rich in grave goods and are clustered in the landscape to an extent that sets them apart from the rest of the country [366]. Also, with very few exceptions, the urn tradition current across the rest of the British Isles is absent. Most striking, however, is the"unparalleled degree of continental influence", mostly from Armorica, which has barrows of very similar construction and contents [367, 366, 254]. This episode of enhanced communication, thought to represent an élite grouping and not the general populace of either area, was possibly of short duration [367, 366], but has nevertheless often dominated much of the

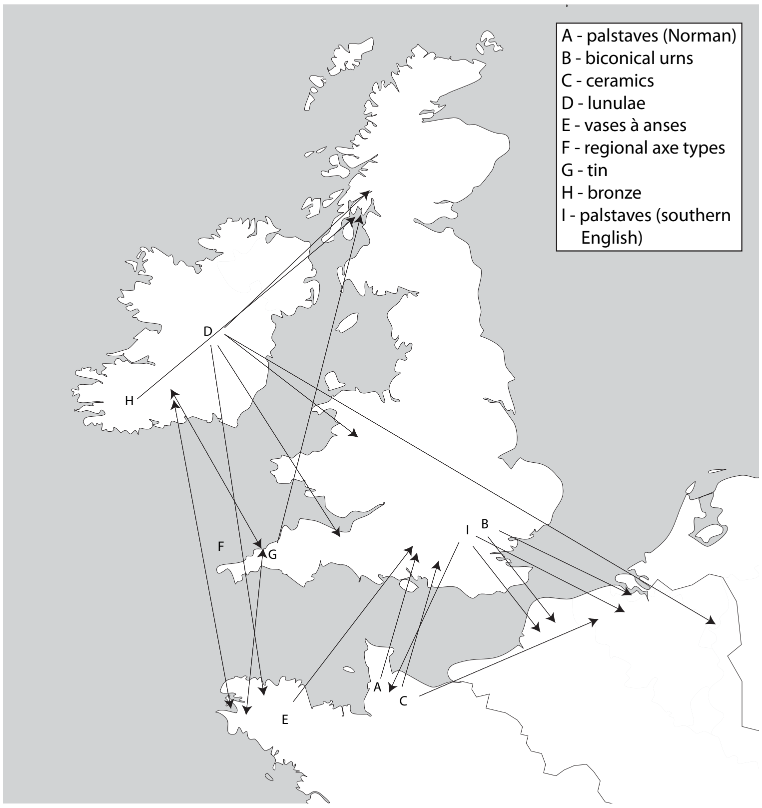


Figure 4.12: **Artefact and metallurgy links** Selected artefacts and commodities displaying connections between the British Isles and the continental coast in the Early and Middle Bronze Age.

English Early Bronze Age scholarship not dedicated to the Beaker phenomenon.

While the strongest and most lasting links between Britain and the continent focus on the above areas – its closest neighbours – central Europe and and more distant Atlantic coastal influences can be distinguished, often indirectly, even in the earlier Bronze Age [363, 163, 264].

The above speaks for the existence of a highly connected Britain in the Early and Middle Bronze Age, even – perhaps especially – latterly, when regionalization of several traditions was ongoing. To what extent there were any sizeable population relocations is, as always, difficult to assess, and is currently but rarely suggested. The exceptions include postulated southern English settlements in the Low Countries (almost universally discredited) and Picardy, and Armorican settlements in Wessex.

Of the changes occurring in Britain upon entering the later Bronze Age (ca. 1500-

1400 BC), the cessation of barrow building is perhaps the most striking, and in many areas of Britain there is little evidence of interment rituals for the next millennium, though barrow building survived into the Iron Age in Ireland [258, 163]. In southern England there are cremation burials associated with a widespread type of pottery (Deverel-Rimbury ware), of which local variants existed, in increasingly smaller barrows or flat graves. Settlement practices likewise saw changes, with field systems, fences, ditches, and stone, earth or timber territorial delineations. Structures become more substantial and though roundhouses are still the standard, though several longhouses similar to Dutch examples are known in southern and east England [258, 163].

Metalwork types continue to be mostly represented by axes, spearheads, and newer swords, of regional but co-evolving types, but the context in which it is found shifts from burial and hoards to increasing numbers of hoards, preferentially in rivers [163]. In general, the ‘dry’ hoards that exist are more utilitarian and/or personal in nature, containing tools and trinkets, as opposed to the weaponry of wet depositions [258, 163]. These too, however, tend to be found in liminal settings. Within this tradition, regional differences can be discerned, with southeast England having a lot of carp’s tongue swords, spearheads and belt fittings (much like northern and western France, see below), while the rest of southern England has mostly spearheads. The Midlands likewise has spearheads, but of a different styles, while in the north swords are most common.

All of the British Isles saw a great increase in the amount of metalwork, if judged by deposition, and at this stage it seems likely that most, if not all, copper was being imported from Alpine sources [258]. The growing importance of warlike expressions (weaponry, armour, and fortifications) may be linked to the population growth and concomitant overexploitation of marginal land and climatic deterioration that the insular Late Bronze Age witnessed [258, 163, 368]. The Late Bronze Age also saw, for the first time in prehistory, the marginalisation of stone (flint) tools [369]. General connections between Britain and its neighbours have been noted above, in connection with the Atlantic Bronze Age (see Chapter 4.1.1.4). Specifically, the Manche-Mer du Nord (Channel-North Sea) culture area continues to show strong cohesion, with

Deverel-Rimbury-like pottery in northern France/southern Low Countries and similarities of fortified sites, coaxial field systems, and comparatively large numbers of imported metal artefacts as well as local imitations of foreign types [362, 163, 257, 370, 259, 258].

Ireland and northern Britain are much more isolated in this period, and such extra-regional connections as are made are mainly between these areas and the rest of highland Britain, especially Ireland and northeast Scotland [359, 263, 163]. Influences of Irish metalwork on the continent appears to have been achieved through England, thence both English and Irish types probably reached the Low Countries via northern France, rather than directly [257, 371, 362].

Weaker connections, most likely sustained by many intermediate links, brought Urnfield influences to Britain, Nordic influences to Ireland, and maintained affinities between the northern and southern Atlantic communities. The Bronze Age as a whole was a period in which Britain was well, and occasionally very intensively, connected with surrounding landmasses. For most of it, indeed, many regions have greater resemblances to overseas neighbours than to more distant parts of the island. From 8th century BC, both these networks and society saw some wide-reaching changes [372].

4.1.2.6 Iron Age

The final flowering of the Bronze Age and the beginnings of the Iron Age in Britain cover the period 800-600 BC. Iron defines the Iron Age to a smaller extent than bronze did the preceding age. This difference is probably due to the lack of equivalent status for the metal in northwest Europe – iron ore is plentiful in most areas, copper and tin distributions are much more restricted and largely non-overlapping. Controlling access to, and transport of, copper and tin for bronze production would likely have enhanced the luxury of bronze in a way not paralleled by iron [265]. Regardless of the reason for the reduction in bronze use and deposition (crisis in supply of, or confidence in, bronze, or the devaluation of it by iron [163, 372, 373]), the Late Bronze Age networks saw an overall reduction and extensive remodelling [240, 372].

The Bronze Age-Iron Age transition is one of broad continuity in the domestic sphere, and innovation in the ritual sphere. Domestic architecture continued to be dominated by round houses and huts, through the specificities of construction varied on a regional basis. Regional distinctions can also be drawn on the basis of enclosure and settlement types, though it is a matter of the frequency of typical forms rather than absolute divisions. In western Britain, both north and south, small settlements, round and defensive in nature, were common [374, 267, 163]. Courtyard houses and rounds are peculiar to Cornwall, though they share features with the rounds of Wales, and the brochs and duns of Atlantic Scotland [374, 375, 267]. Another feature of the Atlantic southwest which it shares with Armorica in particular, but also Ireland and Iberia, is the presence of promontory forts. Enclosures away from the western periphery, containing circular houses where known, vary widely in size, shape and perimeter configuration, and intensity and primary purpose of use. Though enclosures are by and large round, rectilinear enclosures are known in the Welsh Marches and the English Midlands. Multivallate enclosures (earthworks, ditches, pallisades or a combination of these) are also common here. Northern England and Scotland have more pallisaded enclosures, increasingly so with time [374, 267]. Open sites were present all over Britain, and probably more common than currently recognized, but the paucity of enclosures in eastern England compared to the rest of the island means that the predominant settlement type in the east is large and unenclosed. A final zone, where hillforts are the defining element, covers south-central England and includes the Welsh borders. Hillforts developed from Late Bronze Age enclosures, and belong to the extreme end of the enclosure continuum based on larger size and more evidence for ritual practice [375, 163]. In the later pre-Roman Iron Age, nucleated settlements start to appear in the lowland area with strong Continental links (see below), while highland zone settlement displayed a remarkable amount of stability in both form and occupation [374, 375, 376]. The southeast also had settlement features interpreted as storage facilities, with both the remains of post-built structures and pits [375, 374].

While burials are a major source of evidence for much of prehistory, and indeed are still for Iron Age Continental Europe, there is comparatively little evidence for funerary custom in Iron Age Britain [373]. Where regional traditions exist, they appear to be in minority, but habits can broadly be divided into pit burials, possibly with prior excarnation, in south-central England, inhumation in south Dorset, cist inhumation in southwest England,[14] and Arras inhumations in east Yorkshire [377]. Excarnation and non-normative inhumation were likely the most common means of body disposal in the early and middle Iron Age, with a trend to increased cremation in the Late and Roman Iron Age [375, 377]. Both full skeletons and disarticulated remains are also found in contexts that may not seem funerary, such as disused storage pits, settlement ditches or other domestic boundaries, caves, or bodies of water [375, 163, 378]. In fact, much of the Iron Age evidence in the south and east, contrasting the paucity in the north and west, comes from structured deposition of human bone, usually with animal bone and artefacts, in pits and ditches [374, 267]. These types of pit burials are unlikely to be commonplace standard burials, as most remains bear evidence of trauma (B. Cunliffe, pers. comm.).

Two burial rites stand out in the essentially native British context; the Arras burials from the fourth century and the Aylesford-Swarling cremation cemeteries either side of the Lower Thames from the first century BC. Both are notable for their Continental similarities and probable antecedents. The Arras burials are inhumations in mounds surrounded by rectangular ditches, often in larger cemeteries, and may contain carts and harness fittings as well as offerings and objects of personal decoration [267]. These burials are very similar to contemporary ones in northern Gaul (see Figure 4.13), but there are certain differences, for example in Yorkshire the body is often deposited in a crouched position more typical of British tradition than northern French, and while personal ornaments are often obviously Continental in inspiration, others fall into 'native' British categories [267, 163]. The later Aylesford-Swarling burials of the southeast lack the clear-cut geographical area of inspiration of the Arras, but the character is on the whole Gallo-Belgic or South Belgic (see Figure 4.14). The Aylesford-Swarling rituals

---

[14] Formal cemeteries from fourth century BC southwest England are known, similar to ones in Armorica and the Channel Islands, but burials retain a 'British' crouched position [163].

Figure 4.13: **Arras type graves** Distribution of graves nearly identical to those found in East Yorkshire. After Cunliffe 2005 [267].

include cremation burials, with remains placed in urns, or less usually in buckets, and frequently accompanied by Continental type pottery vessels, bronze brooches, and occasionally other personal items [267]. The ritual is generally seen to have developed in the context of Continental influences, but the idea that it owes more to autochthonous trends has also been advanced [377].

Cases for both continuity and shift in the ceramic record can be made, depending on focus [372]. However, for most of the Iron Age, a large portion of Britain displayed the most extreme kind of discontinuity in this respect: Wales, the southwest and Scotland (excepting the northern and western Isles) were virtually devoid of pottery. In these areas, what little pottery is found tends to be both coarse and non-diagnostic and/or indistinguishable from Late Bronze Age ceramic forms [267, 379]. However, when the expectation is that a site should be aceramic, what pottery there is tends to be ignored, and southeast Wales, for example, has a regional version of the Middle Iron Age central/southern British pottery tradition [378]. Within lowland England, pottery fell into several distinguishable but related regional types in the Bronze Age-

Figure 4.14: **Ayleford-Swarling core and South Belgic culture area** After Cunliffe 2005 [267].

Iron Age transitory centuries. Some of these traditions show more external influences, such as the Early All Cannings Cross and Kimmeridge-Caburn groups of Wessex and the south-central/southeast coasts, which display skeuomorphs probably derived from imported metal vessels [267]. At the same time, the last great bronze hoards were deposited. Swords predominate in these, and most examples are of British manufacture, if ultimately of Continental inspiration [267]. Regional groupings based largely on the composition of the hoards provide a complement to those based on pottery, with foci in southeast England, southeast Wales and southwest England, south Wales and into the Midlands, north Wales, Yorkshire and Northumbria, the east coast of northern England and Scotland, and Atlantic Scotland .

If judged by pottery, the period from ca. 600 BC sees further regionalization in southern Britain. The southeastern styles were highly influenced by northern France and the Low Countries wares, especially those of eastern England and the Thames Valley. This influence is most parsimoniously explained by the intensification of exchange networks, though some population movement could be posited [267], particularly in Kent. La Tène stylistic influences stemming from northeastern France can be seen not only in pottery, but also in metalwork such as the numerous fibulae found in Britain

[267]. Despite the above regional distinctions that can be drawn for certain categories of material culture, southeast Britain was of a broadly unified cultural aspect in the Early Iron Age [267].

From 400/300 BC, further general continua of ceramics exist in southern Britain, dominated by saucepan pots in south-central England, decorated bowls in the Midlands, decorated bulbous jars in the east, and decorated neck jars in the southwest; all of which owe something to earlier local traditions. Some pottery was also made in northern England, which may show some links with south-central British and northern French wares, but is mostly of local derivation [267]. Into the proto-historic period, there are also four broad traditions (see Figure 4.15). The distinct pottery of the Aylesford-Swarling culture, whose burials traditions were treated above, spread north and west of its eastern/Thames Valley focus, and has similarities to the pottery of the Belgic Gauls. The south-central area, that was or became Atrebatic, had pottery developed from the local saucepan tradition, though Belgic influence may be discerned. Further west, the 'Durotrigan' and Southwest Cordoned Ware groups were both affected by Armorican traditions.

As in the Bronze Age, flint tools were still in use, but their coarseness and the existence of more diagnostic items means that they are largely ignored [380, 381]. Metal, both bronze and iron, is the primary constituent of surviving tools and weapons. The increasing use of iron freed bronze from utilitarian spheres, and many typically Iron Age examples of metalwork are in fact decorative bronze items. The majority of British Iron Age metalwork finds ultimate inspiration in the Halstatt and La Tène cultures of west-central Continental Europe, such as swords and bronze buckets of the early Iron Age Hallstatt C period, daggers and fibulae of the following Hallstatt D, and swords, brooches, bracelets and torcs of La Tène from the mid-fifth century BC [267, 381]. The local flavour of this metalwork is clear, however, and the main means of disposal sets Britain apart from the Continent: British metalwork is commonly deposited in wet contexts such as rivers, and only exceptionally in graves, cf. Continental traditions [267].

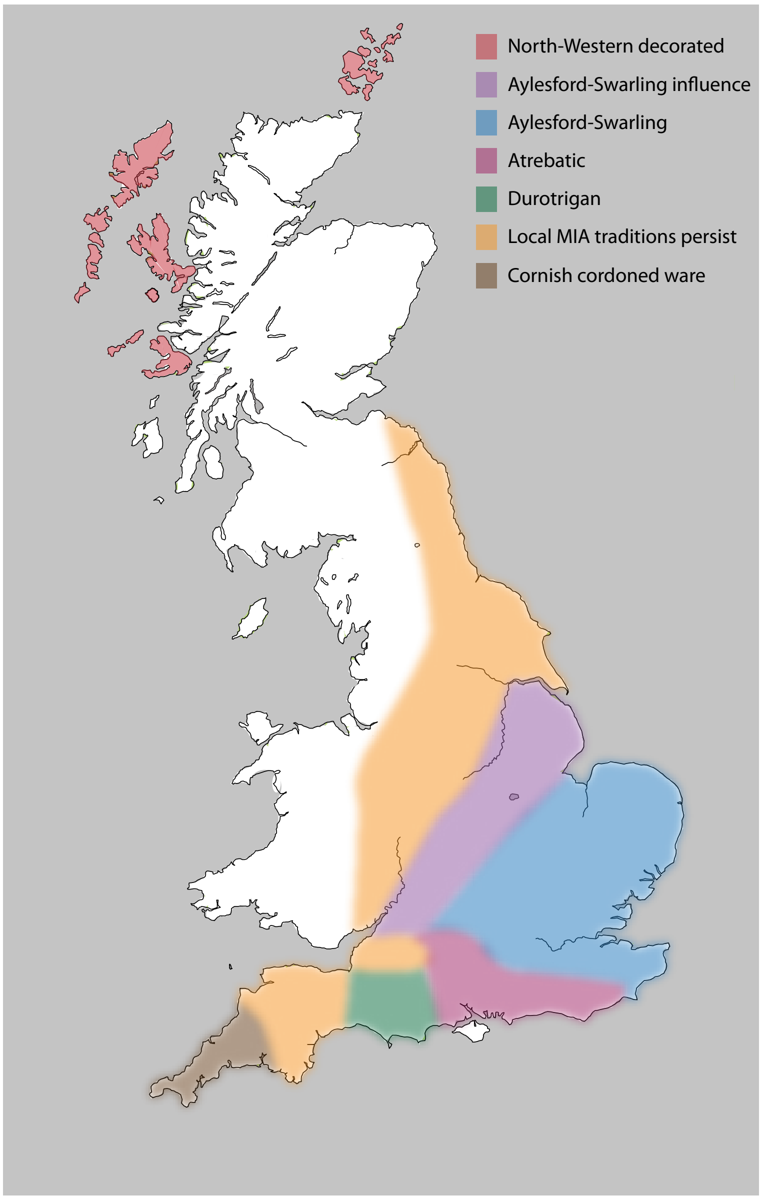


Figure 4.15: **Late Iron Age pottery groups** After Cunliffe 2005 [267].

4.1.2.6.1 Proto-history and the beginnings of Roman influence

In the last century BC, Mediterranean goods and ideas start to reach Britain in greater numbers, initially mainly through a western route via Armorica to ports in Dorset, but later and more voluminously through Belgic Gaul to the Thames Valley [267]. Agglomerated settlements and oppida – very dissimilar to those on the Continent – begin to develop in the south and east [382]. The north and west are largely divorced from these developments [267].

Britain also starts to enter history proper. While the first references to Britain date back to the 4th century BC, specifics of places and people appear only when it enters the Roman sphere [169]. After the Roman army under Caesar had conquered northern Gaul, the British lands across the Channel were seen as an extension of Belgic Gaul (the northernmost of the Gaulish provinces) [267]. By 55-54 BC, when Caesar launched an ultimately failed attempt to conquer southeast England, much was known of the area, including the names of tribal groupings and some of their political affiliations [267]. How far back in time such groups can be traced by archaeological means is disputed, with estimates ranging from vague foreshadowings for some polities in the Middle Iron Age to no time significantly before they are first recorded [267, 169, 382]. In any case, there is a general agreement that the tribal structures were comparatively fluid and, that by the late pre-Conquest period, a number of independent tribes had emerged (see Figure 4.16) [267, 169].

4.1.2.6.2 Contact and isolation

Within Britain, many different cultural zones may be drawn for the Iron Age, depending on the material culture category used, the chronological divisions employed, and the amount of detail considered. The settlement, metalwork, and pottery groups above show differing, if sometimes correlated, trends, and even northern Britain can be divided into eight zones [383]. For practical purposes, a coarse division for the entire period can be posited (see Figure 4.17 [267]):

1. Southwest zone – consisting of Cornwall, Devon, and west Wales

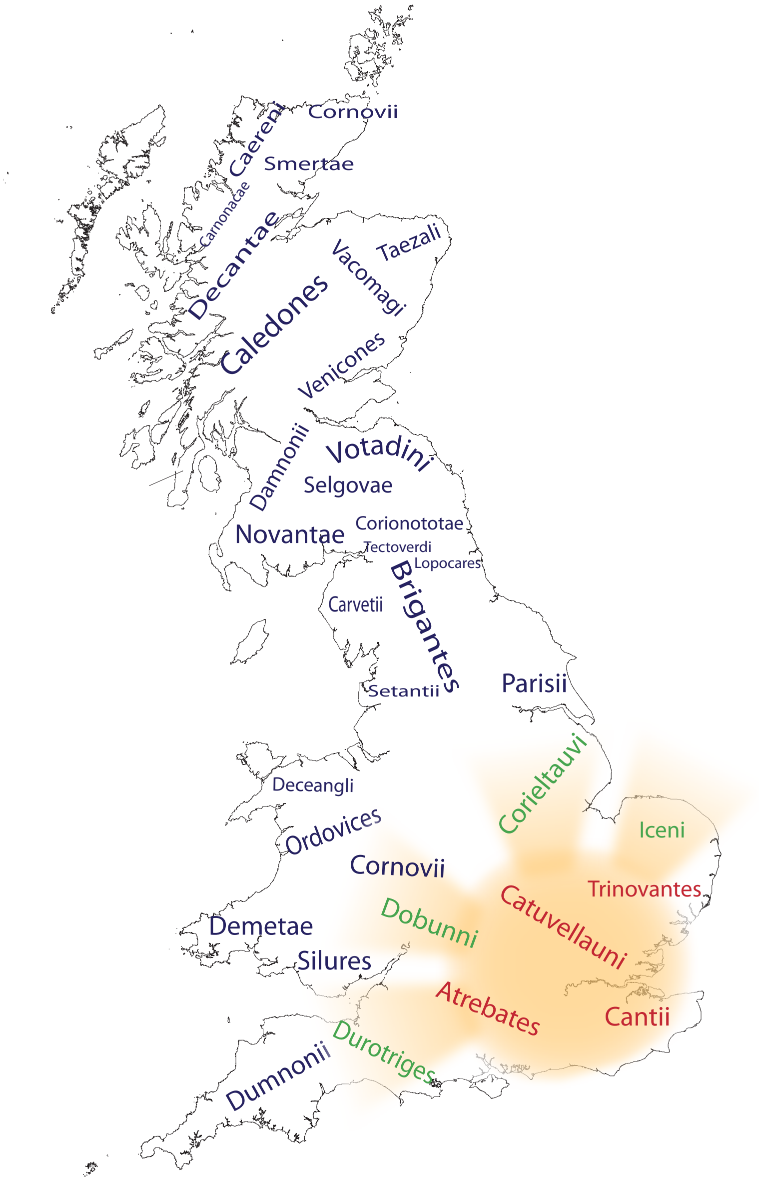


Figure 4.16: **Celtic Tribes** The yellow shading defines the core area of Continental influences. Tribes in red are those in closest contact with the Roman world, those in green have less, and those in blue the least or none. Adapted from Cunliffe 2005 [267].

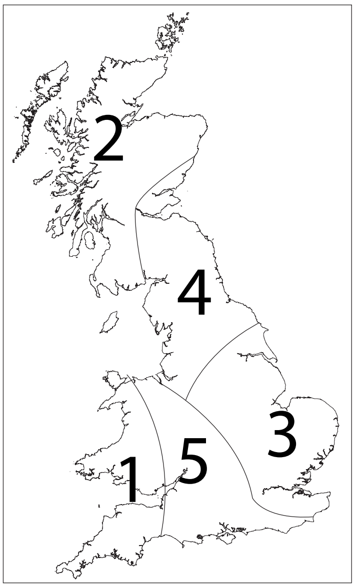


Figure 4.17: **Iron Age cultural zones** 1. southwest zone, 2. northwest zone, 3. east zone, 4. northeast zone, 5. south-central zone. After Cunliffe 2005 [267].

2. Northwest zone – consisting of most of Scotland, including the Northern and Western Isles
3. East zone – from Kent to Yorkshire
4. Northeast zone – from Yorkshire to the Highlands border
5. South-central zone – from the South Channel coast up to the Welsh marches and northeast Wales

Within this scheme, lowland England (approximately the East and South-central zones) was culturally quite homogenous [267, 376], indicating a degree of community interaction. The Southwest zone seems to have had limited contact with the rest of Britain, with the possible exception of coastal trade along Cornwall and Devon into Wessex. There is some evidence of maritime contact within Atlantic Scotland, but the utilitarian nature of its material culture makes seeing any exchanges difficult [376]. Within the north, the Forth is also a consistent cultural boundary [381]. Southwest England and southwest Wales in particular, to the exclusion of north Wales, share settlement similarities [376]. The British west coast as a whole shares many features internally, but there was likely little direct interregional contact [376].

Contact between Britain and the Continent and Ireland waxed and waned over the course of the the pre-Roman Iron Age, but was never completely absent. In broad strokes, Ireland was comparatively isolated, but what links it did have to the wider world were primarily to and through Britain. Occasional imported items, and Irish Hallstatt C and La Tène goods, connect Ireland to south and east Britain, only exceptionally exhibiting direct Continental origins [266]. These links do not appear to have been particularly strong, and were limited to the north and east of Ireland. Promontory forts such as Dún Aonghasa on Inis Mór may be indicative of Atlantic connections to south Wales, southwest England and Armorica [266, 267]. Some also see the settlement and high-status site trends of Britain and Ireland to be part of the same trend, but this is not generally accepted [382, 384].

The north and west of Britain show very little direct or even indirect contact with the Continent. The British northeast has a number of imported northern French torcs and eastern Scotland shows some similarities to southern England and northern France, but the contact – with any region (the Continent, Ireland or the rest of Britain) – appears to have been insubstantial [381, 385]. North Wales was perhaps even more isolated [267].

South Wales and southwest England differ by having persistent links with Armorica [378, 376]. Southwest England and Armorica have particularly strong connections, and show similarities in burials, some settlements and especially pottery; both direct imports and shared cultural trajectories [267, 376, 386]. Both Cornwall and Armorica also had souterrains, unlike their neighbours, but it has been suggested that this construction is too different to warrant considering them part of the same tradition [373, 376].

Finally, southeast Britain, as in the Bronze Age, had intermittently intense contact with the Continent. This is most notable between the southeast coast/Thames estuary area and northern France, just across the Channel [163, 375, 387, 267]. The archaeologies of the Thames estuary and Upper Normandy/Picardy are even considered "very similar – sometimes identical" [275] in the later Iron Age. Similarities between southeast Britain and northern France and sometimes the southern Low Countries range from direct imports (surviving examples being metalwork, pottery, and latterly coins) and domestic copying of these, to shared developments at varying scales including architecture (northern French and southeast England both display ‘intermediate’ structures between typical Insular and Continental fashions), pottery and metalwork developments, burial traditions and cemetery sitings, and ritual activities associated with storage pits [375, 387, 388, 389, 373, 390, 267, 274, 391, 281]. British exports exist, as stray finds, as far away as west central France, Denmark, Germany and Poland [281, 392]. The amount of imported material in southeast England was diachronically inconstant, and if the strength of cross-Channel links are assessed by this metric, they diminished steadily after the Late Bronze Age until ca. 450 BC when there was an ephemeral upsurge associated with the introduction of La Tène metalwork [267]. After this, there was a period of comparative isolation until the first century BC, when extensive trade commenced (see above) [375, 267]. In the late pre-Roman Iron Age, southeast England and northern France had close diplomatic ties as well; Gaulish chieftains had dominions in England and southeast England likely sent mercenaries to serve across the Channel [267].

There must at this time have been some exchange of individuals associated with the stronger and direct connections, both within Britain, and between the island and its neighbours. Known possible influxes sizeable enough to be considered on the scale of group rather than individual migration are limited to two: The Arras culture of east Yorkshire, and a Gallo-Belgic immigration event somewhere in the southeast. The former, originating ca. 400 BC, is most commonly acknowledged to involved substantial movement of people [163, 267], though opinions exist which favour wholesale native adoption of foreign burial rites [375, 393]. While there are some native British elements in the Arras complex (see above), and contemporary houses in the area are round, the similarity of these Yorkshire La Tène burials to those of northern France are so extensive that population movement seems likely [163, 267].[15] A further indication that adoption may have been insufficient to account for all the similarities is the presence,

---

[15] It must be stressed that the difference of opinion in this case is solely a matter of interpretation.

in east Yorkshire, of a tribe called ‘Parisi’ in late Antiquity [267] – the coincidence of both a northeastern French archaeology and an ethnonym is suggestive. The second relates to an incursion of Belgae, also from northern France, documented by Caesar to have happened some time before his expedition to Britain. This mention has often been taken as a reference to the Aylesford-Swarling culture of the southeast. However, few Aylesford-Swarling sites can confidently be said to pre-date Casear, so another possibility is the Solent port area, where material culture indicates a moderately intense and geographically circumscribed introduction of Belgic items and traditions, compatible with a limited but significant population influx [267].

4.1.2.7 Roman Iron Age

The official inclusion of a large part of Britain into the Roman Empire started with the Claudian invasion of the southeast in 43 AD. Over the next half century the remit of Rome expanded west to the Irish sea and northwards to the Forth-Clyde line [394]. The ‘frontier’ at any given time was unstable, subject to the strength of local resistance and extra-insular military concerns – as a peripheral province, troops were often withdrawn from Britain because they were needed in other parts of the Empire, rather than because they were no longer needed in Britain [395, 169]. Despite the constant fluctutation of soldiery in Britain, in particular the exact extent of Imperial conquest in the north, Britain on the whole had a disproportionately large garrison for both the size and probable economic importance of the province (see Figure 4.18) [169]. It is likely that much of lowland Britain was rapidly and permanently[16] brought into the Roman sphere, thanks in part to the installation of client kings acquainted with Rome, and the easy terrain of the southeast. The conquest of Wales, irregularly pursued, was lengthy, lasting nearly three decades until the region was secured ca. 75 AD [394]. While more permanent forts never placed a northern border north of the Forth-Clyde line, campaigns, evidenced by classical texts and temporary marching camps for the army, extended up to Moray along Scotland’s east coast [396, 169]. However, for much of the

---

[16] That is, until the fifth century.

Roman Iron Age, the province of Britannia ended at the Tyne-Solway line, roughly the course of Hadrian's Wall (constructed in the eary second century) [396].

Sites in Roman Britain can be divided into three general categories: rural settlements, forts and military settlements, and towns and villa complexes. The rural communities were on the whole only lightly touched by Rome in terms of changes to their domestic arrangements, and quotidian and ritual habits [169, 397, 398, 399]. Vernacular architecture changed little from the Late Iron Age (local variants of enclosures, roundhouses, etc.), although an increase in rectangular buildings and the use of Roman elements such as tiles can be seen in the southeast [169]. Roman type artefacts, many imported from Gaul, can also be found in rural settlements, but not in great numbers [169, 398].

It cannot be assumed that the lives of the Britons were as undisturbed by the Roman conquest and subsequent occupation as the material culture they continued to produce. A large number (ca. 100,000-250,000) are estimated to have died in the resistance to the initial conquest and in later rebellions, having a demographic impact in a population estimated to have been between 2-3.6 million [169, 398]. The reorganization of power structures and increased ease of communication throughout much of the island thanks to Roman programs of road and bridge construction certainly created new opportunities, but for most of the rural population the coming of Rome seems not have had particularly profound effects on either their composition or their cultural identity [169, 399].

The second category encompasses towns and villas.[17] Both the town and villa definitions form a continuum from large and distinctly Roman to small and less influenced by Roman architecture [169, 399]. At the former extreme lie towns of which many are predecessors of current British cities. A two-fold division can be made between *coloniae* and *civitas* capitals. Coloniae were founded for military veterans, who were Roman citizens. The control of coloniae and their hinterland was therefore more closely under

---

[17] In many syntheses, villas are treated as a rural settlement type, in keeping with their extramural locations. I have chosen to cover them in conjunction with towns in order to highlight the distinctive type of Romanitas that they share – usually different from both the military and rural groups – and the probability that many villa owners also possessed mansions in the towns [400, 169].

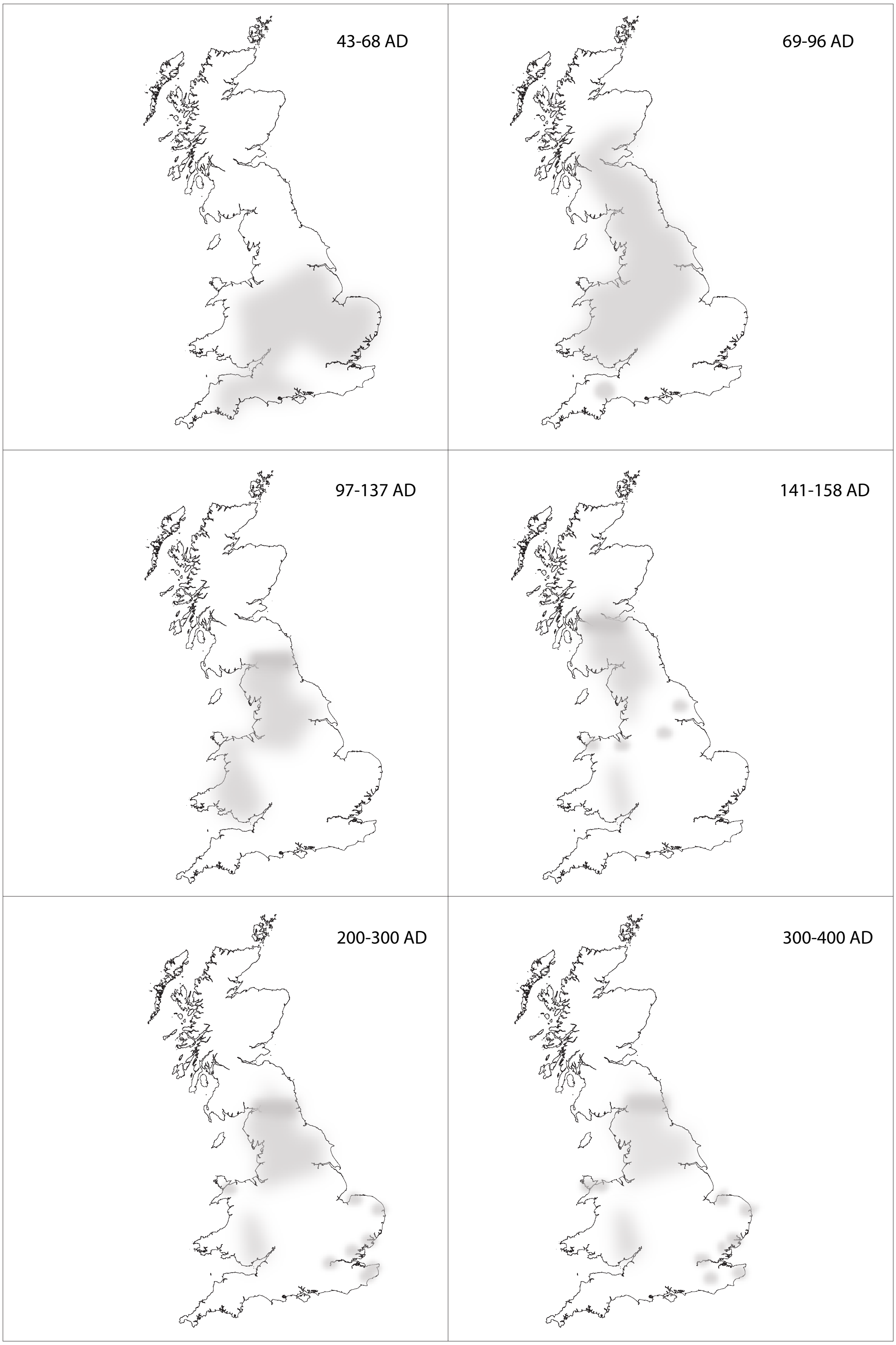


Figure 4.18: **Active garrison forts** Schematic approximation of the varying locations of the bulk of garrison troops in Britain. After Mattingly 2007 [169].

Roman administration [169]. This administration did not distinguish between military and civil affairs – the governor was both the civil head of the province and in charge of military operations [394]. Most civil servants were also part of the army. Gloucester, Lincoln and Colchester were coloniae, and later probably York and London as well. Civitas capitals were the new centres of 'tribal' power and administration, whose development was encouraged by Rome. They were allowed a measure of self-determination, for both the town itself and its surroundings, though both were ultimately answerable to Rome [169].

With coloniae (under Roman law) and civitas centres (following local procedure) controlling certain regions, and the active military garrison regions under martial control, the province of Britannia was a patchwork of different administrative practices and immediate hierarchies (see Figure 4.19) [169].

Coloniae and civitas centres both contained elements crucial to a Roman urban landscape: a more or less quadrilateral street plan, evidence of public services for provision of water and refuse collection, and public buildings and spaces (fora, theatres and temples) [400, 169]. Larger towns like these were a focus for immigrants. The coloniae housed veterans and their families hailing from all corners of the Empire (e.g. North Africa, Syria, and Danubian territories, as well as Gaul, the province of Germania, and Italy), and both large town types attracted merchants to supply the urban markets, primarily from northern Gaul [401, 169]. The presence of foreigners is attested by another feature of urban (and military) life: the practice of erecting inscribed commemorative stones, often detailing both the dedicator and dedicant's names, origin and position in society. Most such monuments are erected by non-natives, and when natives are mentioned, it is often as a wife [401, 169]. In addition to military veterans and foreign traders, towns would have housed native élites and craftsmen, who are less visible due to their failure to adopt the tradition of memorial inscriptions [169].

Towns in Britain are by and large limited to the Lowland Zone and cluster by Roman roads [398, 396]. Villa distribution echoes this, in that the majority of villas are in the hinterlands of towns [398]. Location is not their only commonality. The villa

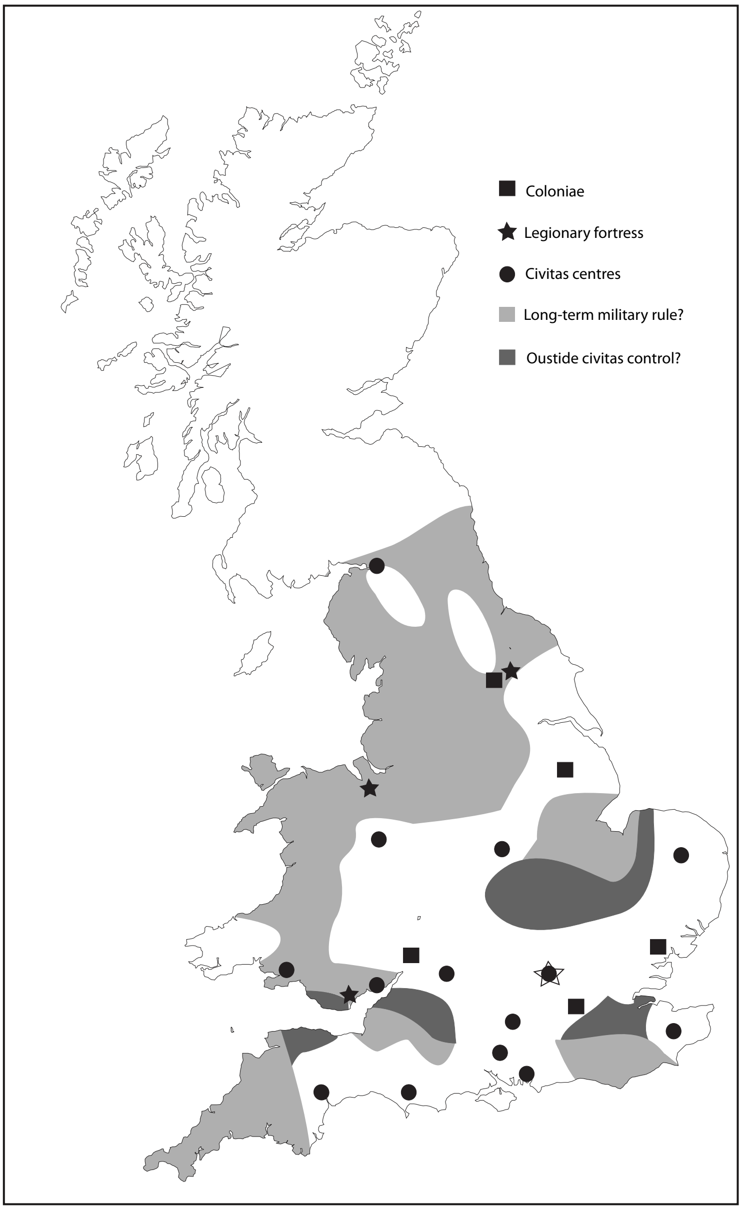


Figure 4.19: **Probable regional control in Roman Britain** After Mattingly 2007 [169].

estate, usually transformed from a native estate, is of Roman inspiration, despite both towns and villas in Britain often having features which set them apart from from their counterparts in other parts of the Empire [299, 399]. Towns and villas show signs of adopting at least some aspects of Roman daily life, with importation of Mediterranean style comestibles such as wine and olive oil, new vegetables and herbs, and novel cooking techniques [399, 398, 402]. The villas – some of them – also stand out for their obvious *Romanitas* in one respect: their mosaics. When representational, as opposed to abstract, in pattern, mosaics are Classical, almost without exception.[18] Villas are usually seen as belonging to the native landowning section of society, but there is also evidence for estates owned by army veterans who upon retirement abandoned much of their military material culture identity [169]. Rome conducted a detailed summary of the quantity and quality of land in Britannia, and the potential for purchasing estates is documented [169]

At the other end of the size/Romanitas scale are small towns and small 'villa' estates, which lack most of the obvious emulations of classical Roman lifestyles, such as street plans and public buildings for towns, and elaborate mosaics and classical layouts for villas [169]. These towns were likely simply local market centres, and the smaller estates home to more modest families, with a certain desire to exhibit 'sophistication' but lacking the wealth or knowledge seen in the more Romanized villa complexes [169]. Smaller settlements of this sort are in an intermediate position between large towns/villas and native rural sites, in much the same way as all towns/villas have a medial level of Romanization between the rural and military settlements. This is not to downplay the differences between a civilian and military experience in the Roman Empire, but in Britain, the garrison forts represent the most thoroughly non-native category of site.

At their peak, forces in Britain probably numbered 40,000-55,000 men [403]. The bulk of the military presence was based at forts, of which most active ones were in the

---

[18] While this indicates at least a certain level of innate or assumed Romanitas on the part of the commissioner, it is not known whether they would have chosen a specific mythological scene or picked it from a stock belonging to the craftsman.

west (Wales) and north (northern England/southern Scotland) after the first century (see Figure 4.18). Evidence for either a military or Romanized civilian presence west of Exeter in the south is very slight, but west Devon and Cornwall were probably under some form of military control [169].

The Roman army consisted of the classic legions, and auxiliary forces. Legionnaires had to be Roman citizens, and this meant that most were from the Italian peninsula until the relaxation of citizenship criteria laws in 212 AD [169]. Auxiliary units, on which Rome increasingly relied, were recruited from all over the Empire. Newly incorporated areas frequently provided the majority of auxiliaries, and they were as a matter of policy stationed in a different province from that of their origin [404]. From the third century AD, both non-Italian legionnaires and locally drawn auxiliaries became more common [403]. Throughout the three and a half centuries of Roman rule, soldiers from Gallia and Germania were the most prominent foreign ethnic presence [405, 169]. The extent to which the Roman army itself provided an ethnic identity is considerable. It imposed a certain uniformity of dress (both uniforms and off-duty clothing set them apart from British civilians), secular and religious habits (e.g. food type and preparation, and votive inscriptions to deities), language (a working knowledge of Latin would be indispensable, and inevitable), and nomenclature (many soldiers took Latin names or Latinized their names on entering the army, at least for the purposes of written records).

Settlements catering to the needs of the military, and taking advantage of its spending power, grew up next to active garrsion forts, and, in most cases, disappeared when the occupying units were redeployed [169]. Such garrison settlements included merchants, craftsmen, wives and families[19], and most probably brothels [169, 403]. The populace of garrison settlements seem to have been more closely identifiable with the army than with the Britons of the area [169]. Overall the army communities were in many ways self-contained and exclusive, though not self-sufficient enough to eliminate

---

[19] Before 197 AD, most ranks of soldier were not permitted to marry, so their conjugal arrangements prior to this date were not legal according to Roman law, though they were undoubtedly contracted unofficially [403].

any dealings with the natives.

The cultural impact on Britain of the Roman occupation probably far outstripped the demographic impact in terms of both immigrants to and emigrants/slaves exported from Britain, though the previously held view of blanket Romanization in the province is very unlikely (see above) [406, 169]. The negative demographic effect impact (emigrants, those dead in rebellions and their repercussions, slaves, and auxiliaries choosing to stay where they were stationed upon discharge) would have been patchy and probably focussed around garrison forts and disobedient tribal territories.

The positive demographic impact, while including individuals from most of Europe and the Mediterranean, would have been mostly from the adjacent provinces of northern Gaul and Germania, both as they provided a large percentage of the army in Britain and the former a large class of merchants settled in British towns. The presence of 40,000-50,000 military men especially, and their spouses (whether local and legal or not), would have produced a sizeable contribution to regional birth rates in garrison areas, so a non-negligable effect must be assumed even if quantification is difficult [403].

The ending of Roman Britain, in the early fifth century, followed an extended period of what can be seen as a decline in Classical Romanitas [301]. Britain was far from isolated from the wider events of the Roman Empire, and the processes that eventualy led to the devolution of the western part of it were if anything exacerbated in the British context. The whole socio-economic structure of roads, long-distance transportation of goods, urban communities, literacy and a monetary economy vanished almost completely when its linchpins – the military salaries, provisioning, and taxation – were removed [301].

Counterintuitively, western Britain, which was much less thoroughly Romanized than the east, was where remnants of classical education and habits seemed to survive the longest into what is commonly called the Dark Ages [169].

4.1.2.8 Anglo-Saxon England/Celtic Britain

4.1.2.8.1 The *Adventus Saxonum* The *Adventus Saxonum*, or the coming of the Saxons to Britain, though occurring in the historical period, has few reliable textual sources covering it. While several writers cover some aspect of the early presence of Germanic settlers in Britain –Ammianus Marcellinus, Zosimus, Procopius, Gildas, Bede, the Anglo-Saxon Chronicle – they all suffer from weaknesses, and none have the stated aim of recounting the event in detail. Contemporary writers were removed by distance from Britain, and largely recorded snippets of hearsay, later writers were distant in time from the migration perod, and Gildas – the only Briton and near-contemporary, probably mid-sixth century – was composing a moralistic tract to frighten his countrymen into mending their wicked ways [407, 408]. Bede, an early eighth century Northumbrian monk, did attempt to write a history of the English people, albeit a religious one (*Historia Ecclesiastica Gentis Anglorum*), but for the early period his main source was Gildas's *De Excidio Britanniae* itself [409, 407]. The textual sources must therefore be treated with extreme caution.

The Saxons were not unfamiliar to eastern Britons prior to the Migration Period. The ethnonym was probably used as a catch-all term for the Germanic sea-raiders operating along the North Sea coasts of the Empire, whose activities led to large shore forts being constructed in the third century [169]. These shore forts formed the military command of the *litus Saxonicum*, or Saxon Shore [409].

Gildas speaks of increased attacks from 'barbarians' on all sides, from Ireland and Pictland as well as the Continent [410], a not wholly unexpected consequence of the removal of imperial military might as a deterrent. The textual sources broadly paint a picture of Continental Germanic immigrants coming into eastern Britain in noteworthy numbers from around the second third of the fifth century, and the Gallic Chronicle states variously that Saxons devastated and took power of Britain around this time [411, 412]. The ethnic background of the incoming people is variously stated to include Saxons, Angles, Jutes, Danes, Frisians, and a number of other probaby Germanic tribes

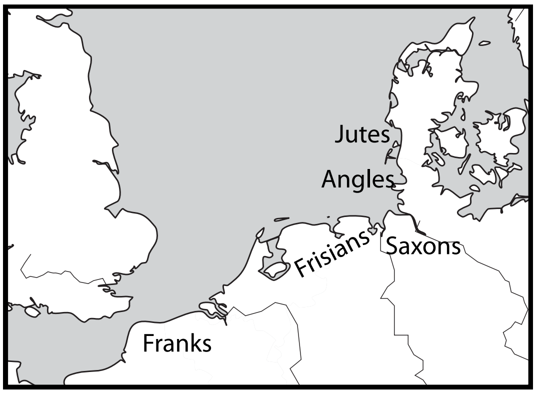


Figure 4.20: **Germanic emigrant homelands** Areas from which putative migrants to Britain are thought to have come in the fifth and sixth centuries.

[409, 412, 172]. The Saxons and Angles dominate accounts, as their names do in the later history of England, but it is important to remember the length of the period over which these names are invoked, and that ethnic units are far from static and impermeable [413]. Overall, the area in which one might *a priori* look for the archaeological traces of the Germanic ancestors of the English would be from the North Sea coast of Holland up to Schleswig-Holstein, and possibly further into Denmark [172] (see Figure 4.20). The Saxons, in the Holstein area in the second century BC, were already spreading south in the third century to between the Elbe and Weser [409], and possibly also settled the sparsely populated Frisia in the fourth and fifth centuries [414].

The limitations of the written evidence are obvious, and while useful for some generics, tend to focus on episodes of violence and at least semi-legendary origin stories, and, whether intentionally or not, usually employ categories of ethnicity that simplify reality [413]. The broader picture, dealing with actual low-status migrants, requires more from archaeology.

In many respects, the literary evidence is borne out by that from archaeology: The earliest 'Anglo-Saxon' material culture – that is, habits and artefact styles originating in Continental Germanic contexts – in Britain does indeed seem to date from around the mid-fifth century, or slightly earlier [414, 172]. Many objects can also be seen to hark back to the specific area of the North Sea coastal region [172]. However, the straight tripartite division described by Bede is impossible to take at face value. His dissertation on his contemporary situation claimed that the East Angles, Middle Angles, Mercians, and Northumbrians were Angles, that the people of Kent and the Isle of Wight with its opposing coast were Jutes, and that the East, West, South and Middle Saxons were, predictably, Saxon [409]. Such clear distinctions are not found in the material culture, especially not in the early Anglo-Saxon migration phase [172, 414].

The majority of Migration Period Anglo-Saxon sites date to 470-520, and are found in eastern Britain (see Figure 4.21) [172]. Most Anglo-Saxon archaeology derives from graves, in contrast to the settlement- and structure-dominated archaeology of the preceding Romano-British period [170]. Two types of burial rites were brought over from the Continent; cremation – this was a reintroduction, as it had gone out of fashion in the later Roman period – and inhumations with grave goods, which had likely declined in Britain with the ascendance of Christianity [172, 415]. The specifics of both of these rites had clear northwest European antecedents [171]. Inhumation without grave goods was also common, as it was in Lowland Zone native British contexts [415].[20] The overall burial traditions nevertheless lack uniformity, and do not bespeak a unified or even regionally subdivided culture coming into claim territories in Britain [416]. Within this scheme, however, cremation cemeteries, which were more common in East Anglia and East Yorkshire, showed a greater degree of uniformity than did the inhumation cemeteries of the southeast [412]. Overall, the artefacts and rites indicate heavy Germanic influence and some definite migration from the Continent, but that the migrants were not a homogenous cultural group - or even three different, homogenous groups [412, 172, 416].

Imported goods and artefact styles evident in early Anglo-Saxon graves point to extensive contact between southeast Britain and Old Saxony, Denmark, and Frisia, but also Thuringia in central Germany and Frankish Gaul [412, 414, 172]. Kent had especially strong links with Frankia, which continued well beyond the Migration Pe-

---

[20] Inhumations without grave goods may, of course, in many cases have had organic grave goods, which only exceptionally survive [171].

Figure 4.21: **Extent of early Anglo-Saxon cemeteries** Within this area, cemeteries are clustered, with notable concentrations in Kent and East Anglia. After Hines 2003 [417].

riod [418]. Among the commonest imports was Frankish pottery, which, along with Anglo-Saxon pottery of Continental Germanic origins almost wholly replaced residual Romano-British domestic ceramics [417].

The fifth century also saw the introduction of a distinctly northwest European structure, the *grubenhaus*, a sunken-floored building type with no obvious antecedents in Britain [170]. Curiously, the commonest house type in the putative Germanic homelands, the longhouse, does not appear to have been used in an insular context [170]. The only other domestic buildings known from Anglo-Saxon contexts are smaller rectangular timber structures, which may owe as much to the Roman-influenced local architecture as it did to Germanic building traditions [301, 170, 414]. One reason for the lack of longhouses may be that one of their functions, stalling animals at one end, was not necessary in the milder climate of southeast Britain [414].

In terms of settlement distribution, the early Anglo-Saxon nucleated settlements clustered around Roman roads and old Roman settlements, and in some few cases estates from the Roman period seemed to pass intact from post-Roman British to Anglo-Saxon material expressions [301, 170] . In addition to this, there are some indications that the earliest Anglo-Saxon farmsteads occupied marginal agricultural areas [301, 172].

Neither the graves of early Germanic settlers and their immediate descendants nor the early settlements show signs that they belonged to a hierarchical society [419]. Most archaeologists of the period would contend rather that they indicate that the settlers were peasant farmers first and foremost [420, 170, 171, 172]. Weapons found in some graves may have been largely symbolic, and when used they were in the hands of opportunistic, rather than professional, warriors [421, 417].

Only in the later sixth century did social hierarchy start to become apparent in the archaeology, and regional identities start to coalesce. The transformation of what had seemed to be a largely egalitarian society is evidenced by the increasing richness of grave goods in a small subset of burials, and the use of burial mounds for certain individuals (some purpose built, some re-used Bronze Age barrows) [415].

Much regionality is defined on the basis of female costumes, as evidenced by the attire they were buried in [415]. What survived was usually metal, amber and glass, and so represent necklaces, and brooches and other clothes fasteners [172]. There were a multitude of brooch shapes, most with clear northern European origins, and wearer configurations of these, which varied according to region. A western Norwegian tradition of wrist clasps in female dress was introduced into Norfolk and Humberside in the late fifth century, and has been used to argue for some very early Norwegian immigrants to Britain [172, 422].[21]

By the late sixth/early seventh centuries the first recognizable Anglo-Saxon kingdoms have emerged, and the broad understandings of Anglian, Saxon, and Kentish/Jutish identities as related by Bede have coalesced [172]. The spread of Anglo-Saxon material culture elements and power structures with self-identified Angles and Saxons at their apex was slow in Britain compared to Germanic conquests in Continental Europe, but

---

[21] Others link these clasps to western Danish traditions [171].

during the course of the seventh century most of lowland Britain and Devon were subject to some form of Anglo-Saxon law (see Figure 4.22) [304, 423]. Anglian influence from Bernicia is also seen in southeast Scotland by this time, evidenced by portable finds as well as *grubenhäuser* [424].

4.1.2.8.2 Native British survival and Anglo-Saxon immigrants Nick Higham's statement about Northumbria can be paraphrased to apply to Anglo-Saxon England as a whole: British survival in Anglo-Saxon England "is not in question; what is, is the nature and extent of that survival" [425] (p.267). There is no doubt that material culture, burial rites and building traditions of Continental Germanic origin made a significant impact on the British landscape in the mid-first millennium. The demographic impact of the Anglo-Saxon migration period, on the other hand, has been a subject of even greater dispute than that of the Neolithic introduction [426]. Following observed traditions (see above), the early students of the period posited a near-complete genocide and/or expulsion of natives in face of a Germanic war machine, and found support of this in Gildas's polemic.[22] The later academic reaction was to nearly do away with any notions of fifth-sixth century immigration. While the idea that the Britons where thoroughly expunged from what became England is now held by few, the number of Germanic immigrants – both in absolute terms and relative to that of the native population – is still hotly debated.

The data used to assess the level of British survival in Anglo-Saxon territories can be divided into archaeological and historical data, and linguistic evidence (see Chapter 4.2), which can be derived from both archaeological and historical sources.

Attempting to find Britons in the context of Anglo-Saxon landscapes is difficult, as, since the decline of Roman Imperial power in the province, they became an 'invisible people'. There is overall very little dateable evidence in the post-Roman period [301, 427]. With the departure of Rome, and the breakdown of the economic systems it

---

[22] This work, as we have seen above, portrays the Anglo-Saxons as an instrument of God sent to punish amoral Britons, and as such is likely to focus rather heavily on violent aspects of British/Anglo-Saxon interactions.

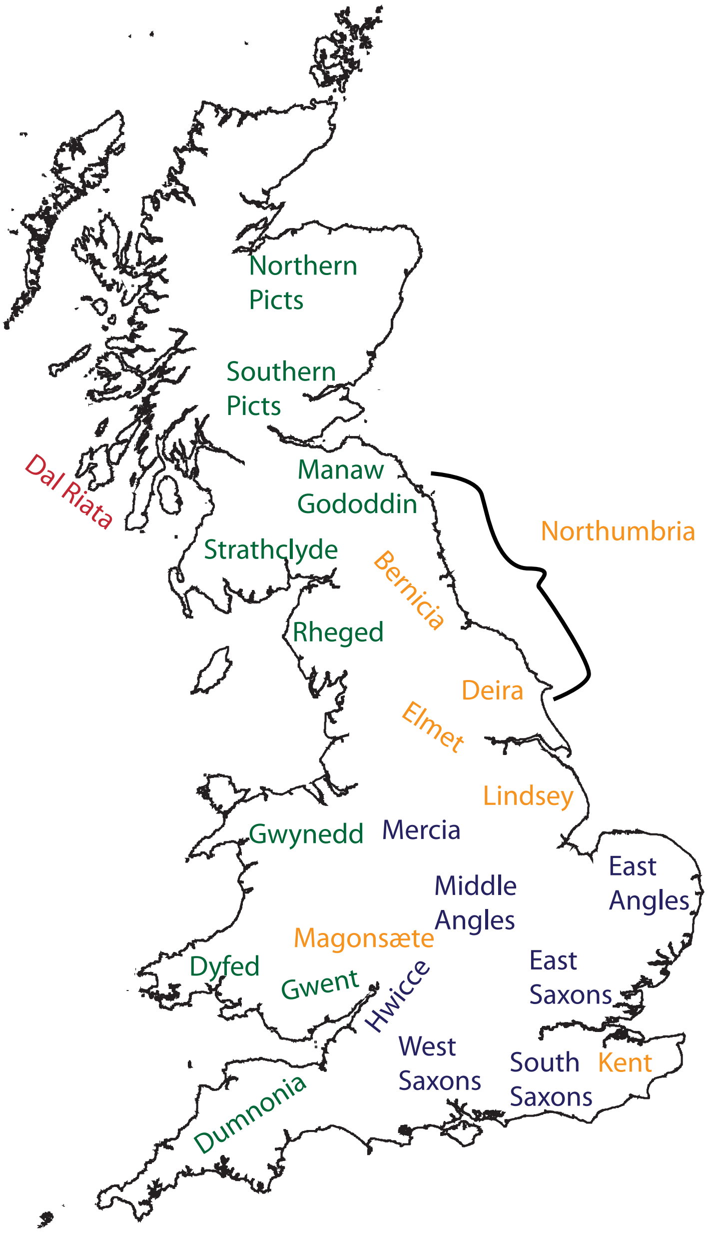


Figure 4.22: **Kingdoms in early Anglo-Saxon Britain** Anglo-Saxon kingdoms with Germanic names are in blue, Anglo-Saxon kingdoms with British or hybrid British-Germanic names in yellow, British kingdoms in green, and Irish/Goidelic in red. Locations after Hunter Blair 1977 [409].

engendered, virtually all of the diagnostic, durable material culture (stone and plaster buildings, coinage, commercial pottery, commemorative sculpture) went with it [301]. There is no evidence of a major population crash or crisis, but the departure of Roman high culture from an area where it had replaced local Iron Age high culture seems to have left almost nothing with which to fill the gap, resulting in the production of quotidian, generic, and non-durable material culture by the vast majority of the populace [301]. There are nevertheless traces of native British within the nascent Anglo-Saxon hegemony, and part of the problem with finding them has been that, as they were expected to be lacking, such traces were often ignored.

Burials provide some of this evidence. The heterogeneity of inhumations and grave goods has been remarked upon, and is key to this discussion. In among 'Anglo-Saxon' graves are inhumations that are more British in character, graves which contain Romano-British artefacts – in some cases treated with 'real' Romanitas, and in others assimilated into a Germanic setting – and one also finds Anglo-Saxon brooches in possible British women's graves, worn in a non-Germanic configuration [172]. It has been suggested that virtually all burials lacking in grave goods may represent British occpancy, but this ignores potential organic grave goods [171], and also assumes that British people would be unusually resistant to re-instating burial with personal objects. In short, there is such variety in burial practices, and the suites and configurations of goods, that there is good reason to believe some of the influence ran from Britons to Anglo-Saxons, and that there was also a good deal of acculturation of Britons to Anglo-Saxon ways in this aspect of life [428].

Much has been made of attempting to discriminate Anglo-Saxons from Britons based on height, but the associated problem of differential diets and their impact on this metric trait make the conclusion drawn very dubious [414]. Genetically determined tooth morphology, which is without known environmental influence, indicates that proximate Anglo-Saxon and British 'cultural' burials often contained biologically related individuals [414]

There also remain pockets where Anglo-Saxon material culture does not appear until the early 7th century, though it is found in the surrounding landscape. These include strongholds of civilian Romano-British power such as Lincoln, St. Albans and Canterbury [423]. The appearance of Anglo-Saxon goods such as pottery in otherwise non-Germanic settlements and the brooches in burials of British appearance may simply be due commercial pressures – 'native' pottery and jewelry industries having largely devolved in the early fifth century [301]. The native craft traditions did not totally disappear, as some metalworking patterns and techniques continued long enough to be absorbed into the emerging British English – as opposed to Continental Germanic – identity [423].

Southwest Britain (excluding Cornwall in the early period), western Mercia and Northumbria maintained a higher degree of British culture than the southeast, despite being under Anglo-Saxon rule. In addition to positive evidence of British traditions continuing (artistic, technological, architectural), grave goods go out of use in Anglo-Saxon contexts, so the burials lose their distinctiveness [420, 423].

The paucity of written sources for the migration period limits what can be inferred about Anglo-Saxon attitudes to Britons at this time, but the slow spread of Anglo-Saxon power north and west in literate times offers some interesting insights and potential parallels. The kingdoms of Wessex, Merica (having absorbed the kingdom of the Middle Angles and other neighbors), and Northumbria provide examples. The Law Code of Ine, king of Wessex, shows that Britons were present in Wessex, albeit in an inferior social position as evidenced by a lower *wergild*[23] attributed to them than to equivalently wealthy Saxons [429]. This law, which appears to reflect actual situations, naturally presupposes that Britons were readily distinguishable from Saxons, which is unlikely to have been the case based on innate physical features [420, 429]. Possible discriminatory characteristics could be material culture and traditions, but language is probably the clearest means of categorization [429]. Mercia, the later kingdom being an agglomeration of many polities, including ones along the Welsh border, appears to have had a substantial British component, though the overarching ethnicity was probably

---

[23] The 'man price' which the family of a murdered man might claim from his killer.

chosen to be Anglian on the basis of dynastic pressures [423]. Northumbria was created from the merger of two Anglo-Saxon kingdoms, Bernicia and Deira, in the early seventh century [423]. The names of both of these kingdoms are nevertheless Brittonic (see Chapter 4.2), and there is written evidence of extensive intermarriage between royal families of Northumbria and neighbouring British kingdoms to the north and west [420].

Within this picture of enroaching Anglo-Saxon people and culture, it must be remembered that political borders (both nominally British and Anglo-Saxon), alliances, and ethnicities fluctuated extensively in the early Medieval period, and that the concord between areas of Anglo-Saxon/British rule, self-identified ethnicities, and actual ancestry proportions was by no means complete.

The demographic impact of Germanic peoples, on the basis of archaeology and what can be gleaned from historical sources, can be assumed to be sizeable, though not catastrophic for the natives. In contrast to the Roman takeover, the Anglo-Saxon invasion may be seen as more insidious – the level of technology was similar, the level of social hierarchy similarly low [301, 304], and, without obvious military advantages, the archaeological impact must be seen as more due to numbers of incomers *contra* the Roman situation. This does not imply that said numbers must be overwhelming, and a reasonable hypothesis appears to be that a perception of social advantage in belonging to the Germanic section of society developed, causing a gradual cultural and lingustic shift (see Chapter 4.2) among the Britons [304].

4.1.2.8.3 Celtic Britain in the Anglo-Saxon period Celtic Britain, in this context, refers to the those parts of the island not subject to Germanic-speaking ruling dynasties, and roughly corresponds to the Highland Zone – Wales, most of Scotland, and, until the eight/ninth centuries Devon and Cornwall. The western part of Britain was nevertheless not free from external influence, and the Irish impacted the west coast from Cornwall to Argyll [430]. In certain aspects the Irish in the west mirrored the Germans in the east: They had been known as raiders[24] in the Roman period, and

---

[24] In this case, at least, raiding was bi-directional across the Irish Sea [409].

shore forts like those on the North Sea shores of Britain and Gaul are known in Cardiff and Lancaster [409]. Their influence is noted in written sources and in inscriptions, and in at least one case a sizeable immigration is possible. Beyond this, the differences are many.

In Wales and the southeast, there is little archaeological evidence for an Irish presence, but references in both Irish and Welsh annals, records of trade, and similiarities of place names (see Chapter 4.2), link east-central Ireland to north Wales, and southeast Ireland to south Wales [431]. A specific type of grass-tempered pottery was once thought to be indicative of Irish in Cornwall, but is now more often seen as an indigenous development with only incidental similarities to Irish tradition. Inscriptions in ogham on commemorative stones are the most telling category of evidence [424]. Ogham was inspired by the Latin alphabet for use in an Irish setting, and may well have been developed in Irish-speaking parts of Wales (see Chapter 4.2 for a discussion of Irish persisting in Wales) [431]. These inscriptions are widespread in Wales and also Dumnonia (Cornwall and Devon), as they are in southern Ireland (see Figure 4.23) . Irish genealogies for ruling houses in Wales likewise speaks for some longterm influence, if possibly only at an élite level [431]. The monastic Christian church in Ireland also highlights contact across the Irish Sea, in that its roots are clearly Welsh [431].The Irish impact on Scotland had more longevity than that of southwest Britain, but the evidence is enigmatic. Written sources recount invasions by the *Scotti* – a word whose earliest use was in reference only to people from Ireland – to Argyll from the sixth century, in much the same way as the coming of the Anglo-Saxons is mythologized [432]. Apart from these sources, the only reason to posit an Irish invasion is language (see Chapter 4.2). There seems to be no archaeological trace of incoming Irish, whether in terms on settlement patterns, burial patterns or even a shift in portable object fashions [433, 424]. Whatever its origin, the kingdom of Dal Riada in Scotland, supposed to be an offshoot of the Irish Dal Riada from Antrim, was quite spectacularly successful in spreading its influence throughout the ninth century, to absorb the Pictish kingdom and form Alba [432].

Figure 4.23: **Ogham stones in southern Britain** The dots provide the locations of known ogham stones in southern Britain.

Material remnants of Irish incomers to Britain, then, is thin on the ground, compared with the wealth of novelties attending the arrival of the Anglo-Saxons in the east. Some of this may possibly be explained by a greater inherited similarity and longer familiarity between western Britons and the Irish, and so a more open disposition among the Irish to immediately adopt British material culture upon taking up residence. Nevertheless, Irish influence rests almost wholly on linguistic grounds, be they inscriptions, personal names or place names.

The material culture of Highland Britain continued to a great degree along the trajectories set in the Iron Age. Some areas of the erstwhile Empire had a veneer of lingering *Romanitas* in patches, but the nature of settlement, when it can be discovered, is not much changed from earlier centuries [424]. The majority of domesic structures were still round, though rectangular buildings became more common throughout the period, probably stemming from Romano-British traditions rather than Germanic ones [417, 424]. The rectangular timber halls in Scotland's southeast may also owe something to Anglian involvement [417, 424]. Some hillforts in England were re-occupied, and continued to be so into a time when Anglo-Saxon artefacts were available to the inhabitants, and some hillforts in southern Scotland were probably never abandoned [424]. The round stone-built houses of Atlantic Scotland underwent evolution of form and elaboration, but the basic principles remained the same, as was the case for its pottery traditions [424]. Anglo-Saxon material culture made itself felt in northern and western Britain only peripherally [424].

4.1.2.8.4 Anglo-Saxon demographic impact The end result of the Anglo-Saxon invasion of England can be compared to situations on the Continent, especially that of northern Gaul. Northern Gaul and southeast Britain shared many features; each were rather distant Roman provinces, with Celtic-speaking natives, that were ‘overrun’ by Germans in the fifth century. In both, German tradition, such as intrusive burial rites, are evident [299]. The differences are equally striking: In contrast to the dominant pattern in England, the Franks in northern Gaul adopted Gallic customs in addition to contributing their own. They likewise adopted some of what remained of Gallo-Roman civil institutions and the Gallo-Roman church [299]. Perhaps most importantly, they shifted to the Romance language and used mostly native place-names, while retaining and bequeathing to the Gauls their ethnonym and personal names [434].

There is no pressing reason to suppose that the relative quantity of immigrant Germans in Gaul was materially smaller than in Britain, and the primary reasons for the different outcomes may probably be found in the background of both the invaders and the invaded, and their reverence, or lack of it, for Roman culture. As the German kingdoms of Europe became more Roman the closer to Rome they were situated (see Chapter 4.1.1.6), so the relative Romanitas of the Franks, the more northerly Germanic tribes such as Angles and Saxons, Britons and northern Gauls is likely important. The Franks’ long familiarity with Rome probably predisposed them to respect and be willing to adopt the remnants of Romanized ways and the common vulgar Latin vernacular of the Gauls in northern France [434]. In contrast, the probable majority of German immigrants to Britain were northern and had been buffered from intense contact with the Empire, and met a less thoroughly Romanized southeast Britain, where Roman élite ways had virtually disappeared. The British language probably held no attractions, and British culture no perceived superiority – the Angles and Saxons would have had little incitement to actively adopt either [434].

The quantitative demographic impact has variously been estimated to be between ten and two hundred thousand immigrants,[25] into a population numbering an estimated minimum of 800,000 to a maximum 4 million [434, 435]. The arguments for a massive influx has, in addition to Gildas's questionable testimony, usually focussed on the linguistic results of the invasion (see below). The material culture changes *could* have been introduced by rather few people, and be easily taken up by the Britons – this is not so much the case with the language.

The best explanation for the observed Anglo-Saxon legacy is most probably a scenario where there was a sizeable immigration into southeast Brition, followed by a spreading perception of German ethnicity being higher status than British,[26] leading to the gradual spread of self-identified Anglo-Saxon ethnicity.[27]

4.1.2.9 Viking Age

In the latter part of the eighth century, a new phase of Germanic raiding in Britain commenced, this time from Norway, and subsequently Denmark [436]. Written records mostly stress the violent nature of these Norsemen and Danes,[28] as they were usually written in monastic settings, and monasteries were concentrations of wealth targeted by raids [307].

In addition to raiding, the Scandinavians also engaged in widespread trading (see Chapter 4.1.1.7) and eventually settled some areas of Britain [311]. The nature of this

---

[25] While, as a mass, these estimates may sound large, they equate only to an average of one hundred to two thousand incomers per year over the century or so of the migration period.

[26] Possibly aided by social repercussions in Anglo-Saxon controlled territories, such as mentioned in Ine's law. Such ethnic differentiation in terms of social standing was paralleled in the *Lex Salica*, where the Franks were superior to the Gauls, so this may reflect a standard Germanic practice [429]. *Lex Salica* covered an area of northern Gaul where Frankish actually did replace Gallo-Romance [304], in another parallel to Britain.

[27] This was naturally in conjunction with successes in contest for land in the expansion of early kingdoms, which was an activity not monopolized by Anglo-Saxon potentates [409].

[28] These terms were often used interchangeably, without regard for actual geographical origin, and 'heathen' was also used synonymously [437].

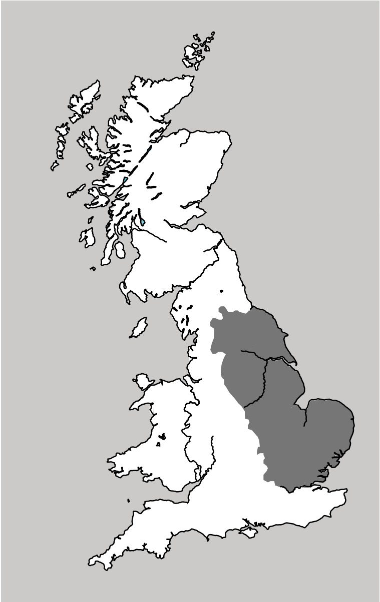


Figure 4.24: **The Danelaw** The approximate area of the Danelaw is shaded in grey.

settlement varied widely. In the wake of the 'Great Army' of, mostly, Danes in 873-874, about a third of English territory fell under the rule of Danes and their puppet kings, forming the *Danelaw* (see Figure 4.24) [311]. 'Norse Vikings', in the main, were from Norway, and had settlements in northern Scotland and the Western Isles, and northwest England. This sketch is naturally an oversimplification of the actual situation – the settlements in the northwest of England and western Scotland, for example, were probably that of Hiberno-Norse,[29] rather than Norwegians straight from Norway, and the extent of settlement in many cases is very uncertain [311, 438, 439].

Much evidence for Scandinavian settlement, in fact, is neither historical nor archaeological, but based on place names (see Chapter 4.2). The historical records refer to Danish partitioning and dominion of land, and occasional settling, but they are too patchy to give a comprehensive account [440, 311, 441]. By archaeological evidence alone, the question of an immigrant community would, for most areas, not even be mooted. Scandinavian-style bowed longhouses were not introduced in the Danelaw and very few Viking type burials with grave goods have been found there, most in

---

[29] People of mixed Irish/Gaelic and Norse Stock; the *Gall-Gaedhil* of Irish textual sources.

the north, and general suites of Scandinavian artefacts were not introduced in bulk [442, 443, 311, 437].

In northwest England several viking hoards have been found, containing the hack-silver traditionally used for currency, as have a number of pagan Scandinavian burials, and possible influence on domestic buildings [436, 444, 443, 437].[30] In northern England especially, one further material culture category speaks for Scandinavian influence, though in this case favouring hegemony rather than lower-status migration. This was the development of Anglo-Scandinavian religious sculpture, which created a tradition showing both Insular and Scandinavian artistic influences, and whose dual religious imagery could include both Christian scenes and ones from Norse mythology [437].

Northern Scotland and the Northern and Western Isles are the only areas of Britain that show very extensive material evidence for a Scandinavian presence. Norse long-houses were built here, ofte replacing native circular structures, a significant number of Norse burials (boat burials, female graves with typically Norse costumes, etc.) have been found here, as have soapstone objects originating in Norway [311, 445, 446]. Further signs of immigration include textile types found in Orkney and the Western Isles that are native to, and otherwise restricted to, western Norway, and a great increase in the consumption of fish in Orkney [447, 445]. Orkney fell under the Norwegian crown, and was only relinquished to Scotland in the late fifteenth century [448], so the official connections between Orkney and Norway were of much longer duration than those between the latter and other parts of Scotland.

The general lack of positive material evidence does not completely negate the possibility of some substantial immigration from Denmark and Norway into northen and eastern England and southern Scotland. The written and toponymic evidence makes the proposition worth taking seriously, and Viking Scandinavian settlement in other parts of Europe appears to show a willingness to conform to local tradition [438]. This is the case in the apparently frictionless adoption of Christianity by many Danes and

---

[30] These buildings' similarities to Scandinavian ones may be due to geographic determination [437].

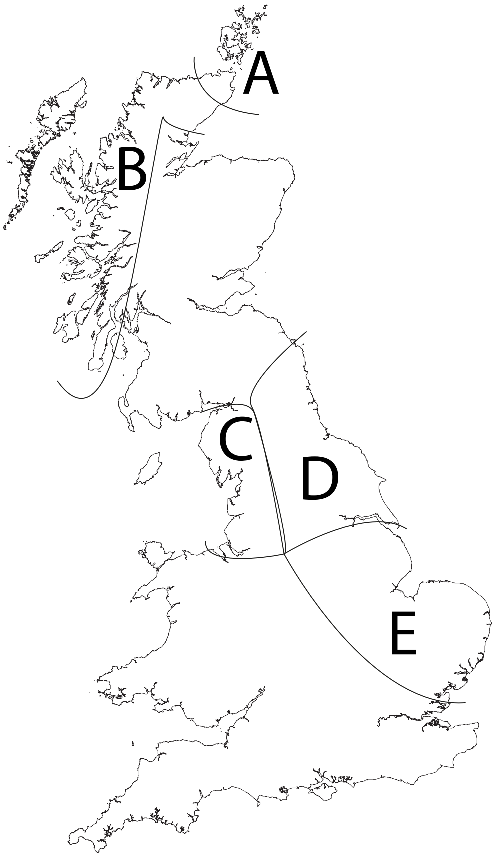


Figure 4.25: **Viking settlement zones** Lines define the areas of heaviest Viking settlement in Britain. A - Orkney and northernmost Scotland, B - northern Scotland and Western Isles, C - northwest England, D - northern Danelaw, E - southern Danelaw

Norsemen [449].[31]

Overall, the Viking presence in Britain appears to have been interstitial in most affected areas (see Figure 4.25). Excepting the Northern Isles, the most Scandinavian area in Britain was a belt from Danish Yorkshire across to the more Norse-Gaelic Cumbria, and the Scottish Viking west was largely a joint Norse-Gaelic endeavour as well. The combined evidence of archaeology, historical and toponymic studies indicates a sizeable presence of Norwegians and Danes in these areas.

---

[31] This was probably a matter of pragmatism for most – being Christian facilitated interaction on an even footing with local native élites. Also, there are hints that baptism was not necessarily taken seriously by first generation converts. There are accounts of Vikings undergoing repeated christenings in order to obtain gifts of money and clothes from the Church [307].

4.1.2.10 Later Medieval and Modern

The last exogenous invasion of Britain came in the form of the Norman Conquest in 1066, though it is questionable whether invasion is the correct term to use. William of Normandy came with as legitimate a claim to the throne of England as many others, and the previous contact between Normandy and England was probably more intense than some intra-Britain links [450]. Overall, the impact of the Norman Conquest was focussed heavily on the upper levels of the power structures in Britain - there was a replacement of the ruling class, who often retained properties in France and spent the majority of their time there, and the newcomers introduced new types of élite architecture, but the majority of the population was unaffected [309, 451, 452] . Nor were all those who took over English properties Normans – many Bretons and Flemmings supported William I, and were rewarded for their support with land, as did a handful of Poitevins, Aquitanians, Danes etc. [309, 453, 450]. In the wake of the greater lords who got estates in England came vassals of lower social orders, and more French merchants set up in English towns [450]. There is evidence in the Domesday book of settlers from a social class below land owners, but these do not appear to have been numerous enough to change the character of local populations [450]. After the first generation, many Norman lords could only properly said to be half Norman (or Breton, or Flemish), as a common, expedient and pacifcying way of transferring English estates to Continental owners was marryinng the daughters of the Anglo-Saxon incumbents [450].

In the later Medieval and early modern periods, written records are sufficiently detailed to render negligible any archaeological insight into intra- and extra-insular connectivity. Most movement continued to be on a local scale, and the bulk of 'foreigners' in a given locality usually originated from adjacent regions, whether said regions belonged to the same 'nation state' or not. Irish migrants are preferentially found in Wales and the west, Scottish migrants in northern England, and French, Dutch, Flemish in eastern and central England [454, 431, 455]. Overall, the proprtion of immigrants in the population in 15th century England may have been as high as 1%

[454], which provides a reasonable high-end estimate for Britain overall. There is an indubitable concentration of these migrants in urban areas, but they are not limited to them [454, 455].

4.2 Languages and linguistics

4.2.1 Languages spoken in Britain

In discussing the history of languages in Britain, the limit of study scarcely exceeds 2000 years, and for the first eight hundred or so years of this period the records are very sparse. Almost all the languages *known* to have been spoken are of the Indo-European family. The single possible exception, Pictish, is treated below.

The first known linguistic stratum consists of the Brittonic dialects, part of the so-called P-Celtic branch of the greater Celtic language family (see Figure 4.26 for a language tree). Brittonic was used throughout much of Britain, from Cornwall and Kent to the Forth-Clyde valley, and possibly even further north [167].[32] There are few, if any, contemporary inscriptions in Brittonic from this early period, and what little is known is derived almost solely from names in Latin or Greek sources, though some is from early native coinage [457]. However, it is certainly Celtic – anthroponyms, ethnonyms, and its descendants Welsh, Breton, and the extinct Cornish and Cumbric (which died out at the latest by the 19th and 12th century, respectively) testify to this [458, 459]. Tacitus also claims that the language of the Britons "differs but little" from that of the Gauls [4].[33]

When southern Britain was incorporated into the Roman Empire in the 1st century AD, Latin was introduced. Latin would have been used by the governing classes, and Vulgar Latin spoken by most of the the troops, whether they were legionaries or auxiliaries [167, 460]. There is some evidence that at least the upper echelons of Roman

---

[32] This is generally held to be the case, but Peter Schrijver prefers a distinction between 'highland' and 'lowland' British, and argues on phonological grounds that lowland British was more similar to Irish than to later highland British [168, 456].

[33] Modern linguistic research fortunately confirms the Gaulish-Brittonic connection, relieving the need to rely heavily on the judgement of Tacitus regarding comparative philology.

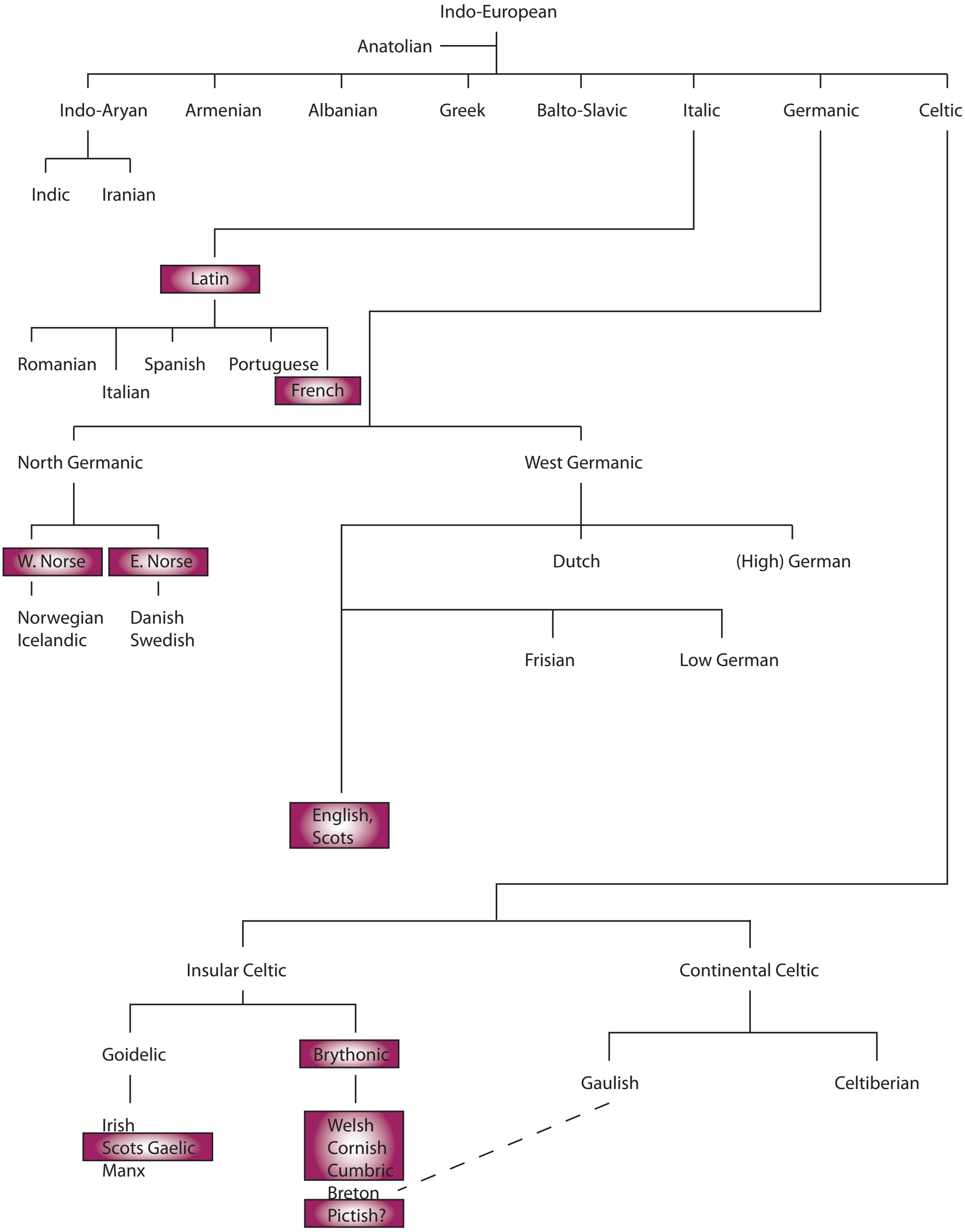


Figure 4.26: **Indo-European Languages** Approximate relationships of key living and dead Indo-European languages. Those known to have been spoken in Britain are in maroon.

Britain's native people used Latin regularly, if not exclusively, by the 5th century when the Roman Empire abandoned Britain [167, 461]. Latin certainly had a massive lexical impact on the surviving British languages, as exemplified by Welsh. The greatest impact was in the sphere of Christian religion, but less specialized words were also incorporated, e.g. *plant* being borrowed for Welsh 'child' [462].

In the late Roman period, a group called by the Romans *Picti* emerged north of of the Wall [417, 169]. It seems likely that the Picts included primarily the Caledoniae and Maetae [169], earlier known tribes of eastern Scotland. The nature of Pictish society, on the whole, is poorly known, and the interest in them is partially sustained by this fact. The evidence for a Pictish language is sparse and contradictory, and written only in ogham,[34] an Irish script, so Irish/Gaelic intrusions are common [19]. Some 'Pictish' is obviously Brittonic, other examples are probably not even Indo-European [19]. The most parsimonious interpretations of the inscriptions, Picitish personal names and literary references to who required interpreters to speak to Picts, is that there were in fact two different Pictish languages – one Brittonic and one not [19]. This dichotomy may well be mirrored by the bipartite nature of the Pictish kingdom and the tribes from which they arose.

The Goidelic branch of Celtic, including Irish, Scots Gaelic and the recently deceased Manx, is usually thought to have been an early Medieval introduction to western Scotland from Ireland (but see below, Chapter 4.2.2) [432]. Eventually, most of Scotland, except the lowlands and the Northern Isles, converted to Gaelic speech [432]. The similarities between Scots Gaelic and Irish are so great that no substantial period of isolation can be posited [463]. The speech communities in western Scotland and northern Ireland must have been in close contact throughout their existence, regardless of how long there have been Gaelic speakers in the former. The speech that Gaelic supplanted was, in addition to some possible small Norse/Hiberno-Norse enclaves, in the main Brittonic and Pictish, in so far as 'northern Pictish' may not be Brittonic [463, 445]. In a lesser

---

[34] There are very few exceptions in Latin script, and some Pictish names are recorded in non-Pictish sources.

way, Irish was also most likely spoken in a limited way in Irish enclaves in Wales from the fourth to sixth centuries, as evidenced by ogham inscriptions [167].

Several Germanic dialects were introduced into Britain with the Anglo-Saxon invasions.[35] English, the descendant of the introduced dialects, is a language of the western branch of Germanic, along with Dutch, Frisian, and Low and High German. It is further classed with Frisian and Low German in the Ingvaeonic or North Sea sub-branch. Early sources of the nascent English language are few, consisting mostly of very short runic inscriptions [464, 465]. Longer texts were composed or recorded in Anglo-Saxon monasteries and King Alfred (the Great) of Wessex initiated a program of translation into 'Ænglisc' as he despaired of his fellow countrymen's tenuous grasp of Latin [466]. Thanks to the latter, the late Old English corpus is quite abundant, but likewise, our impression of the speech of England at the time is heavily skewed towards the West Saxon dialect [467]. The Anglian dialect is less well attested, but it has a more generally northern character than Saxon. It appears to share some features with the North Germanic languages of Scandinavia, which is perhaps not surprising given the proximity of the historical Angeln to areas of proto-Norse speech [465].

Of the modern Germanic languages, Frisian is the one that appears closest to English, and this has been seen as confirmatory for a sizeable number of early Medieval Frisian immigrants to Britain [468, 469]. However, there are reasons to believe that the similarities between English and Frisian owe more to maintained Germanic archaisms (not indicative of a close genetic relationship) than to shared Anglo-Frisian innovations (the main indicator of relatedness) [470, 47].

The complete eradication of Celtic languages in most of England, having been interpreted as evidence for genocide, has undergone revision attempts in concert with the theories of acculturation: The search for effects of Celtic speech on the English language has been extensive, but so far it has also been largely inconclusive [471]. Few examples, be they Migration Period loanwords, syntactical transfers, or phonological traits, are

---

[35] Germanic languages would not have been completely unknown in Britain, as many Late Roman auxiliary units were drawn from Teutonic peoples within or on the fringes of the Empire (see Chapter 4.1.2.7).

generally accepted [471, 472]. There are several candidates (see e.g. Filppula [472] references therein), but overall there cannot have been a very large Celtic component in English.

A North Germanic language, Norse, was in all likelihood also spoken in Britain, as a result of immigrant Scandinavian communities in the Viking Age. Geographical and linguistic identities must not be confused: The term 'Norse vikings' is used to denote Viking Age Norwegians, in opposition to Danes and Swedes. The term Norse in a linguistic setting, however, refers to the Viking Age language which was common to all three Scandinavian countries and their dominions. Dialectal differences between Old Danish and Old Norwegian did exist [473], however, and this point will be important in the discussion of toponyms and the provenance of their coiners.

Dialectal Middle and Modern English reveal strong remnants of Norse in the Danelaw and northwest England, particularly in a belt across Britain from Yorkshire to Cumbria [474]. The most obvious effects are a large number of loanwords from Norse, and English words that took on the meaning of their Norse cognates, e.g *dream*, which in Old English meant *(sounds of) joy* [475]. In Orkney, there are even stronger Norse vestiges in the English spoken, as here, a Norse-derived language was the common vernacular until the early modern period, while in Britain, Norse was only spoken in pockets, and probably did not survive beyond the twelfth century [475].

Norman French was introduced with the new aristocracy after 1066, but remained an élite language, and after a few generations it ceased to be the native tongue even of the landed classes [475]. Its lasting effect on English is mostly limited to a few specific areas, largely of upper class concerns, and situations in which the French word would disambiguate a statement. A simple example of both situation are the 'meat' words (pork, beef, mutton), which all come from the French for the animal whose meat is referred to by the Germanic version – only the upper class could afford to eat meat on a regular basis, and the words served to differentiate the organism from the food it provided [476]. Furthermore, in support of the limited and élite nature of French influence, only a very small and circumscribed number of French personal names were introduced, compared with a wealth of both male and female Norse names during the Viking Age [436, 441].

4.2.2 Toponyms

Several toponymic strata are preserved in Britain. The oldest of these consists mostly of river names [477]. River names, especially those of large rivers, are among the most conservative name categories [17]. Many river names in Britain belong to the so-called 'Old European' family of hydronyms, a suite of appelations which recur throughout much of Europe and into West and South Asia [477]. The linguistic affiliation of these names is thought to be either Indo-European or an Indo-European/non-Indo-European hybrid, but to belong to an older level of the language family than any recorded [477, 478]. This stratum is possibly due to older forms that remained productive in the specific function of naming rivers, having lost their standard lexical values [477]. Nothing detailed, therefore, can be said of the pre-Celtic languages of Britain, beyond a probable Indo-European affiliation, if the Old European hydronyms (e.g. Derwent, Thames, and related names) were more than fossils.

The second stratum includes, for most of Britain, both settlement and river names of an identifiably Brittonic form. Many of them are recorded in Ptolemy's first century AD information [479]. A cluster of probable exceptions to the general Brittonic picture of pre-Anglo-Saxon Britain exists in eastern Scotland; these names have no obvious Indo-European affiliation at all [480]. Indeed, data from Ptolemy's *Geography* places southern Britain especially firmly in the heartland of Celtic-looking toponyms (see Figure 4.27).

Later layers of toponyms are testament to Medieval incursions from non-Brittonic speaking lands. Indeed, for some, they are prime evidence for the Irish impact on western Britain in late Roman/post-Roman period as it is not easy to detect in the archaeology [424, 431]. Irish place-names are found in both Wales, especially the southwest, and Cornwall (e.g. *Llŷn*, from the same name as Leinster) [431]. Both Irish forms[36]

---

[36] As Irish and Welsh/Cornish are both Celtic, it is sometimes difficult to differentiate between

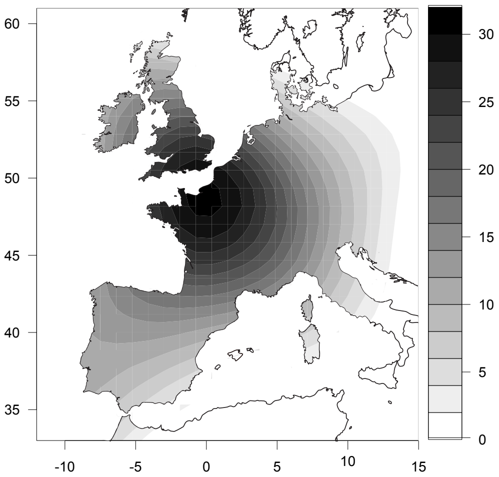


Figure 4.27: **Core of Celtic-looking placenames in Late Antiquity** Longitude and latitude are given by the x- and y-axes, respectively. Data from Sims-Williams 2005 [479].

and names referring to Irish people and places are known [431]. In western Scotland, the language and place names are the primary reasons for supposing a large Irish invasion into the area [432]. However, due to the lack of supporting material evidence, it has plausibly been suggested that western Scotland had long been part of the Goidelic language area, and that the apparent Brittonic character of earlier names recorded for its inhabitants was the result of Brittonic speaking informants transmitting said names [432]. In either case, the expansion of the Gaelic speakers' power in the ninth century resulted in the replacement of many British, and possibly non-British Pictish names in the rest of Scotland [477, 432, 424].

Place names in much of England are, overwhelmingly, English [481]. The spreads of Anglo-Saxon material culture and names show a good deal of congruity. The proportion of English place names decreases in the north and west, and increases over time in English-held territories [425, 420, 423]. The attrition of non-English names in favor of English names in areas where early Anglo-Saxon occupation was contemporary with written records implies that the prevalence of English place names in the earlier settled

---

nativized loanwords, cognates, and foreign-affected cognates.

areas too might have been a gradual process. The predominance of English place names was used, in concert with the appearance of Anglo-Saxon archaeology, to support the idea of a decimation of the British population. There are, on the contrary, a number of clues that indicate the continued presence of Britons. The first is that pre-Anglo-Saxon names were adopted by the Anglo-Saxons. These names were mostly those of significant landmarks, such as rivers – both 'Old European' and Celtic – and large Roman power centres such as London, Lincoln and York [481]. Kent, the site of some of the earliest Germanic incomers, likewise has a pre-Germanic name [423]. Some later Anglo-Saxon kingdoms even had Brittonic names, such as Bernicia and Deira, which later formed the majority of Northumbria [423]. All names incorporating the Latin elements *'eccles'*, *'camp'*, *'funta'*, *'port'* and *'vicus'* would have come via the British as well [481]. The second is the prevalence of names such as Walton, Wallingford, etc. The Wal(l)- element derives from *wealh*, whose meaning switched from 'foreigner' or 'Latin speaker', to mean British,[37] and latterly also 'slave' [482]. Names with this element have been taken to attest the tolerated presence of identifiably British communities among those who were hereditarily or adoptedly Germanic [481]. This is not to downplay the considerable impact the English language had on the landscape, but the fact that not all English names were first generation coinings must be kept in mind when attempting to infer the demographic impact needed to effect the observed changes.

The last great impact on the toponymic landscape was the Scandinavian settlement in Britain. The effect on place names was such that settlement is thought to be quite extensive, though there is rather little evidence of it in other spheres. Scandianvian place names are scattered throughout most areas of supposed Scandinavian settlement in Britain, but in the northern tip of Scotland and the Northern Isles over 95% of toponyms on record are Norwegian [446]. The different levels and types of Scandinavian impact are reflected in the distribution of Danish/Norwegian names and the toponymic milieu in which they exist (see Figure 4.28).

Most Norse names are compounds, but the elements most frequently used vary by

---

[37] Whence 'Welsh'.

region. In the Danelaw, particularly those areas highlighted in Figure 4.28, *-by* and *-thorpe* are most common [436]. Many of these are combined with Norse personal names [436, 440]. In other areas of Scandinavian settlement, by and large those with Norwegian immigrants, the repertoire of generics is larger, including *-holm*, *-nes*, *-sætr* and *-skál*, but the combination with personal names is less predominant [440, 314, 483].

While it has been pointed out that any generic element might have been taken up by the local non-Norse vernacular and that not all people with Norse names were Scandinavian by either birth or heritage, most Scandinavian place names were probably nevertheless coined by actual Norse speakers (immigrants or their immediate descendants). Many names contain properly declined Norse elements, and the declention paradigms would be unlikely to survive transferral into another, albeit closely related, language [441].

4.2.3 Languages and peoples

The different languages known to have been spoken in Britain can only strictly be informative on some aspect of their speakers for the last two thousand years – including Old European hydronyms one may perhaps add some centuries. Nevertheless, linguistic information is absent or highly speculative for most of the post-Younger Dryas time during which Britain has been inhabited. On this linguistic palimpsest, the oldest layers are completely eradicated, but the importance of languages in determining community identities in later periods reveals how much we are missing.

Whether an incoming language ultimately failed to take root (like, after several centuries, Latin, Irish in Wales, and Norse) or persisted (like English and various Celtic languages), the geographic patterns they left have a strong influence on how the history of their speakers is currently viewed.

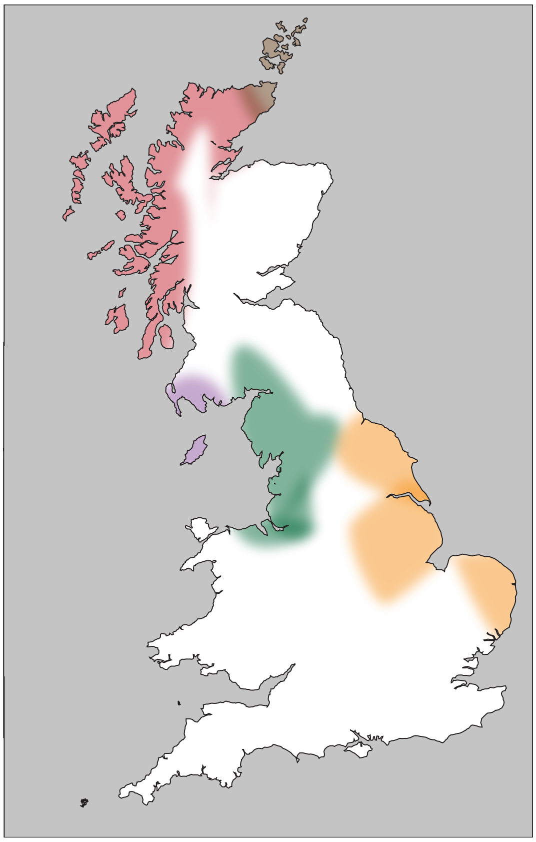


Figure 4.28: **Scandinavian place names** After Crawford 2003 [311]. Red – Norwegian-Gaelic, purple – Norwegian-Danish-Gaelic, green – Danish-Norwegian-Anglo-Saxon, yellow – Danish-Anglo-Saxon, brown – almost wholly Norwegian.

4.3 Population history and genetic data

4.3.1 Genetic patterning in Europe

A prerequisite of being able to interpret genetics in a historical framework is, clearly, knowledge of the genetic structure of populations in the area under consideration. Ideally, a diachronic perspective could also be introduced by the use of ancient DNA, but the issues of insufficient sample sizes, modern contamination, and difficulties of provenance[38] combine to limit the current usefulness of ancient samples.

The genetic patterning in Europe is among the best known in the world, in terms of both the resolution of sampling and genotyping, though the coverage within Europe is unequal. Blood groups and other classical markers have long been known to display frequency gradients across Europe, most strongly along a southeast-northwest axis [22]. Y chromosome haplogroups also show very distinct geographic patterns in Europe, e.g. the very high levels of R1b in Atlantic Europe, high levels of I central Europe, and higher levels of R1a in Scandinavia and the northeast [484, 485]. Mitochondrial haplogroups show much less patterning on a European scale, though some haplogroups are, like their Y chromosome counterparts, more typical for certain areas than others. Notably, haplogroups thought to be indigenous to Europe (U5, V, H1 and H3) typically have higher frequencies on the western fringe, and U4 is more typical of northeast Europe [485, 486].

Larger scale SNP-based studies, using various eigen-analysis approaches, have managed to broadly recapitulate nation-state, or even regional, geographical affiliation of European individuals solely on the basis of the genetic information, and show that there are two main axes of decreasing genetic diversity in Europe; primarily, southeast-northwest, and secondarily, southwest-northeast, so that the populations of the north generally contain only subsets of the variation present in the south (e.g. [487, 488, 489, 30, 490] .

---

[38] Provenance here meaning not necessarily the archaeological provenance, but where individuals and their immediate ancestors came from in life, and there are difficulties with trying to assess how representative these individuals were of their broader community.

4.3.2 The peopling of Europe

Ever since genetic patterning was discovered, it has been interpreted in a historical framework. Genetic data fairly conclusively support the palaeoanthropological hypothesis that modern humans developed in Africa, to which a minor amendement is the recent discovery that non-African human populations all have ca. 1-4% Neanderthal-like DNA [16, 491]. This implies a limited amount of hybridisation with Neanderthals among the first non-African peoples, probably in the Middle East. Beyond such basics of human development, many questions regarding the peopling of Europe are framed in terms of archaeology. From skeletal remains we know the length of anatomically modern humans' tenure of Europe, and from archaeology we have hypotheses regarding subsequent movements. Examples are studies seeking and/or finding evidence for post-glacial expansion from refugia, and particularly attempts to quantify the relative demographic impact of Middle/Near Eastern farmers during the spread of the Neolithic to and through Europe.

At a coarse-grained level, genetic data is not incompatible with two main refugia, in southern France/northern Iberia and the Ukraine/Balkans, with a potential, less important one in the Italian peninsula [486, 492, 493, 494, 495, 496, 485, 497]. Much prior work has focused specifically on the Franco-Cantabrian refuge as the the source of 'Palaeolithic' mitochondrial DNA haplogroups such as H1, H3, U5 and V, and Y haplogroup I1a, while the eastern refuge is invoked to explain the phylogeography of most of Y haplogroup I1b and possibly mtHG U4 [486, 485, 495]. The SNP based axis of decreasing diversity from southwest to north (see above) may also contribute to the conviction that the glacial refuge was a true source of post-glacial genetic material for northern Europe, as well as for material culture (see Chapter 4.1.1.1).

The question of the relative genetic contribution to modern Europeans of Palaeolithic Europeans and peoples from the Near East arriving in the context of the Neolithic is also much covered in the scientific literature, perhaps more than any other similar issue. Early on Cavalli-Sforza *et al.* suggested that their synthetic maps of classical gene frequencies in Europe might be showing the serial dilution of 'Neolithic' genetic material as one moves north and west [22].[39]

Attempts to apportion modern European mitochondrial and Y chromosome DNAs to ultimately 'Neolithic' and 'Palaeolithic' sources have resulted in wildly differing estimates, not only between but also among these uniparental markers. The methodology used is to assign different haplogroups to Palaeolithic European versus Neolithic Near Eastern categories, a procedure of which even some of its proponents admit the heuristic character [485].

The refinement of haplogrouping is of the utmost importance here, and many earlier studies do not reach fine levels of distinction – the common (ca. 10-15%) mitochondrial haplogroup J used to be seen as a 'Neolithic' marker, now the designation is thought to be limited, potentially, to the minor J2a1a subgroup [485]. Similar reassessments have also been undertaken for other haplogroups, causing estimates of Neolithic input to be reduced, in some cases drastically.

Age estimates of the spread of haplogroups and their subgroups are crucial to interpretation, and the uncertainty involved in dating by 'molecular clocks' is – though improving – one of the pitfalls of the broader approach [497, 499, 485, 500].

Estimates of 'Neolithic' ancestry, which rarely take into account probable pre-Neolithic genetic heterogeneity in Europe, cover a large range (15-80%), but some consensus on a smaller interval (15-25%) exists [501, 502, 503, 504, 123, 484, 485].

Ancient DNA (aDNA), in Europe, has been used primarily to address the Neolithic question. However, a number of problems currently limit its usefulness. Most aDNA studies, like many modern ones, have only addressed mitochondrial DNA.[40] More seriously, the number of samples assessed is inevitably very small, and often the sample selection means the sets are unsuitable for comparison (e.g. individuals from the same cemetery are used without controlling for close familial relationships,

---

[39] Though see Currat and Excoffier for frequency gradients being equally probable under cultural diffusion [498].

[40] The uniparental marker has an additional advantage in the context of aDNA, as there are multiple copies of the mitochondrial genome in each cell, which increases the chances of DNA recovery.

or grouping of 'Neolithic' people across many hundreds of years [505, 506]). Tentative results show not only diversity among geographically different Neolithic populations, but occasional discontinuities between Mesolithic, Neolithic, and modern groups [507, 508, 509, 506, 510, 511]. One analysis of central Europe suggests that even admixture between Mesolithic and Neolithic (Levantine-derived) groups alone is insufficient to explain most modern European variation [507].

Overall comparisons between Mesolithic, Neolithic and modern DNA, plagued by issues of suitable and adequate sampling and modelling of lineage (haplogroup) fluctuations, has yielded results that are either conflicting, or highlight the differing proportions of Neolithic demic and cultural influence in different areas of Europe [511].

4.3.3 Genetic patterning and the peopling of Britain

That there are genetic differences within Britain has long been known, first in the form of blood groups (e.g. [512]). Later, using a set of 14 classical markers including blood groups, Falsetti and Sokal [513] attempted to define genetically determined boundaries within the British Isles, and identified the following: North West Highlands of Scotland vs. Outer Hebrides, Northern Ireland vs. Scotland, Ireland vs. England, Isle of Man vs. the Midlands, Ireland vs. Wales, Cornwall, East Anglia, Yorkshire and Humberside vs. the East Midlands and East Anglia, Shetlands vs. Orkneys, Orkneys vs. the Scottish mainland, eastern edge of the Highlands, and Wales vs. England. In general, they also found tendencies towards broadly north-south frequency gradients.

Mitochondrial DNA is comparatively poorly elucidated at a fine geographical scale in Britain, but the overall northwestern European/Atlantic character of the island's mtDNA seems established [514, 515].

The best known genetic system for Britain is the Y chromosome. There is notable internal structure, mostly due to the differing proportions of the three best represented haplogroups (R1b, R1a and I). Broadly speaking, the western and northern extremities have higher proportions of R1b and are closer to the Irish in a principal components analysis, the Western Isles and Orkney especially have higher R1a levels and pull towards a PCA pole representing Norway, and eastern Britain has more instances of haplogroup I which pull this area toward a Danish/German component of a PC plot [516].

Previous studies using large numbers of autosomal SNPs, but not careful geographic sampling schemes, have noted north-south and northwest-southeast axes of variation within Britain, and place Scotland and Wales in an intermediate position between southern England and Ireland [517, 36, 518].

In terms of interpretation, the overall 'Neolithic' input to the British population is estimated to be substantial, by uniparental markers, and comparable to the European average (15-20%), and so far no genetic evidence has been interpreted as supporting any Iron Age 'Celtic' influx from Central Europe [484]. The size of the Anglo-Saxon invasion of England, however, has garnered the most attention in British genetic studies. Estimates of the proportion of Y chromosomes in modern Britain that entered the island vary, but are invariably high: 40-100% [519, 516]. This level of Anglo-Saxon input may be moderated by several circumstances: The comparatively low resolution of the Y STR systems used, a potentially confounding Danish Viking input which is unlikely to be distinguishable from the Anglo-Saxon, and the assumption – probably unwarranted – that the Y chromosome landscape of Britain and Ireland was homogenous prior to 400 AD [484]. Additionally, an "apartheid-like" social structure, suggested by Anglo-Saxon law codes, might well act to amplify the 'Anglo-Saxon' component [520, 521]. In this scenario, Anglo-Saxon males would have a higher level of reproductive success, due to higher social standing, that would, in successive generations, increase the number of their Y chromosomes.

From the current genetic perspective, then, Britain has been shaped largely by three events: The post-glacial re-population in the Palaeolithic, an influx of xenochthonous peoples in the Neolithic, and finally, the early Medieval *Adventus Saxonum.*

Chapter 5

Materials and Methods – People of the British Isles and associated data

5.1 Materials

5.1.1 DNA samples

5.1.1.1 People of the British Isles

The People of the British Isles sampling scheme was carefully designed to be able to detect fine-scale local variation [522]. Accepted donors normally had four grandparents born in the same area. This area was usually defined as a distance not exceeding 30-40 miles between grandparental birthplaces. Samples were geographically coded according to sampling location. Volunteers were mostly sampled in county-based cohorts (see Figure 5.1 for sampling locations and Table 5.1 for the distribution of sample numbers). By only sampling individuals with only local connections, going as far back in time as is practical for large groups of people, we hope to ensure that said individuals are as far as possible genetically representative of any given area. Major urban areas were avoided, on the assumption that rural communities provide a more stable demographic setting.

The author's contribution to the project consists of assisting with collections, giving

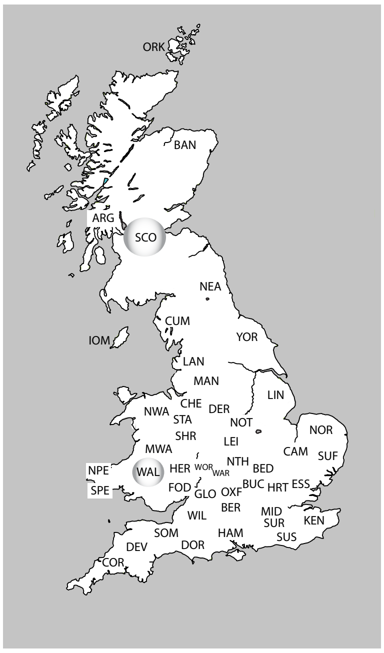


Figure 5.1: **PoBI sampling locations** See Table 5.1 for keys to the region codes. Northern Ireland (NIR) and the Republic of Ireland (RIR) not shown. SCO and WAL, respectively, are samples from Scotland and Wales without specified region of origin within the country.

public outreach lectures regarding the project, and processing of samples (lymphocyte separation and DNA isolation, see Chapters 5.2.1.1 and 5.2.1.2), and organizing and conducting sample collections in Norway (see below). In addition, I sourced DNA samples from Brittany, performed mitochondrial genotyping of all available DNA samples, and conducted background reasearch into candidate genes for normal variation of several genetically determined traits.

5.1.1.2 European samples

A set of samples from Northwest Europe were collected/obtained for comparative purposes.

5.1.1.2.1 Norwegian samples Norwegian samples were collected by the author in three regions of the country; Nordfjord, Ottadalen and Setesdal (see Figure 5.2). The same selection criteria were employed as for the PoBI project, except for occasional

Table 5.1: **PoBI sampling areas and sampling numbers** The geographic code refers to the area where an individual's grandparents self-reportedly come from. Included are samples collected up to January 2011.

Code	**Area**	**N(samples)**	**Code**	**Area**	**N(samples)**
ARG	Argyll, Bute, Western Isles	87	MWA	Middle Wales	91
BAN	Banff and Buchan, Eastern Scotland	76	NEA	Northeast England	259
BED	Bedfordshire	3	NIR	Northern Ireland	166
BER	Berkshire	27	NOR	Norfolk	145
BUC	Buckinghamshire	2	NOT	Nottinghamshire	88
CAM	Cambridgeshire	16	NPE	North Pembrokeshire	85
CHE	Cheshire	61	NTH	Northamptonshire	67
COR	Cornwall	301	NWA	North Wales	134
CUM	Cumbria	307	ORK	Orkney	208
DER	Derbyshire	17	OXF	Oxfordshire	157
DEV	Devon	124	RIR	Republic of Ireland	12
DOR	Dorset	54	SCO	Scotland	55
ESS	Essex	12	SHR	Shropshire	7
FOD	Forest of Dean	61	SOM	Somerset	20
GLO	Gloucestershire	55	SPE	South Pembrokeshire	54
HAM	Hampshire	53	STA	Staffordshire	28
HER	Herefordshire	36	SUF	Suffolk	122
HRT	Hertfordshire	1	SUR	Surrey	4
IOM	Isle of Man	3	SUS	Sussex	95
KEN	Kent	89	WAL	Wales	47
LAN	Lancashire	61	WAR	Warwickshire	16
LEI	Leicestershire	97	WIL	Wiltshire	20
LIN	Lincolnshire	169	WOR	Worcestershire	44
MAN	Manchester	1	YOR	Yorkshire	270
MID	Middlesex	2			

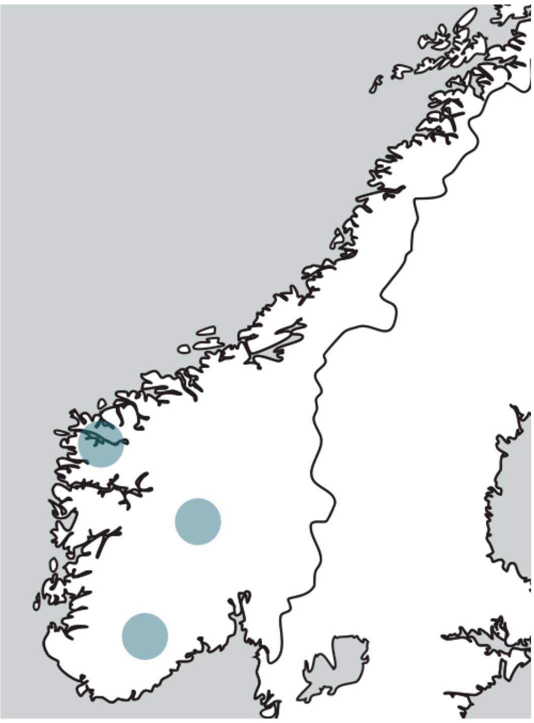


Figure 5.2: **Norwegian sampling locations** Nordfjord in the west, Ottadalen in the centre, and Setesdal in the south.

extension of the geographical distance criterion to allow for the sparseness of the population when compared to the population density of Britain. These three areas were chosen on the basis of differing levels of and ease of communication. Setesdal is a historically self-contained valley with a high degree of endogamy. Ottadalen, in the central massif of the country, is as far from the sea as it is possible to get in Norway. Nordfjord is very mountainous, and overland travel is difficult. It has no historical commercial ports, like Bergen or Trondheim, but communication was preferentially by boat until the early/mid-twentieth century. A total of 87 samples were collected from Norway; 38 from Nordfjord, 22 from Ottadalen, 22 from Setesdal, and 2 of mixed Norwegian ancestry.

5.1.1.2.2 Breton samples 760 DNA samples (380 males and 380 females) from Finistère in Brittany were kindly provided by Claude Ferec of the University of Brest. Similar selection criteria to those in use in the PoBI project were employed, i.e. unrelated individuals for which all four grandparents of each subject were from Finistère.

5.1.1.2.3 Dutch samples 96 Dutch DNA samples were provided by Peter de Knijff of the University of Utrecht. Individuals were selected on the basis of having at least three grandparents from the Netherlands.

5.2 Methods

5.2.1 Sample preparation and DNA analysis

PoBI and Norwegian samples were treated as detailed below.

5.2.1.1 Sterile separation of lymphocytes from blood samples

Lymphocytes were separated from the blood samples for transformation into cell lines, in order to provide renewable sources of DNA.

Equal amounts of whole blood and RPMI-1640 (PAA Laboratories GmBH, Pasching, Austria) were mixed at room temperature. In a fresh tube, 25ml of blood/RPMI mixture was layered on to 15ml of LymphoPrep^{TM}(Axis-Shield Diagnostics, Dundee, Scotland) using a pipette. This was centrifuged at 1800 rpm for 20 minutes, with no post-spin brake applied. The cells at the LymphoPrep^{TM}/RPMI-1640 interface were transferred to a fresh tube, and an equal volume of RPMI-1640 was added. The blood residue was kept for DNA extraction. The lymphocytes in RPMI-1640 were centrifuged at 2300 rpm for 10 minutes, and the supernatant discarded. The pellet was resuspended in 0.75-2.0ml ice-cold 10% DMSO in foetal bovine serum (Sigma-Aldrich Co. LLC, St. Louis, MO, USA), and kept on dry ice for transfer to liquid nitrogen.

5.2.1.2 DNA isolation

DNA was isolated from the residual blood, after lymphocyte separation, using the QIAamp DNA Blood Maxi kit from Qiagen (GmbH, Hilden, Germany). Unless otherwise specified, all reagents are from this kit. Following the manufacturer protocol, 750μl of

Qia protease was added to the blood sample, and the volume adjusted to 10ml with PBS. 12ml of Buffer AL was added, the tube inverted 15 times and shaken for 1 minute, and incubated for 20 minutes at 70 °C. 10ml ethanol (96%) was then added, and mixed by inversion. Half of the resulting solution was applied to the Qiagen spin column, and centrifuged at 3000 rpm for 3 minutes, and the filtrate discarded. This step was repeated with the second half of the solution. 7.5ml AW1 buffer was added to the column and spun at 4000 rpm for 2 minutes, then 7.5ml AW2 buffer was added to the column and spun at 4000 rpm for 30 minutes. The filtrate was discarded, and the DNA was eluted from the column membrane in 1ml distilled water after incubation for 5 minutes at room temperature, and centrifugation at 4000 rpm for 4 minutes.

5.2.1.3 Mitochondrial DNA typing

4855 samples (3913 PoBI, 87 Norwegian, 95 Dutch, and 760 Breton) were genotyped for mitochondrial DNA variation using mitochondrial SNPs on the Illumina 1M-Duo and Affymetrix 6.0 arrays and/or protocols modified from Köhnemann *et al.*, Grignani *et al.* (2006), Grignani *et al.* (2009) and Quintans *et al.* [523, 524, 525, 526]. 2912 of the PoBI samples were run on the SNP arrays (see Chapter 5.2.3); the remaining samples and those not fully genotyped to the specified level by the SNPs on these arrays were genotyped using the above modified protocols. If DNA amplification for a given sample failed after three attempts, the said sample was excluded from the analysis.

Köhnemann *et al.* describes a multiplex of 20 amplified fragments corresponding to diagnostic SNPs at the following rCRS sites: 1719, 1811, 3010, 3915, 4580, 5004, 6365, 6776, 7028, 8251, 8697, 9055, 11251, 12372, 12858, 13708, 14470, 14766, 14798, and 15904. These sites define most major European haplogroups (see Table 5.2 and Figure 5.5). All primer sequences were from Köhnemann *et al.* [525], and the multiplex PCR was performed in a total volume of 10μl per sample, with 5ng of template DNA, 1μl of a balanced primer mix (each primer at 2μM), 4μl dH$_2$O and 5μl Multiplex PCR Master Mix from Qiagen (GmbH, Hilden, Germany). PCR conditions were as follows: Initial denaturing at 95 °C for 10 minutes; 3 cycles of 95 °C for 30 seconds, 55 °C for 45 seconds and 72 °C for 30 seconds; 19 cycles of 95 °C for 30 seconds, 55 °C +0.2 °C per cycle for 45 seconds, and 72 °C for 30 seconds; 11 cycles of 95 °C for 30 seconds, 60 °C for 45 seconds and 72 °C for 30 seconds, followed by the final extension at 72 °C for 7 minutes. Post-PCR cleanup was performed using 2.5μl shrimp alkaline phosphatase (1U/μl) from USB Corporation (Cleveland OH, USA) or Thermo Scientific (Loughborough, Leicestershire, UK), and 0.05μl Exonuclease I (20U/μl) from Thermo Scientific, incubation at 37 °C for 75 minutes, and denaturing at 75 °C for 15 minutes.

Single base extension of the amplified fragments was performed using the SNaPshot Mulitplex Kit (Applied Biosystems, Foster City CA, USA), with 1μl of the multiplex PCR product, 0.5μl 10x AmpliTaq Gold PCR buffer (Applied Biosystems), 6μl balanced extension primer mix (see Köhnemann *et al.* [525]), and 2.5μl SNaPshot Multiplex Mix. Extension was carried out over 25 cycles of 96 °C for 10 seconds, 50 °C for 5 seconds and 60 °C for 30 seconds. Post-SBE clean-up consisted of 1μl shrimp alkaline phosphatase incubated at 37 °C for 45 minutes, and denatured at 75 °C for 15 minutes.

All steps were carried out in ThermoFast® 96 well PCR plates from Thermo Scientific on a DNA Engine Tetrad (Gradient Cycler) from MJ Research (Waltham MA, USA).

Further dissection of the major haplogroup H was performed using diagnostic SNPs at rCRS sites 456, 4745, 4769, 13020 and 13101, using primer sequences and SBE conditions from Grignani *et al.* (2006) and Quintans *et al.* [523, 526], with the same multiplex PCR and post-PCR protocol and conditions as above. Single base extension of the amplified fragments was performed using the SNaPshot Mulitplex Kit (Applied Biosystems), with 1μl of the multiplex PCR product, 3.3μl 10x AmpliTaq Gold PCR buffer (Applied Biosystems), 0.7μl balanced extension primer mix, and 2μl SNaPshot Multiplex Mix.

U, K, J and T haplogroup dissection was conducted using diagnostic SNPs 980, 1189, 3197, 3348, 6047, 9698, 9716, 13104, 14094, 14139, 14182, 14793 for U/K and 7476, 12633 and 14233 for J/T. Multiplex PCR and post-PCR protocol and conditions were as above. Primer sequences and SBE protocol were from Grignani *et al.* (2009)[524].

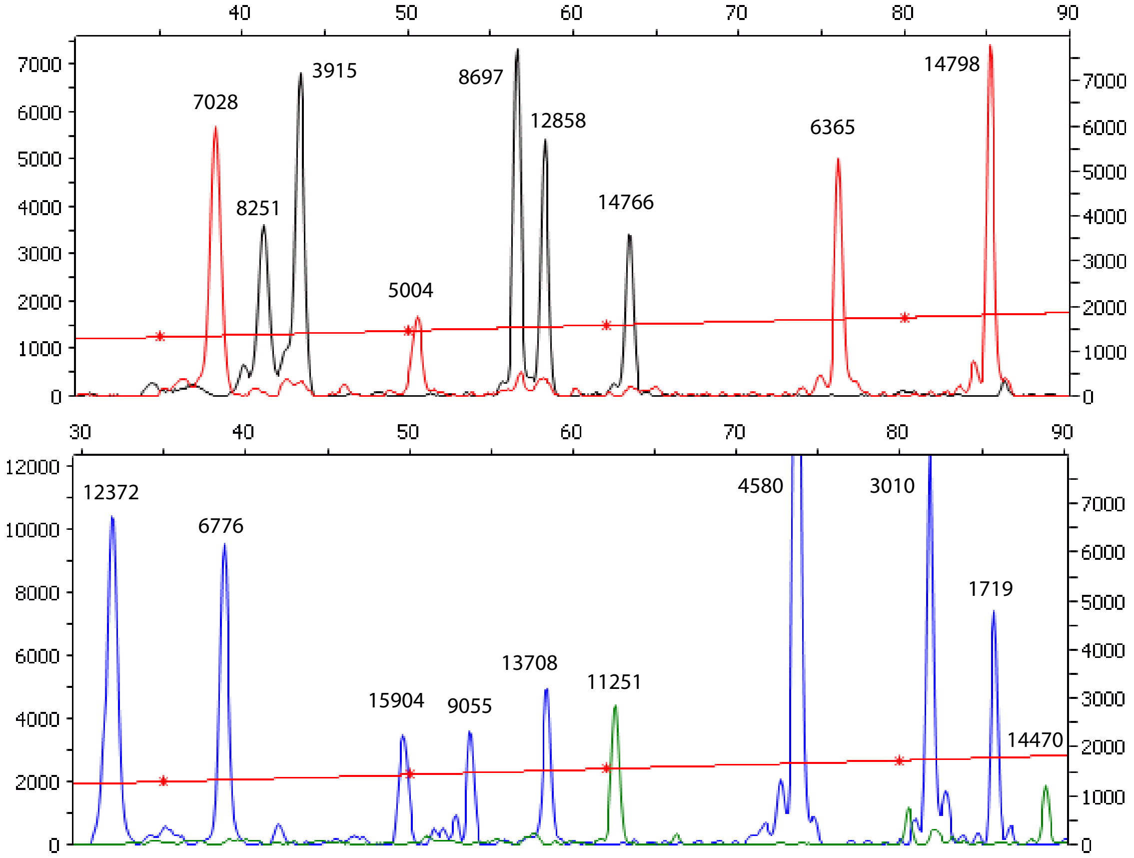


Figure 5.3: **Electropherogram of 20plex mitochondrial fragment analysis** An example of the results from the SNaPshot procedure. The upper panel shows the red and yellow fluorescence signals, the bottom the green and blue, for the same sample. All peaks are labelled according to the SNP they represent. The peak colour gives the identity of the nucleotide at the indicated position, red -T, black (from yellow fluorescence) - C, green - A, blue - G.

SBE fragment lengths and final nucleotide identities (i.e. allele identity) were assessed by a 3730 DNA Analyzer and/or a 3130 DNA Analyzer (both Applied Biosystems) using default settings and Peak Scanner Software 2.0 (Applied Biosystems), see Figures 5.3 and 5.4. Each electropherogram was scored in two independent rounds, and the haplogroup assignments were confirmed by comparison with the results of a number of random samples that had also been assessed by HVS-I sequencing and SNP arrays (Illumina Human1M-Duo DNA Analysis BeadChip and Affymetrix Genome-Wide Human SNP Array 6.0).

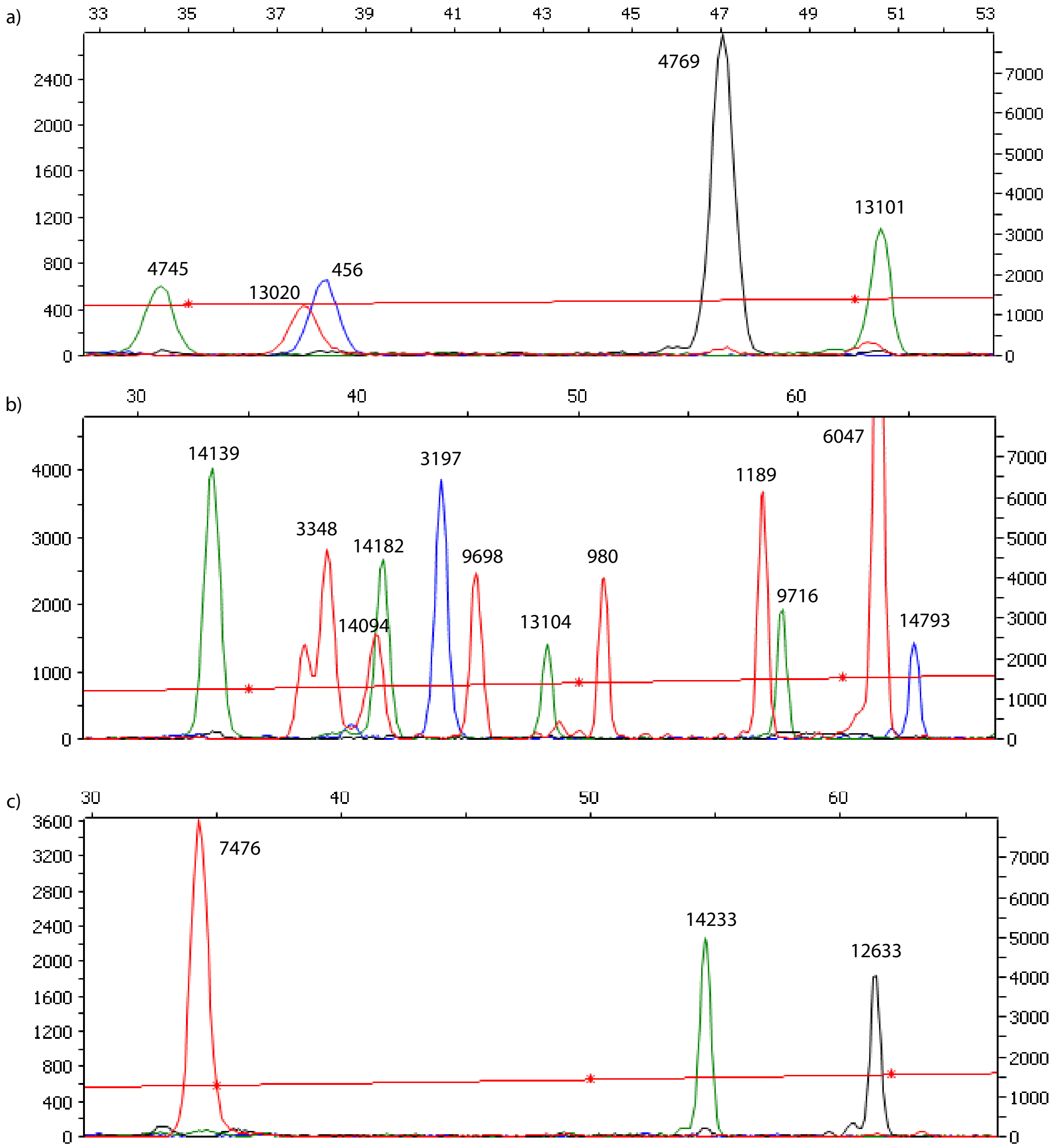


Figure 5.4: **Electropherograms of lower level mitochondrial haplogroup analysis** a) Subgroups of haplogroup H; b) Subgroups of haplogroup U, c) Subgroups of haplogroups J and T. All peaks are labelled according to the SNP they represent.The peak colour gives the identity of the nucleotide at the indicated position, red -T, black (from yellow fluorescence) - C, green - A, blue - G.

5.2.2 Mitochondrial haplogroup determination

Haplogroup determination for all the samples was based on published data from Herrnstadt *et al.*, Quintans *et al.*, Achilli *et al.*, Alvarez-Iglesias *et al.*, van Oven *et al.*, and two papers by Grignani *et al.* [527, 526, 486, 523, 524, 492]. Table 5.2 gives the SNPs used in this study, which haplogroups they in combination define, and in which assay (SNaPshot or SNP array) they occur.

The focus was on typical European haplogroups, i.e. several subgroups of superhaplogroup N (N1, W, X and most particularly haplogroup R). Figure 5.5 shows the relevant mitochondrial genome phylogenetic tree.

Samples were classified as 'other' if they showed none of the variants in Table 5.2, if they were were classified into a very minor group (e.g. L3*(xN), R*(xR0, JT, U)), or if they showed SNP variant patterns at odds with given haplogroup specifications that could not be resolved.

Additionally, mitochondrial haplogroup frequencies from surrounding northwest European countries were obtained from the literature [528, 529, 530, 531].

Table 5.2: **Mitochondrial haplogroup diagnostic SNPs** Major – the first-pass SNaPshot reaction, UK/JT – SNaPshot reaction for further dissection of UK and JT haplogroups, H – SNaPshot reaction for further dissection of haplogroup H, 6.0 – the Affymetrix chip, 1M-Duo – the Illumina chip. a) The position of the SNP in the mitochondrial chromosome. b) The haplogroup defined by the given SNP. c) The base present at the given locus for the given haplogroup. An X indicates that the assay provides a genotype for the given SNP.

SNP[a]	**Haplogroup**[b]	**Variant**[c]	**Major**	**UK/JT**	**H SNaPshot**	**6.0**	**1M-Duo**
456	H5	T			X		X
769	L3	A				X	
980	U7	C		X			
1189	K1	C		X			X
1719	N1/H7a/X2	A	X				X
1811	U2-4,7-9	G	X			X	

Continued on the next page

– continued from the previous page

SNP[a]	**Haplogroup**[b]	**Variant**[c]	**Major**	**UK/JT**	**H SNaPshot**	**6.0**	**1M-Duo**
3010	H1/J1	A	X			X	X
3197	U5	C		X		X	X
3348	U6	G		X			X
3720	U2	G					X
3915	H6a	A	X				X
4580	V	A	X				
4745	H13	G			X		
4769	H2	G			X		X
5004	H4	C	X				X
5046	W	A					X
5390	U2	G					X
5460	W	A				X	X
6047	U4	G		X			
6152	U2	C					X
6365	H1a1	C	X				
6776	H3	C	X				X
7028	H	C	X			X	
7476	J2	T		X			
8251	I/W/N1e	A	X			X	
8697	T	A	X				X
9055	K	A	X				
9123	H4	G				X	
9540	N	C				X	
9698	U8	C		X		X	X
9716	K2	C		X			X
10238	I	G				X	X
10398	J	G				X	X
10463	T	T				X	X
10550	K	A				X	X
11251	JT	G	X			X	X
11299	K	T				X	
11674	W	C				X	
11719	R0	C				X	
12308	U	A				X	
12372	U	A	X				X
12633	T1	A		X		X	
12705	R	G				X	X
12858	H1c	T	X				
13020	H9	C			X		
13101	H8	C			X		
13104	U1	G		X			

Continued on the next page

– continued from the previous page

SNP^{a}	Haplogroup^{b}	Variant^{c}	Major	UK/JT	H SNaPshot	6.0	1M-Duo
13708	X2b'd/J	A	X			X	
14094	U9	C		X			
14139	U3	G		X			
14182	U5b	C		X			
14233	T2	T		X			X
14470	X	C	X			X	
14470	H10	A	X				
14766	HV	C	X				
14793	U5a	G		X		X	
14798	K/J1c	C	X				X
14905	T	G				X	
15607	T	A				X	
15904	V	T	X				X
16162	H1a	G					X
16163	T1	G					X

5.2.3 SNP arrays

The Human1M-Duo DNA Analysis BeadChip from Illumina and the Affymetrix Genome-Wide Human SNP Array 6.0 aim to give comprehensive cover of genome variation. The former has 1,049,008 SNP probes, and the latter 906,000 SNP probes. 2912 PoBI samples were run on the Human1M-Duo chip, and 2931 on the Affymetrix 6.0 array as part of the Wellcome Trust Case Control Consortium 2. A further 283 PoBI samples were run on the Omniexpress 750K SNP chip from Illumina. The Norwegian samples and 192 Breton samples were also run on the latter chip.

Data from these arrays were used for partial haplogroup determination of mitochondrial DNA (see Chapters 5.2.1.3 and 5.2.2), and for markers below (see Chapters 5.2.4, 5.2.5, and 5.2.6).

5.2.4 Classical markers and lactase persistence

Several classical markers were chosen for separate analysis to compare with the overall analysis done using fineSTRUCTURE. The markers chosen were ones for which there was independent evidence for interesting population differences with the aim of looking for possible effects of natural selection. Loci were taken from Cavalli-Sforza *et al.* [22],

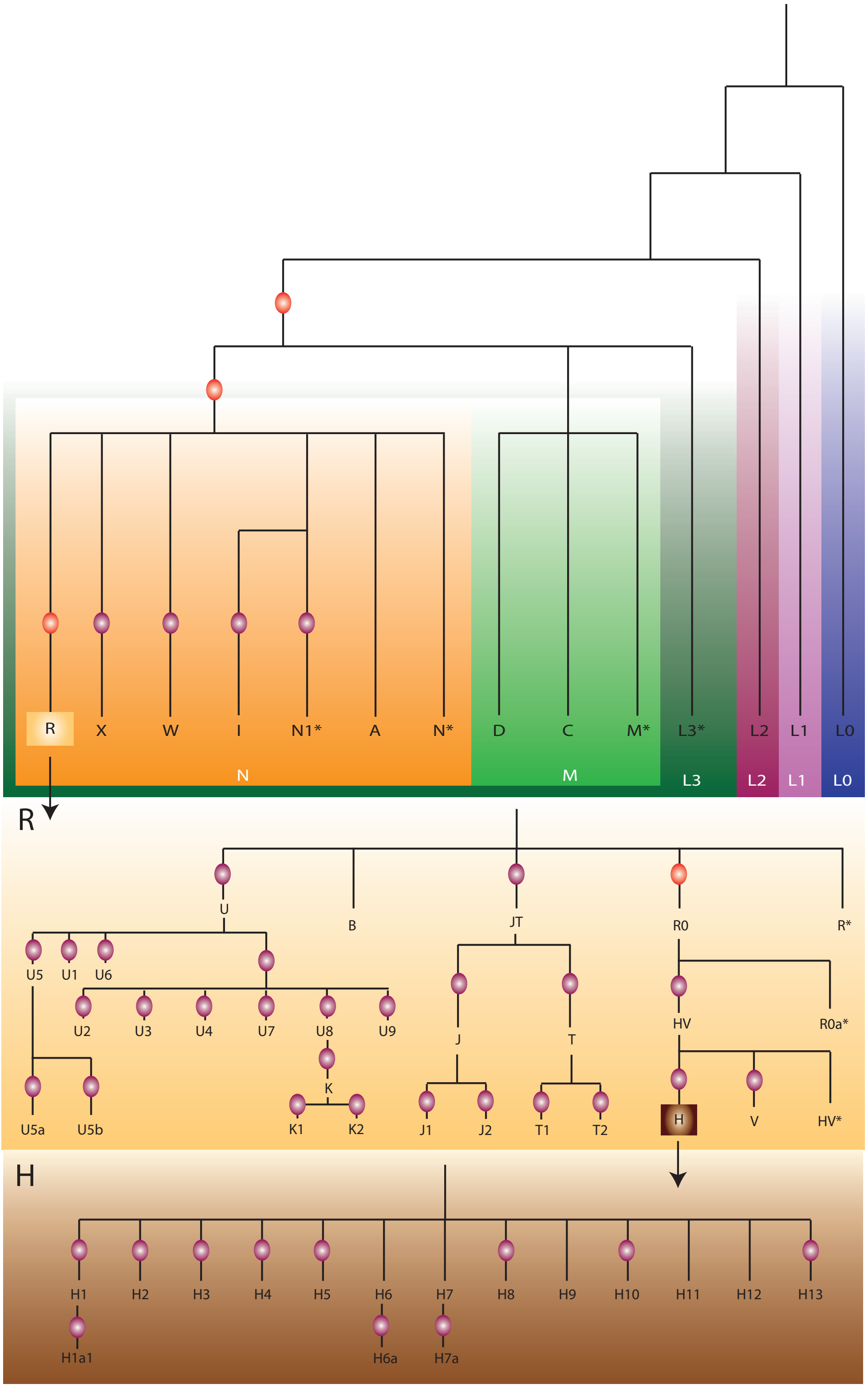


Figure 5.5: **Phylogenetic tree of mitochondrial genomes** A purple dot on a branch indicates that SNPs defining this haplogroup have been assessed for all samples; a red dot that SNPs defining these haplogroups were only found in the SNP arrays.

and the SNPs from their corresponding genes were discovered from the literature (see Table 5.3). Known single nucleotide polymorphisms within these genes and associated data were extracted using the UCSC Genome Table Browser (http://genome.ucsc.edu), variation and repeats group, table SNP(130) for the March 2006 human genome assembly and dbSNP-Q (https://cgsmd.isi.edu/dbsnpq). RefSeq sequences for any given gene were selected as default, and accession numbers were noted when there was more than one.

Thirty-seven classical marker genes had SNPs present on the Illumina 1M-Duo chip, and all of these genes were represented by at least one SNPs, the majority of which are in introns (see Table 5.4).

In addition, a SNP 22,018bp upstream of the lactase gene (*LCT*) was assessed. Lactase allows the digestion of the milk sugar lactose [532]. This SNP, C/T-13910 (rs4988235), is completely associated with lactose intolerance [533]. The converse, lactase persistence beyond infancy, is a recently selected trait; the probable selective pressure is the introduction of pastoral farming. The frequency of lactase persistence in a population has long been known to vary according to the geographical origin of that population [532]. For all PoBI individuals typed on the Illumina 1M-Duo platform, the rs4988235 genotype was extracted.

5.2.5 Pigmentation

5.2.5.1 Phenotype information

For 951 PoBI volunteers, from January 2010 to March 2012, phenotype information on hair colour (self-defined) and skin tone (by hand-held spectrophotometer) was collected by the project.

5.2.5.2 Genetic variants and phenotype prediction

Genes and genetic variants that have shown association with pigmentation phenotypes were collected from the literature, many from whole genome association studies. For

Table 5.3: **Classical Markers**

Locus	**Name**	**HUGO symbol**
ABO	ABO blood group	ABO
ACP1	acid phosphatase	ACP1
ADA	adenosine deaminase	ADA
AK1	adenylate kinase	AK1
ALPP	alkaline phosphatase, placental	ALPP
BF	glycine-rich beta glycoprotein	CFB
C3	complement protein	C3
CHE1	cholinesterase 1	
CHE2	cholinesterase 2	CHE2
CP	ceruloplasmin	CP
DI	Diego blood group	SLC4A1
ESD	esterase D	ESD
FY	Duffy blood group	DARC
G6PD	glucose-6-phosphate dehydrogenase	G6PD
GC	group-specific component (binds vitamin D)	GC
GLO1	glyoxalase-locus 1	GLO1
GPT	glutamate-pyruvate transaminase	GPT
HLAA	major histocompatibility complex, class I A	HLA-A
HLAB	major histocompatibility complex, class I B	HLA-B
HP	haptoglobin	HP
IGHG1G3	immunoglobulin heavy chain constant gamma 1	IGHG1
JK	Kidd blood group	SLC14A1
KEL	KELL blood group	KEL
KM	kappa light chain immunoglobulin	IGKC
LE	Lewis blood group	FUT3
LPA	beta lipoprotein	LPA
LU	Lutheran blood group	BCAM
MDH1	malate dehydrogenase 1, NAD	MDH1
MNS	glycophorin A	GYPA
MNS	glycophorin B	GYPB
P1	P1 blood group	P1
PEPA	peptidase A1	CNDP2
PGD	6-phosphogluconate dehydrogenase	PGD
PGK1	phosphoglycerate kinase	PGK1
PGM1	phosphoglucomutase	PGM1
PI	antitrypsin	SERPINA1
PTC	taste receptor type 2, member 38 (phenylthiocarbamide)	TAS2R38
RH	RH blood group C, E	RHCE
RH	RH blood group D	RHD
SE	secretor	FUT2
SOD1	superoxide dismutase 1	SOD1
TF	transferrin	TF

Table 5.4: **Classical marker gene SNPs** SNPs – The number of SNPs in the given gene present in the Illumina 1M-Duo array. The remaining columns give the number of SNPs in each category; NS – non-synonymous coding, S – synonymous coding, Near 5'/3' – within 2000bp of transcribed region.

Gene	**SNPs**	**NS**	**S**	**Intron**	**5'-UTR**	**3'-UTR**	**Near 5'**	**Near 3'**
ABO	107	11	7	81	0	2	6	0
ACP1	8	2	1	7	0	0	3	2
ADA	19	0	1	17	0	0	0	1
AK1	23	1	0	17	0	1	4	0
ALPP	2	0	0	0	0	0	2	0
BCAM	42	8	3	23	0	4	4	3
C3	127	6	8	110	0	0	3	1
CFB	18	5	5	5	0	3	0	4
CNDP2	29	1	2	21	1	4	2	1
CP	23	0	1	18	0	1	3	0
DARC	5	2	0	1	1	0	3	1
ESD	27	0	0	23	0	0	3	1
FUT2	37	3	5	12	0	11	3	1
FUT3	32	5	2	13	4	1	10	0
G6PD	14	2	1	11	1	0	2	0
GC	27	3	1	21	1	0	1	0
GLO1	16	0	0	9	1	4	2	0
GPT	6	1	0	1	0	1	3	2
GYPA	16	0	0	11	0	3	0	2
GYPB	5	0	0	5	0	0	0	0
HLAA	4	0	0	3	0	1	0	0
HLAB	29	6	1	10	0	2	6	3
HP	20	0	2	13	0	0	5	1
KEL	45	6	3	29	0	0	7	1
MDH1	7	2	0	5	0	0	1	0
PGD	10	0	0	9	0	0	1	0
PGK1	2	0	0	2	0	0	0	0
PGM1	31	3	0	27	1	2	1	0
PGM2	26	1	1	21	0	1	2	0
RHCE	26	1	0	21	0	3	2	2
RHD	18	1	0	15	0	1	1	0
SERPINA1	19	3	1	11	0	3	2	1
SLC14A1	161	5	6	128	3	5	25	1
SLC4A1	34	5	5	19	0	4	2	1
SOD1	6	0	0	4	1	1	0	0
TAS2R38	2	2	0	0	0	0	0	0
TF	102	5	3	81	2	0	11	0

overall pigmentation, those genes that were found to significantly impact pigment phenotypes in three or more studies were included. Known single nucleotide polymorphisms within these genes and associated data were extracted using the UCSC Genome Table Browser (http://genome.ucsc.edu) , variation and repeats group, table SNP(130) for the March 2006 human genome assembly and dbSNP-Q (https://cgsmd.isi.edu/dbsnpq). RefSeq sequences for any given gene were selected as default, and accession numbers were noted when there was more than one.

Within these genes, SNPs that showed strong association with a specific hair or eye colour were identified, see Table 5.5. The latter associated SNPs were then used to predict regional phenotype proportions based on genotypes extracted from the Illumina 1M-Duo PoBI dataset.

Table 5.5: **Phenotype-genotype correlations**

Gene	**SNP**	**Allele**	**aa change**	**Genotype**	**Phenotype**[a]	**Freq.**[b]
MC1R	rs1805007	T	R>C	homozygotes/compound	red hair	61.7%
MC1R	rs1805008	T	R>W	heterozygotes		
MC1R	rs1805005	T	V>L	compound heterozygotes		10.8%
MC1R	rs2228479	A	V>M	with rs1805007 or	red hair	
MC1R	rs885479	A	R>Q	rs1805008		
SLC45A2	rs16891982	C	F>L	hetero- or homozygotes	black hair	47.0%
MC1R	rs1805007	T	R>C	heterozygotes[c]	blonde hair	23.5%
KITLG	rs12821256	C	near gene	hetero- or homozygotes	blond hair	25.0%
HERC2	rs12913832	G	intron	homozygotes	light eyes	100%

---

[a] Phenotype associated with genotype

[b] Frequency of phenotype given genotype

[c] Not in combination with any other given *MC1R* variants

The estimated frequencies of phenotype given genotype were either taken directly from the publications or back-calculated from given odds ratios and sample numbers [534, 78, 84, 535, 86]. For the former, the probability of observing a given phenotype given the indicated genotype was stated explicitly. For the latter, a matrix of the form seen in Table 5.6 was created. The probability of observing the phenotype given the genotype is then $\frac{X}{X+Y}$. $X+Y, A+B, X+A$ and $Y+B$ were available in the relevant papers, as was the odds ratio (OR), given by $\frac{XB}{YA}$. $X+Y, A+B, X+A$ and $Y+B$ were then expressed in terms of X, and these expressions replaced A, B and Y in the $\frac{XB}{YA} = OR$ equation. The resulting quadratic equation was solved for X, permitting the calculation of $\frac{X}{X+Y}$.

Table 5.6: **Odds ratio matrix** Binary separation of relevant phenotypes and genotypes are given by j/k, and l/m, respectively.

		Phenotype		
		j	k	N
Genotype	l	X	Y	X+Y
	m	A	B	A+B
	N	X+A	Y+B	

5.2.5.3 Pigmentation genes

Forty-two genes associated with pigmentation, and melanin synthesis and distribution, were selected (see above, Chapter 5.2.5.2) [111, 78, 79, 81, 76, 536, 537, 538, 539, 540, 541, 542, 543, 544]. These genes were represented by 0-441 SNPs, the majority of which are intronal (see Table 5.7), and for fifteen of the genes no non-synonymous coding SNPs were present.

5.2.6 Facial morphology candidate gene selection

Little is known of the molecular basis underlying normal craniofacial variation (see Chapter 2). Therefore, the first step in identifying candidate regions/genes/SNPs involved in normal variation was to examine the molecular basis of facial development. Two approaches were taken when looking for potential targets. First, genes known to be involved in the development of the face in vertebrate model organisms were included. A search was performed in the PubMed database to identify human homologues of genes experimentally shown to be instrumental in the development of facial or analogous structures in model organisms. Second, the genetic basis of *abnormal* variation was sought. Genes which, when seriously disrupted, cause abnormal facial variation

Table 5.7: **Pigmentation gene SNPs** SNPs – The number of SNPs in the given gene present in the Illumina 1M-Duo array. The remaining columns give the number of SNPs in each category; NS – non-synonymous coding, S – synonymous coding, Near 5'/3' – within 2000bp of transcribed region.

Gene	**SNPSs**	**NS**	**S**	**Intron**	**5'-UTR**	**3'-UTR**	**Near 5'**	**Near 3'**
ADAM17	22	0	1	18	0	2	1	0
ASIP	6	2	0	2	0	0	2	0
ATRN	79	8	3	55	0	10	2	1
BNC2	233	3	1	222	0	7	0	0
CLMP	54	1	0	51	0	1	0	1
CLTC	31	2	0	26	0	1	2	0
CORIN	71	4	2	61	0	2	0	2
CTNS	14	0	0	7	1	5	1	0
DCT	19	0	0	17	0	0	1	1
DEFB103A	1	0	0	1	0	0	0	0
DRD2	36	0	2	30	0	2	2	0
DTNBP1	9	4	1	0	0	4	0	0
EDN3	26	2	0	12	1	9	2	0
EDNRB	21	5	2	8	1	4	1	0
EGFR	127	7	5	110	0	2	2	1
ETS1	62	0	1	56	0	4	1	0
EXOC2	113	0	1	107	0	3	1	1
HERC2	88	0	2	83	0	2	0	1
HPS6	6	2	2	0	0	1	0	1
IRF4	16	0	0	7	0	6	2	1
KITLG	28	1	0	20	0	6	1	0
LYST	101	3	6	90	0	1	1	0
MC1R	24	7	2	0	8	5	2	0
MITF	60	0	0	54	0	2	4	0
MLANA	14	0	1	10	0	1	2	0
MLPH	42	6	1	32	0	2	1	0
MYO5A	87	3	4	72	0	7	1	0
NRG1	441	3	0	431	0	1	6	0
OCA2	143	2	4	134	0	1	1	1
OPRM1	0	0	0	0	0	0	0	0
PAX3	56	0	0	49	0	2	4	1
PMEL	13	2	2	4	0	0	5	0
POMC	6	0	0	6	0	0	0	0
RAB27A	31	1	1	23	0	4	2	0
SLC24A4	6	1	3	0	0	2	0	0
SLC24A5	11	1	0	9	0	0	1	0
SLC45A2	37	2	1	30	0	3	0	1
TPCN2	26	3	1	14	0	7	1	0
TRPM1	62	6	6	44	0	1	4	1
TYR	36	3	1	31	0	0	1	0
TYRP1	16	0	0	10	1	3	2	0
UBASH3B	100	1	1	90	1	6	1	0

are likely good candidates for causing normal variation due to non-pathogenic variants. There are a plethora of genetic syndromes in humans that cause facial dysmorphia, see e.g. Winter [545]. A search was performed in the PubMed database to discover genes that are causative, when mutated, of such syndromes. Nearly two hundred such genes/regions were found that were causative of, or good causative candidates for, these genetic syndromes, and/or are crucial to craniofacial development (see Appendix H for final candidates).

Of these genes, 54% are known to have functions involved in development and morphogenesis as determined by functional annotation at DAVID (http://david.abcc.ncifcrf.gov). Known single nucleotide polymorphisms within these genes and associated data were extracted using the UCSC Genome Table Browser (http://genome.ucsc.edu), variation and repeats group, table SNP(130) for the March 2006 human genome assembly and dbSNP-Q (https://cgsmd.isi.edu/dbsnpq). RefSeq sequences for any given gene were selected as default, and accession numbers were noted when there was more than one.

197 genes had SNPs present on the Illumina 1M-Duo chip, the majority of which are intronal (see Appendix L).

5.2.7 Statistical analyses

Fisher's exact tests for significant differences in mitochondrial haplogroup frequencies was performed in R v.2.9.2 [546].

For the classical marker genes, non-synonymous coding SNPs present on the Illumina 1M-Duo array were selected for analysis (see Table5.4). Twelve of 37 classical marker genes had no non-synonymous SNPs present. For the pigmentation genes and facial feature candidate genes, all SNPs present on the array were used (see Table 5.7 and Appendix L).

In order to test for differentiation, population pairwise F_{ST} for all above markers were estimated according to Weir and Cockerham [547] using the R package *Geneclust* or custom-written C code.[1] F_{ST} for pairs of PoBI groups for all SNPs present on the

---

[1] The output of this code was checked to ensure it provided the same results as published software.

Illumina array, and for the subset of all non-synonymous SNPs, were calculated as well. These measures serve as controls representing the expected average F_{ST} level for the pigmentation and facial feature candidate genes (all SNPs), and the classical marker genes (non-synonymous SNPs).

A two-sided t-test for melanin index distributions was performed in Microsoft Excel for Windows 7. All principal components analyses and multidimensional scalings were performed in R v.2.9.2 using default settings.

Plots of the extent of homozygosity surrounding SNPs with very high pairwise population F_{ST} values was conducted as described in Tishkoff *et al.* [548].

5.2.8 Mapping

For frequency distribution maps of alleles, population samples were given latitude and longitude values based on the geographic centre of their sampling location.[2] The level of resolution is one degree square (see Figure 5.6 for grid definitions within Britain). The longitudinal and latitudinal positions of sampling locations were determined using iTouchMap on GoogleMaps, available from: *http://itouchmap.com/latlong.html.* To construct a frequency surface, a simple first-order polynomial regression was performed on the data, using the *loess* function in R v.2.9.2, with a scaling parameter alpha = 0.45, chosen to retain some local detail without displaying large fluctuations due to data variability [546, 128]. Vector maps were obtained from the Bodleian Map Room, and the Blighty package in R.

---

[2] i.e. the region given in Table 5.1

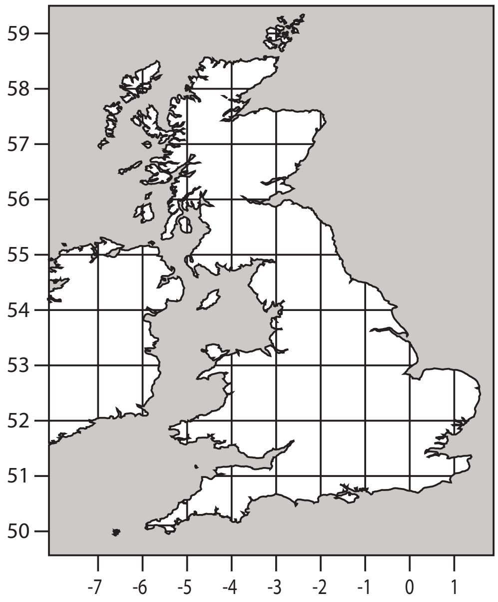


Figure 5.6: **Square degree grid of Britain and Northern Ireland** Longitude is on the x-axis, latitude on the y-axis.

Chapter 6

Results

The following chapter presents the results for all PoBI-based analyses on geographic trends in the genetic data. Interpretation and discussion will follow in Chapter 7.

Two schemes for grouping individuals are employed for the results. The first is geographic and combines counties to create roughly equally-sized regions, based on the sampling locations of the individuals, i.e. the county given as their grandparents' birthplace, see Table 6.1.

The second is on the basis of genetic clustering using fineSTRUCTURE (provisional, performed by G. Hellenthal, see Chapter 5 and Appendix K, and Table K.1). This clustering is based solely on the genetic profile of individuals, and not on their geographic affiliations.

6.1 Mitochondrial DNA

Of the 4855 samples typed, 127 were excluded by the SNP array data as duplicates or by reason of close genetic relationships (close relatives who had different mitochondrial DNA types, and were therefore not maternally related, were not excluded). A further 61 samples were excluded when consistently poor DNA quality prevented them from being reliably genotyped. The vast majority of samples could be classified into one of the major west Eurasian haplogroups; HV, JT, U, W, or N1e-I. 58 samples could not

Table 6.1: **Geographic grouping scheme**

Region	**Constituents**	**N=**
Central England	Bedfordshire, Northamptonshire, Oxfordshire	227
East-central England	Leicestershire, Lincolnshire, Nottinghamshire	354
Eastern England	Cambridgeshire, Essex, Norfolk, Suffolk	295
Eastern Scotland	Banff and Buchan	76
Middle Wales	Montgomeryshire, Radnorshire, Cardiganshire	91
Northeast England	Durham, Northumberland	259
Northwest England	Cumbria, Lancashire	368
North Wales	North Ceredigion, North Powys, Merioneth, Caernarfonshire, Denbighshire	134
Orkney	Orkney	208
South-central England	Gloucestershire, Hampshire, Somerset, Wiltshire	148
Southeast England	Kent, Sussex	183
South Wales	Pembrokeshire	139
Southwest England	Cornwall, Devon	425
Western England	Herefordshire, Staffordshire, Warwickshire, Worcestershire	124
Western Scotland	Argyll and Bute	87
Yorkshire	Yorkshire	270

be positively classified, either because they showed no diagnostic changes for the SNPs tested (n=33), or because of the presence of contradictory SNP variants (n=25), most probably indicating recurrent or back mutations, such as are comparatively common in the mitochondrial genome [525].

Due to the non-overlapping nature of the SNaPshot protocols and the SNP array markers, an ‘other’ category was created, containing samples that were classified by SNP array if the same haplogroups were not classifiable by SNaPshot. 36 of 66 ‘other’ mtDNAs are in this category. Of these, 12 were classified by SNP array, five were L3*(xN),[1] one was N*(xR, N1, W), five were R*(xR0, JT, U), and one sample was not within the dominant extra-African L3 clade. The 24 unclassified by SNaPshot may well have a similar distribution. The remaining 30 samples are classed as ‘other’ due to SNP patterns at odds with standard classifications. Four of these had incongruous variants in SNPs that are known to be liable to mutation (3010 and 13708), the rest had

---

[1] This nomenclature entails that the haplogroup in parentheses, after the x, is excluded from its parent clade – the one outside the parenthesis – in the given grouping. In this case, it should be read as ‘all members of haplogroup L3 except those belonging to its subhaplogroup N’.

incongruent variants at 1189 (n=2), 1811 (n=3), 3348 (n=2), 3915 (n=1), 6465 (n=1), 7028 (n=1), 9055 (n=12), 12372 (n=1), 14470 (n=3), 14766 (n=2), and 14788 (n=1). The number of unclassifiables due to an additional 9055A implies that this locus may also be prone to recurrent mutation.

Data for Denmark and northern France were collected from the literature (see Chapter 5), and data for Ireland and Norway as a whole were partially derived from the literature, depending on the refinement of haplogrouping. 12 Irish samples were typed in house, 300 were from published data, and 558 published Norwegian samples were added to the 84 typed in house when the genotyping level in the published data permitted such a combination.

Table 6.2 shows the relative frequencies of the major higher level haplogroups present in Britain and proximate countries. Tables 6.3, 6.4, 6.5, and 6.6 give the frequencies of the lower level haplogroup assignments. Maps showing the frequency gradients of the haplogroups in Britain can be found in Figures 6.1, 6.2, and 6.3, and in Appendix I, and the regional data for Norway in Appendix J.

Major haplogroup frequencies lie within those expected for Europe (see e.g. Achilli *et al.*[529]), with haplogroup H predominating, followed by JT and U. Within Britain, the highest frequency of H haplogroups is found in middle Wales (51.1%), and the lowest in Orkney (35.8%) and Holland (35.6%). For JT, the highest is 30.6% in eastern Scotland, and the lowest 12.3% in Brittany. U frequencies peak at 28.7% in Ulster, and are lowest in north Wales at 14%.

At a finer level, among H haplogroups, H1, H3 and unassigned H-subgroups predominate, and H subhaplogroups frequencies are also broadly similar between the north-western European regions. Unassigned H-subgroups include all those not in H1-5, H6a, H7a, H8-10, and H13, and their frequency ranges from 12.1% in southeast England to 5.9% in western England.[2] H1 ranges from 25% in middle Wales, to 9.3% in north Wales. H3 ranges from 11.6% in southeast England to 2.2% in south Wales. Average frequencies for H2, H4, H5 and H6a lie between 1-4%, but north Wales presents

---

[2] All percentages given are of the entire number of mitochondrial DNA samples assessed.

Table 6.2: **Major haplogroup frequencies in Britain and northwest Europe** N is the total number of individuals typed for the region. Regions outside Britain are below the line.

Region	**H**	**JT**	**U**	**Non-major**	**N=**
Central England	45.2	20.3	20.3	14.1	241
Eastern-central England	44.6	22.2	20.8	12.4	361
Eastern England	47.3	19.4	19.8	13.4	283
Eastern Scotland	45.8	30.6	16.7	6.9	72
Middle Wales	51.1	18.2	17.0	13.6	88
Northeast England	44.0	26.0	22.8	7.2	250
North Wales	46.5	25.6	14.0	9.3	129
Northwestern England	41.2	24.9	23.7	10.2	354
Orkney	35.8	25.1	21.9	17.1	187
South Wales	37.0	20.7	25.9	16.3	135
South-central England	43.1	19.2	23.5	14.2	260
Southeast England	45.1	20.2	19.7	15.0	173
Southwest England	50.4	17.9	20.4	11.3	407
Western England	41.7	21.9	17.6	18.7	187
Western Scotland	39.8	27.7	16.9	15.6	83
Yorkshire	39.2	23.2	23.6	14.1	263
Brittany	42.8	12.3	28.1	16.8	750
Denmark	45.9	22.5	22.1	9.6	9251
Holland	35.6	21.8	21.8	20.7	87
Ireland	43.6	18.3	24	14.1	312
Northern France	44	17.6	26.9	11.4	193
Norway	49.5	19	21.5	10.5	642
Ulster	36.6	24.4	28.7	10.4	164

an exceptional 10% of H4, and other outliers include an H2 frequency of 7% in Norway (distributed throughout the country), with the second highest occurrences being only 2.8% and 2.7% in Wales and Orkney, respectively. H7a, H8 and H9 are all very rare, with most groups harbouring no examples of these haplogroups. H7a, present in Britain only in one individual each from Devon, Cornwall, Dorset and Somerset, is outside Britain only seen in Brittany (0.4%). Only one H9 carrier was observed, in east-central England. H10 and H13 are also quite uncommon, the highest observed frequency for H10 being 1.8%, and only one group having a higher H13 frequency than 1.1%: eastern Scotland at 8.3%

Table 6.3: **H haplogroup frequencies in Britain and northwest Europe** N is the total number of individuals typed for the region, the frequencies given are percentages of the total for all mitochondrial haplogroups. 'H*' encompasses any H haplogroup except 1-5, 6a, 7a, 8-10, and 13.

Region	**H***	**H1**	**H2**	**H3**	**H4**	**H5**	**H6a**	**H7a**	**H8**	**H9**	**H10**	**H13**	**N=**
Central England	10.0	15.8	2.5	3.7	2.9	6.6	2.9	0.0	0.0	0.0	0.0	0.8	241
Eastern-central England	9.4	16.3	2.2	4.2	3.0	4.4	1.7	0.0	0.3	0.3	1.7	1.1	361
Eastern England	12.0	15.9	1.8	6.7	1.1	2.8	4.2	0.0	0.4	0.0	1.8	0.7	283
Eastern Scotland	11.1	11.1	1.4	6.9	2.8	1.4	1.4	0.0	1.4	0.0	0.0	8.3	72
Middle Wales	10.2	25.0	3.4	9.1	0.0	0.0	2.3	0.0	0.0	0.0	1.1	0.0	88
North Wales	6.2	9.3	1.6	7.8	10.1	3.9	2.3	0.0	3.9	0.0	1.6	0.0	129
Northeast England	6.4	18.8	2.0	5.2	2.4	2.0	5.2	0.0	0.0	0.0	1.2	0.8	250
Northwest England	8.8	13.8	3.7	5.9	0.8	3.1	2.8	0.0	0.0	0.0	1.1	1.1	354
Orkney	8.0	11.8	2.7	3.7	1.6	4.3	1.1	0.0	1.1	0.0	1.1	0.5	187
South central England	6.9	20.4	1.9	3.8	3.1	2.7	1.5	0.8	0.0	0.0	1.2	0.8	260
South Wales	6.7	13.3	3.7	2.2	3.7	2.2	4.4	0.0	0.0	0.0	0.7	0.0	135
Southeast England	12.1	16.2	1.2	11.6	1.2	1.2	1.7	0.0	0.0	0.0	0.0	0.0	173
Southwest England	12.0	14.0	1.2	9.1	3.7	4.4	3.4	0.5	0.0	0.0	1.2	0.7	407
Western England	5.9	15.5	2.1	7.0	3.2	3.2	4.3	0.0	0.0	0.0	0.5	0.0	187
Western Scotland	6.0	19.3	0.0	6.0	3.6	1.2	2.4	0.0	0.0	0.0	1.2	0.0	83
Yorkshire	11.0	15.2	1.5	4.2	0.4	3.8	1.1	0.0	0.0	0.0	1.5	0.4	263
Brittany	8.4	15.9	1.9	6.8	3.2	2.7	2.7	0.4	0.3	0	0.4	0.3	750
Holland	13.8	11.5	1.1	4.6	1.1	1.1	1.1	0	0	0	1.1	0	87
Norway	9.3	19.8	7	4.7	1.2	15.1	1.2	0	0	0	0	1.2	86
Ulster	8.5	11	1.8	7.3	3.7	1.8	0	0	1.8	0	0	0.6	164

Of the JT subhaplogroups, J1 and T2 were the most common. J1, of which most are J1c, peaks at 22.2% in eastern Scotland, and is lowest in eastern and southeast England, at 6.4%. J2 is absent from several groups (eastern Scotland, Orkney and middle and south Wales), and peaks at 3.7% in western England, with a comparable 3.5% in Norway. J*(x1,2) is absent in most groups, and is nowhere more frequent than 1.6% (north Wales). T1 peaks at 5.9% in Orkney, and is completely absent from north Wales and Holland. T2 finds its highest frequency in west Scotland (14.5%)

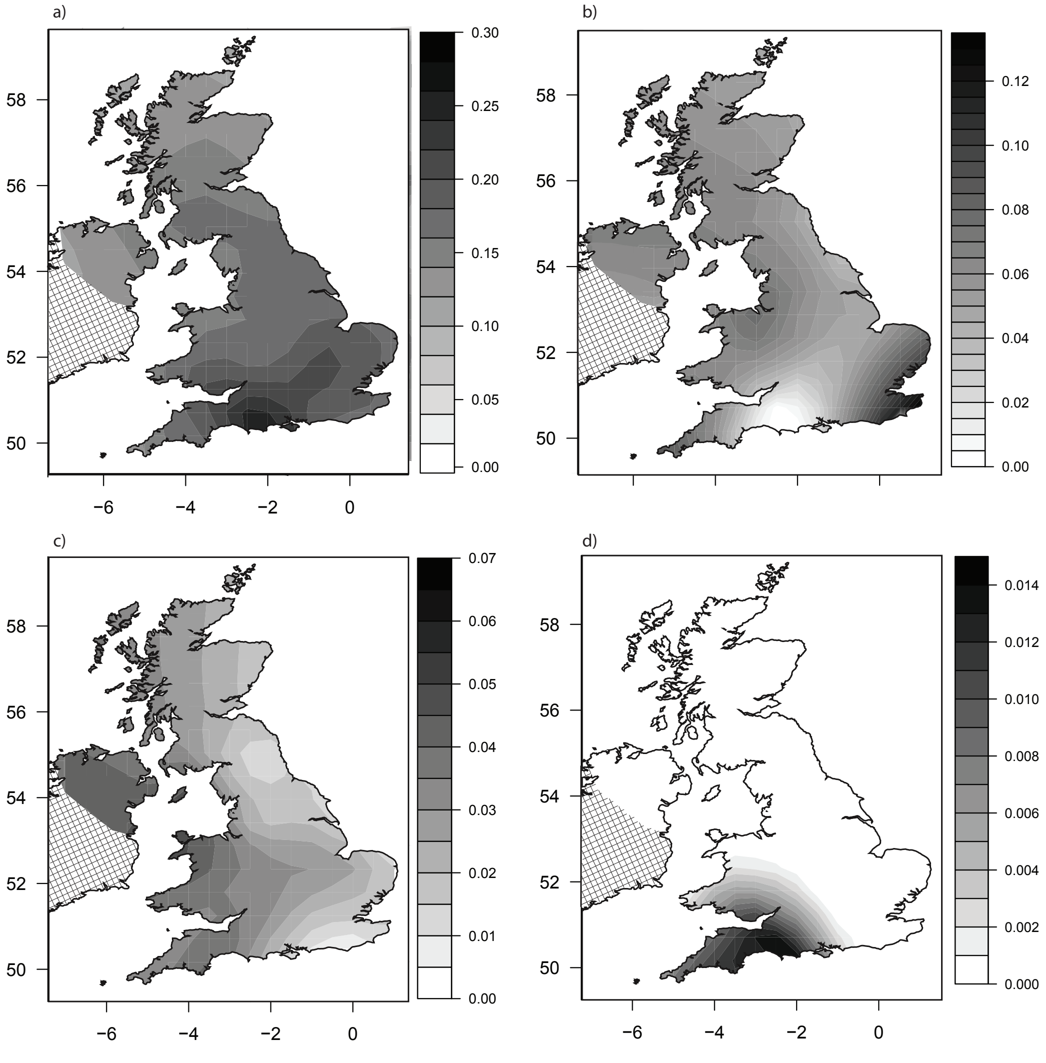


Figure 6.1: **H sub-haplogroup frequency distributions** a) H1, b) H3, c) H4, d) H7a. Note that the frequency scale varies.

and its lowest in middle Wales (3.4%), and Brittany (2.5%). For most groups T2 outnumbers T1 mtDNAs by 1 to 3-4, but Orkney has nearly equal levels, with T1 actually predominating over T2 (5.9 versus 4.3%). T*(x1,2) is absent from most groups, and only present in central and south central England, and south Wales where it reaches 3%. JT mtDNAs belonging neither to J nor T were very rare, with no group having an occurrence above 0.6%.

Table 6.4: **J/T haplogroups frequencies in Britain** N is the total number of individuals typed for the region, the frequencies given are percentages of the total for all mitochondrial haplogroups.

Region	**JT***	**J***	**J1**	**J2**	**T***	**T1**	**T2**	**N=**
Central England	0.0	0.0	10.0	1.2	0.8	0.8	7.5	241
Eastern-central England	0.3	0.0	8.3	2.8	0.0	1.4	9.4	361
Eastern England	0.0	0.0	6.4	2.1	0.0	3.2	7.8	283
Eastern Scotland	0.0	0.0	22.2	0.0	0.0	0.0	8.3	72
Middle Wales	0.0	1.1	10.2	0.0	0.0	3.4	3.4	88
North Wales	0.0	1.6	17.8	0.8	0.0	0.0	5.4	129
Northeast England	0.0	0.0	12.0	1.2	0.0	2.0	10.8	250
Northwest England	0.3	0.0	8.5	1.1	0.0	4.0	11.0	354
Orkney	0.0	0.0	15.0	0.0	0.0	5.9	4.3	187
South-central England	0.4	0.0	10.4	1.5	0.4	0.8	5.8	260
South Wales	0.0	0.0	7.4	0.0	3.0	0.7	9.6	135
Southeast England	0.6	0.0	6.4	1.7	0.0	2.9	8.7	173
Southwest England	0.2	0.2	9.1	0.5	0.0	1.5	6.4	407
Western England	0.0	0.0	9.1	3.7	0.0	1.1	8.0	187
Western Scotland	0.0	0.0	9.6	2.4	0.0	1.2	14.5	83
Yorkshire	0.0	0.0	10.6	1.9	0.0	2.7	8.0	263
Brittany	0	0	7.2	1.2	0	1.3	2.5	750
Denmark	na	12.6			na	9.9		9251
Holland	0	0	9.2	1.1	1.1	0	10.3	87
Northern France	na	7.3				10.4		193
Norway	0	0	9.3	3.5	0	1.2	4.7	86
Ulster	0.6	0	12.8	3	0.6	1.8	5.5	164

Among the thirteen subgroups of haplogroup U, only five reached a frequency of over 2% in any group. K*(x1,2), U*(x1,3-8), U1, U3, U5*(x5a,5b), U6, U7 and U8*(xK) have very low frequencies, 0 or less than 1% for most groups. K1 frequencies range from 3.5% (southeast England) to 10.4% (northeast England), and K2 from 0% (Orkney) to 4.5% (middle Wales). U4 frequencies are low but largely uniform; it is completely absent from Wales, but otherwise is present at levels between 1.1-3.7% in Britain, with the greatest density in southwest England. Holland, northern France and Ulster have higher U4 levels, of 5.7, 4.1 and 4.3%, respectively. Most U5 mtDNAs belong to either U5a or U5b. U5a is at its lowest in north Wales (0.8%) and its highest in northwestern England (7.3%), while U5b is at its lowest in eastern Scotland (1.4%) and Holland (1.1%), and highest in Orkney (8%).

Table 6.5: **U and K haplogroups frequencies in Britain and northwest Europe** N is the total number of individuals typed for the region, the frequencies given are percentages of the total for all mitochondrial haplogroups.

Region	K*	K1	K2	U*	U1	U3	U4	U5*	U5a	U5b	U6	U7	U8*	N=
Central England	0.0	7.1	3.7	0.8	0.4	0.0	1.2	0.0	4.6	2.5	0.0	0.0	0.0	241
Eastern-central England	0.0	6.4	0.6	0.6	0.0	0.0	1.9	0.0	6.1	5.3	0.0	0.0	0.0	361
Eastern England	0.0	6.0	0.7	0.7	0.0	1.1	3.5	0.0	4.9	2.5	0.4	0.0	0.0	283
Eastern Scotland	0.0	6.9	1.4	0.0	0.0	1.4	2.8	0.0	2.8	1.4	0.0	0.0	0.0	72
Middle Wales	1.1	5.7	4.5	0.0	0.0	1.1	0.0	0.0	1.1	3.4	0.0	0.0	0.0	88
North Wales	0.0	3.9	1.6	0.8	0.0	0.0	0.0	0.0	0.8	6.2	0.0	0.8	0.0	129
Northeast England	0.0	10.4	0.8	0.8	0.0	0.0	1.6	0.0	6.4	2.4	0.0	0.0	0.4	250
Northwest England	0.0	7.3	2.0	1.1	0.0	2.0	1.1	0.0	7.3	2.3	0.3	0.0	0.3	354
Orkney	0.0	7.0	0.0	0.0	0.0	0.0	2.1	0.0	4.8	8.0	0.0	0.0	0.0	187
South central England	0.0	6.9	1.2	0.8	0.0	0.8	1.5	0.0	6.5	5.8	0.0	0.0	0.0	260
South Wales	0.0	9.6	3.7	0.7	0.0	0.7	0.0	0.0	5.2	5.2	0.0	0.0	0.7	135
Southeast England	0.0	3.5	2.3	0.0	0.6	0.6	2.9	0.6	5.8	3.5	0.0	0.0	0.0	173
Southwest England	0.2	5.2	2.7	0.2	0.0	0.2	3.7	0.0	2.7	4.9	0.2	0.0	0.2	407
Western Engalnd	0.0	6.4	1.1	0.0	0.0	0.0	1.1	0.0	4.3	4.3	0.0	0.5	0.0	187
Western Scotland	0.0	3.6	1.2	1.2	0.0	0.0	1.2	1.2	4.8	3.6	0.0	0.0	0.0	83
Yorkshire	0.4	7.2	0.4	1.1	0.0	1.1	1.5	0.4	4.9	4.6	0.4	0.4	1.1	263
Brittany	0	8.7	3.5	0.3	0	2	1.7	0	5.3	6.4	0.3	0	0	750
Denmark	6.2				15.9									9251
Holland	0	10.3	2.3	0	0	0	5.7	1.1	1.1	1.1	0	0	0	87
Ireland	12.2			1.3	0	1	1.3	0	3.5	4.8	0	0	0	312
Northern France	10.4			1.6	0	1	4.1	7.8			na	0	na	193
Norway	5			0	0.2	1.2	2.8	0	6.7	5.1	0	0	0.5	642
Ulster	0	6.1	1.2	3	0	1.8	4.3	0	6.7	4.9	0.6	0	0	164

HV mtDNAs not belonging to either H or V are rare, with no examples in Orkney, and 3.8% in Yorkshire.[3] Eastern Scotland and south central England had no examples of haplogroup W, and the distribution peak was in Holland at 5.7%. N1e mtDNAs, most of which belong to haplogroup I, were most common in south Wales (11.9%) and least common in western Scotland (0%). Haplogroup V frequencies ranged from 0% in eastern Scotland and Orkney, to 5.4% north Wales. Haplogroup X has no examples in

---

[3] If one excludes Ireland, northern France and Norway on the assumption that, V not being typed in the sources, their HV* cohort includes V individuals. Among the 86 Norwegians typed for this study, only two V individuals were observed (2.3%), both from Ottadalen. V reaches a frequency of 5.4% in Norway in Passarino *et al.* [549]. These frequencies – if they are representative of the true Norwegian V level – would account for the vast majority of HV* mtDNAs, and the case is likely the same for northern France and Ireland.

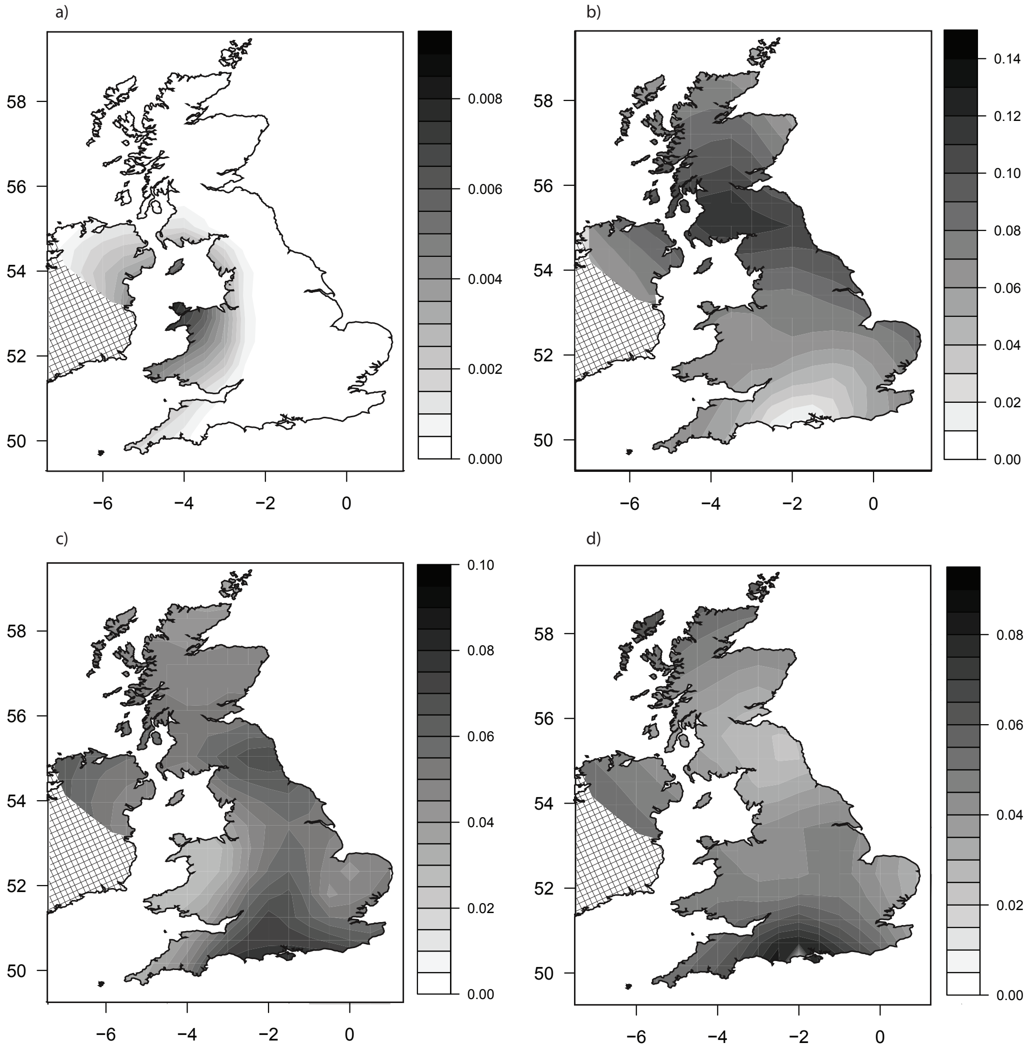


Figure 6.2: **Haplogroup frequency distributions** a) J*(x1,2), b) T2, c) U5a, d) U5b. Note that the frequency scale varies.

middle Wales, but reaches 7.5% in Orkney, with western England following, at 3.2%. Unresolved haplogroups were absent in central and southeast England, eastern Scotland and all of Wales, and infrequent in all other groups, with southwest England having the most at 2.2%. All but one of the unresolved chromosomes was found in Cornwall, and the majority appeared to belong to haplogroup U, with incongruous additional variants present.

Table 6.6: **Non-major haplogroup frequencies in Britain** N is the total number of individuals typed for the region, the frequencies given are percentages of the total for all mitochondrial haplogroups. HV* - HV*(xH,V).

Region	**HV***	**W**	**N1e-I**	**other**	**V**	**X**	**N=**
Central England	1.7	2.1	4.1	0.0	5.0	1.2	241
East-central England	2.5	1.9	3.6	1.1	1.4	1.9	361
Eastern England	1.4	3.5	3.2	1.1	3.5	0.7	283
Eastern Scotland	1.4	0.0	4.2	0.0	0.0	1.4	72
Middle Wales	1.1	2.3	6.8	0.0	3.4	0.0	88
North Wales	2.3	3.1	1.6	0.0	5.4	0.8	129
Northeast England	2.0	0.0	2.8	0.8	0.4	1.6	250
Northwest England	2.0	1.4	1.7	0.8	3.1	1.1	354
Orkney	0.0	1.1	7.5	1.1	0.0	7.5	187
South-central England	1.2	0.0	3.8	1.2	3.8	0.8	260
South Wales	2.2	2.2	11.9	0.0	4.4	2.2	135
Southeast England	1.7	4.6	4.0	0.0	2.9	1.7	173
Southwest England	1.0	1.2	2.5	2.2	2.7	1.7	407
Western England	2.1	1.6	7.0	1.1	3.7	3.2	187
Western Scotland	3.6	3.6	0.0	1.2	4.8	2.4	83
Yorkshire	3.8	0.8	2.7	0.8	5.3	0.8	263
Brittany	1.3	3.5	4.9	3.3	3.1	0.7	750
Denmark	na	0.4	na	1.4	4.5	na	9251
Holland	1.1	5.7	2.3	6.9	2.3	2.3	87
Ireland	6.7	na	3.2	2.9	na	1	312
Northern France	6.2	3.6	0.5	0	na	1	193
Norway	3.4	na	2	4.2	na	0.3	642
Ulster	1.8	1.8	1.8	0	3.7	1.2	164

Fisher's exact tests for haplogroup frequencies between populations[4] showed only low levels of differentiation overall. Of H subhaplogroups, H4 showed north Wales to be different from east England, northwest England and Yorkshire, and H5 differentiated Norway from Brittany, Ulster, north Wales and southeast, south-central, northeast and

---

[4] Only those typed for this study.

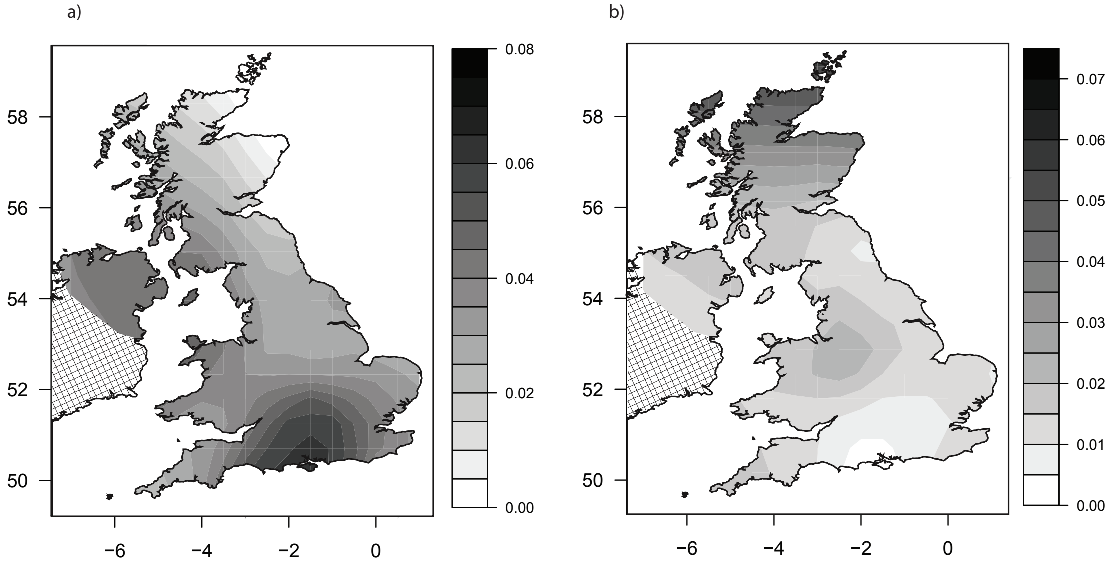


Figure 6.3: **Haplogroup frequency distributions** a) V, b) X. Note that the frequency scale varies.

east England. East Scotland's high H13 frequency serves to separate it from Brittany only. The comparatively high level of unattributed T mtDNAs (i.e. non-T1/2) in south Wales separates it from Brittany, and the low T2 frequency in the latter separates it from west Scotland, and east-central, northeast and northwest England. The very high X frequency in Orkney sets it apart from Yorkshire, northwest and east England, Brittany and Norway. The fact that Brittany had by far the largest regional sample size may account for the many differences between it and other areas that reached a level of significance.

Principal components analysis of mitochondrial haplogroup frequencies roughly supports these findings, particularly with respect to outliers. Figure 6.4 shows plots of PCs 1-4, in which PC1 accounts for 23% of the variance, PC2 15%, PC3 13%, and PC4 11% of the variance in the data. Incuded in the plot are data from Brittany, Holland, and Norway. The former two show some tendency to separate from Britain and Brittany for PCs 1 and 2, but less so for Holland than for Ulster, Orkney, north Wales and east Scotland. Brittany, notably, fits squarely into the main British cluster.

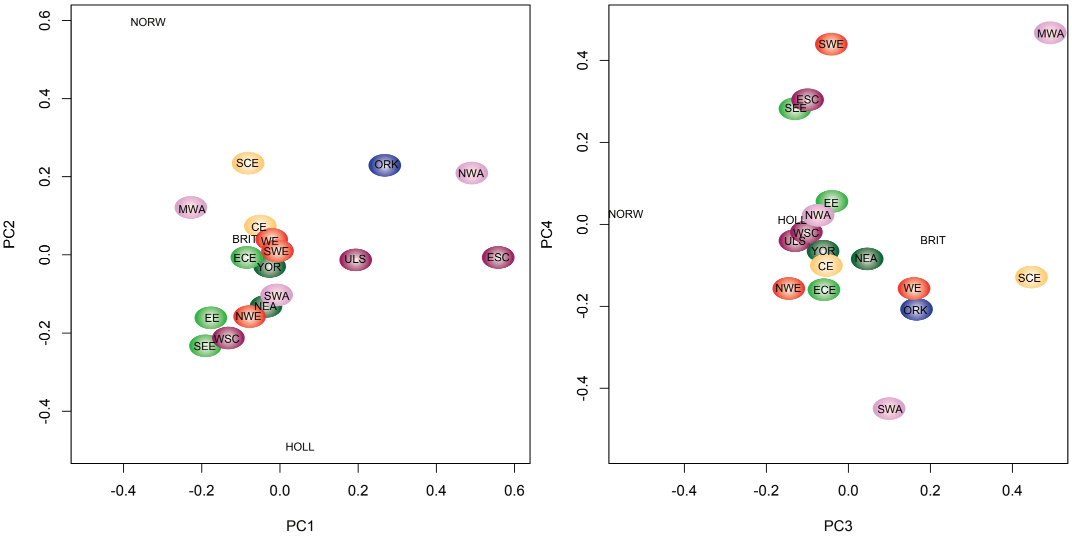


Figure 6.4: **Principal components plots of mitochondrial variation** Left: PC1 by PC2, accounting for 23% and 15%, respectively. Right: PC3 by PC4, accounting for 13% and 11%, respectively. BRIT – Brittany, CE – Central England, ECE – East-central England, EE – Eastern England, ESC – Easter Scotland, HOLL – Holland, MWA – Middle Wales, NEA – Northeast England, NORW – Norway, NWE – Northwest England, NWA – North Wales, ORK – Orkney, SCE – South-central England , SEE – Southeast England, SWA – South Wales, SWE – Southwest England, ULS – Ulster, WE – Western England, WSC – Western Scotland, YOR – Yorkshire. Colour-coding implies larger regions; red – western England, yellow – central England, light green – eastern England, dark green – northeast England, blue – Orkney, violet – Wales, and maroon – Scotland and northern Ireland.

6.2 Genome-wide SNP array data

The groupings based on a provisional fineSTRUCTURE analysis (see Appendix K), which are used for this section, can be found in Table K.1.

Table 6.7: **Cluster-based grouping**

Cluster	Major constituents
COR	89% of Cornish samples
DEV	84% of Devon samples
ENGY	Majority of all Lowland British samples
NENG	86% of Cumbrian, and 62% of northeast English samples
ORK	99% of Orcadian samples
SCO	98% of eastern Scottish, 100% of western Scottish, 100% of northern Irish, 9% of northeast English, and 6% of Cumbrian samples
WAL	100% of south Welsh, 99% of north Welsh, and 15% of Herefordshire samples
WEM	98% of Forest of Dean, 78% of Herefordshire, 33% of Worcestershire, 30% of Cheshire, 19% of Lancashire, and 17% of Gloucestershire samples

6.3 Average F_{ST} between population groupings

Background F_{ST} levels between population groupings were obtained calculating the measure for each SNP on the Illumina array (1 million), and the subset of non-synonymous coding SNPs (18,500), in order to assess departures from the average for genetic loci of interest. A graphical representation of the genome-wide averages is given in Figure 6.5 (the pattern for non-synonymous SNPs is nearly identical, with only the magnitude of differences varying, and the same basic pattern was also repeated for all the genetic subcategories investigated). Overall, the larger groupings (ENGY and NENG) show smaller F_{ST} values, and Orkney stand and Wales out with elevated average pairwise values compared to other groups.

A multidimensional scaling analysis of F_{ST} values was also performed, to give some indication of what magnitude and direction of population pairwise distances might be expected. The results are plotted in Figure 6.6. This method shows WAL, ORK, and COR at extreme edges of the plot, with WEM and SCO approaching WAL, DEV very

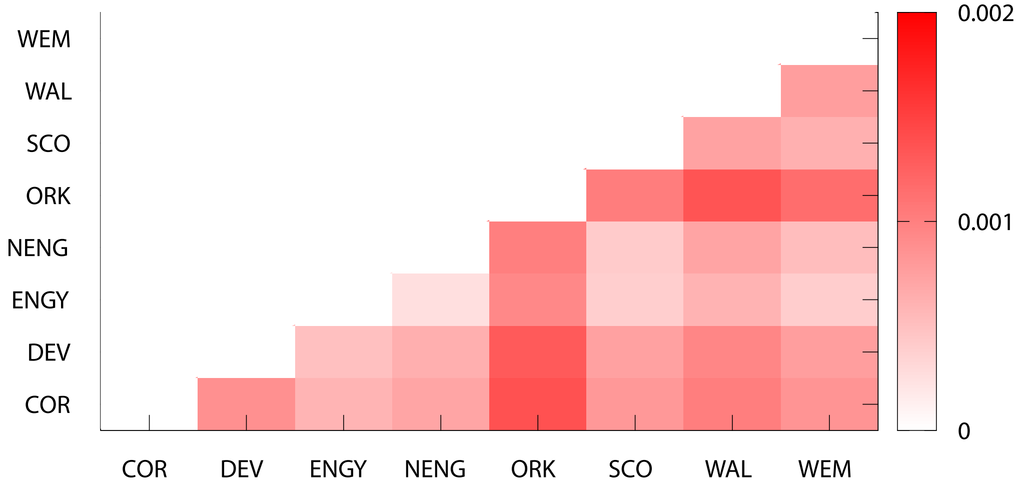


Figure 6.5: **Average F_{ST} between population groupings** Estimates of average F_{ST} between population groups, based on all available SNPs, genome-wide.

close to COR, and ENGY/NENG clustered near COR/DEV.

6.4 Classical marker genes and lactase persistence

Twenty-five of the original 37 classical marker genes had non-synonymous coding SNPs present on the Illumina 1M-Duo array, and so were included for analysis. The overall picture of differentiation – by F_{ST} – is similar to that for the genome-wide SNP and overall non-synonymous SNP data, with ORK showing the highest population pairwise values, followed by WAL and SCO. A multidimensional scaling plot of the values demonstrates the outlying position of ORK and WAL especially, and a closer association of SCO and NENG (see Figure 6.7).

A number of genes exhibited F_{ST} SNP values that exceeded the 95th quantile distribution of the genome-wide non-synonoymous SNP background for population pairwise comparisons. The distributions of both categories are illustrated in Figures 6.8 and 6.9, which include the greatest outlier proportions (not necessarily the most extreme F_{ST}

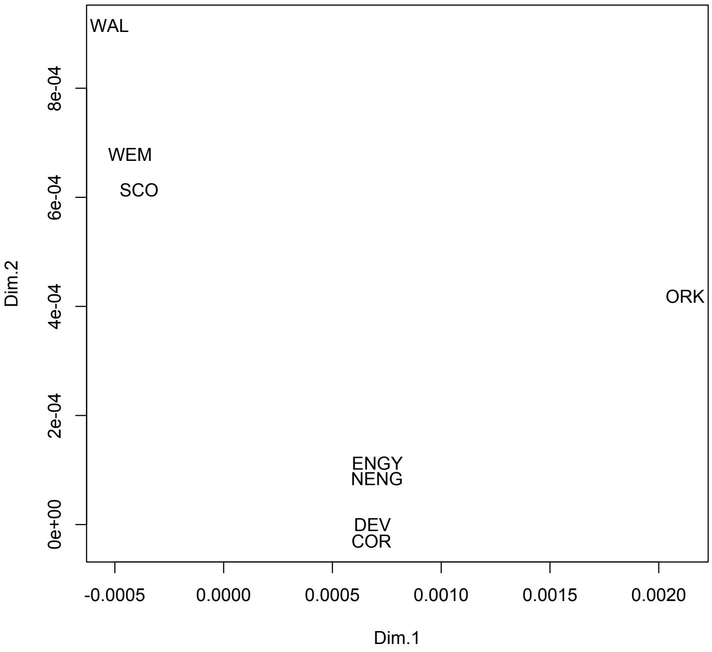


Figure 6.6: **Multidimensional scaling plot of genome-wide $\mathbf{F}_{ST}$ values.**

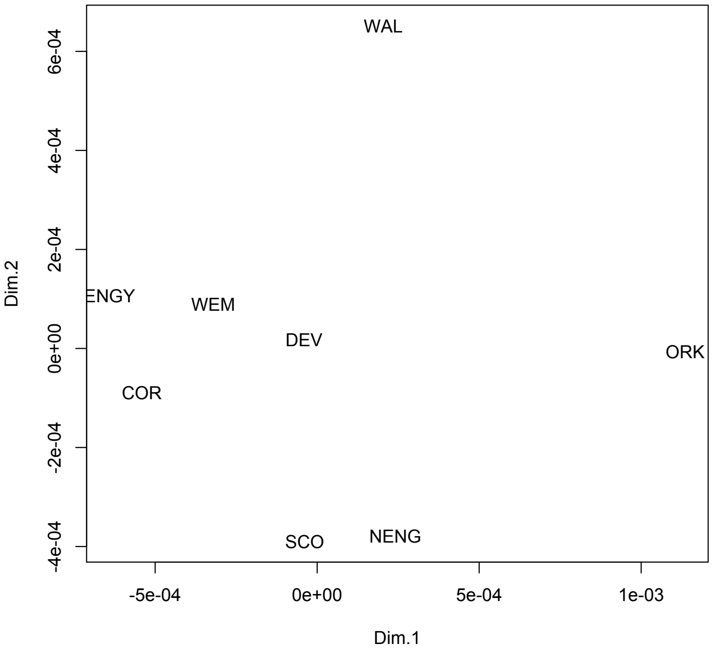


Figure 6.7: **Multidimensional scaling plot of classical marker gene $\mathbf{F}_{ST}$ values.**

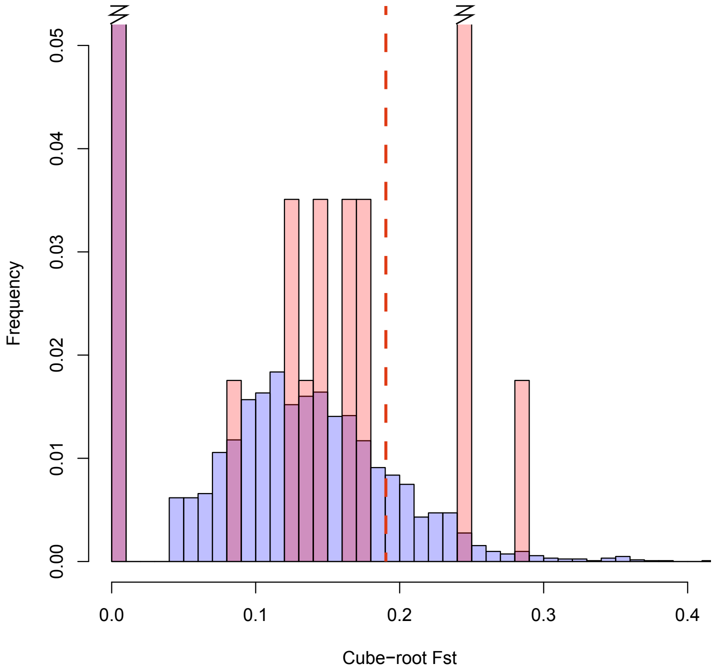


Figure 6.8: **ENGY vs. ORK F$_{ST}$ values** Frequency-scaled histograms of cube-root transformed F$_{ST}$ values for the ENGY-ORK pair, where genome-wide values are in blue and values for analysed genes are in pink, and the overlap is purple. The red dashed line indicates the limit of 95% CI for the background (blue) non-synonymous SNP distribution. The y-axis is scaled down to show variation, frequencies over 5% are indicated by a chevron over the bin. Note that the vast majority of values for both distributions are contained in the first, purple bin with F$_{ST}$ values approximating zero.

measures).

All the genes are ranked in order of increasing number of F$_{ST}$ values over the 95th quantile of genome-wide non-synonymous SNP values for 28 population pairwise comparisons (see Table 6.8), where *FUT2* and *TF* rank the highest with – by this token – significant values in 7 different population group comparisons. Of the top five ranked genes, *SERPINA1* separates ENGY from DEV, WAL, and WEM, DEV from ORK, and NENG from WEM. All high *SLC14A1* values separate SCO from both its neighbours, COR, DEV and WAL. *CFB* largely separates Wales and its hinterland from other groups (WAL from COR, DEV and SCO, NENG from ENGY and WEM, and DEV from WEM). *FUT2* separates the Scottish/Northern Irish group from all other

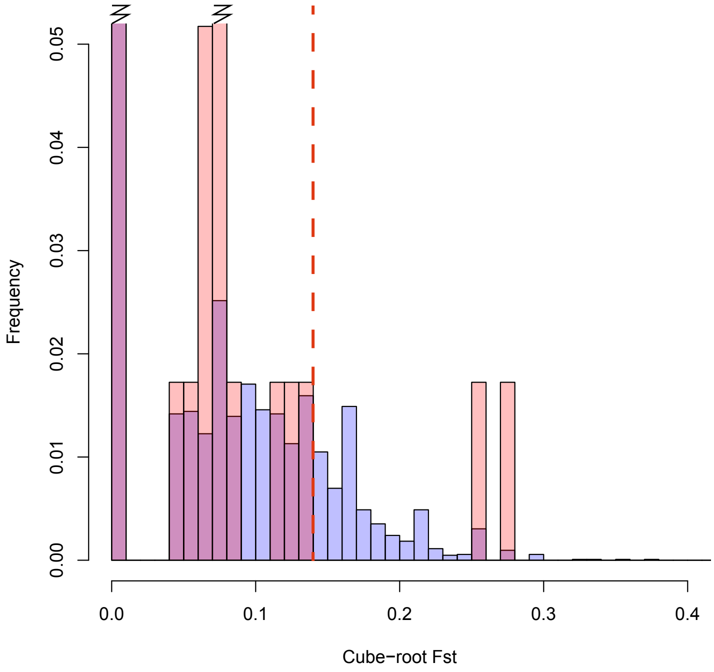


Figure 6.9: **ENGY vs. SCO F$_{ST}$ values** Frequency-scaled histograms of cube-root transformed F$_{ST}$ values for the ENGY-SCO pair, where genome-wide values are in blue and values for analysed genes are in pink, and the overlap is purple. The red dashed line indicates the limit of 95% CI for the background (blue) non-synonymous SNP distribution. The y-axis is scaled down to show variation, as the majority of values are proximate to zero, frequencies over 5% are indicated by a chevron over the bin. Note that the vast majority of values for both distributions are contained in the first, purple bin with F$_{ST}$ values approximating zero.

populations, and *TF* COR from NENG, ORK and SCO, ENGY from NENG and WAL, NENG from WEM, and SCO from WAL.

Table 6.8: **Classical marker genes by high $\mathbf{F}_{ST}$** 95th Q – the number of population pairwise comparisons (of a total 28) in which the gene listed has at least one SNP whose F_{ST} measure exceeds the non-synonymous SNP 95th quantile background threshold.

Gene	95th Q	Gene	95th Q
ACP1	0	BCAM	2
GC	0	DARC	2
GPT	0	HLAB	2
KEL	0	PGM1	2
MDH1	0	CNDP2	3
PGM2	0	ABO	4
RHD	0	C3	4
SLC4A1	0	SERPINA1	5
TAS2R38	0	SLC14A1	5
AK1	1	CFB	6
FUT3	1	FUT2	7
RHCE	1	TF	7

The top five hits in Table 6.8 (*TF, FUT2, CFB, SLC14A1, SERPINA1*) were chosen to examine further, and the highest ranking SNP present in each gene was chosen to investigate the extent of homozygosity around it. The SNP ranking within each gene was a product of the number of population pairwise comparisons for which the SNP exceeded the 95th quantile background threshold and the magnitude of the F_{ST} measures themselves. The selected SNPs were: *TF – rs1049296, FUT2 – rs602662, CFB – rs4151659, SLC14A1 – rs1058396, SERPINA1 – rs1303.* The homozygosity plots for a 2.5Mb window around the given SNPs can be found in Figure 6.22.

Of the classical marker genes, only *FUT2* shows any indication of differing extents of homozygosity surround the selected SNP, with the derived allele having a slightly extended runs of homozygosity compared to the ancestral allele. The number of individuals heterozygotic for the *CFB* derived allele (n=2) is too small for comparison.

The allele that is co-occurrent with lactase persistence in Europeans is present in our samples at a maximum of 86% (Ulster) and a minimum of 65.5% (Suffolk).

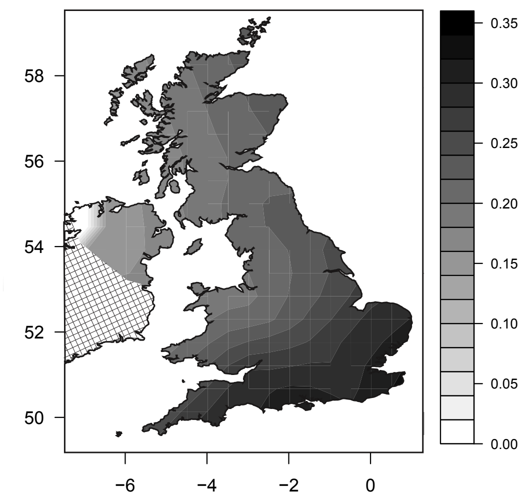


Figure 6.10: **Lactose intolerance allele frequency** The frequency surface shows the negative of the lactase persistence allele, i.e. lactose intolerance is increased in the south and east.

It has a striking distribution, where the south and east of Britain has the highest levels of predicted lactose intolerance, and the northwest the least (see Figure 6.10). A chi square heterogeneity test shows significant differences between ENGY and the northern/western extremes of Britain (including Ulster), i.e. SCO, WAL and even WEM (see Table 6.9). Orkney, in the extreme north, breaks the overall pattern with a frequency of 74%.

Table 6.9: **Lactase persistence differences** Chi square p-values for the lactase persistence variant. Significant differences are in bold.

	COR	DEV	ENGY	NENG	ORK	SCO	WAL	WEM
COR	9.0e-1	8.0e-1	4.1e-1	7.6e-1	6.9e-1	3.8e-2	8.9e-2	1.0e-1
DEV		9.0e-1	7.6e-1	4.3e-1	1.0	1.2e-2	3.3e-2	4.1e-2
ENGY			9.7e-1	2.4e-2	8.7e-1	**1.3e-5**	**3.1e-4**	**1.0e-3**
NENG				9.4e-1	3.1e-1	2.2e-2	7.5e-2	9.4e-2
ORK					9.1e-1	4.9e-3	1.7e-2	2.3e-2
SCO						9.2e-1	8.3e-1	8.9e-1
WAL							9.1e-1	9.4e-1
WEM								9.0e-1

F_{ST} departures from the average for this allele likewise separate SCO, WAL, and WEM very positively from the rest of the populations present (see Figure 6.11), while

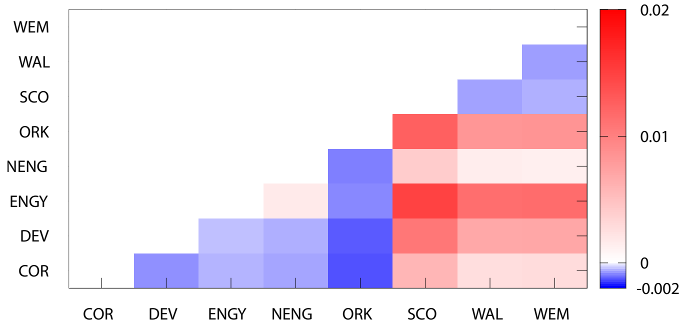


Figure 6.11: **Lactose intolerance allele $\mathbf{F}_{ST}$ differences** Pairwise F_{ST} depatures from the inferred genomic averages, the bar to the right of the plot gives the magnitude of the depature.

the remainder show a mild reduction in average pairwise F_{ST}.

6.5 Pigmentation: phenotypes and genotypes

To examine pigmentation across Britain, I first explore some historical surveys of phenotype, and link these results to contemporary regional genotypes, before examining wider genetic data.

6.5.1 Observed phenotypes

6.5.1.1 Late 19th-early 20th century phenotypes

Hair and eye colour for 10,207 rural people people from all over Britain were taken from Parsons 1920 [550], which comprises observations collected in the late 19th/early 20th century. Eye colour was divided into two categories, light and dark, and hair colour divided into five shades: red, blonde, brown, dark, and black. These data were used both to test if there were regional differences in hair and eye colour pigmentation in Britain, and to function as a proxy for the accuracy of phenotype prediction based on genotypes from PoBI individuals from the same areas. Britain was separated into fifteen regions for this purpose (see Table 6.10 for details).

Table 6.10: **Regions for 1920 [550] observed phenotypes**

Region	**Constituents**
Central England	Bedfordshire, Northamptonshire, Oxfordshire
East-central England	Leicestershire, Lincolnshire, Nottinghamshire
Eastern England	Cambridgeshire, Essex, Norfolk, Suffolk
Eastern Scotland	Perthshire
Northeast England/ Southeast Scotland	Durham, Northumberland, Berwickshire, Roxburgshire, Midlothian
Northwest England	Cumbria, Lancashire
North Scotland	Ross and Cromarty
North Wales	North Ceredigion, North Powys, Merioneth, Caernarfonshire, Denbighshire
South-central England	Gloucestershire, Hampshire, Somerset, Wiltshire
Southeast England	Kent, Sussex
South Wales	Brecon and Radnorshire, Glamorganshire, Carmarthenshire, Pembrokeshire
Southwest England	Cornwall, Devon
Western England	Herefordshire, Staffordshire, Warwickshire, Worcestershire
Western Scotland	Argyll, Wigtownshire
Yorkshire	Yorkshire

Observed frequencies of black hair were lowest in northeast England and Yorkshire, at 2.1 and 2.5%, respectively, and highest in Wales and southwest England, at 14.2 and 12.5%, respectively (see Figure 6.12).[5] Dark hair was also most common in Wales, southwest England and Scotland, with a maximum frequency of 43% in south Wales and western Scotland, and a minumum of 24-25% in northwest and northeast England. Brown and blonde hair frequency distributions show a reverse trend from black and dark hair. Brown hair is the least common in south Wales (27%) and most common in east and southeast England (44.5%), and fair hair is also least common in south Wales (9%) and southwest England (12%), and most common in northwest and northeast England (26-27%). Red hair was the rarest colour category, and the least different from region to region, with a minimum of 3% in eastern and western England, and

---

[5] Wales is here represented only by south Wales, as the data for north Wales does not distinguish dark from black hair. However, the combined dark/black frequency for north Wales (78%) is noticeably higher than for south Wales and southwest England (57% and 55%, respectively), and it is therefore likely that the real frequency of black hair in Wales as a whole is even higher.

7-8% in Yorkshire and eastern Scotland, respectively.

For all shades of hair colour barring red, Wales and southwest England, with their high incidence of black and dark hair and low incidence of blonde hair, were significantly different from most other regions, in particular from areas in eastern England (see Figure 6.13). Dark hair also separates western Scotland from some areas of England, mostly eastern (east-central, northeast, northwest, and southeast England and Yorkshire), and northeast England is significantly different from western and central England and eastern Scotland as well. The high level of blonde hair in northeast England distinguishes it from western and eastern Scotland, central, south central, and eastern and western England, as well as from Wales and southwest England. For red hair, the areas with highest red hair frequency (eastern Scotland, northeast England and Yorkshire) were significantly different only from a few groups, western and eastern England. Northern Scotland is poorly represented compared with the other regions, with only fifty samples, and is therefore less likely to show significant differences from any other group.

Eye colour, divided into light (blue, green, grey, etc.) and dark (brown), distributions revealed a broadly similar pattern to that of the major hair colors (see Figure 6.12). North Wales had the lowest frequency of light eyes at 48.6%, followed by south Wales (58.3%) and northern Scotland the highest, 82.4%. North Wales is significantly different from all other areas but south Wales (see Figure 6.13), and south Wales/southwest England are significantly different from eastern Scotland, northeast and northwest England and Yorkshire. In addition, northeast England is distinguished from central, south central, east central, eastern and western England by a high frequency of light eyes (75.7%). Northern Scotland and northwest England have an even higher frequency of light eyes, but both have comparatively small samples sizes.

6.5.1.2 PoBI phenotypes

Of the five hair colour shades from the observed late 19th-early 20th century dataset, only three had sufficiently strong associated genetic variants to attempt phenotype pre-

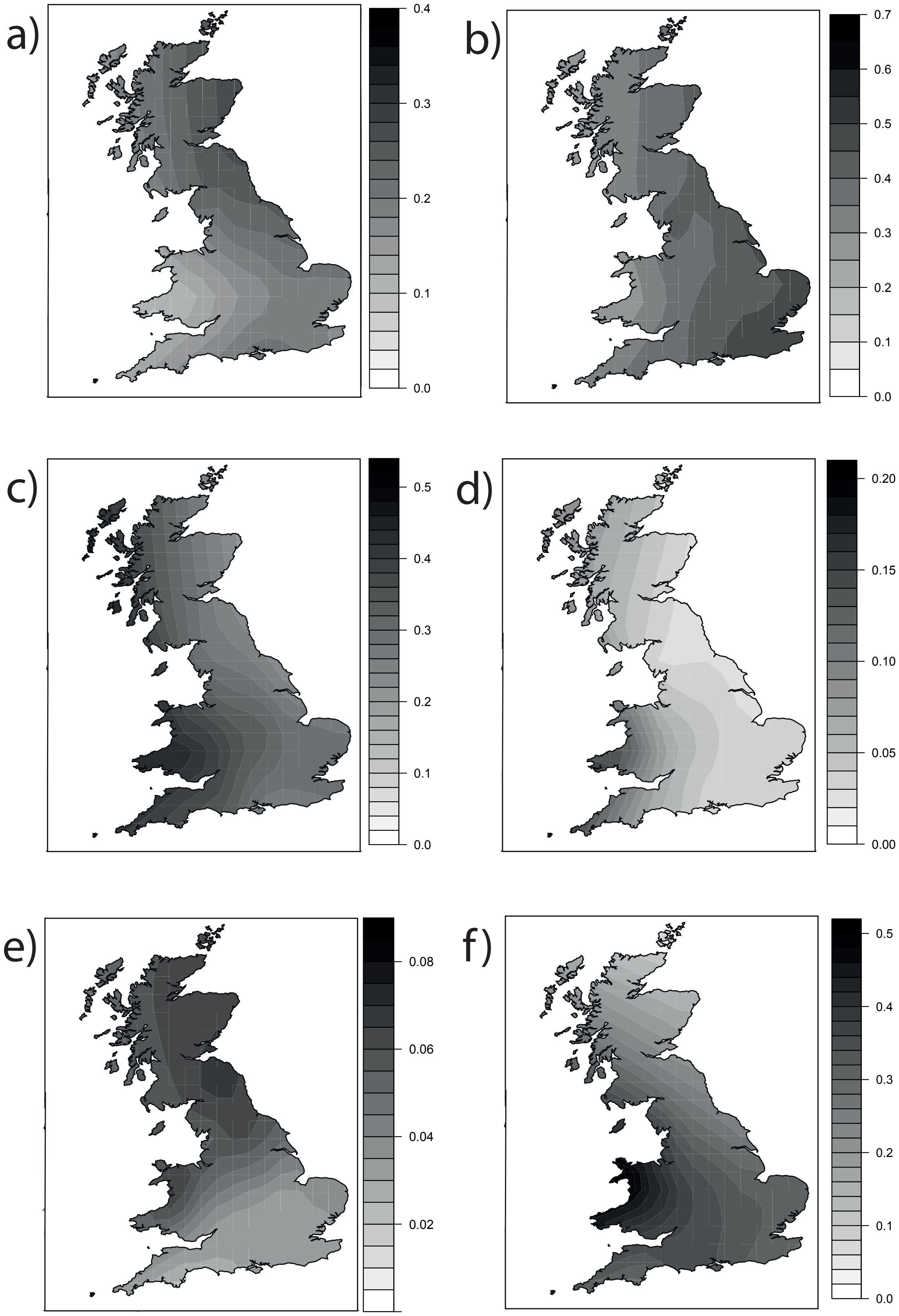


Figure 6.12: **Eye and hair colours observed in late-nineteenth/early-twentieth century Britain [550]** Frequency distributions of a) fair, b) brown, c) dark, d) black and e) red hair, and f) brown eyes. Note that the frequency scale varies.

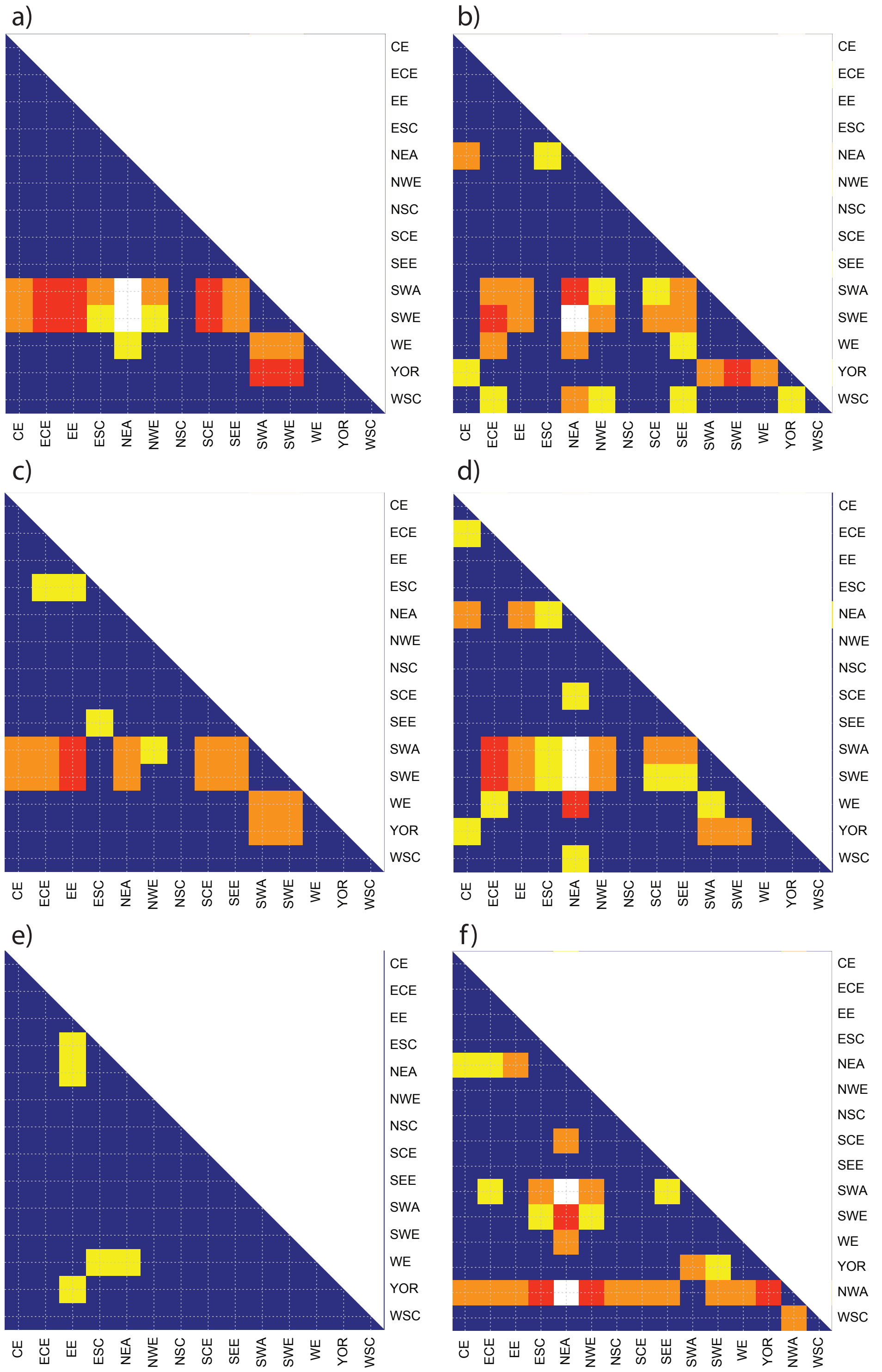


Figure 6.13: **Hair and eye colour differentiation** Heatmap of p-values for pairwise cluster comparisons of colour frequencies; a) black hair, b) dark hair, c) brown hair, d) fair hair, e) red hair, f) light eyes. All colours but blue are below the Bonferroni-corrected significance threshhold; for a-e): yellow > 1.0e-5, orange < 1.0e-5, red < 1.0e-10, white < 1.0e-20; for f) white < 1.0e-15.

diction based on genotypes from the PoBI dataset (see Table 5.5, Chapter 5). In this section I will consider these *estimates*: predictions of regional phenotype frequencies from genetic data. These were red, blonde and black. Light eye frequency was estimated as well, having a near perfect association with the given genetic variant (Table 5.5, Chapter 5). Estimated values of colour frequencies given genotype are presented in Table 6.11, as are the observed values. The colour frequency estimates based on genotypes from the PoBI cohort match the observed frequencies from the 1920 cohort well for light versus dark eyes and for red hair (differences assessed by Fisher's exact tests, with Bonferroni-corrected significance thresholds, see Table 6.11 and Figure 6.14).

For most regions, the frequency of light eyes was underestimated with regards to the observed dataset, but the only significant departure for eye colour was in northwest England. Estimated frequencies for light eyes in Britain ranged between 51.6% (north Wales) to 70.4% (northeast England). In comparison, Ulster had a congruent level of 59%, while Norway and Brittany formed very different extremes, with 90% and 34%, respectively.

The frequency of red hair was also usually underestimated, but never significantly so. Black hair estimates likewise frequently fell short of the observed values, but significantly so only in two cases, south Wales and southwest England (there were no comparative observed data from north Wales). For southwest England and south Wales, the estimates for black hair are comparable to those for the rest of Britain, most between 1.3 and 3.6%[6] but the 1920 observed frequencies, as seen above, are very much higher (12.5% and 14.2%, respectively), and are significantly different from the rest of the island. Estimations of blonde hair frequencies were very often significantly lower than the observed values; underestimations sometimes being of the order of 20%. Among the few populations not showing high observed-estimated differences, only south Wales and southwest England had estimated values within 1.4-4.6 of the observed ones, such intervals as were common among the other phenotype/genotype comparisons discussed.

Skin colour, as expressed by the melanin index of the inner arm, was collected for 865

---

[6] The west Scotland estimate is an outlier at 6%.

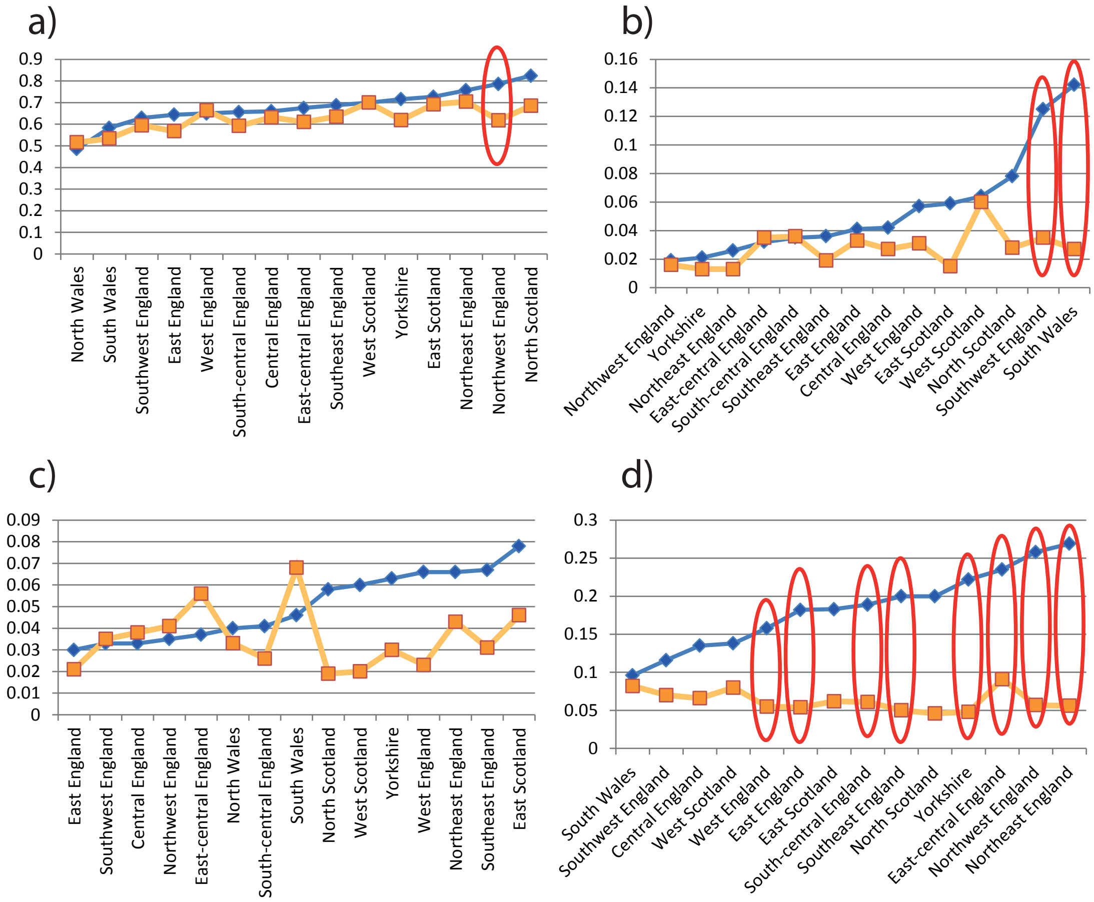


Figure 6.14: **Observed versus estimated pigmentation frequencies** Observed frequencies are in blue, estimated frequencies in yellow. Significant differences between observed and estimated frequencies are circled in red. a) Light eyes, b) black hair, c) red hair, and d) blond hair. Frequency scale is on the left of each plot.

Table 6.11: **Observed (1920) and estimated (PoBI) frequencies** Significant differences between observed and predicted frequencies are in underlined bold, based on Bonferroni-corrected p-values from Fisher's exact tests. Note that for '1920', Northeast England includes Southeast Scotland; and Northern Scotland in '1920' refers to the mainland, but for 'PoBI' it is represented by Orkney.

	Light eyes		**Black hair**		**Blonde hair**		**Red hair**	
Region	1920	PoBI	1920	PoBI	1920	PoBI	1920	PoBI
Central England	66.0	63.2	4.2	2.7	13.5	6.6	3.3	3.8
East Central England	67.5	61.0	3.2	3.5	**23.5**	**9.1**	3.7	5.6
East England	64.4	56.8	4.1	3.3	**18.2**	**5.4**	3.0	2.1
East Scotland	72.7	69.2	5.9	1.5	18.3	6.2	7.8	4.6
Northeast England	75.7	70.4	2.6	1.3	**26.9**	**5.6**	6.6	4.3
North Scotland	82.4	68.5	7.8	2.8	20.0	4.6	5.8	1.9
North Wales	48.6	51.6		2.2		5.5	4.0	3.3
Northwest England	**78.6**	**61.8**	1.9	1.6	**25.8**	**5.7**	3.5	4.1
South Central England	65.6	59.2	3.5	3.6	**18.9**	**6.1**	4.1	2.6
Southeast England	68.8	63.5	3.6	1.9	**20.0**	**5.0**	6.7	3.1
South Wales	58.3	53.4	**14.2**	**2.7**	9.6	8.2	4.6	6.8
Southwest England	62.9	59.5	**12.5**	**3.5**	11.6	7.0	3.3	3.5
West England	65.0	66.4	5.7	3.1	**15.8**	**5.5**	6.6	2.3
West Scotland	70.0	70.0	6.4	6.0	13.8	8.0	6.0	2.0
Yorkshire	71.6	61.9	2.1	1.3	**22.2**	**4.8**	6.3	3.0

individuals of the PoBI cohort, which were divided into five groups: Cornish (n=200), English without Cornish (n=452), Orcadians (n=86), Scots (n=54), and Welsh (n=73). Figure 6.15 shows the distribution of melanin index values in these groups.

Orcadians and Scots are significantly paler than the Cornish and other English[7], and Orcadians are also significantly paler than the Welsh (see Table 6.12).

Table 6.12: **Melanin index p-values** P-values for t-test comparing melanin index distributions in Britain. Numbers in bold indicate significant differences.

	Cornwall	England	Orkney	Scotland	Wales
Cornwall					
England	0.2910				
Orkney	**4.36e-11**	**1.96e-8**			
Scotland	**0.0001**	**0.0008**	0.6703		
Wales	0.0194	0.1340	**0.0018**	0.0203	

---

[7] Northern England was very poorly represented in these samples, and therefore could not form a separate group.

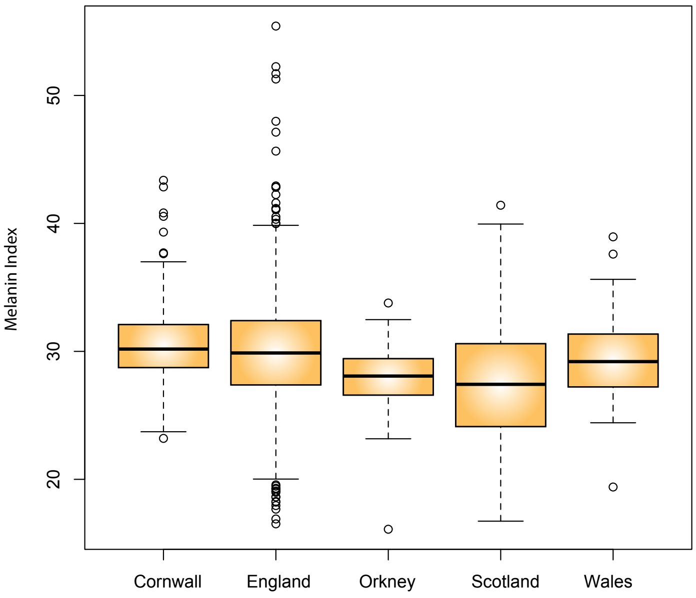


Figure 6.15: **Melanin index distributions** Boxplots showing the spread of melanin index measurements in Cornwall, England excepting Cornwall, Orkney, Scotland, and Wales.

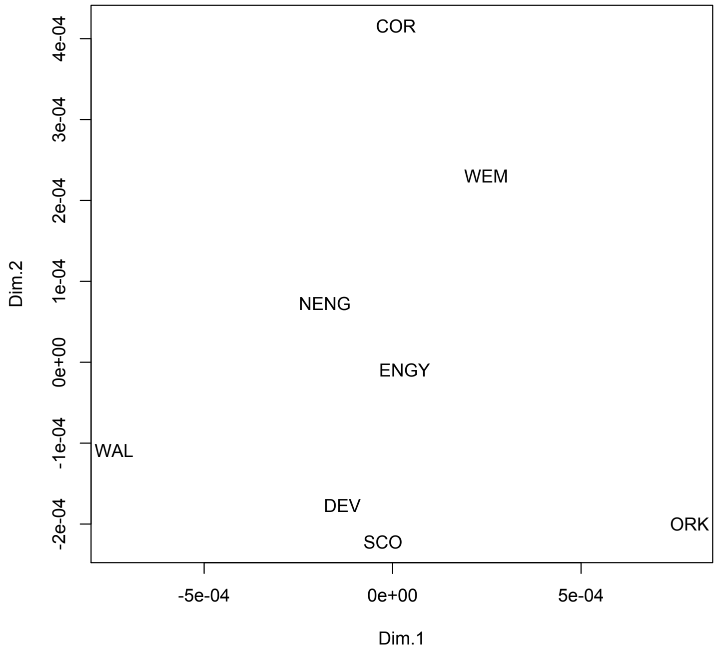


Figure 6.16: **Multidimensional scaling plot of pigmentation gene $\mathbf{F}_{ST}$ values**.

6.5.2 Pigmentation genes

The genetic distances between population groups for selected pigmentation- and melanin-related genes gave the same general pattern as the genome-wide analysis, with Orkney standing out as on average more different from other groups, followed by the Welsh group. A visual rendering of a multdimensional scaling of these distances demonstrates these 'outlier' groups, while also separating DEV from COR, and placing the former closer to SCO. NENG, ENGY and WEM occupy the middle of the plot (see Figure 6.16).

A number of genes exhibited F_{ST} SNP values that exceeded the 95 quantile distribution of the genome-wide background for population pairwise comparisons. The distributions of both categories are illustrated in Figures 6.17 and 6.18, which include the greatest outlier proportions (not necessarily the most extreme F_{ST} measures).

All the genes are ranked in order of increasing number of F_{ST} values over the 95th

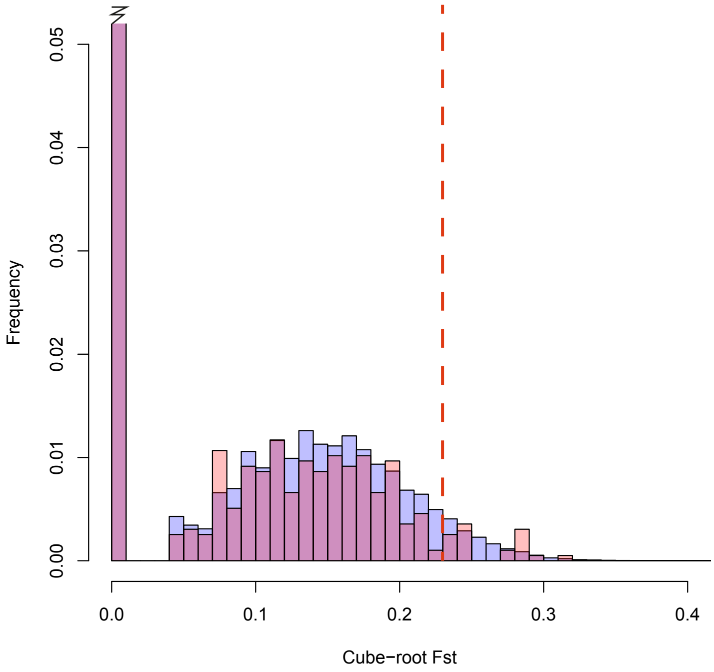


Figure 6.17: **COR vs. WEM F$_{ST}$ values** Frequency-scaled histograms of cube-root transformed F$_{ST}$ values for the COR-WEM pair, where genome-wide values are in blue and values for analysed genes are in pink, and the overlap is purple. The red dashed line indicates the limit of 95% CI for the background (blue) distribution. The y-axis is scaled down to show variation, as the majority of values are proximate to zero, frequencies over 5% are indicated by a chevron over the bin. Note that the vast majority of values for both distributions are contained in the first, purple bin with F$_{ST}$ values approximating zero.

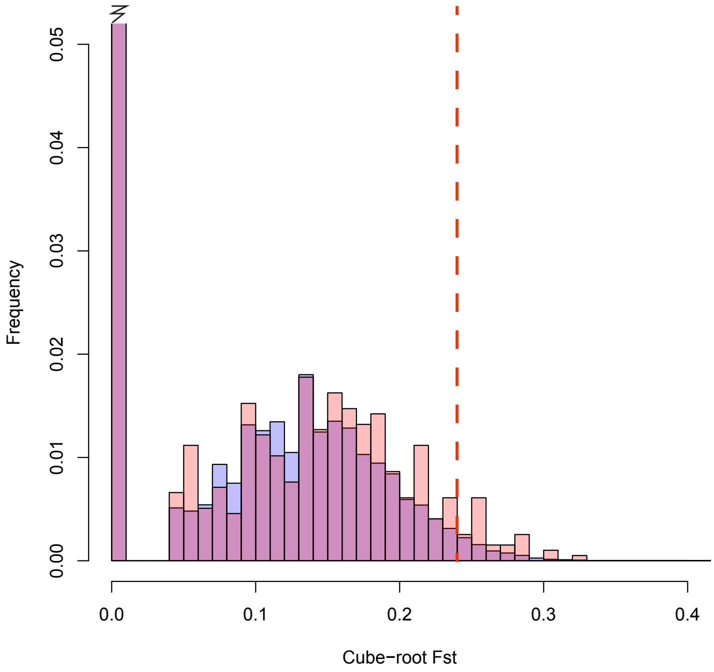


Figure 6.18: **WAL vs. WEM F$_{ST}$ values** Frequency-scaled histograms of cube-root transformed F$_{ST}$ values for the WAL-WEM pair, where genome-wide values are in blue and values for analysed genes are in pink, and the overlap is purple. The red dashed line indicates the limit of 95% CI for the background (blue) distribution. The y-axis is scaled down to show variation, as the majority of values are proximate to zero, frequencies over 5% are indicated by a chevron over the bin. Note that the vast majority of values for both distributions are contained in the first, purple bin with F$_{ST}$ values approximating zero.

quantile of genome-wide values for 28 population pairwise comparisons (see Table 6.13). The top five genes include only one 'classical' pigmentation gene, *OCA2*, which separates all population pairs from each other, barring ORK from COR and DEV. The next highest ranked gene, *PAX3*,[8] similarly separates all population pairs, this time with the exception of WEM from ENGY, NENG and WAL.

Table 6.13: **Pigmenation genes by high $\mathbf{F}_{ST}$** 95th Q – the number of population pairwise comparisons (of a total 28) in which the gene listed has at least one SNP whose F_{ST} measure exceeds the 95th quantile background threshold.

Gene	95th Q	Gene	95th Q
DTNBP1	0	SLC45A2	8
CLTC	1	TYR	8
DEFB103A	1	ASIP	10
POMC	1	TPCN2	10
EDNRB	2	MYO5A	11
SLC24A4	2	LYST	12
DRD2	3	MITF	12
PMEL	3	ETS1	14
SLC24A5	3	IRF4	14
DCT	4	CORIN	15
HPS6	4	HERC2	15
MLANA	4	CLMP	16
TYRP1	4	EXOC2	21
MLPH	5	EGFR	23
KITLG	6	TRPM1	23
ADAM17	7	UBASH3B	23
ATRN	7	BNC2	24
EDN3	7	NRG1	24
CTNS	8	PAX3	25
MC1R	8	OCA2	26

The top five hits in Table 6.13 (*OCA2, PAX3, NRG1, BNC2, UBASH3B*) were chosen to examine further, and the highest ranking SNP present in each gene was chosen to investigate the extent of homozygosity around it. The SNP ranking within each gene was a product of the number of population pairwise comparisons for which the SNP exceeded the 95th quantile background threshold and the magnitude of the F_{ST} measures themselves. The selected SNPs were: *OCA2 – rs7179994, PAX3 – rs12694578, NRG1 – rs10095694, BNC2 – rs1927631, UBASH3B – rs11218790.* The homozygosity plots for a 2.5Mb window around the given SNPs can be found in Figure 6.22.

---

[8] Also present in the facial feature candidate gene set.

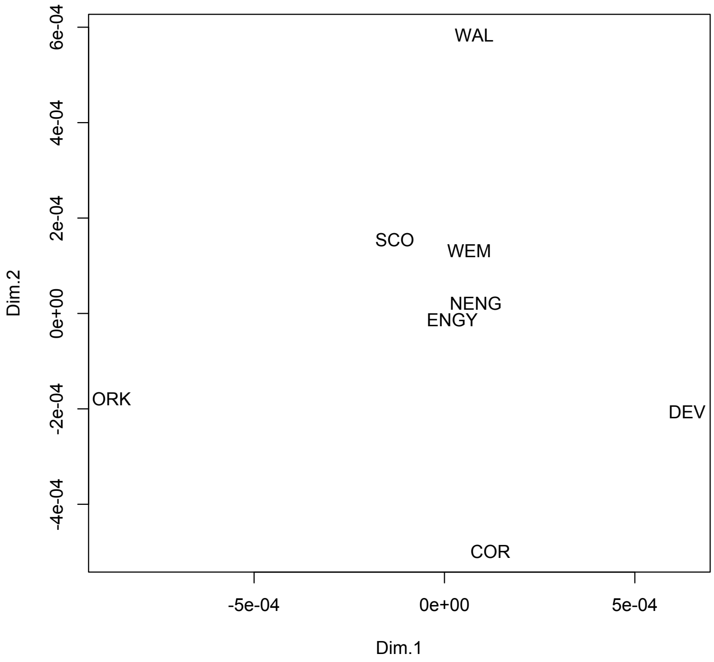


Figure 6.19: **Multidimensional scaling plot of facial feature candidate gene $\mathbf{F}_{ST}$ values**.

None of the pigmentation genes shows differing extents of homozygosity surrounding the selected SNPs, but it is worth noting that for two of them (*OCA2* and *UBASH3B*), the frequency of the derived allele far outstrips that of the ancestral one.

6.6 Facial genetics

Genetic distances between the facial feature candidate genes again place ORK and WAL at the periphery of a multidimensional scaling plot (see Figure 6.19), with COR and DEV approaching each other in a third corner of a notional triangle, and the residual English populations cluster with SCO in the middle.

As for the other gene categories, a number of facial feature candidate genes exhibited F_{ST} SNP values that exceeded the 95 quantile distribution of the genomic background for population pairwise comparisons. The distributions of both categories are illustrated

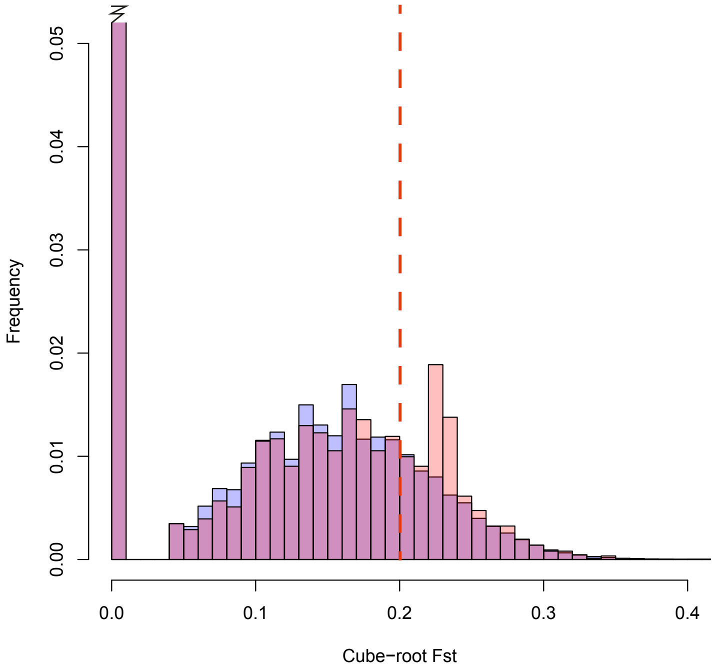


Figure 6.20: **DEV vs. ORK F$_{ST}$ values** Frequency-scaled histograms of cube-root transformed F$_{ST}$ values for the DEV-ORK pair, where genome-wide values are in blue and values for analysed genes are in pink, and the overlap is purple. The red dashed line indicates the limit of 95% CI for the background (blue) distribution. The y-axis is scaled down to show variation, as the majority of values are proximate to zero, frequencies over 5% are indicated by a chevron over the bin. Note that the vast majority of values for both distributions are contained in the first, purple bin with F$_{ST}$ values approximating zero.

in Figures 6.20 and 6.21, which include some of the greatest outlier proportions (not necessarily the most extreme F$_{ST}$ measures).

All the genes are ranked in order of increasing number of F$_{ST}$ values over the 95th quantile of genome-wide values for 28 population pairwise comparisons SNPs which differentiate between all possible population pairs, and a further two omit significant differences between only one pair. Both of these, *CACNA1C* and *NKAIN2*, fail to distinguish between neighbouring pairs COR-DEV, and WAL-WEM, respectively.

The top five hits in Table 6.14 (*PDE4D, CNTN4, NKAIN2, CACNA1C, SMYD3*) were chosen to examine further, and the highest ranking SNP present in each gene was chosen to investigate the extent of homozygosity around it. The SNP ranking

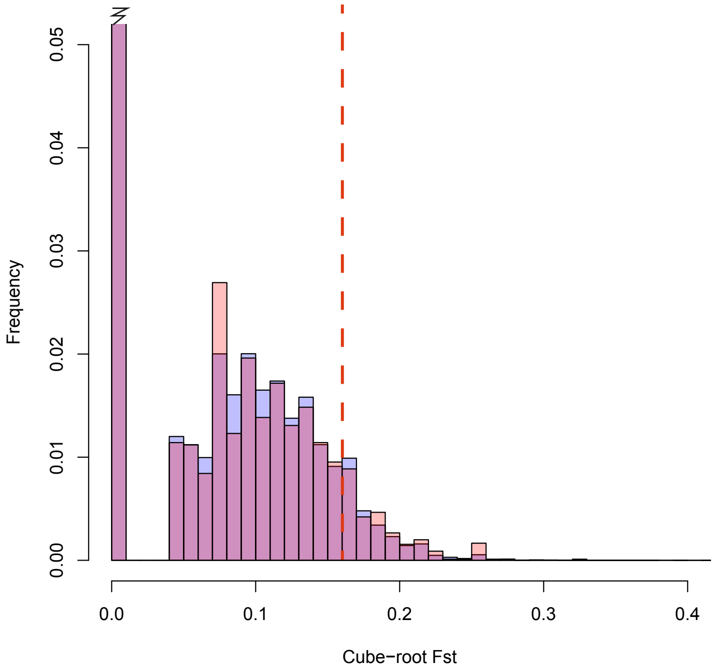


Figure 6.21: **ENGY vs. SCO F$_{ST}$ values** Frequency-scaled histograms of cube-root transformed F$_{ST}$ values for the ENGY-SCO pair, where genome-wide values are in blue and values for analysed genes are in pink, and the overlap is purple. The red dashed line indicates the limit of 95% CI for the background (blue) distribution. The y-axis is scaled down to show variation, as the majority of values are proximate to zero, frequencies over 5% are indicated by a chevron over the bin. Note that the vast majority of values for both distributions are contained in the first, purple bin with F$_{ST}$ values approximating zero.

Table 6.14: **Facial feature candidate genes by high $\mathbf{F}_{ST}$** 95th Q – the number of population pairwise comparisons (of a total 28) in which the gene listed has at least one SNP whose F_{ST} measure exceeds the 95th quantile background threshold.

Gene	**95th Q**	**Gene**	**95th Q**	**Gene**	**95th Q**
ACTA1	0	PCDH18	5	ELN	12
BMP4	0	TBCE	5	MITF	12
C17orf69	0	ZNF238	5	TRPS1	12
DLX2	0	CD300A	6	KRAS	13
DLX3	0	FGF8	6	MAP2K1	13
DLX6	0	GSC	6	MAPK8IP1	13
FOXG1	0	MAP2K2	6	PCNT	13
FOXL2	0	PIGV	6	RAI1	13
GJA1	0	PRKAR1A	6	AKT3	14
GLI1	0	SUMO1	6	GRB10	14
NOG	0	ALDH1A2	7	TCF4	14
SOX10	0	B4GALT7	7	BLM	15
SOX9	0	CRHR1	7	BMP7	15
TWIST1	0	EDN3	7	COL11A2	15
CTSK	1	EDNRA	7	EDN1	15
PITX1	1	FSTL3	7	FGF9	15
SALL1	1	KAT6B	7	FRZB	15
SMC3	1	KIAA1267	7	GHR	15
SOX2	1	PPPDE1	7	H19	15
TFAP2A	1	RARA	7	TGFB2	15
TWIST2	1	RILP	7	ALX3	16
BMP2	2	C1orf101	8	FGFR2	16
CDKN1C	2	DHODH	8	MAPT	16
CRK	2	ERCC8	8	PAX7	16
CRLF1	2	GEMIN2	8	RCAN1	16
EDNRB	2	IMP5	8	TGFB3	16
FGF4	2	PEX16	8	CHD7	17
MYO1C	2	RET	8	ERCC3	17
SNAI2	2	ADSS	9	FLNB	17
STH	2	ALX4	9	IGF2	17
ASXL1	3	COL1A1	9	TCOF1	17
DLX5	3	DNASE1	9	TNKS	17
IGFBP1	3	ERCC6	9	UBE3A	17
IRF6	3	FGFR3	9	COL9A1	18
MSX1	3	FOXC1	9	CTBP2	18
NIPBL	3	JAG1	9	FREM1	18
NOTCH2	3	LETM1	9	SKI	18
PRDX1	3	NSD1	9	VPS13B	18
PRPF8	3	PMP22	9	ZEB2	18
PTPN11	3	TMCO1	9	CTNNB1	20
SHH	3	ZBTB24	9	EVC2	20
SHOC2	3	ALCAM	10	ROR2	20
SPRY2	3	DHCR7	10	SDCCAG8	20
TBX1	3	EHMT1	10	EXT1	21
WNT5A	3	EXT2	10	GLI2	21
YWHAE	3	KCNQ1OT1	10	GLI3	21
ZIC2	3	MYBPC3	10	KCNQ1	22
DLX1	4	PTCH1	10	NRXN1	22
DNMT3B	4	SHANK3	10	WRN	22
ESCO2	4	STK11	10	EVC	24
INPP5K	4	ABCC9	11	PLD5	24
KCNJ2	4	COL11A1	11	SETBP1	24
LHX8	4	ERCC5	11	TP63	24
PITPNA	4	FGFR1	11	NLGN1	25
PITX2	4	KCNH2	11	PAX3	25
RAD21	4	SH3BGR	11	PRKD1	25
CYP1B1	5	BRAF	12	SMYD3	26
HRAS	5	C1orf100	12	CACNA1C	27
MLL2	5	CBLB	12	NKAIN2	27
MSX2	5	COL2A1	12	CNTN4	28
PAX9	5	CREBBP	12	PDE4D	28

within each gene was a product of the number of population pairwise comparisons for which the SNP exceeded the 95th quantile background threshold and the magnitude of the F_{ST} measures themselves. The selected SNPs were: *PDE4D – rs1439, CNTN4 – rs1178495, NKAIN2 – rs1008742, CACNA1C – rs4370987, SMYD3 – rs10924373*. The homozygosity plots for a 2.5Mb window around the given SNPs can be found in Figure 6.22. Of the facial feature candidate genes, only *PDE4D* shows any sign of having differing extent of homozygosity between the ancestral and the derived alleles, with the minority derived allele homozygotes having a greater average range than the ancestral homozygotes.

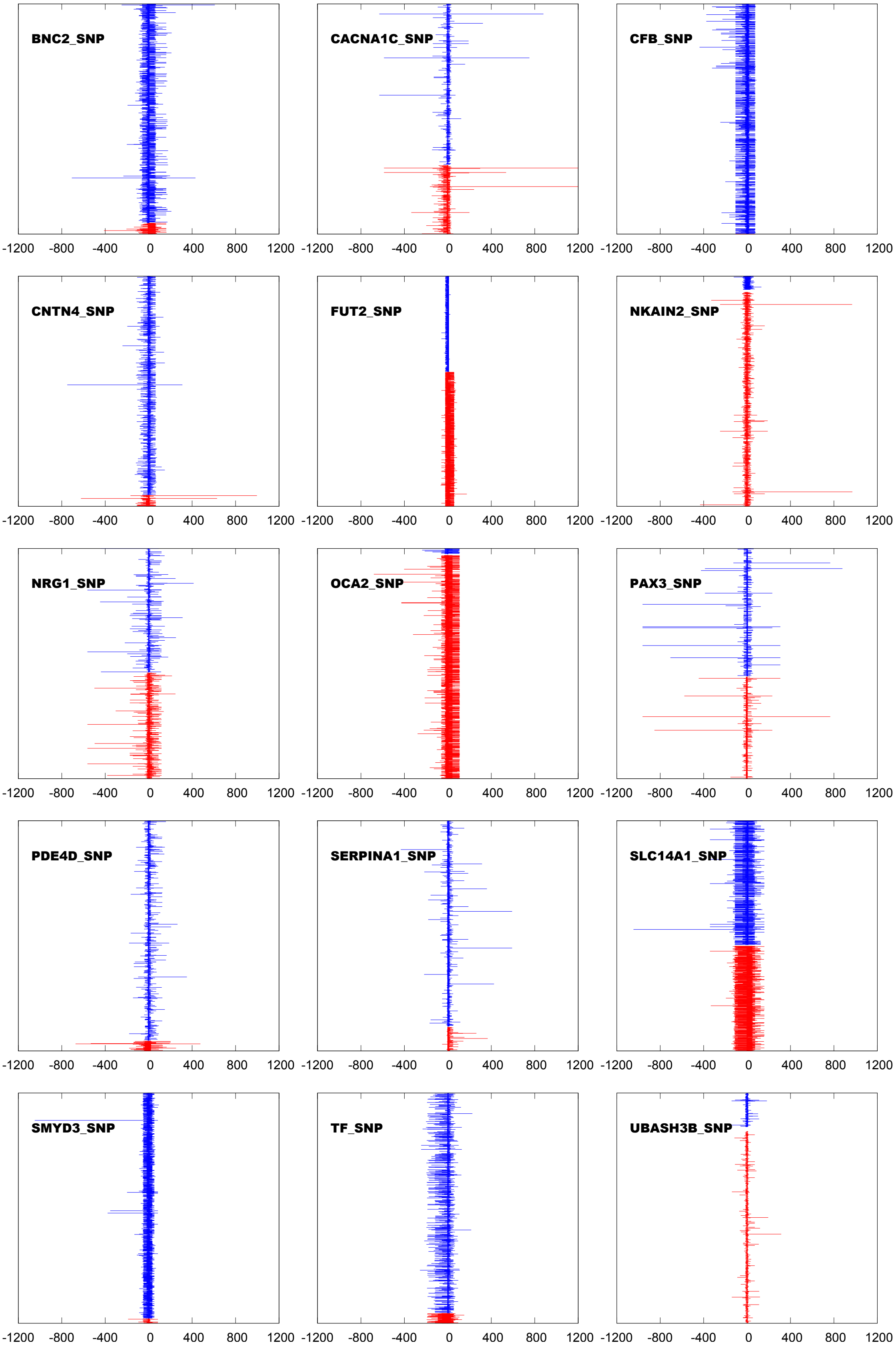


Figure 6.22: **Homozygosity plots** Each line is an individual, blue – central ancestral allele, red – central derived allele. The x-axis gives the number of SNPs in the 2.5Mb window.

Chapter 7

Discussion

7.1 Archaeological synthesis – the long durée of British connectivity and its implications

As seen in Chapter 4, Britain has rarely been isolated from its surrounding landmasses, and often, cultural discontinuities have been more noticeable between parts of the island than across comparatively narrow waterways to adjacent lands. The cultural links and continua over the past dozen millennia may briefly be summarized (see Figure 7.1): Lowland Britain formed a certain unit throughout the period, and was more or less continuously in contact with the region on the opposite side of the Channel. Occasionally, a greater area of the western North European Plain may be included. Interplay along the Atlantic coasts was irregular, if sporadically intense. Connections between Britain and Ireland were for the most part infrequent, with occasional contact between southwest England and southern Ireland, and between Wales and eastern Ireland, but southwest Scotland and northeast Ireland are an exception, and formed a fairly continuous zone of exchange. Southwest England also had a near-continuous relationship with Brittany. Northeastern Britain also formed something of a unit, and direct extra-British influence was rare. The same goes for the northernmost extremities of Britain, but for the notable event of the Viking hegemony in the area. Within Britain there were naturally no unsurpassable barriers, but by and large, there seems to have been not only a classical Highland-Lowland divide, but a broad east-west one. For the latter, the eastern portion was more internally connected, leaving the western portion as semi-isolated sections (the southwest, north/northwest Scotland, and particularly Wales).

The implications of this scenario as to what one might expect from a genetic portrait of the British populations are straightforwardly qualifiable, if not easily quantifiable in a system where all differences are very small: Extensive similarity of the Lowland Zone, perhaps extending some way up into northern Britain, and noticeable similarities to northern France, the Low Countries, into northern Germany and possibly Denmark. The above isolated sections of western and northern Britain are likely to be less varied than lowland Britain, and form internally more coherent wholes. The English southwest and possibly also southern Wales might be expected to be more similar to Brittany than any other part of the continent, and southwest Scotland to northeast Ireland. Orkney, due to the heavy late Medieval Norse Viking incursions, is expected to have at least some genetic input from Norway.

The latter supposition, of course, has already been vindicated by the use of uniparental markers, especially the Y chromosome [55, 516, 551]. As for most of the other predictions, suitable data is just now becoming available (see Appendix K).

7.2 Genetics

7.2.1 Mitochondrial DNA

The mitochondrial landscape of Britain, represented in this work by an unprecedented number of samples, fits firmly within the northwest European, and broader European, sphere. The differences in haplogroup frequencies are generally, and not unexpectedly, very small.

'Peripheral' groups like the Orcadians and the Welsh seem to hold the some of the

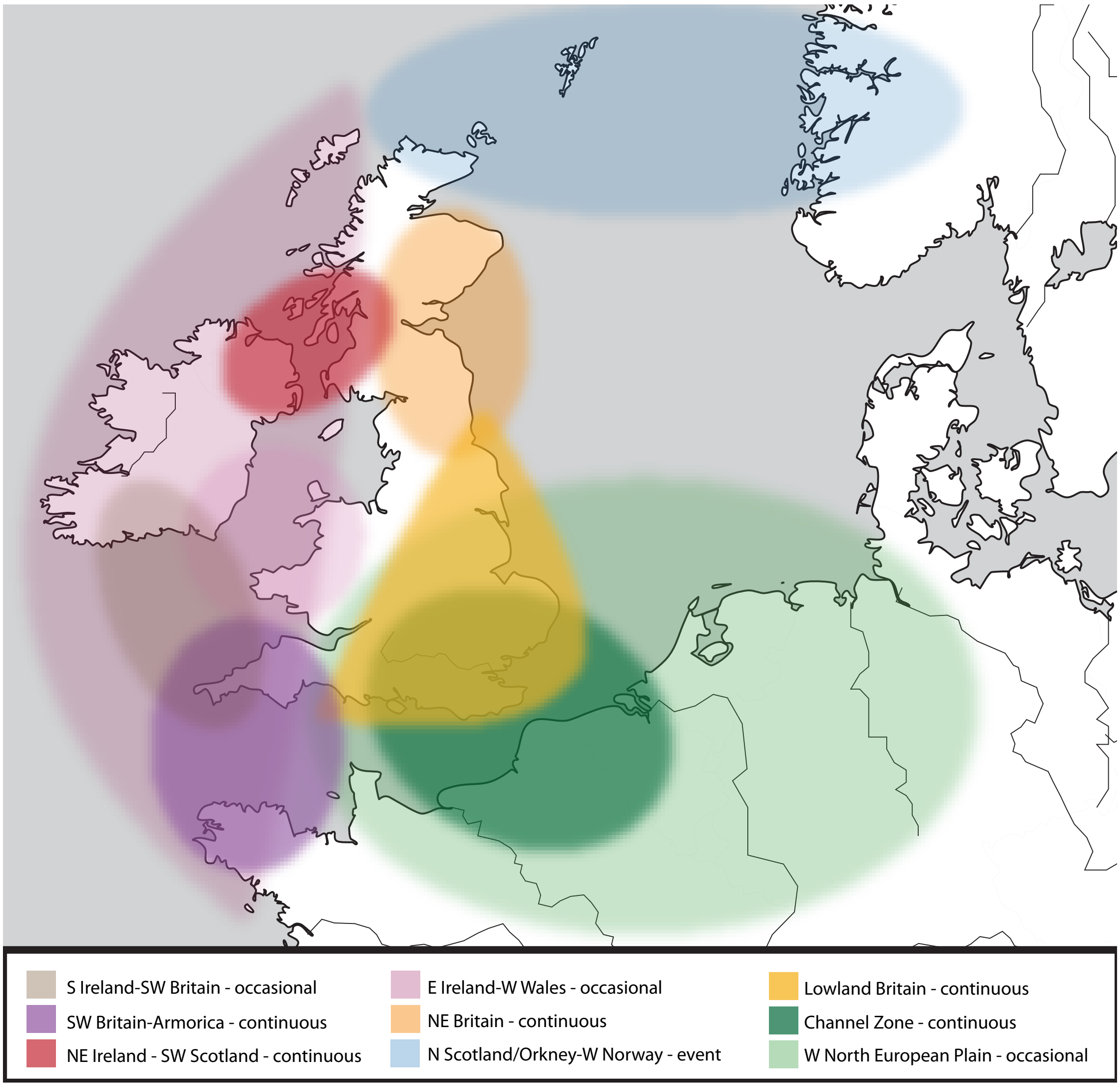


Figure 7.1: **Regions of contact** Schematic representation of connectivity, stronger colours indicate more intense and longer-lasting links. The pale purple area covering Ireland, Brittany and Atlantic Britain represents intermittent Atlantic connections.

most extreme haplogroup frequencies. The latter group is notable for comparatively very low levels of U5a, high levels of H4 and J haplogroups outside of the more common J1 and J2. Eastern Scotland, also, is notable for its very high H13 frequency. For Orkney, the most distinctive feature is the high haplogroup X frequency (7.5%), and Norwegian immigration – an obvious cause of discrepancies between Orkney and Scotland – is almost certainly not the source of this, as the haplogroup is exceedingly rare in Norway (0.3% of 642). Perhaps the higher level of isolation in these peripheral groupings has caused some acceleration of genetic drift, leading to what in the context of the broader island are extreme frequencies. However, given that most visually obvious differences between population groups are to be found in overall low frequency haplogroups, these differences rarely reach statistical significance.

Oddities within the data set include an apparent negative correlation between the two most frequent H haplogroups, H1 and H3, in Somerset, Dorset and Wiltshire. H1 is unusually frequent among these 92 samples (24%), and H3 is completely absent. Haplogroup V, as a presumptive marker of post-LGM recolonization from the Franco-Cantabrian refuge, might be predicted to be more common in the more peripheral areas of Britain, but this haplogroup shows no very striking patterning beyond a complete absence in several such places as one might expect it to be highest, Orkney and northeast Scotland.

The overall picture given by mitochondrial haplogroups is certainly not one of fine geographic structure – outliers in the PCA plot (Figure 6.4) are largely groups with small sample sizes (<100) and thus the frequencies, especially for minor haplogroups, may be quite different from the real frequencies in these populations.

Finally, there is one mitochondrial pattern that deserves mention, although its determinant was only found in seven individuals in the whole dataset: that of haplogroup H7a. All seven were from either the English southwest, or Brittany. While providing single, strong explanations for weak observed patterns is dangerous, it is tempting to hypothesize that this haplogroup is a representative of the longstanding interaction between these two otherwise comparatively isolated peninsulas.

7.2.2 Classical markers and lactase

Classical markers have previously shown at least some structuring within Britain [513], however, the approach used in this work demonstrated comparatively little. This may most likely be accounted for the fact that variant determinants (SNPs or others) within the classical marker genes were most often absent from the Illumina 1M-Duo SNP array, and therefore the genetic data would naturally be unable to recapitulate previously observed patterns. ABO blood group SNPs, for example, exhibited only a limited amount of between-group discrimination.

Nevertheless, a plot of genetic distances (Figure 6.7) displays some differentiation consistent with geographical sample location. Lowland Britain, represented perfectly by the ENGY group, lies at one pole, flanked by peripheral edges of England (Devon, Cornwall and the Welsh border area), and with northern England and Scotland/northern Ireland closer to Orkney, and Wales separate.

Two genes, *FUT2* and *TF* stand out, with the former differentiating between the SCO group and all other groups, and the latter COR from northern populations, and lowland England and northern English inhabitants from a small array of others. The remaining twenty genes included showed much smaller differences. *FUT2* and *TF*, like most of the classical markers, are either blood/serum/immune system proteins or involved in their synthesis and transport. The SNPs causing the higher level of differentiation usually have very high (sometimes nearly 50%) minor allele frequencies – this is also the case for the highly differentiating pigmentation and facial feature candidate gene loci (see below). No explicit evidence for selection in human transferrin (*TF*), which is involved in transport of iron ions in the blood and haeme degradation, has been reported [552]. In the case of *FUT2*, the product of which is necessary to create secretory A/B antigens, this high MAF is likely due to the long-term balancing selection to which it has been subjected [553, 554, 555]. The *FUT2* SNP selected by the F_{ST} analysis has been shown to be associated with both Crohn's disease and vitamin B12 deficiency, of which the latter is probably most relevant in an evolutionary sense

[556, 557]. The slightly increased extent of homozygosity around the derived allele of the examined SNP in this study may speak to the aforementioned balancing selection, and the very short range of homozygosity around both the ancestral and derived alleles may be due to the length of time this selection has been acting.

The remaining top-five population-differentiating genes (*CFB, SLC14A1, SERPINA1*) have also been suggested as targets of selection, in keeping with their demonstrated or suggested roles in the immune system: *CFB* is a component of the complement system, *SERPINA1* may have a host-pathogen interaction role, and *SLC14A1* is the determinant gene of the the Kidd blood group [554, 558, 552]. In fact, SLC14A1 SNP highlighted by the F_{ST} analysis is the very one causative of the Kidd blood group type transition (D280N) [559]. The *CFB* SNP with the highest F_{ST} impact has shown strong associations with type 1 diabetes [560].

While the classical marker genes were in a sense pre-selected for displaying substantial between-population differentiation, both because they have long been used for this purpose and because of the use of solely non-synonymous coding SNPs, the ability of the F_{ST}-based analysis to pick out SNPs with these demonstrated functional and clinical roles is noteworthy.

The lactase persistence allele is an interesting case because its frequency expansion is most likely linked to the introduction of dairying in Europe, and so may be considered 'Neolithic' in terms of spread. The lack of observed examples of this allele in early Neolithic samples from Europe [34] – though few – indicate that this expansion took place subsequent to the first introduction of domesticated bovids. There is a clear increase in lactase persistence in the northwest of Britain (Figure 6.10), compared to the south, Scotland, northern Ireland, Wales and the Welsh Marches stand out particularly. If lactase persistence can be seen as a Neolithic trait, and was introduced to the British Isles along with cattle, this pattern would run contrary to a standard east to west infiltration of the new economy. However, if, as posited in Chapter 4, England in fact received fewer Neolithic migrants than Atlantic Britain, this might provide an explanation. Such a hypothesis is likely untenable in view of the pan-

European distribution of the lactase persistence allele in question. The British pattern is an extension of this, with a general trend of increasing frequencies from southeast to northwest [561]. This is the exact opposite of what would be expected had the persistence variant been highly selected for in early Neolithic populations and diffused into the European peninsula with them.

Perhaps a more likely alternative is that there was a higher continuous selection pressure for the persistence allele in the north and west, which, being less suited to arable crops than the Lowland Zone, was more dependent on pastoral farming. This latter hypothesis may be more in keeping with a Bronze Age intensification of dairying in many parts of Europe, during a proposed 'secondary products revolution', though the evidence for such an increase is not unequivocal [562, 563, 564].

7.2.3 Pigmentation

Pigmentation has been treated in this work in three separate but interconnected ways: in terms of phenotypes alone, genotypes for pigmentation genes alone, and assessment of phenotype-genotype correlations based on current understandings of associations.

7.2.3.1 Phenotypes

Phenotypes considered were hair and eye colour frequencies collected in Britain a century and more ago, and skin color measurements collected by the PoBI project.

The results indicate that there are indeed real differences of integument pigmentation across Britain, along several different axes. For hair colour, the axis runs east-west, with fair and brown hair predominating in eastern Britain, and dark and black hair in western Britain. Statistical significance for these differences, however, is mostly limited to Wales and southwest England versus other groups. Wales and the southwest also stand out in terms of high numbers of dark-eyed individuals, and in this case, the above east-west trend is broken by the very high numbers of blue eyed individuals in western Scotland. Red hair, often idealized as a 'Celtic'[1] trait, shows little significant

---

[1] i.e. particularly Irish

patterning, but where it does reach higher levels it is in eastern Scotland and northeast England, and not traditionally 'Celtic' areas.

Skin color differences appear to lie along a north-south axis, with the northern populations (Scots and Orcadians) being the palest, Wales intermediate, and the English and Cornish the darkest. The most obvious explanation for this trend is the latitude itself. As detailed in Chapter 3, paler skin tones are correlated with increasing latitude, a probable adaptation to decreased UV radition further from the equator and the need for such radiation to synthesize vitamin D. While it may be surprising to be able to observe this effect on such a small geographic scale as the island of Britain, there is a geophysical discontinuity between England and Scotland: Scotland lies just beyound the latitudinal limit at which natural sunlight alone can provide provide sufficient yearly vitamin D, even for the palest skinned of peoples [565]. Paler skin, even than the English, would likely help the inhabitants of Scotland and its islands take optimal advantage of their seasonally very limited UV exposure.

7.2.3.2 Genotypes and Phenotypes

Given that hair, eye and skin colour are fully genetically controlled traits (see Chapter 2), it is theoretically possible to predict an individual's phenotype based on their genotypes at functionally relevant loci. So far, a model combining acceptable sensitivity and specificity has remained elusive (see e.g. Branicki *et al.* [89]). The broad-brush attempt in this work to predict phenotype based on genotype used population groups, rather than individuals, and previously published data on regional phenotypes [550]. Potential problems with this approach naturally include different opinions of where different shades of particularly hair colour begin and end, but as only very strong reported associations of genotype with a specific pigmentation phenotype were included, this left out 'brown', likely to be the most variable category.

Blonde hair, for which the reported association was weakest, accordingly showed the poorest correspondence between the genotype-estimated frequency of blondes, and the presumed 'true' observed frequency. For red hair, the normal pigmenation phenotype that has been most extensively studied, the simple model performed quite well in predicting the observed values, and performance was even better for light versus dark eyes – this is a further vindication of the near-perfect association between a single intronal SNP and light eyes reported in Eiberg *et al.* [86]. The phenotype frequency prediction based on a black hair-associated genetic variant (rs16891982 in *SLC45A2*) presents an interesting dichotomy. For most of Britain, the predicted values were very similar, and certainly not significantly different from, those in the directly observed dataset. The exceptions were Wales and southwest England, for which significantly higher frequencies of black hair were observed (averaging 13.4% versus the rest of Britain's 4.2%), but not predicted (average of 3.1% versus the rest of Britain's 2.7%). This indicates that the allele which causes black in eastern Britain, and in the discovery dataset of samples collected in Poland [535], may be different that/those carried by the majority of black-haired individuals in western Britain.[2] As such, this/these variants are likely to constitute a genetic element capable of distinguishing between the east and west of Britain.

Cultural cohesion is, circularly, caused by and cemented by commonalities, and visual cues as to who belong to the ingroup and who might be considered 'other' or 'foreign' are vital. Many of these visually available features are determined by cultural norms themselves (clothing, hairstyles, personal ornamentation, etc). Occasionally, however, genetically determined physical features can be of importance, most noticeably in cases of non-overlapping palettes of skin tones – until recently, cultural groups would only exceptionally contain people of different races.[3] Within Britain, there is no evidence for such radical differences as would constitute a perception of complete separation of regional groupings, but anecdotally mean differences in pigmentation have been noticed, in terms of regional stereotypes. Tacitus recounts some of these when describing northern and southeastern Britons as red- and fair haired, and ascribing a

---

[2] Northwest England and western Scotland also have higher black hair frequencies unaccounted for by their levels of rs16891982, but this difference fails to reach statistical signficance.

[3] In this context, 'race' is used to refer to a suite of mean physical characteristics including skin colour, hair colour and texture, and eye colour variation.

dark complexion to the inhabitants of Wales [4]. Such stereotypes, probably the result of perceived differences in *mean* appearance, rather than extremes, still exist. The results presented here go some way to vindicating these perceptions – darker in the southwest, fairer in the north and east – though stereotypical representations will tend to extrapolate a rule from a trend. Facial features are another key way of recognizing group affiliation, but the variation is less easy to describe, especially in anecdotal terms and over small geographical scales. While a genetically-determined 'typical' face of any given region may exist (see below), facial features are less likely to be used to either construct or assign micro-scale identities.

7.2.3.3 Pigmentation genes

For *MC1R*, the highest pairwise chi square allele frequency difference was between WEM and NENG, probably due to the former's notably higher frequency of allele R160Q (12.3% versus 7%), as well as a discrepancy in V60L levels (5.6% and 16.4%, respectively). This pattern is rather counterintuitive when considered in relation to the distribution patterns of red hair in Britain, of which R160Q is one of the major determinants, along with R151C. Though red hair levels are quite low and approximately uniform, Figure 6.12 shows the Welsh border area to be among the areas with lowest red hair frequencies, and northern England the highest. Quite the converse would be indicated by these allele frequencies, given that these groups' frequencies of R151C are comparable. In general, however, *MC1R* variants are uninformative regarding regional British variation, in sharp contrast to the high discriminating power its non-synonymous coding SNPs have for Eurasian population in general (see Chapter 3).

Another pigmentation gene outlier relates to a single SNP in *TPCN2*, rs382941, of which ENGY, NENG, SCO and WAL have 42-47%. This SNP's frequencies in COR, DEV, and WEM are rather lower, but Orkney is a clear outlier with only 29%. As this SNP is associated with shifts between brown and blond hair [79], perhaps the strong historical link with the anecdotally blond Scandinavia may account for this difference.

Of the top five genes in the F_{ST} analysis, the highest ranked – *OCA2* – is the one known to be most strongly involved in pigmentation, as intimated by its name, oculocutaneous albinism 2 (see Chapter 2). As such, it is a prime selection candidate, and the highest ranked SNP's minor allele is the ancestral allele, as opposed to the more common situation where the ancestral allele is in majority. This is also the case for the fifth-ranked gene's (*UBASH3B*) highest ranked SNP. The next two genes, *PAX3* and *NRG1*, have important developmental roles, which would make them sensitive to functional perturbation, and indeed, their high-ranking SNPs in the present study are not functional ones. In addition to their separate links with various diseases (e.g. Waardenburg syndrome, Hirchsprung disease, psychiatric disorders), they also both have associations with melanoma and normal roles in pigment cell development and function [566, 567, 568, 569]. *BNC2*, ranked fourth, is a transcription factor specific to skin keratinocytes in adults, and has strong effects on both basal pigmentation, and is a strong candidate for skin colour saturation in Europeans [570, 571]. The highest ranking SNP for this gene is again a non-synonymous coding one, and so is unlikely in itself to change the efficiency of its protein product's action. None of the above genes showed any signal of selection for derived alleles on the basis of extent of homozygosity around the highest population group differentiating SNP.

A final point to bear in mind is that only SNPs in or within 2kb of any given gene were used for the analysis, which may have eliminated variants pertinent to the observed geographical differences in phenotype (see above). Regulatory elements in or near the pigmentation genes may account for much of the variation, as is the case with the 'blue eye' variant, rs12913831, which is located in an intron of *HERC2* and downregulates the expression of *OCA2*.

7.2.4 Facial candidate genes

The identification of genes that determine facial features falls outside the scope of this thesis, and would be hampered by the some of the same considerations as reviewed in the previous section. Regulatory elements, more than non-synonymous coding SNPs, are likely to play a huge role in facial determination, as much of facial develpment depends on dosage gradients. This has already been demonstrated for one gene (see Chapter 2). Furthermore, the involvement of key development genes in structuring the face, means that any change to the coding sequence is far more likely to be lethal than equivalent mutations in, for example, most pigmentation genes.

The facial feature gene cohort was far larger than the others examined in the present study, and as such it is to be expected that more genes with high population-pairwise discriminatory effects may be observed. This was indeed the case, and the top five ranked genes showed high differences between either all or all but 1-2 population pairs. Among these, all the top-ranked SNPs were in introns. As the skull bones, and thereby many facial features, are formed based on the forebrain (see Chapter 2), most of the candidate genes are developmental and therefore mutations in them are associated with a whole host of pathologies, often, but not solely, relating to brain function. Two of these top-ranked genes have turned up as hits in the same genome-wide association study, not as relevant to facial features, but to neuroticism [572].

Facial feature candidate genes were selected as a system of study on the basis that, if they truly did affect normal facial variation, they might be a useful group of markers to investigate for population differentiation due to the complete heritability and regional differences of facial features (cf. Chapter 2). While the population pairwise differences were on the whole more marked for this category of genes than for the classical marker and pigmentation genes, the pattern of differentiation was very similar across all genetic subsets, as it was to the genome-wide distribution of genetic variation, both for the global and non-synonymous SNP sets.

7.3 Genetics, archaeology, linguistics and the peoples of Britain – concluding remarks

The interplay between biological and cultural features is undoubtedly complicated, and the different systems are prone to different types of hybridization, different time scales for variation to arise and dissipate, and different obligate means of transmission. The more transitory nature of cultural trends (including long-lasting languages) versus the long-term nature of biological descent implies that one-to-one relationships between genetically coherent groups and any given coherent cultural packages are unlikely to endure for a very long time. Innovation, in the genetic context, is unlikely to play a very large role on any historical – as opposed to a general 'evolutionary' – timescale,[4] whereas it is absolutely crucial to the changing nature of cultural features. Exchange, both genetic and cultural, will affect populations, but the exchange of culture can be more wholesale than that of genetic material, which is strongly dependent on the number of exchanged individuals and often less so on the socio-economic status they/their cultural attributes are accorded. This complex nature of culture-population dynamics is important to bear in mind when attempting to synthesize trends from the component subjects considered in this work.

The genetic results from this study confirm that there exists population structure within the UK, even if most of it, using these genetic systems, is modest. Orkney is the most consistently different for the different marker systems, which is not inconsistent with its archaeologically and linguistically attested high level of Norse Viking immigration. Certain Orcadian features, however, such as the islands' abnormally high level of mitochondrial haplogroup X, cannot be suggested to be due to Norwegian influence, and part of Orkney's distinctiveness may be attributable to its peripheral location. Wales, too, possibly shows signs of archaeologically inferred isolation with many of its mitochondrial haplogroups lying at the extreme ends of the overall frequency distribu-

---

[4] Recently acquired genetic variants take a long time to be established in a population of any size, unless there is very strong selection acting on them. The latter case is unusual, as most *de novo* mutations are neutral or harmful.

tions, and its position on the MDS plots that show it to be removed from England. In terms of pigmentation Wales also occupies a genetically determined outpost, with its high levels of dark hair and eyes, though the exact genetic cause of the former trait remains unknown.

Scotland was grouped with northern Ireland by unsupervised clustering of individuals (Appendix K), which may have exaggerated the occasional differences seen in the parts of this study which used the genetically determined groups. Without too far presuming on the final results of the fineSTRUCTURE analysis of the PoBI data, this provisional grouping of Scottish and northern Irish samples can be seen as a vindication of the archaeological and lingustic evidence which point to long periods of intense contact across the narrow sea which separates southwest Scotland from Ulster. There can also be no doubt that the 17th century mass population movement from Scotland and England known as the Ulster Plantation had a strong homogenizing effect on the genetic signatures of northern Ireland and northern Britain: By 1641, an estimated one-fifth of the entire population of Ulster was of recent British origin [573, 574, 575].

England remains on the whole homogenous in this study, apart from vagaries of mitochondrial haplogroup frequencies of uncertain relevance, especially the Lowland area, as determined for the 'ENGY' analysis group. Whether this uniformity is a result primarily of the ease of communication within this area, or owes a heavy debt to near-continual contributions from the near Continent remains to be seen. The latest, and certainly the most archaeologically and linguistically visible, of these contributions, the early medieval Germanic peoples are still strong contenders for a further homogenizing influence. The general uniformity of England is broken in a very minor and interesting way by the tantalizing mitochondrial link of the southwest with Britanny, which the longstanding archaelogical and recent linguistic links suggest may reflect genuine biological exchanges.

We have seen that an amalgamation of archaeological, linguistic and genetic data provides several consistent messages regarding the population history of Britain and the British Isles. Distributions of several gene variants, particularly those potentially involved in facial determination, tie in with the broad trends expected from archaeology while also providing intriguing information about finer-scale processes. Substantial links have been found between genetic determinants of human appearance and measured phenotype distributions throughout Britain. In addition, genetic data herein serve to consistently place Britain within the context of its geographical surroundings. Each of these facets of investigation provides strong support for the hope that forthcoming work on the PoBI project will successfully elucidate further details of the population history of Britain.

References

[1] Stringer, C. *Homo Britannicus: The Incredible Story of Human Life in Britain.* Allen Lane, London (2006)

[2] Nennius. *Historia Brittonum.* D.S. Brewer, Cambridge (1985)

[3] Buchanan, G. *Rerum Scoticarum Historia.* Francofurti ad Moenum, Frankfurt (1638)

[4] Cornelius Tacitus, P. Agricola & Germania (1898)

[5] Barbujani, G. and Belle, E. M. Genomic boundaries between human populations. *Hum Hered*, **61** (1): 15–21 (2006)

[6] Delisle, R. The Biology/Culture Link in Human Evolution, 1750-1950: The problem of Integration in Science. *Stud Hist Phil Biol & Biomed Sci*, **31** (4): 531–556 (2000)

[7] Cavalli-Sforza, L. L. Archaeology, Genes and Language: Reflecting on Five Decades of Human Genetics. In M. Jones, editor, *Traces of Ancestry: Studies in Honour of Colin Renfrew.* MacDonald Institute of Archaeological Research, Cambridge (2004)

[8] Cavalli-Sforza, L. L. The DNA revolution in population genetics. *Trends Genet*, **14** (2): 60–5 (1998)

[9] Renfrew, C. From molecular genetics to archaeogenetics. *Proc Natl Acad Sci U S A*, **98** (9): 4830–2 (2001)

[10] Darwin, C. *The descent of man, and selection in relation to sex.* John Murray, London (1871)

[11] Jones, W. Anniversary Discourse to the Asiatic Society, On the Hindus. In *The works of Sir William Jones.* Printed for G. G. and J. Robinson; and R. H. Evans successor to Mr. Edwards, London (1799)

[12] McMahon, A. *Understanding Language Change.* Cambridge University Press, Cambridge (1994)

[13] Comrie, B. Is there a single time depth cut-off point in historical linguistics? In C. Renfrew, A. McMahon, and L. Trask, editors, *Time Depth in Historical Linguistics.* McDonald Institute for Archaeological Research, Cambridge (2000)

[14] Kupfermann, H., Mayer, W. E., O'hUigin, C., Klein, D., and Klein, J. Shared polymorphism between gorilla and human major histocompatibility complex DRB loci. *Hum Immunol*, **34** (4): 267–78 (1992)

[15] McMahon, A. and McMahon, R. Linguistics, genetics and archaeology: internal and external evidence in the Amerind controversy. *Transactions of the Philological Society*, **93**: 125–225 (1995)

[16] Jobling, M., Hurles, M., and Tyler-Smith, C. *Human Evolutionary Genetics: Origins, Peoples and Disease.* Garland Publishing, New York (2004)

[17] Campbell, L. *Historical Linguistics.* Edinburgh University Press, Edinburgh (1998)

[18] Harpending, H., Sherry, S. T., Rogers, A., and Stoneking, M. The Genetic Structure of Ancient Human Populations. *Curr Anthropol*, **34** (4): 483–496 (1993)

[19] Jackson, K. The Pictish Language. In F. Wainwright, editor, *The Problem of the Picts.* Nelson, Edinburgh (1980)

[20] Schrijver, P. More on non-Indo-European surviving in Ireland in the first millennium AD. *Ériu*, **55**: 137–144 (2005)

[21] Popper, K. *The Poverty of Historicism.* Routledge, London (1960)

[22] Cavalli-Sforza, L. L., Menozzi, P., and Piazza, A. *The History and Geography of Human Genes.* Princeton University Press, Princeton, NJ (1994)

[23] Hunley, K. L., Healy, M. E., and Long, J. C. The global pattern of gene identity variation reveals a history of long-range migrations, bottlenecks, and local mate exchange: Implications for biological race. *Am J Phys Anthropol* (2009)

[24] Tishkoff, S. A., Dietzsch, E., Speed, W., Pakstis, A. J., Kidd, J. R. et al. Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science*, **271** (5254): 1380–7 (1996)

[25] Thomas, M. G., Parfitt, T., Weiss, D. A., Skorecki, K., Wilson, J. F. et al. Y chromosomes traveling south: the Cohen modal haplotype and the origins of the Lemba – the "Black Jews of Southern Africa". *Am J Hum Genet*, **66** (2): 674–86 (2000)

[26] Hellenthal, G., Auton, A., and Falush, D. Inferring human colonization history using a copying model. *PLoS Genet*, **4** (5): e1000078 (2008)

[27] Zalloua, P. A., Platt, D. E., El Sibai, M., Khalife, J., Makhoul, N. et al. Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. *Am J Hum Genet*, **83** (5): 633–42 (2008)

[28] Jobling, M. A. and Tyler-Smith, C. The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet*, **4** (8): 598–612 (2003)

[29] Barbujani, G. and Bertorelle, G. Genetics and the population history of Europe. *Proc Natl Acad Sci U S A*, **98** (1): 22–5 (2001)

[30] Novembre, J., Johnson, T., Bryc, K., Kutalik, Z., Boyko, A. R. et al. Genes mirror geography within Europe. *Nature*, **456** (7218): 98–101 (2008)

[31] Reich, D., Price, A. L., and Patterson, N. Principal component analysis of genetic data. *Nat Genet*, **40** (5): 491–2 (2008)

[32] Mielke, J. and Fix, A. The Confluence of Anthropological Genetics and Anthropological Demography. In M. H. Crawford, editor, *Anthropological Genetics*. Cambridge University Press, Cambridge (2007)

[33] Cavalli-Sforza, L. L., Menozzi, P., and Piazza, A. Response. *Science*, **261** (5128): 1508 (1993)

[34] Burger, J., Kirchner, M., Bramanti, B., Haak, W., and Thomas, M. G. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. *Proc Natl Acad Sci U S A*, **104** (10): 3736–41 (2007)

[35] Frost, P. European hair and eye color - A case of frequency-dependent sexual selection? *Evol Hum Behav*, **27**: 85–103 (2006)

[36] McEvoy, B. P., Montgomery, G. W., McRae, A. F., Ripatti, S., Perola, M. et al. Geographical structure and differential natural selection amongst North European populations. *Genome Res* (2009)

[37] Trigger, B. *A History of Archaeological Thought*. Cambridge University Press, Cambridge (1989)

[38] Trigger, B. *Time and Traditions: Essays in Archaeological Interpretation*. Edinburgh University Press, Edinburgh (1978)

[39] Harris, M. *The Rise of Anthropological Theory*. Thomas Y. Crowell Company, New York (1968)

[40] Childe, V. *The Danube in Prehistory*. Oxford University Press, Oxford (1929)

[41] Renfrew, C. and Bahn, P. *Archaeology: Theories, Methods and Practice*. Thames & Hudson, Ltd, London (2000)

[42] Thomason, S. *Language Contact: An Introduction*. Georgetown University Press, Washington D.C. (2001)

[43] Wardhaugh, R. *Languages in Competition*. Basil Blackwell Ltd, Oxford (1987)

[44] Campbell, L. Time perspective in Linguistics. In C. Renfrew, A. McMahon, and L. Trask, editors, *Time Depth in Historical Linguistics*. McDonald Institute for Archaeological Research, Cambridge (2000)

[45] *Oxford English Dictionary*. Oxford University Press, Oxford (1999)

[46] Mallory, J. P. Time Perspective and Proto-Indo-European Culture. *World Archaeology*, **8** (1): 44–56 (1976)

[47] Trask, L. *Historical linguistics*. Arnold, London (1996)

[48] Chapman, J. and Hamerow, H. On the move again - Migrations and invasions in archaeological explanation. In J. Chapman and H. Hamerow, editors, *Migrations and invasions in archaeological explanation.* Archaeopress, Oxford (1997)

[49] Börglum, A. D., Vernesi, C., Jensen, P. K., Madsen, B., Haagerup, A. et al. No signature of Y chromosomal resemblance between possible descendants of the Cimbri in Denmark and Northern Italy. *Am J Phys Anthropol*, **132** (2): 278–84 (2007)

[50] Anthony, D. Migration in archaeology: the baby and the bathwater. *American Anthropologist*, **92**: 895–914 (1990)

[51] Rouse, I. The inference of migration from anthropological evidence. In R. Thompson, editor, *Migrations in New World Culture History.* University of Arizona Press, Tucson (1958)

[52] Anthony, D. Prehistoric Migration as a Social Process. In J. Chapman and H. Hamerow, editors, *Migrations and Invasions in Archaeological Explanation.* Archaeopress, Oxford (1997)

[53] Robb, J. Random Causes with Directed Effects - the Indo-European Language Spread and the Stochastic Loss of Lineages. *Antiquity*, **65** (247): 287–291 (1991)

[54] Hallgrimsson, B., Ó Donnabhain, B., Walters, G. B., Cooper, D. M., Gudbjartsson, D. et al. Composition of the founding population of Iceland: biological distance and morphological variation in early historic Atlantic Europe. *Am J Phys Anthropol*, **124** (3): 257–74 (2004)

[55] Helgason, A., Hickey, E., Goodacre, S., Bosnes, V., Stefansson, K. et al. mtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *Am J Hum Genet*, **68** (3): 723–37 (2001)

[56] Helgason, A., Sigurethardottir, S., Nicholson, J., Sykes, B., Hill, E. W. et al. Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet*, **67** (3): 697–717 (2000)

[57] Relethford, J. H. The Use of Quantitative Traits in Anthropological Genetic Studies of Population Structure and History. In M. H. Crawford, editor, *Anthropological Genetics.* Cambridge University Press, Cambridge (2007)

[58] De Hoz, J., Lujn, E., and Sims-Williams, P. *New approaches to Celtic place-names in Ptolemy's Geography.* Ediciones Clásicas, Madrid (2005)

[59] Strachan, T. and Read, A. *Human Molecular Genetics 2.* BIOS Scientific Publishers, Oxford (1999)

[60] www.ncbi.nlm.nih.gov/projects/snp

[61] www.census.gov

[62] Nickerson, D. A. Next Generation Human Genetics (ESHG Yearly Meeting) (2010)

[63] Chowdhury, R., Bois, P., Feingold, E., Sherman, S., and Cheung, V. Genetic Analysis of Variation in Human Meiotic Recombination. *PLoS Genet*, **5** (9) (2009)

[64] Reich, D. E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P. C. et al. Linkage disequilibrium in the human genome. *Nature*, **411** (6834): 199–204 (2001)

[65] Nordborg, M. and Tavaré, S. Linkage disequilibrium: what history has to tell us. *Trends Genet*, **18** (2): 83–90 (2002)

[66] Hartl, D. *A Primer of Population Genetics*. Sinauer Associates, Sunderland, Mass. (2000)

[67] Page, R. and Holmes, E. *Molecular Evolution – A Phylogenetic Approach*. Blackwell Science, Oxford (1998)

[68] Mayr, E. *Systematics and the Origin of Species*. Columbia University Press, New York (1942)

[69] Holsinger, K. E. and Weir, B. S. Genetics in geographically structured populations: defining, estimating and interpreting Fst. *Nat Rev Genet*, **10** (9): 639–50 (2009)

[70] Bodmer, W. F. Genetic Markers, Evolution and Selection. *Transfusion Immunol*, **6**: 35–46 (1975)

[71] Millot, J. Blood-groups and Race. *Antiquity*, **9** (36): 399–409 (1935)

[72] Avise, J. *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Mass. (2000)

[73] Myles, S., Stoneking, M., and Timpson, N. An assessment of the portability of ancestry informative markers between human populations. *BMC Med Genomics*, **2** (1): 45 (2009)

[74] Pareek, C., Smoczynski, R., and Tretyn, A. Sequencing technologies and genome sequencing. *J Appl Genet*, **52**: 413–35 (2011)

[75] Sturm, R., Box, N., and Ramsay, M. Human pigmentation genetics: the difference is only skin deep. *Bioessays*, **20** (9) (1998)

[76] Sturm, R. A. Molecular genetics of human pigmentation diversity. *Hum Mol Genet*, **18** (R1): R9–17 (2009)

[77] Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E. et al. Genome-wide detection and characterization of positive selection in human populations. *Nature*, **449** (7164): 913–8 (2007)

[78] Sulem, P., Gudbjartsson, D. F., Stacey, S. N., Helgason, A., Rafnar, T. et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet*, **39** (12): 1443–52 (2007)

[79] Sulem, P., Gudbjartsson, D. F., Stacey, S. N., Helgason, A., Rafnar, T. et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet*, **40** (7): 835–7 (2008)

[80] Williamson, S. H., Hubisz, M. J., Clark, A. G., Payseur, B. A., Bustamante, C. D. et al. Localizing recent adaptive evolution in the human genome. *PLoS Genet*, **3** (6): e90 (2007)

[81] Han, J., Kraft, P., Nan, H., Guo, Q., Chen, C. et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet*, **4** (5) (2008)

[82] Hearing, V. Determination of melanin synthetic pathways. *J Invest Dermatol*, **131** (E1) (2011)

[83] Valverde, P., Healy, E., Jackson, I., Rees, J. L., and Thody, A. J. Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nat Genet*, **11** (3): 328–30 (1995)

[84] Beaumont, K. A., Shekar, S. N., Newton, R. A., James, M. R., Stow, J. L. et al. Receptor function, dominant negative activity and phenotype correlations for MC1R variant alleles. *Hum Mol Genet*, **16** (18): 2249–60 (2007)

[85] Voisey, J., Gomez-Cabrera Mdel, C., Smit, D. J., Leonard, J. H., Sturm, R. A. et al. A polymorphism in the agouti signalling protein (ASIP) is associated with decreased levels of mRNA. *Pigment Cell Res*, **19** (3): 226–31 (2006)

[86] Eiberg, H., Troelsen, J., Nielsen, M., Mikkelsen, A., Mengel-From, J. et al. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. *Hum Genet*, **123** (2): 177–87 (2008)

[87] Stringer, S., Wray, N., Kahn, R., and Derks, E. Underestimated effect sizes in GWAS: fundamental limitations of single SNP analysis for dichotomous phenotypes. *PLoS One*, **6** (11) (2011)

[88] Kayser, M., Liu, F., Janssens, A. C., Rivadeneira, F., Lao, O. et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet*, **82** (2): 411–23 (2008)

[89] Branicki, W., Liu, F., van Duijn, K., Draus-Barini, J., Pospiech, E. et al. Model-based prediction of human hair color using DNA variants. *Hum Genet* (2011)

[90] Bouchard, T. Twins Reared Together and Apart: What They Tell Us About Human Diversity. In S. Fox, editor, *Individuality and Determinism – Chemical and Biological Bases*. Plenum Press, New York (1984)

[91] Coussens, A. K. and van Daal, A. Linkage disequilibrium analysis identifies an FGFR1 haplotype-tag SNP associated with normal variation in craniofacial shape. *Genomics*, **85** (5): 563–73 (2005)

[92] Kang, E. H., Yamaguchi, T., Tajima, A., Nakajima, T., Tomoyasu, Y. et al. Association of the growth hormone receptor gene polymorphisms with mandibular height in a Korean population. *Arch Oral Biol*, **54** (6): 556–562 (2009)

[93] Young, N. M., Chong, H. J., Hu, D., Hallgrimsson, B., and Marcucio, R. S. Quantitative analyses link modulation of sonic hedgehog signaling to continuous variation in facial growth and shape. *Development*, **137** (20): 3405–9 (2010)

[94] Ermakov, S., Rosenbaum, M. G., Malkin, I., and Livshits, G. Family-based study of association between ENPP1 genetic variants and craniofacial morphology. *Ann Hum Biol*, **37** (6) (2010)

[95] Paternoster, L., Zhurov, A., Toma, A., Kemp, J., St. Pourcain, B. et al. Genome-wide Association Study of Three-Dimensional Facial Morphology Identifies a Variant in PAX3 Associated with Nasion Position. *Am J Hum Genet*, **90**: 478–485 (2012)

[96] Liu, F., van der Lijn, F., Schurmann, C., Zhu, G., Chakravarty, M. et al. A Genome-Wide Association Study Identifies Five Loci Influencing Facial Morphology in Europeans. *PLoS Genet*, **8** (9) (2012)

[97] Marcucio, R. S., Cordero, D. R., Hu, D., and Helms, J. A. Molecular interactions coordinating the development of the forebrain and face. *Dev Biol*, **284** (1): 48–61 (2005)

[98] Meikle, M. *Craniofacial Development, Growth and Evolution*. Bateson Publishing, Bressingham (2002)

[99] Richman, J. M. and Lee, S. H. About face: signals and genes controlling jaw patterning and identity in vertebrates. *Bioessays*, **25** (6): 554–68 (2003)

[100] Gitton, Y., Heude, E., Vieux-Rochas, M., Benouaiche, L., Fontaine, A. et al. Evolving maps in craniofacial development. *Semin Cell Dev Biol*, **21** (3) (2010)

[101] Creuzet, S., Couly, G., Vincent, C., and Le Douarin, N. Negative effect of Hox gene expression on the development of the neural crest-derived facial skeleton. *Development*, **129** (18) (2002)

[102] Krader, L. *Peoples of Central Asia*. Indiana University Publications, Bloomington (1966)

[103] Bregel, Y. *An Historical Atlas of Central Asia*. Brill, Leiden (2003)

[104] Wells, R. S., Yuldasheva, N., Ruzibakiev, R., Underhill, P. A., Evseeva, I. et al. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci U S A*, **98** (18): 10244–9 (2001)

[105] Moore, J. P., Trkola, A., and Dragic, T. Co-receptors for HIV-1 entry. *Curr Opin Immunol*, **9** (4): 551–62 (1997)

[106] Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B. et al. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell*, **85** (7): 1135–48 (1996)

[107] Stephens, J. C., Reich, D. E., Goldstein, D. B., Shin, H. D., Smith, M. W. et al. Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am J Hum Genet*, **62** (6): 1507–15 (1998)

[108] Agrawal, L., Jin, Q., Altenburg, J., Meyer, L., Tubiana, R. et al. CCR5Delta32 protein expression and stability are critical for resistance to human immunodeficiency virus type 1 in vivo. *J Virol*, **81** (15): 8041–9 (2007)

[109] Limborska, S. A., Balanovsky, O. P., Balanovskaya, E. V., Slominsky, P. A., Schadrina, M. I. et al. Analysis of CCR5Delta32 geographic distribution and its correlation with some climatic and geographic factors. *Hum Hered*, **53** (1): 49–54 (2002)

[110] Savage, S. A., Gerstenblith, M. R., Goldstein, A. M., Mirabello, L., Fargnoli, M. C. et al. Nucleotide diversity and population differentiation of the melanocortin 1 receptor gene, MC1R. *BMC Genet*, **9**: 31 (2008)

[111] Norton, H. L., Kittles, R. A., Parra, E., McKeigue, P., Mao, X. et al. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Mol Biol Evol*, **24** (3): 710–22 (2007)

[112] Roberts, D. F. and Kahlon, D. P. Environmental correlations of skin colour. *Ann Hum Biol*, **3** (1): 11–22 (1976)

[113] Jablonski, N. G. and Chaplin, G. The evolution of human skin coloration. *J Hum Evol*, **39** (1): 57–106 (2000)

[114] Comas, D., Plaza, S., Wells, R. S., Yuldasheva, N., Lao, O. et al. Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. *Eur J Hum Genet*, **12** (6): 495–504 (2004)

[115] Quintana-Murci, L., Chaix, R., Wells, R. S., Behar, D. M., Sayar, H. et al. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am J Hum Genet*, **74** (5): 827–45 (2004)

[116] Yuasa, I., Umetsu, K., Harihara, S., Kido, A., Miyoshi, A. et al. Distribution of the F374 allele of the SLC45A2 (MATP) gene and founder-haplotype analysis. *Ann Hum Genet*, **70** (6): 802–811 (2006)

[117] Pamjav, H., Zalan, A., Beres, J., and Nagy, Y., M.Chang. Genetic structure of the paternal lineage of the Roma people. *Am J Phys Anthropol*, **145** (1): 21–29 (2011)

[118] Ochildiev, D., Pinkhasov, R., and Kalontarov, I. *A history and culture of the Bukharian Jews*. Club Roshnoyi Light, New York (2007)

[119] Bartlett, S., Straub, J., Tonks, S., Wells, R. S., Bodmer, J. G. et al. Alkaline-mediated differential interaction (AMDI): a simple automatable single-nucleotide polymorphism assay. *Proc Natl Acad Sci U S A*, **98** (5): 2694–7 (2001)

[120] Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M. et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*, **23** (2): 147 (1999)

[121] Maddison, W. and Maddison, D. Interactive analysis of phylogeny and character evolution using the computer program MacClade. *Folia Primatol (Basel)*, **53** (1-4): 190–202 (1989)

[122] Richards, M. B., Macaulay, V. A., Bandelt, H. J., and Sykes, B. C. Phylogeography of mitochondrial DNA in western Europe. *Ann Hum Genet*, **62** (Pt 3): 241–60 (1998)

[123] Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B. et al. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet*, **67** (5): 1251–76 (2000)

[124] Tambets, K., Rootsi, S., Kivisild, T., Help, H., Serk, P. et al. The western and eastern roots of the Saami – the story of genetic "outlier" told by mitochondrial DNA and Y chromosomes. *Am J Hum Genet*, **74** (4): 661–82 (2004)

[125] Derenko, M., Malyarchuk, B., Denisova, G. A., Wozniak, M., Dambueva, I. et al. Contrasting patterns of Y-chromosome variation in South Siberian populations from Baikal and Altai-Sayan regions. *Hum Genet*, **118** (5): 591–604 (2006)

[126] Yao, Y. G., Kong, Q. P., Wang, C. Y., Zhu, C. L., and Zhang, Y. P. Different matrilineal contributions to genetic structure of ethnic groups in the Silk Road region in China. *Mol Biol Evol*, **21** (12): 2265–80 (2004)

[127] Excoffier, L., Laval, G., and Schneider, S. Arlequin ver 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioniformatics Online*, **1**: 47–50 (2005)

[128] Cleveland, W., Grosse, E., and Shyu, W. Local regression models. In J. Chambers and T. Hastie, editors, *Statistical Models*. Wadsworth & Brooks/Cole, Pacific Grove, CA (1992)

[129] Comas, D., Calafell, F., Mateu, E., Perez-Lezaun, A., Bosch, E. et al. Trading genes along the silk road: mtDNA sequences and the origin of central Asian populations. *Am J Hum Genet*, **63** (6): 1824–38 (1998)

[130] Frachetti, M. Migration Concepts in Central Eurasian Archaeology. *Annu Rev Anthropol*, **40**: 191–212 (2011)

[131] Zhong, H., Shi, H., Qi, X. B., Duan, Z. Y., Tan, P. P. et al. Extended Y-chromosome investigation suggests post-Glacial migrations of modern humans into East Asia via the northern route. *Mol Biol Evol* (2010)

[132] Chaix, R., Austerlitz, F., Hegay, T., Quintana-Murci, L., and Heyer, E. Genetic traces of east-to-west human expansion waves in Eurasia. *Am J Phys Anthropol*, **136** (3): 309–17 (2008)

[133] Abdulla, M. A., Ahmed, I., Assawamakin, A., Bhak, J., Brahmachari, S. K. et al. Mapping human genetic diversity in Asia. *Science*, **326** (5959): 1541–5 (2009)

[134] Frenzel, B., Pecsi, M., and Velichko, A. *Atlas of Paleoclimates and Paleoenvironments of the Northern Hemisphere: Late Pleistocene-Holocene*. Hungarian Academy of Sciences, Budapest (1992)

[135] Hummel, S., Schmidt, D., Kremeyer, B., Herrmann, B., and Oppermann, M. Detection of the CCR5-Delta32 HIV resistance gene in Bronze Age skeletons. *Genes Immun*, **6** (4): 371–4 (2005)

[136] Sabeti, P. C., Walsh, E., Schaffner, S. F., Varilly, P., Fry, B. et al. The case for selection at CCR5-Delta32. *PLoS Biol*, **3** (11): 378 (2005)

[137] Galvani, A. P. and Novembre, J. The evolutionary history of the CCR5-Delta32 HIV-resistance mutation. *Microbes Infect*, **7** (2): 302–9 (2005)

[138] Novembre, J., Galvani, A. P., and Slatkin, M. The geographic spread of the CCR5 Delta32 HIV-resistance allele. *PLoS Biol*, **3** (11): e339 (2005)

[139] Bollback, J. P., York, T. L., and Nielsen, R. Estimation of 2Nes from temporal allele frequency data. *Genetics*, **179** (1): 497–502 (2008)

[140] Novembre, J. and Han, E. Human population structure and the adaptive response to pathogen-induced selection pressures. *Philos Trans R Soc Lond B Biol Sci*, **367**: 878–8865 (2012)

[141] Johansson, A., Ingman, M., Mack, S. J., Erlich, H., and Gyllensten, U. Genetic origin of the Swedish Sami inferred from HLA class I and class II allele frequencies. *Eur J Hum Genet*, **16** (11): 1341–9 (2008)

[142] Salmela, E., Lappalainen, T., Fransson, I., Andersen, P. M., Dahlman-Wright, K. et al. Genome-wide analysis of single nucleotide polymorphisms uncovers population structure in Northern Europe. *PLoS ONE*, **3** (10): e3519 (2008)

[143] Lappalainen, T., Koivumaki, S., Salmela, E., Huoponen, K., Sistonen, P. et al. Regional differences among the Finns: a Y-chromosomal perspective. *Gene*, **376** (2): 207–15 (2006)

[144] Huyghe, J. R., Fransen, E., Hannula, S., Van Laer, L., Van Eyken, E. et al. A genome-wide analysis of population structure in the Finnish Saami with implications for genetic association studies. *Eur J Hum Genet* (2010)

[145] Martinez-Cruz, B., Vitalis, R., Segurel, L., Austerlitz, F., Georges, M. et al. In the heartland of Eurasia: the multilocus genetic landscape of Central Asian populations (2011)

[146] Flanagan, N., Healy, E., Ray, A., Philips, S., Todd, C. et al. Pleiotropic effects of the melanocortin 1 receptor (MC1R) gene on human pigmentation. *Hum Mol Genet*, **9** (17): 2531–7 (2000)

[147] Bastiaens, M. T., ter Huurne, J. A., Kielich, C., Gruis, N. A., Westendorp, R. G. et al. Melanocortin-1 receptor gene variants determine the risk of nonmelanoma skin cancer independently of fair skin and red hair. *Am J Hum Genet*, **68** (4): 884–94 (2001)

[148] Jablonski, N. G. The Evolution of Human Skin and Skin Color. *Annu Rev Anthropol*, **33**: 585–623 (2004)

[149] John, P. R., Makova, K., Li, W. H., Jenkins, T., and Ramsay, M. DNA polymorphism and selection at the melanocortin-1 receptor gene in normally pigmented southern African individuals. *Ann N Y Acad Sci*, **994**: 299–306 (2003)

[150] Makova, K. D., Ramsay, M., Jenkins, T., and Li, W. H. Human DNA sequence variation in a 6.6-kb region containing the melanocortin 1 receptor promoter. *Genetics*, **158** (3): 1253–68 (2001)

[151] Rogers, A., Iltis, D., and Wooding, S. Genetic Variation at the MC1R Locus and the Time since Loss of Human Body Hair. *Curr Anthropol*, **45** (1) (2004)

[152] Harding, R. M., Healy, E., Ray, A. J., Ellis, N. S., Flanagan, N. et al. Evidence for variable selective pressures at MC1R. *Am J Hum Genet*, **66** (4): 1351–61 (2000)

[153] Rana, B. K., Hewett-Emmett, D., Jin, L., Chang, B. H., Sambuughin, N. et al. High polymorphism at the human melanocortin 1 receptor locus. *Genetics*, **151** (4): 1547–57 (1999)

[154] Lao, O., de Gruijter, J. M., van Duijn, K., Navarro, A., and Kayser, M. Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. *Ann Hum Genet*, **71** (Pt 3): 354–69 (2007)

[155] Edwards, M., Bigham, A., Tan, J., Li, S., Gozdzik, A. et al. Association of the OCA2 polymorphism His615Arg with melanin content in east Asian populations: further evidence of convergent evolution of skin pigmentation. *PLoS Genet*, **6** (3) (2010)

[156] Mundy, N. I. and Kelly, J. Evolution of a pigmentation gene, the melanocortin-1 receptor, in primates. *Am J Phys Anthropol*, **121** (1): 67–80 (2003)

[157] Cavalli-Sforza, L. L. Some old and new data on the genetics of human populations. *Ala J Med Sci*, **3** (4): 376–81 (1966)

[158] Lewontin, R. C. and Krakauer, J. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74** (1): 175–95 (1973)

[159] Xue, Y., Wang, Q., Long, Q., Ng, B. L., Swerdlow, H. et al. Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Curr Biol*, **19** (17): 1453–7 (2009)

[160] Lalueza-Fox, C., Rompler, H., Caramelli, D., Staubert, C., Catalano, G. et al. A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. *Science*, **318** (5855): 1453–1455 (2007)

[161] Coles, B. Doggerland: A Speculative Survey. *Proceedings of the Prehistoric Society*, **64**: 45–81 (1998)

[162] Weninger, B., Schulting, R. J., Bradtmüller, M., Clare, L., Collard, M. et al. The catastrophic final flooding of Doggerland by the Storegga Slide tsunami. *Documenta Praehistorica*, **35**: 2–24 (2008)

[163] Bradley, R. *The Prehistory of Britain and Ireland.* Cambridge University Press, Cambridge (2007)

[164] Rouse, I. *Migrations in Prehistory: Inferring Population Movement from Cultural Remains.* Yale University Press, New Haven (1986)

[165] Sherratt, A. The genesis of megaliths: monumentality, ethnicity and social complexity in Neolithic north-west Europe. *World Archaeology*, **22**: 147–167 (1990)

[166] Chapman, R. Ten years after - megaliths, mortuary practices and the territorial model. In L. Beck, editor, *Regional Approaches to Mortuary Analysis.* Plenum Press, New York (1995)

[167] Jackson, K. *Language and history in early Britain : a chronological survey of the Brittonic languages, first to twelfth century A.D.* Edinburgh University Press, Edinburgh (1953)

[168] Schrijver, P. The rise and fall of British Latin: Evidence from English and Brittonic. In M. Filppula, J. Klemola, and H. Pitknen, editors, *The Celtic roots of English.* Joensuu University Press, Joensuu (2002)

[169] Mattingly, D. *An Imperial Possession. Britain in the Roman Empire.* Penguin, London (2007)

[170] Hamerow, H. Migration Theory and the Anglo-Saxon Identity Crisis. In J. Chapman and H. Hamerow, editors, *Migrations and Invasions in Archaeological Explanation.* Archaeopress, Oxford (1997)

[171] Crawford, S. Britons, Anglo-Saxons and the Germanic Burial Ritual. In J. Chapman and H. Hamerow, editors, *Migrations and Invasions in Archaeological Explanations.* BAR International Series 664, Oxford (1997)

[172] Fleming, R. *Britain after Rome: the fall and rise, 400-700.* Allen Lane, London (2010)

[173] Barton, N. *The Late Glacial in north-west Europe: human adaptation and environmental change at the end of the Pleistocene.* Council for British Archaeology, London (1991)

[174] Clark, P., Dyke, A. S., Shakun, J., Carlson, A., Clark, J. et al. The Last Glacial Maximum. *Science*, **325**: 710–714 (2009)

[175] Banks, W. E., d'Errico, F., Peterson, A. T., Vanhaeren, M., Kageyama, M. et al. Human ecological niches and ranges during the LGM in Europe derived from an application of eco-cultural niche modeling. *Journal of Archaeological Science*, **35** (2): 481–491 (2008)

[176] Gamble, C. The archaeological and genetic foundations of the European population during the Late Glacial: implications for 'agricultural' thinking. *Camb. Archaeol. J.*, **15**: 193–223 (2005)

[177] Djindjian, F. Identité, chronologie et territoires du Magdalénien en Europe occidentale: questions posées. In *Le Paléolithique supérieur récent: nouvelles données sur le peuplement et l'environment.* Société Préhistorique Française (2000)

[178] Breest, K. and Veil, S. The Late Upper Palaeolithic site of Schweskau, Ldkr Lücknow-Dannenberg (Germany), and some comments on the relationship between the Magdalenian and the Hamburgian. In N. Barton, A. Roberts, and D. Roe, editors, *The Late Glacial in north-west Europe: human adaptation and environmental change at the end of the Pleistocene*, CBA Research Report. Council for British Archaeology, London (1991)

[179] Dolukhanov, P. Foraging and Farming Groups in Northeastern and Northwestern Europe – Identity and Interaction. In J. Chapman and P. Dolukhanov, editors, *Cultural Transformations and Interactions in Eastern Europe.* Aldershot, Avebury (1993)

[180] Jacobi, R. The Creswellian, Creswell and Cheddar. In N. Barton, A. Roberts, and D. Roe, editors, *The Late Glacial in north-west Europe: human adaptation and environmental change at the end of the Pleistocene*, CBA Research Reports. Council for British Archaeology, London (1991)

[181] Jacobi, R. M. and Higham, T. F. G. The Late Upper Palaeolithic Recolonisation of Britain - New Results from AMS Radiocarbon Dating. *Developments in Quarternary Science*, **14**: 201–225 (2011)

[182] Barton, R. N. E., Jacobi, R. M., Stapert, D., and Street, M. J. The late-glacial reoccupation of the British Isles and the Creswellian. *Journal of Quaternary Science*, **18** (7): 631–643 (2003)

[183] Barton, N. Technological innovation and continuity at the end of the Pleistocene in Britain. In N. Barton, A. Roberts, and D. Roe, editors, *The Late Glacial in north-west Europe: human adaptation and environmental change at the end of the Pleistocene*, CBA Research Report. Council for British Archaeology, London (1991)

[184] Fragnart, J.-P. New observations on the Late Upper Palaeolithic site of Belloy-sur-Somme (Somme, France). In N. Barton, A. Roberts, and D. Roe, editors, *The Late Glacial in north-west Europe: human adaptation and environmental change at the end of the Pleistocene*, CBA Research Report. Council for British Archaeology, London (1991)

[185] Newell, R. R. and Constandse Westermann, T. S. The use of ethnographic analyses for researching Late Palaeolithic settlement systems, settlement patterns and land use in the Northwest European Plain. *World Archaeology*, **27** (3): 372–388 (1996)

[186] Verhart, L. New Developments in the Study of the Mesolithic of the Low Countries. In G. Bailey and P. Spikins, editors, *Mesolithic Europe*. Cambridge University Press, Cambridge (2008)

[187] Reynier, M. *Early Mesolithic Britain: Origins, developments and directions*. Archaeopress, Oxford (2005)

[188] Bailey, G. Mesolithic Europe: Overview and New Problems. In G. Bailey and P. Spikins, editors, *Mesolithic Europe*. Cambridge University Press, Cambridge (2008)

[189] Valdeyron, N. The Mesolithic in France. In G. Bailey and P. Spikins, editors, *Mesolithic Europe*. Cambridge University Press, Cambridge (2008)

[190] Zvelebil, M. Innovating Hunter-Gatherers. In G. Bailey and P. Spikins, editors, *Mesolithic Europe*. Cambridge University Press, Cambridge (2008)

[191] Bocquet-Appel, J. and Demars, P.-Y. Population Kinetics in the Upper Palaeolithic in Western Europe. *J Arch Sci*, **27**: 551–570 (2000)

[192] Price, T. D. Europe's first farmers: an introduction. In T. D. Price, editor, *Europe's First Farmers*. Cambridge University Press, Cambridge (2000)

[193] Tringham, R. Southeastern Europe in the transition to agriculture in Europe: bridge, buffer, or mosaic. In T. D. Price, editor, *Europe's First Farmers*. Cambridge University Press, Cambridge (2000)

[194] Bogucki, P. How agriculture came to north-central Europe. In T. D. Price, editor, *Europe's First Farmers*. Cambridge University Press, Cambridge (2000)

[195] Gronenborn, D. Beyond the models: 'Neolithisation' in Central Europe. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe*. Oxford University Press, Oxford (2007)

[196] Barnett, W. Cardial pottery and the agricultural transition in Mediterranean Europe. In T. D. Price, editor, *Europe's First Farmers*. Cambridge University Press, Cambridge (2000)

[197] Robb, J. and Miracle, P. Beyond 'migration' versus 'acculturation': new models for the spread of agriculture. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe*. Oxford University Press, Oxford (2007)

[198] van Berg, P. Céramique du Limbourg et néolithisation en Europe du nord-ouest. In D. Cahen and M. Otte, editors, *Rubané et Cardial. Actes du colloque international de Liège 1988*. Université de Liège, Liège (1990)

[199] Allard, P. The Mesolithic-Neolithic transition in the Paris Basin: a review. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe.* Oxford University Press, Oxford (2007)

[200] Zvelebil, M. Plant use in the Mesolithic and the implications for the transitions to farming. *Proc. Prehist. Soc.*, **60**: 35–75 (1994)

[201] Larsen, C. Biological Changes in Human Populations with Agriculture. *Annu Rev Anthropol*, **24**: 185–213 (1995)

[202] Jackes, M., Lubell, D., and Meiklejohn, C. Healthy but mortal: human biology and the first farmers of western Europe. *Antiquity*, **71**: 639–658 (1997)

[203] Roberts, C. and Cox, M. *Health and Disease in Britain. From Prehistory to the Present Day.* Sutton Publishing, Stroud (2003)

[204] Wittwer-Backofen, U. and Tomo, N. From Health to Civilization Stress? In Search for Traces of a Health Transition During the Early Neolithic in Europe. In J. Bocquet-Appel and O. Bar-Yosef, editors, *The Neolithic Demographic Transition and its Consequences.* Springer, Berlin (2008)

[205] Mummert, A., Esche, E., Robinson, J., and Armelagos, G. Stature and robusticity during the agricultural transition: Evidence from the bioarchaeological record. *Econ Hum Biol*, **9** (3): 284–301 (2011)

[206] Louwe Kooijmans, L. The gradual transition to farming in the Lower Rhine Basin. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe.* Oxford University Press, Oxford (2007)

[207] Larsson, L. Mistrust traditions, consider innovations? The Mesolithic-Neolithic transition in southern Scandinavia. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe.* Oxford University Press, Oxford (2007)

[208] Binder, D. Mesolithic and Neolithic interaction in southern France and northern Italy: new data and current hypotheses. In T. D. Price, editor, *Europe's First Farmers.* Cambridge University Press, Cambridge (2000)

[209] Forenbaher, S. and Miracle, P. The spread of farming in the Eastern Adriatic. *Antiquity*, **79**: 514–528 (2005)

[210] Guilaine, J. and Manen, C. From Mesolithic to Early Neolithic in the western Mediterranean. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe.* Oxford University Press, Oxford (2007)

[211] Zilhao, J. From the Mesolithic to the Neolithic in the Iberian peninsula. In T. D. Price, editor, *Europe's First Farmers.* Cambridge University Press, Cambridge (2000)

[212] Zvelebil, M. Farmers, our ancestors and the identity of Europe. In P. Graves-Brown, C. Gamble, and S. Jones, editors, *Cultural identity and archaeology : the construction of European communities*. Routledge, London (1996)

[213] Edwards, C. J., Bollongino, R., Scheu, A., Chamberlain, A., Tresset, A. et al. Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proc Biol Sci*, **274** (1616): 1377–85 (2007)

[214] Larson, G., Albarella, U., Dobney, K., Rowley-Conwy, P., Schibler, J. et al. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc Natl Acad Sci U S A*, **104** (39): 15276–81 (2007)

[215] Stock, F., Edwards, C. J., Bollongino, R., Finlay, E. K., Burger, J. et al. Cytochrome b sequences of ancient cattle and wild ox support phylogenetic complexity in the ancient and modern bovine populations. *Anim Genet*, **40** (5): 694–700 (2009)

[216] Bollongino, R. and Burger, J. Neolithic cattle domestication as seen from ancient DNA. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe*. Oxford University Press, Oxford (2007)

[217] Götherström, A., Anderung, C., Hellborg, L., Elburg, R., Smith, C. et al. Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proc Biol Sci*, **272**: 2345–2350 (2005)

[218] Bollongino, R., Elsner, J., Vigne, J., and Burger, J. Y-SNPs do not indicate hybridisation between European aurochs and domestic cattle. *PLoS One*, **10** (2008)

[219] Whittle, A. *Europe in the Neolithic: the creation of new worlds*. Cambridge University Press, Cambridge (1996)

[220] Scarre, C. Traditions and tranformations in Neolithic France (Review article). *Antiquity*, **72**: 945–948 (1998)

[221] Vanmontfort, C. Bridging the Gap. The Mesolithic-Neoltihic transition in a frontier zone. *Documenta Praehistorica*, **34**: 105–118 (2007)

[222] Scarre, C. Modelling prehistoric populations : the case of Neolithic Brittany. *Journal of Anthropological Archaeology*, **20** (3): 285–313 (2001)

[223] Scarre, C., Laporte, L., and Joussaume, R. Long mounds and megalithic origins in western France : recent excavations at Priss-la-Charrire. *Proceedings of the Prehistoric Society*, **69**: 235–251 (2003)

[224] Childe, V. The Origin of Neolithic Culture in Northern Europe. *Antiquity*, **23** (91): 129–135 (1949)

[225] Hodder, I. Burials, houses, women and men in the European Neolitihic. In D. Miller and C. Tilley, editors, *Ideology, power and prehistory.* Cambridge University Press, Cambridge (1984)

[226] Scarre, C. Contexts of monumentalism: Regional diversity at the Neolithic transition in north-west France. *Oxford Journal of Archaeology*, **21**: 23–61 (2002)

[227] Scarre, C. The Early Neolithic of western France and Megalithic origins in Atlantic Europe. *Oxford Journal of Archaeology*, **11**: 121–154 (1992)

[228] Constantin, C., Mordant, D., and Simonin, D. *La Culture de Cerny. Nouvelle économie, nouvelle société au Néolithique.* Association pour la promotion de la recherche archéologique en Ile de France, Nemours (1997)

[229] Cassen, S. Material culture and chronology of the middle Neolithic of western France. *Oxford Journl of Archaeology*, **12**: 197–208 (1993)

[230] Gleser, R. Periodisierung, Verbreitung und Entstehung der älteren Michelsberger Kultur. In J. Biel, H. Schlichtherle, M. Strobel, and A. Zeeb, editors, *Die Michelsberger Kultur und ihre Randgebiete - Probleme der Entstehung, Chronologie und des Siedlungswesens.* Konrad Theiss Verlag, Stuttgart (1998)

[231] Dubouloz, J. Réflexions sur le Michelsberg ancien en Bassin parisien. In J. Biel, H. Schlichtherle, M. Strobel, and A. Zeeb, editors, *Die Michelsberger Kultur und ihre Randgebiete - Probleme der Entstehung, Chronologie und des Siedlungswesens.* Konrad Theiss Verlag, Stuttgart (1998)

[232] Jeunesse, C. Pour une origine occidentale de la culture de Michelsberg? In J. Biel, H. Schlichtherle, M. Strobel, and A. Zeeb, editors, *Die Michelsberger Kultur und ihre Randgebiete - Probleme der Entstehung, Chronologie und des Siedlungswesens.* Konrad Theiss Verlag, Stuttgart (1998)

[233] Raemaekers, D. *The Articulation of a 'New Neolithic' - The meaning of the Swifterbant Culture for the process of neolithisation in the western part of the North European Plain (4900-3400 BC).* Faculty of Archaeology, University of Leiden, Leiden (1999)

[234] Hibbs, J. The Neolithic of Brittany. In C. Scarre, editor, *Ancient France. Neolithic societies and their landscapes 6000-2000 BC.* Edinburgh University Press, Edinburgh (1983)

[235] Giot, P.-R. Qu'y-a-t-il de Chasséen en Bretagne? In A. Beeching, D. Binder, J.-C. Blanchet, C. Constantin, J. Dubouloz, R. Martinez, D. Mordant, J.-P. Thevenot, and J. Vaquer, editors, *Identité du Chasséen.* Association pour la promotion de la recherche archéologique en Ile de France, Nemours (1991)

[236] Chancerel, A. and Billard, C. Le Chasséen en Normandie. In A. Beeching, D. Binder, J.-C. Blanchet, C. Constantin, J. Dubouloz, R. Martinez, D. Mordant, J.-P. Thevenot, and J. Vaquer, editors, *Identité du Chasséen.* Association pour la promotion de la recherche archéologique en Ile de France, Nemours (1991)

[237] L'Helgouach, J. Néolithique moyen armoricain, es-tu Chasséen? (Chasséen et megaliths). In A. Beeching, D. Binder, J.-C. Blanchet, C. Constantin, J. Dubouloz, R. Martinez, D. Mordant, J.-P. Thevenot, and J. Vaquer, editors, *Identité du Chasséen.* Association pour la promotion de la recherche archéologique en Ile de France, Nemours (1991)

[238] Sherratt, A. The Emergence of Elites: Earlier Bronze Age Europe, 2500-1300 BC. In B. Cunliffe, editor, *The Oxford Illustrated History of Prehistoric Europe.* Oxford University Press, Oxford (1994)

[239] Case, H. Beakers and Beaker Culture. In C. Burgess, P. Topping, and F. Leach, editors, *Beyond Stonehenge: Essays on the Bronze Age in honour of Colin Burgess.* Oxbow Books, Oxford (2007)

[240] Cunliffe, B. *Facing the ocean : the Atlantic and its peoples, 8000 BC-AD 1500.* Oxford University Press, Oxford (2001)

[241] Needham, S. Transforming Beaker Culture in North-West Europe; Processes of Fusion and Fission. *Proceedings of the Prehistoric Society*, **71**: 171–217 (2005)

[242] Vander Linden, M. What linked the Bell Beakers in third millennium BC Europe? *Antiquity*, **81**: 343–352 (2007)

[243] Pearson, M. P. The Earlier Bronze Age. In J. Hunter and I. Ralston, editors, *The Archaeology of Britain.* Routledge, Oxford (1999)

[244] Harrison, R. *The Beaker Folk: Copper Age Archaeology in Western Europe.* Thames and Hudson, London (1980)

[245] Vander Linden, M. For Whom the Bell Tolls: Social Hierarchy vs Social Integration in the Bell Beaker Culture of Southern France (Third Millennium BC). *Cambridge Archaeological Journal*, **16** (3): 317–332 (2006)

[246] Brodie, N. New Perspectives on the Bell-Beaker Culture. *Oxford Journal of Archaeology*, **16** (3): 297–314 (1997)

[247] Shennan, S. Bell Beakers and their context in central Europe. In J. Lanting and J. van der Wals, editors, *Glockenbecher Symposion. Oberried 1974.* Fibula-Van Dishoeck, Haarlem (1976)

[248] Vander Linden, M. *Le phénomène campaniforme dans l'Europe du 3ème millénaire avant notre ère. Synthèse et nouvelles perspectives.* Archaeopress, Oxford (2006)

[249] Harding, A. *European societies in the Bronze Age.* Cambridge University Press, Cambridge (2000)

[250] Coles, J. and Harding, A. *The Bronze Age in Europe: An introduction to the prehistory of Europe c.2000-700 BC.* Methuen & Co. Ltd., London (1979)

[251] Billard, C., Blanchet, J.-C., Talon, M., Billand, G., Bulard, A. et al. Origine et composantes de l'âge du Bronze ancien dans le nord-ouest de le France . In C. Mordant and O. Gaiffe, editors, *Cultures et sociétés du Bronze ancien en Europe*. CTHS, Paris (1992)

[252] Theunissen, L. La Culture de Hilversum: Les sociétés de l'âge du Bronze aux Pays-Bas. In *Actes du 117eme Congrès Sociétés Savantes (Clermont-Ferrant, 1992)*. CTHS, Paris (1996)

[253] Warmenbol, E. Les débuts de l'Âge du Bronze en Belgique. In *Actes du 117eme Congrès Sociétés Savantes (Clermont-Ferrant, 1992)*. CTHS, Paris (1996)

[254] Marcigny, C., Ghesquiere, E., and Kinnes, I. 'Bronze Age Cross Channel Relations. The Lower Normandy (France) Example: Ceramic Chronology and First Reflections. In C. Burgess, P. Topping, and F. Lynch, editors, *Beyond Stonehenge: Essays on the Bronze Age in Honour of Colin Burgess*. Oxbow Books, Oxford (2007)

[255] Garwood, P. Before the hills in order stood: chronology, time and history in the interpretation of Early Bronze Age round barrows. In J. Last, editor, *Beyond the Grave: New Perspectives on Barrows*. Oxbow Books, Oxford (2007)

[256] Burgess, C. The bronze age. In C. Renfrew, editor, *British prehistory*. Duckworth, London (1974)

[257] O'Connor, B. *Cross-channel relations in the later Bronze Age : relations between Britain, North-Eastern France, and the low countries during the later Bronze Age and the early Iron Age, with particular reference to the metalwork*. BAR, Oxford (1980)

[258] Champion, T. The Later Bronze Age. In J. Hunter and I. Ralston, editors, *The Archaeology of Britain*. Routledge, Oxford (1999)

[259] Warmenbol, E. Le Bronze Final atlantique entre Côte et Escaut. In C. Chevillot and A. Coffyn, editors, *L'Âge du Bronze Atlantique*. A.MU.SA., Dordognee (1991)

[260] Kristiansen, K. *Europe before History*. Routledge, London (1998)

[261] Brun, P. Le Bronze atlantique et ses subdivisions culturelles : essai de définition. In C. Chevillot and A. Coffyn, editors, *L'Âge du Bronze Atlantique*. A.MU.SA., Dordogne (1991)

[262] Cunliffe, B. Looking forward: maritime contacts in the first millennium BC. In P. Clark, editor, *Bronze Age Connections: Cultural Contact in Prehistoric Europe*. Oxbow Books, Oxford (2009)

[263] Eogan, G. Ideas, People and Things: Ireland and the External World during the Later Bronze Age. In J. Waddell and E. Shee Twohig, editors, *Ireland in the Bronze Age*. Stationary Office, Dublin (1995)

[264] Roussout-Larroque, J. Bronze moyen et récent du Médoc et middle Bronze Age II: des connections atlantiques. In C. Burgess, editor, *Beyond Stonehenge: Essays on the Bronza Age in Honour of Colin Burgess.* Oxbow Books, Oxford (2007)

[265] Cunliffe, B. *The Ancient Celts.* Penguin Books, London (1997)

[266] Raftery, B. *Pagan Celtic Ireland.* Thames and Hudson, London (1994)

[267] Cunliffe, B. *Iron Age Communities in Britain.* Routledge, Abingdon, fourth edition (2005)

[268] Härke, H. Transformation and collapse: Bronze Age to Iron Age settlement in west central Europe. In M. Sørensen and R. Thomas, editors, *The Bronze Age - Iron Age transition in Europe: Aspects of Continuity and Change in European societies c. 1200-500 B.C.* Archaeopress, Oxford (1989)

[269] Parzinger, H. The Beginning of La Tène culture in Central Europe. In J. Hill and C. Cumberpatch, editors, *Different Iron Ages: Studies on the Iron Age in Temperate Europe.* Tempus Reperatum, Oxford (1995)

[270] Haselgrove, C. Rethinking Earlier Iron Age settlement in the eastern Paris Basin. In C. Haselgrove, editor, *The Earlier Iron Age in Britain and the Near Continent.* Oxbow Books, Oxford (2007)

[271] Büschenschütz, O. Urbanisme et oppida. *Rev archéol Ouest*, **Suppl.3** (1990)

[272] Brun, P. Oppida and social complexification in France. In J. Hill and C. Cumberpatch, editors, *Different Iron Ages: Studies on the Iron Age in Temperate Europe.* Tempus Reperatum, Oxford (1995)

[273] Wells, P. Boundaries and identity in Early Iron Age Europe. In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent.* Oxbow Books, Oxford (2007)

[274] Audouze, F. and Büschenschütz, O. *Towns, Villages and Countryside of Celtic Europe.* B.T. Batsford, London (1991)

[275] Haselgrove, C. The age of enclosure: Later Iron Age settlement and society in northern France. In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond.* Oxbow Books, Oxford (2007)

[276] Roymans, N. Understanding social change in the Late Iron Age Lower Rhine region. In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond.* Oxbow Books, Oxford (2007)

[277] Julius Caesar, G. and Edwards, H. *The Gallic War.* Harvard University Press, London (1994)

[278] Lhuyd, E. *Archaeologica Britannica [...] of the Languages, Histories and Customs of the Original Inhabitants of Great Britain: from Collections and Observations in Travels through Wales, Cornwal, Bas-Bretagne, Ireland and Scotland. Vol. I. Glossography.* The Theater, Oxford (1707)

[279] Isaac, G. The nature of origins of the Celtic languages: Atlantic seaways, Italo-Celtic and other paralinguistic misapprehensions. *Studia Celtica*, **38**: 49–58 (2004)

[280] Collis, J. The origin and spread of the Celts. *Studia Celtica*, **30**: 17–34 (1996)

[281] Haselgrove, C. Iron Age Britain and its European setting. In J. Collis, editor, *Society and settlement in Iron Age Europe.* J.R. Collis Publications, Sheffield (2001)

[282] Le Bihan, J.-P., Bardel, J.-P., Menez, Y., and Tanguy, D. Les établissements ruraux du second âge du Fer en Armorique. *Rev archéol Ouest*, **Suppl.3** (1990)

[283] Giraud, P., Rohmer, M.-A., and Cocollos, A. *Gaulois sous les pommiers: découvertes de l'âge du Fer en Basse-Normandie, IXe-Ie siècle av. J.-C.* Cahier du Temps Editions, Caborg (2009)

[284] Menez, Y., Daire, M.-Y., Hyvert, J., Langouet, L., Le Bihan, J.-P. et al. Les Bâtiments de l'âge du Fer en Armorique. *Rev archéol Ouest*, **Suppl.3** (1990)

[285] Gerritsen, F. Familiar landscapes with unfamiliar pasts? Bronze Age barrows and Iron Age communities in the southern Netherlands. In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent.* Oxbow Books, Oxford (2007)

[286] Fontijn, D. and Fokkens, H. The emergence of early Iron Age 'chieftains' graves' in the southern Netherlands: reconsidering transformations in burial and depositional practices . In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent.* Oxbow Books, Oxford (2007)

[287] Verney, A. Les nécropoles de l'âge du Fer en Basse-Normandie. *Rev archéol Ouest*, **Suppl.6** (1993)

[288] Tanguy, D., Briard, J., Hyvert, J., Le Bihan, J.-P., and Menez, Y. Sépultres et habitats en Armorique à l'âge du Fer. *Rev archéol Ouest*, **Suppl.3** (1990)

[289] Pion, P. and Guichard, V. Tombes et nécropoles en France et au Luxembourg entre le IIIème et le Ier siècles avant J.-C. Essai d'inventaire. *Rev archéol Ouest*, **Suppl.6** (1993)

[290] Diepeveen-Jansen, M. Early La Tène burial practices and social (re)constructions in the Marne-Moselle region. In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent.* Oxbow Books, Oxford (2007)

[291] Metzler, J. Les sépultres de l'aristocratie en Gaule Belgique. *Rev archéol Ouest*, **Suppl.6** (1993)

[292] Carter, S. and Hunter, F. An Iron Age chariot burial from Scotland. *Antiquity*, **77**: 531–535 (2003)

[293] Røyrvik, E. Western Celts? A genetic impression of Britain in Atlantic Europe. In B. Cunliffe and J. Koch, editors, *Celtic from the West - Alternative Perspectives from Archaeology, Genetics, Language and Literature.* Oxbow Books, Oxford (2010)

[294] Lorrio, A. and Ruiz Zapatero, G. The Celts in Iberia: An Overview. *e-Keltoi: Journal of Interdisciplinary Celtic Studies*, **6** (2005)

[295] Herodotus and Godley, A. *Book II.* Harvard University Press, London (1926)

[296] Eska, J. and Ellis Evans, D. Continental Celtic. In M. Ball and J. Fife, editors, *The Celtic Languages.* Routledge, London (1993)

[297] Remy-Watte, M. Une nécropole de la fin de La Tène ancienne à Bordeaux-Saint-Clair (Seine-Maritime). *Rev archéol Ouest*, **55** (1998)

[298] De Hoz, J. Lepontic, Celt-Iberian, Gaulish and the Archaeological Evidence. *Etudes Celtiques*, **29**: 223–239 (1992)

[299] King, A. *Roman Gaul and Germany.* British Museum Publications, London (1990)

[300] Drinkwater, J. *Roman Gaul: the three provinces, 58 BC-AD 260.* Croom Helm, London (1983)

[301] Esmonde Cleary, S. *The Ending of Roman Britain.* Batsford, London (1989)

[302] Becker, K. Iron Age Ireland – finding an invisible people. In G. Cooney, K. Becker, J. Coles, M. Ryan, and S. Sievers, editors, *Relics of old decency: archaeological studies in later prehistory.* Wordwell, Dublin (2009)

[303] Wood, I. The Final Phase. In M. Todd, editor, *A Companion to Roman Britain.* Blackwell Publishing, Oxford (2003)

[304] Woolf, A. Apartheid and Economics in Anglo-Saxon England. In N. Higham, editor, *Britons in Anglo-Saxon England.* The Boydell Press, Woobridge (2007)

[305] James, E. *The Franks.* Blackwells, Oxford (1988)

[306] Sawyer, P. The Viking expansion. In K. Helle, editor, *The Cambridge History of Scandinavia.* Cambridge University Press, Cambridge (2008)

[307] Bates, D. *Normandy before 1066.* Longman, London (1982)

[308] Ridel, E. The Linguistic Heritage of the Scandinavians in Normandy. In J. Adams and K. Holman, editors, *Scandinavia and Europe 800-1350 - Contact, Conflict and Coexistence.* Brepols, Turnhout (2004)

[309] Brown, R. *The Normans.* Boydell Press, Woodbridge (1984)

[310] Ó Corráin, D. Vikings in Ireland and Scotland in the Ninth Century

[311] Crawford, B. The Vikings. In W. Davies, editor, *From the Vikings to the Normans.* Oxford University Press, Oxford (2003)

[312] Valante, M. Reassessing the Irish ‘Monastic Town’. *Irish Historical Studies*, **31** (121): 1–18 (1998)

[313] Wallace, P. Dublin in the Viking Age. In M. Ryan, editor, *The Illustrated Archaeology of Ireland*. Country House, Dublin (1991)

[314] Fellows-Jensen, G. Scandinavian settlement in the British Isles and Normandy: what the place-names reveal. In J. Adams and K. Holman, editors, *Scandinavia and Europe 800-1350 - Contact, Conflict and Coexistence*. Brepols, Turnhout (2004)

[315] Fox, C. *The personality of Britain: its influence on inhabitant and invader in prehistoric and early historic times*. National Museum of Wales and by the press board of the University of Wales, Cardiff (1943)

[316] Barton, N. The Lateglacial or Late and Final Upper Palaeolithic colonization of Britain. In J. Hunter and I. Ralston, editors, *The Archaeology of Britain*. Routledge, Oxford (1999)

[317] Pryor, A. Following the Fat: Food and Mobility in the European Upper Palaeolithic 45,000 to 18,000 BP. *Archaeological Review from Cambridge*, **23** (2): 161–179 (2008)

[318] Fischer, A. Pioneers in deglaciated landscapes: the expansion and adaptation of Late Palaeolithic societies in southern Scandinavia. In N. Barton, A. Roberts, and D. Roe, editors, *The Late Glacial in north-west Europe: human adaptation and environmental change at the end of the Pleistocene*, CBA Research Report. Council for British Archaeology, London (1991)

[319] Cook, J. and Jacobi, R. A reindeer antler or 'Lyngby' axe from Northamptonshire and its context in the British Late Glacial. *Proceedings of the Prehistoric Society*, **60**: 75–84 (1994)

[320] Ballin, T. B., Saville, A., Tipping, R., and Ward, T. An Upper Palaeolithic Flint and Chert Assemblage from Howburn Farm, South Lanarkshire, Scotland: First Results. *Oxford Journal of Archaeology*, **29** (4): 323–360 (2010)

[321] Blockley, S. M. Two hiatuses in human bone radiocarbon dates in Britain (17,000 to 5000 cal BP). *Antiquity*, **79** (305): 505–513 (2005)

[322] Mithen, S. Hunter-gatherers of the Mesolithic. In J. Hunter and I. Ralston, editors, *The Archaeology of Britain*. Routledge, Oxford (1999)

[323] Tolan-Smith, C. Mesolithic Britain. In G. Bailey and P. Spikins, editors, *Mesolithic Europe*. Cambridge University Press, Cambridge (2008)

[324] Jacobi, R. Britain inside and outside Mesolithic Europe. *Proceedings of the Prehistoric Society*, **42**: 67–84 (1976)

[325] Edmonds, M. *Stone tools and society : working stone in Neolithic and Bronze Age Britain* . Batsford, London (1995)

[326] Price, T. The Mesolithic of Northern Europe. *Annu Rev Anthropol*, **20**: 211–233 (1991)

[327] Thomas, J. Current debates on the Mesolithic-Neolithic transition in Britain and Ireland. *Documenta Praehistorica*, **31**: 113–120 (2004)

[328] Pailler, Y. and Sheridan, A. Everything you always wanted to know about... La Néolithisation de la Grande-Bretagne et d'Irlande. *Bulletin de la Société Préhistorique Française*, **106**: 25–56 (2009)

[329] Whittle, A. The Neolithic period, c. 4000-2500/2200 BC: Changing the World. In J. Hunter and I. Ralston, editors, *The Archaeology of Britain*. Routledge, Oxford (1999)

[330] Whittle, A. The temporality of transformation: dating the early development of the southern British Neolithic. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe*. Oxford University Press, Oxford (2007)

[331] Thomas, J. Thoughts on the 'Repacked' Neolithic Revolution. *Antiquity*, **77**: 67–75 (2004)

[332] Whittle, A. The Coming of Agriculture - People, Landscape and Change ca. 4000-1500 BC. In P. Slack and R. Ward, editors, *The Peopling of Britain*. Oxford University Press, Oxford (2002)

[333] Noble, G. *Neolithic Scotland: timber, stone, earth and fire*. Edinburgh University Press, Edinburgh (2006)

[334] Callaghan, R. and Scarre, C. Simulating the Western Seaways. *Oxford Journal of Archaeology*, **28** (4): 357–372 (2009)

[335] Childe, V. *The Dawn of European Civilization*. Paladin, St. Albans (1973)

[336] Piggott, S. *The neolithic cultures of the British Isles : a study of the stone-using agricultural communities of Britain in the second millennium B.C.* . Cambridge University Press, Cambridge (1954)

[337] Case, H. Neolithic Explanations. *Antiquity*, **43**: 176–186 (1969)

[338] Sheridan, A. From Picardie to Pickering and Pencraig Hill? New information on the 'Carinated Bowl Neolithic' in northern Britain. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe*. Oxford University Press, Oxford (2007)

[339] Garrow, D. and Sturt, F. Grey waters bright with Neolithic argonauts? Maritime connections and the Mesolithic-Neolithic transition within the 'western seaways' of Britain, c. 5000-3500 BC. *Antiquity*, **85**: 59–72 (2011)

[340] Tresset, A. French Connections II: of cows and men. In I. Armit, E. Murphy, E. Nelis, and D. Simpson, editors, *Neolithic Settlement in Ireland and Western Britain*. Oxbow Books, Oxford (2003)

[341] Schulting, R. J. and Richards, M. Finding the coastal Mesolithic in southwest Britain: AMS dates and stable isotope results on human remains from Caldey Island, south Wales. *Antiquity*, **76**: 1011–1025 (2002)

[342] Richards, M. P. and Schulting, R. J. Touch not the fish: the Mesolithic-Neolithic change of diet and its significance. *Antiquity*, **80** (308): 444–456 (2006)

[343] Thomas, J. Mesolithic-Neolithic transitions in Britain: from essence to inhabitation. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe.* Oxford University Press, Oxford (2007)

[344] Collard, M., Edinborough, K., Shennan, S., and Thomas, M. Radiocarbon evidence indicates that migrants introduced farming to Britain. *Journal of Archaeological Science*, **37**: 866–870 (2010)

[345] Bronk Ramsey, C. Radiocarbon Dating: Revolutions in Understanding. *Archaeometry*, **50**: 249–275 (2008)

[346] Lemmen, C., Gronenborn, D., and Wirtz, K. A simulation of the Neolithic transition in Western Eurasia. *Journal of Archaeological Science*, **doi: 10.1016/j.jas.2011.08.008** (2010)

[347] Whittle, A., Bayliss, A., and Healy, F. The Timing and Tempo of Change: Examples from the Fourth Millennium cal. bc in Southern England. *Cambridge Archaeological Journal* (2008)

[348] Cummings, V. From midden to megalith? The Mesolithic-Neolithic transition in western Britain. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe.* Oxford University Press, Oxford (2007)

[349] Hey, G. and Barclay, A. The Thames Valley in the late fifth millennium cal BC: the appearance of domestication and the evidence for change. In A. Whittle and V. Cummings, editors, *Going Over. The Mesolithic-Neolithic Transition in North-West Europe.* Oxford University Press, Oxford (2007)

[350] Thomas, J. The Return of the Rinyo-Clacton Folk? The Cultural Significance of the Grooved Ware Complex in Later Neolithic Britain. *Cambridge Archaeoligcal Journal*, **20** (1): 1–15 (2010)

[351] Sheridan, A. Megaliths and Megalomania: An Account, and Interpretation, of the Development of Passage Tombs in Ireland. *The Journal of Irish Archaeology*, **3**: 17–30 (1986)

[352] Bradley, R. Position and Possession: Assemblage variation in the British Neolithic. *Oxford Journal of Archaeology*, **1** (1): 27–38 (1982)

[353] Gelling, M. and Cole, A. *The Landscape of Place-Names.* Shaun Tyas, Donington (2003)

[354] Roberts, B. Migration, Craft Expertise and Metallurgy: Analysing the 'Spread' of Metal in Western Europe. *Archaeological Review from Cambridge*, **23** (2): 27–45 (2008)

[355] Budd, P., Millard, A., Chenery, C., Lucy, S., and Roberts, C. Investigating population movement by stable isotope analysis: a report from Britain. *Antiquity*, **78** (299): 127–141 (2004)

[356] Pollard, A. Isotopes and impact: A cautionary tale. *Antiquity*, **85**: 631–638 (2011)

[357] Price, T. D., Knipper, C., Grupe, G., and Smrcka, V. Strontium Isotopes and Prehistoric Human Migration: The Bell Beaker Period in Central Europe. *European Journal of Archaeology*, **7** (1): 9–40 (2004)

[358] Evans, J. and Chenery, C. Bronze Age childhood migration of individuals near Stonehenge, revealed by strontium and oxygen isotope enamel analysis. *Archaeometry*, **48** (2): 309–321 (2006)

[359] Waddell, J. The Irish Sea in Prehistory. *The Journal of Irish Archaeology*, **6**: 29–40 (1992)

[360] Kavanagh, R. The Encrusted Urn in Ireland. *Proceedings of the Royal Irish Academy. Section C: Archaeology, Celtic Studies, History, Linguistics, Literature*, **73**: 507–617 (1973)

[361] Mordant, C. and Gaiffe, O. *Cultures et sociétés du Bronze ancien en Europe*. CTHS, Paris (1992)

[362] Bourgeois, J. and Talon, M. From Picardy to Flanders: Transmanche connections in the Bronze Age. In P. Clark, editor, *Bronze Age Connections: Cultural Contact in Prehistoric Europe*. Oxbow, Oxford (2009)

[363] Briard, J. L'Âge du Bronze Ancien Atlantique. In C. Mordant and O. Gaiffe, editors, *Cultures et sociétés du Bronze ancien en Europe*. CTHS, Paris (1992)

[364] Theunissen, L. British immigrants killed abroad in the seventies: the rise and fall of a Dutch culture. In P. Clark, editor, *Bronze Age Connections: Cultural Contact in Prehistoric Europe*. Oxbow Books, Oxford (2009)

[365] Fokkens, H. Bronze Age settlements in the Netherlands. In C. Chevillot and A. Coffyn, editors, *L'Age du Bronze Atlantique*. A.MU.SA., Dordogne (1991)

[366] Burgess, C. Urns, Culture du Wessex et la transition Bronze ancien – Bronze moyen en Grande-Bretagne. In C. Mordant and O. Gaiffe, editors, *Cultures et sociétés du Bronze ancien en Europe*. CTHS, Paris (1996)

[367] Coles, J. The Wessex culture: a minimal view. *Antiquity*, **45**: 6–14 (1971)

[368] Cooney, G. and Grogan, E. *Irish Prehistory: a social perspective*. Wordwell, Dublin (1994)

[369] Butler, C. The demise of the flint tool industry. In P. Clark, editor, *Bronze Age Connections: Cultural Contact in Prehistoric Europe*. Oxbow, Oxford (2009)

[370] Needham, S. Encompassing the Sea: “Maritories” and Bronze Age Maritime Interactions. In P. Clark, editor, *Bronze Age Connections: Cultural Contact in Prehistoric Europe*. Oxbow Books, Oxford (2009)

[371] Fontijn, D. Land at the other end of the sea? Metalwork circulation, geographical knowledge and the significance of British/Irish imports in the Bronze Age of the Low Countries. In P. Clark, editor, *Bronze Age Connections: Cultural Contact in Prehistoric Europe*. Oxbow Books, Oxford (2009)

[372] Needham, S. and Giardino, C. From Sicily to Salcombe: a Mediterranean Bronze Age object from British coastal waters. *Antiquity*, **82**: 60–72 (2007)

[373] Haselgrove, C. and Pope, R. Characterising the Earlier Iron Age. In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent*. Oxbow Books, Oxford (2007)

[374] Hill, J. The Iron Age in Britain and Ireland (c. 800 BC-AD 100): an overview. *Journal of World Prehistory*, **9**: 47–98 (1995)

[375] Haselgrove, C. The Iron Age. In J. Hunter and I. Ralston, editors, *The Archaeology of Britain*. Routledge, Oxford (1999)

[376] Henderson, J. The Atlantic West in the Early Iron Age. In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent*. Oxbow Books, Oxford (2007)

[377] Carr, G. Excarnation to cremation: continuity or change? In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond*. Oxbow Books, Oxford (2007)

[378] Gwilt, A. Silent Silures? Locating people and places in the Iron Age of south Wales. In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond*. Oxbow Books, Oxford (2007)

[379] Davies, M. Dominated by unenclosed settlement? The Later Iron Age in eastern Scotland north of the Forth. In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond*. Oxbow Books, Oxford (2007)

[380] Humphrey, J. Simple tools for tough tasks or tough tools for simple tasks? Analysis and experiment in Iron Age flint utilization. In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent*. Oxbow Books, Oxford (2007)

[381] Hunter, F. Artefacts, regions and identities in the northern British Iron Age. In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond*. Oxbow Books, Oxford (2007)

[382] Hill, C., Soares, P., Mormina, M., Macaulay, V., Clarke, D. et al. A mitochondrial stratigraphy for island southeast Asia. *Am J Hum Genet*, **80** (1): 29–43 (2007)

[383] Harding, D. Redefining the Northern British Iron Age. *Oxford Journal of Archaeology*, **25** (1): 61–82 (2006)

[384] Armit, I. Social landscapes and identities in the Irish Iron Age. In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond.* Oxbow Books, Oxford (2007)

[385] Ralstson, I. and Ashmore, P. The character of Earlier Iron Age societies in Scotland. In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent.* Oxbow Books, Oxford (2007)

[386] Cripps, L. Re-situating the Later Iron Age in Cornwall and Devon: new perspectives from the settlement record . In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond.* Oxbow Books, Oxford (2007)

[387] Hills, C. Anglo-Saxon Attitudes. In N. Higham, editor, *Britons in Anglo-Saxon England.* The Boydell Press, Woodbridge (2007)

[388] Hamilton, S. Cultural choices in the 'British Eastern Channel Area' in the Late Pre-Roman Iron Age. In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond.* Oxbow Books, Oxford (2007)

[389] O'Connor, B. Llyn Fawr metalwork in Britain: a review. In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent.* Oxbow Books, Oxford (2007)

[390] Haselgrove, C. and Moore, T. New narratives of the Later Iron Age. In C. Haselgrove and T. Moore, editors, *The Later Iron Age in Britain and Beyond.* Oxbow Books, Oxford (2007)

[391] Peuchet-Geilenbrügge, C. and Morzadec, H. The Iron Age in Lower Normandy: a definition through ceramics. In J. Collis, editor, *Society and settlement in Iron Age Europe.* J.R. Collis Publications, Sheffield (2001)

[392] Guillaumet, J.-P. and Eugene, A. À propos des trois objets métalliques, témoins des relations entre le pays éduen et les Iles Britanniques. In G. Cooney, K. Becker, J. Coles, M. Ryan, and S. Sievers, editors, *Relics of old decency: archaeological studies in later prehistory.* Wordwell, Dublin (2009)

[393] Giles, M. Refiguring rights in the Early Iron Age landscapes of East Yorkshire . In C. Haselgrove and R. Pope, editors, *The Earlier Iron Age in Britain and the Near Continent.* Oxbow Books, Oxford (2007)

[394] Todd, M. The Claudian Conquest and its Consequences. In M. Todd, editor, *A Companion to Roman Britain.* Blackwell Publishing, Oxford (2003)

[395] Hanson, W. Roman Britain: the military dimension. In J. Hunter and I. Ralston, editors, *The Archaeology of Britain.* Routledge, Oxford (1999)

[396] Crow, J. The Northern Frontier of Britain from Trajan to Antoninus Pius: Roman Builders and Native Britons. In M. Todd, editor, *A Companion to Roman Britain.* Blackwell Publishing, Oxford (2003)

[397] Manning, W. The Conquest of Wales. In M. Todd, editor, *A Companion to Roman Britain*. Blackwell Publishing, Oxford (2003)

[398] Hingley, R. Rural Settlement in Northern Britain. In M. Todd, editor, *A Companion to Roman Britain*. Blackwell Publishing, Oxford (2003)

[399] Esmonde Cleary, S. Roman Britain: civil and rural society. In J. Hunter and I. Ralston, editors, *The Archaeology of Britain*. Routledge, Oxford (1999)

[400] Jones, M. Cities and Urban Life. In M. Todd, editor, *A Companion to Roman Britain*. Blackwell Publishing, Oxford (2003)

[401] Allason-Jones, L. The Family in Roman Britain. In M. Todd, editor, *A Companion to Roman Britain*. Blackwell Publishing, Oxford (2003)

[402] Cramp, L., Evershed, R., and Eckhardt, H. What was a mortarium used for? Organic residues and cultural change in Iron Age and Roman Britain. *Antiquity*, **85** (2011)

[403] Scheidel, W. Marriage, families, and survival: demographic aspects. In P. Erdkamp, editor, *A Companion to the Roman Army*. Blackwell Publishing, Oxford (2007)

[404] Wesch-Klein, G. Recruits and Veterans. In P. Erdkamp, editor, *A Companion to the Roman Army*. Blackwell Publishing, Oxford (2007)

[405] Southern, P. The Army in Late Roman Britain. In M. Todd, editor, *A Companion to Roman Britain*. Blackwell Publishing, Oxford (2003)

[406] Fulford, M. Economic Structures. In M. Todd, editor, *A Companion to Roman Britain*. Blackwell Publishing, Oxford (2003)

[407] Sims-Williams, P. The Settlement of England in Bede and the 'Anglo-Saxon Chronicle'. *Anglo-Saxon England*, **12**: 1–41 (1983)

[408] Higham, N. Historical Narrative as Cultural Politics: Rome, 'British-ness' and 'English-ness'. In N. Higham, editor, *Britons in Anglo-Saxon England*. The Boydell Press, Woodbridge (2007)

[409] Hunter Blair, P. *An Introduction to Anglo Saxon England*. Cambridge University Press, Cambridge (1977)

[410] Gildas and Stevenson, J. *Gildas de excidio Britanniae, recens. J. Stevenson.* London, London (1838)

[411] Jones, M. and Casey, J. The Gallic Chronicle Restored: A Chronology for the Anglo Saxon Invasions and the End of Roman Britain. *Britannia*, **19**: 367–398 (1988)

[412] Wood, I. Before and after the migration to Britain. In J. Hines, editor, *The Anglo-Saxons from the Migration Period to the eighth century - An ethnographic perspective*. The Boydell Press, San Marino (1997)

[413] Pohl, W. Ethnic names and identities in the British Isles: A comparative perspective. In J. Hines, editor, *The Anglo-Saxons from the Migration Period to the eighth century - An ethnographic perspective.* The Boydell Press, San Marino (1997)

[414] Hills, C. *Origins of the English.* Duckworth, London (2003)

[415] Dickinson, T. Overview: Mortuary Ritual. In H. Hamerow, D. Hinton, and S. Crawford, editors, *Britain after Rome: the fall and rise, 400-700.* Oxford University Press, Oxford (2011)

[416] Hills, C. Overview: Anglo-Saxon Identity. In H. Hamerow, D. Hinton, and S. Crawford, editors, *Britain after Rome: the fall and rise, 400-700.* Oxford University Press, Oxford (2011)

[417] Hines, J. Society, community and identity. In T. Charles-Edwards, editor, *After Rome.* Oxford University Press, Oxford (2003)

[418] Welch, M. The Mid-Saxon 'Final Phase'. In H. Hamerow, D. Hinton, and S. Crawford, editors, *Britain after Rome: the fall and rise, 400-700.* Oxford University Press, Oxford (2011)

[419] Ulmschneider, K. Settlement Hierarchy. In H. Hamerow, D. Hinton, and S. Crawford, editors, *Britain after Rome: the fall and rise, 400-700.* Oxford University Press, Oxford (2011)

[420] Faull, M. British Survival in Anglo-Saxon Northumbria. In L. Laing, editor, *Studies in Celtic Survival.* BAR, Oxford (1977)

[421] Härke, H. "Warrior Grave"? The Background of the Anglo-Saxon Weapon Burial Rite. *Past & and Present*, **126**: 22–43 (1990)

[422] Owen-Crocker, G. Dress and Identity. In H. Hamerow, D. Hinton, and S. Crawford, editors, *Britain after Rome: the fall and rise, 400-700.* Oxford University Press, Oxford (2011)

[423] Loveluck, C. and Laing, L. Britons and Anglo-Saxons. In H. Hamerow, D. Hinton, and S. Crawford, editors, *Britain after Rome: the fall and rise, 400-700.* Oxford University Press, Oxford (2011)

[424] Laing, L. *The Archaeology of Celtic Britain and Ireland.* Cambridge University Press, Cambridge (2006)

[425] Higham, N. *The northern counties to AD 1000.* Longman, London (1986)

[426] Brugmann, B. Migration and Endogenous Change. In H. Hamerow, D. Hinton, and S. Crawford, editors, *Britain after Rome: the fall and rise, 400-700.* Oxford University Press, Oxford (2011)

[427] Härke, H. Invisible Britons, Gallo-Romans and Russians: Perspectives on culture change. In N. Higham, editor, *Britons in Anglo-Saxon England.* The Boydell Press, Woodbridge (2007)

[428] Yorke, B. *The Anglo-Saxons*. Stroud, Sutton (1999)

[429] Grimmer, M. Britons in Early Wessex: The evidence of the Law Code of Ine. In N. Higham, editor, *Britons in Anglo-Saxon England*. The Boydell Press, Woodbrige (2007)

[430] Chadwick, N. *Celtic Britain*. Thames and Hudson, London (1963)

[431] Mac Cana, P. Ireland and Wales in the Middle Ages: an overview. In K. Jankulak and J. Wooding, editors, *Ireland and Wales in the Middle Ages*. Four Courts Press, Dublin (2007)

[432] Campbell, E. Were the Scots Irish? *Antiquity*, **75** (288): 285–292 (2001)

[433] Alcock, L. Was there an Irish-Sea Culture-Province in The Dark Ages? In D. Moore, editor, *The Irish Sea Province on archaeology and history*. Cambrian Archaeological Association, Cardiff (1970)

[434] Ward-Perkins, B. Why Did The Anglo-Saxons not Become More British? *English Historical Review*, **115** (462): 513–533 (2000)

[435] Higham, N. Britons in Anglo-Saxon England: An Introduction . In N. Higham, editor, *Britons in Anglo-Saxon England*. The Boydell Press, Woodbridge (2007)

[436] Fellows-Jensen, G. The Vikings in England. A Review. *Anglo-Saxon England*, **4** (1975)

[437] Richards, J. Anglo-Scandinavian Identity. In H. Hamerow, D. Hinton, and S. Crawford, editors, *The Oxford Handbook of Anglo-Saxon Archaeology*. Oxford University Press, Oxford (2011)

[438] Griffiths, D. Settlement and acculturation in the Irish Sea Region. In J. Hines, A. Lane, and M. Redknap, editors, *Land, sea and home : proceedings of a conference on Viking-period settlement, at Cardiff, July 2001*. Maney, Leeds (2004)

[439] Higham, N. Viking-age settlement in the north-western countryside: lifting the veil? In J. Hines, A. Lane, and M. Redknap, editors, *Land, sea and home : proceedings of a conference on Viking-period settlement, at Cardiff, July 2001*. Maney, Leeds (2004)

[440] Fellows-Jensen, G. Scandinavian settlement in the British Isles and Normandy : what the place-names reveal. In J. Adams and K. Holman, editors, *Scandinavia and Europe 800-1350 - Contact, Conflict and Coexistence*. Brepols, Turnhout (2001)

[441] Abrams, L. and Parsons, D. Place-names and the history of Scandinavian settlement in England. In J. Hines, A. Lane, and M. Redknap, editors, *Land, sea and home : proceedings of a conference on Viking-period settlement, at Cardiff, July 2001*. Maney, Leeds (2004)

[442] Richards, J. Identifying Anglo-Scandinavian settlements. In D. Hadley and J. Richards, editors, *Cultures in contact : Scandinavian settlement in England in the ninth and tenth centuries*. Brepols, Turnhout (2000)

[443] Graham-Campbell, J. Pagan Scandinavian burial in the central and southern Danelaw. In J. Graham-Campbell, R. Hall, J. Jesch, and D. Parsons, editors, *Vikings and the Danelaw: Papers from the Proceedings of the Thirteenth Viking Congress, Nottingham and York, 21st-30th August 1997.* Oxford Books, Oxford (2001)

[444] Halsall, G. The Viking presence in England? In D. Hadley and J. Richards, editors, *Cultures in contact: Scandinavian settlement in England in the ninth and tenth centuries.* Brepols, Turnhout (2000)

[445] Barrett, J. Beyond war or peace: The study of culture contact in Viking-age Scotland. In J. Hines, A. Lane, and M. Redknap, editors, *Land, sea and home : proceedings of a conference on Viking-period settlement, at Cardiff, July 2001.* Maney, Leeds (2004)

[446] Owen, O. The Scar boat burial - and the missing decades of the early Viking age in Orkney and Shetland. In J. Adams and K. Holman, editors, *Scandinavia and Europe 800-1350 - Contact, Conflict and Coexistence.* Brepols, Turnhout (2001)

[447] Henry, P. Changing weaving styles and fabric types: the Scandinavian influence. In J. Hines, A. Lane, and M. Redknap, editors, *Land, sea and home : proceedings of a conference on Viking-period settlement, at Cardiff, July 2001.* Maney, Leeds (2004)

[448] Lynch, M. *Scotland – A New History.* Pimlico, London (1991)

[449] Abrams, L. Conversion and assimilation. In D. Hadley and J. Richards, editors, *Cultures in contact : Scandinavian settlement in England in the ninth and tenth centuries.* Brepols, Turnhout (2000)

[450] Golding, B. *Conquest and Colonisation - the Normans in Britain, 1066-1100.* Palgrave, New York (2001)

[451] Beeler, J. Castles and Strategy in Norman and Early Angevin England. *Speculum*, **31** (4): 581–601 (1956)

[452] Bony, J. French Influences on the Origins of English Gothic Architecture. *Journal of the Warburg and Courtauld Institutes*, **12**: 1–15 (1949)

[453] Keats-Rohan, K. The Bretons and Normans of England 1066-1154: the family, the fief and the feudal monarchy. *Nottingham Mediaeval Studies*, **36**: 42–78 (1992)

[454] http://www.englandsimmigrants.com

[455] Hinde, A. *England's population : a history since the Domesday survey.* Hodder Headline, London (2003)

[456] Schrijver, P. What Britons Spoke around 400 AD. In N. Higham, editor, *Britons in Anglo-Saxon England.* The Boydell Press, Woodbridge (2007)

[457] Horst Schmidt, K. Insular Celtic, P- and Q-Celtic. In M. Ball and J. Fife, editors, *The Celtic Languages.* Routledge, London (1993)

[458] George, K. Cornish. In M. Ball and J. Fife, editors, *The Celtic Languages.* Routledge, London (1993)

[459] Price, G. Cumbric. In G. Price, editor, *Encyclopedia of the Languages of Europe.* Blackwell, Blackwell Reference online (2000)

[460] Russell, P. Latin and British in Roman and post-Roman Britain: Methodology and Morphology. *Transactions of the Philological Society*, **109** (2): 138–1157 (2011)

[461] McManus, D. Linguarum diversitas: Latin and the vernaculars in early medieval Britain. *Peritia*, **3**: 151–188 (1984)

[462] Goodward, A. *A Dictionary of Welsh Etymology.* Self-published, Derby (2006)

[463] Gillies, W. Scottish Gaelic. In M. Ball and J. Fife, editors, *The Celtic Languages.* Routledge, London (1993)

[464] Henriksen, C. and van der Auwera, J. The Germanic Languages. In E. König and J. van der Auwera, editors, *The Germanic Languages.* Routledge, London (1994)

[465] Hines, J. The runic inscriptions of Early Anglo-Saxon England. In A. Bammesberger and A. Wollmann, editors, *Britain 400-600 : language and history.* Carl Winter, Heidelberg (1990)

[466] Greenfield, S. and Calder, D. *A New Critical History of Old English Literature.* New York University Press, New York (1986)

[467] van Kremenade, A. Old and Middle English. In E. König and J. van der Auwera, editors, *The Germanic Languages.* Routledge, London (1994)

[468] Hoekstra, J. and Tiersma, P. Frisian. In J. van der Auwera and E. König, editors, *The Germanic Languages.* Routledge, London (1994)

[469] Hoad, T. Preliminaries: Before English. In L. Mugglestone, editor, *The Oxford History of English.* Oxford University Press, Oxford (2006)

[470] Bremmer, R. The Nature of the Evidence for a Frisian Participation in the *Adventus Saxonum*. In A. Bammesberger and A. Wollman, editors, *Britain 400-600: Language and history.* Carl Winter, Heidelberg (1990)

[471] Benskin, M. Present Indicative Plural Concord in Brittonic and Early English. *Transactions of the Philological Society*, **109** (2): 158–185 (2011)

[472] Filppula, M. and Klemola, J. Special issue on Re-evaluating the Celtic hypothesis. *English Language and Linguistics*, **13**: 155–161 (2009)

[473] Faarlund, J. T. Old and Middle Scandinavian. In E. König and J. van der Auwera, editors, *The Germanic Languages.* Routledge, London (1994)

[474] Samuels, M. The Great Scandinavian Belt. In M. Laing, editor, *Middle English Dialectology: Essays on some Principles and Problems.* Aberdeen University Press, Aberdeen (1989)

[475] Townend, M. Contacts and Conflicts: Latin, Norse and French. In L. Mugglestone, editor, *The Oxford History of English*. Oxford University Press, Oxford (2006)

[476] Rothwell, W. Anglo-French Lexical Contacts, Old and New. *The Modern Language Review*, **74**: 287–296 (1979)

[477] Kitson, P. British and European river-names. *Transactions of the Philological Society*, **94**: 73–118 (1996)

[478] De Bernardo Stempel, P. More on Ptolemy's evidence for Celtic Ireland. In J. De Hoz, E. Luján, and P. Sims-Williams, editors, *New Approaches to Celtic Place-Names in Ptolemy's Geography*. Ediciones Clásicas, Madrid (2005)

[479] Sims-Williams, P. Measuring Celticity from Wales to the Orient. In J. De Hoz, E. Lujn, and P. Sims-Williams, editors, *New Approaches to Celtic Place-Names in Ptolemy's Geography*. Ediciones Clasicas, Madrid (2005)

[480] Isaac, G. Scotland. In J. De Hoz, E. Luján, and P. Sims-Williams, editors, *New Approaches to Celtic Place-Names in Ptolemy's Geography*. Ediciones Clásicas, Madrid (2005)

[481] Gelling, M. Place-Names and Archaeology. In H. Hamerow, D. Hinton, and S. Crawford, editors, *The Oxford Handbook of Anglo-Saxon Archaeology*. Oxford University Press, Oxford (2011)

[482] Lutz, A. Celtic influence on Old English and West Germanic. *English Language & Linguistics*, **13** (2): 227–249 (2009)

[483] Corrigan, L. Hunting the Vikings in South Cumbria from Ambleside to Haverbrack. In N. Higham and M. Ryan, editors, *Place-names, Language and the Anglo-Saxon Landscape*. The Boydell Press, Woodbridge (2011)

[484] Richards, M., Capelli, C., and Wilson, J. F. Genetics and the origins of the British population. In *Encyclopaedia of Life Sciences*. John Wiley & Sons, Chichester (2008)

[485] Soares, P., Achilli, A., Semino, O., Davies, W., Macaulay, V. et al. The archaeogenetics of Europe. *Curr Biol*, **20** (4): R174–83 (2010)

[486] Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A. et al. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet*, **75** (5): 910–8 (2004)

[487] Seldin, M., Shigeta, R., Villoslada, P., Selmi, C., Tuomilehto, J. et al. European Population Substructure: Clustering of Northern and Southern Populations. *PLoS One*, **2** (9) (2006)

[488] Bauchet, M., McEvoy, B., Pearson, L. N., Quillen, E. E., Sarkisian, T. et al. Measuring European population stratification with microarray genotype data. *Am J Hum Genet*, **80** (5): 948–56 (2007)

[489] Lao, O., Lu, T. T., Nothnagel, M., Junge, O., Freitag-Wolf, S. et al. Correlation between genetic and geographic structure in Europe. *Curr Biol*, **18** (16): 1241–8 (2008)

[490] O'Dushlaine, C., McQuillan, R., Weale, M. E., Crouch, D. J., Johansson, A. et al. Genes predict village of origin in rural Europe. *Eur J Hum Genet* (2010)

[491] Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U. et al. A draft sequence of the Neandertal genome. *Science*, **328** (5979): 710–22 (2010)

[492] Alvarez-Iglesias, V., Mosquera-Miguel, A., Cerezo, M., Quintans, B., Zarrabeitia, M. T. et al. New population and phylogenetic features of the internal variation within mitochondrial DNA macro-haplogroup R0. *PLoS ONE*, **4** (4): e5112 (2009)

[493] Behar, D. M., Harmant, C., Manry, J., van Oven, M., Haak, W. et al. The Basque Paradigm: Genetic Evidence of a Maternal Continuity in the Franco-Cantabrian Region since Pre-Neolithic Times. *Am J Hum Genet*, **90** (3): 486–493 (2012)

[494] Pala, M., Achilli, A., Olivieri, A., Kashani, B. H., Perego, U. A. et al. Mitochondrial haplogroup U5b3: a distant echo of the epipaleolithic in Italy and the legacy of the early Sardinians. *Am J Hum Genet*, **84** (6): 814–21 (2009)

[495] Rootsi, S., Magri, C., Kivisild, T., Benuzzi, G., Help, H. et al. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am J Hum Genet*, **75** (1): 128–37 (2004)

[496] Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S. et al. The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science*, **290** (5494): 1155–9 (2000)

[497] Torroni, A., Bandelt, H. J., Macaulay, V., Richards, M., Cruciani, F. et al. A signal, from human mtDNA, of postglacial recolonization in Europe. *Am J Hum Genet*, **69** (4): 844–52 (2001)

[498] Currat, M. and Excoffier, L. The effect of the Neolithic expansion on European molecular diversity. *Proc Biol Sci*, **272** (1564): 679–88 (2005)

[499] Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T. et al. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet*, **84** (6): 740–59 (2009)

[500] Busby, G. B., Brisighelli, F., Sanchez-Diz, P., Ramos-Luis, E., Martinez-Cadenas, C. et al. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. *Proc Biol Sci* (2011)

[501] Chikhi, L., Nichols, R., Barbujani, G., and Beaumont, M. A. Y genetic data support the Neolithic demic diffusion model. *PNAS*, **99** (17): 11008–13 (2002)

[502] Balaresque, P., Bowden, G. R., Adams, S. M., Leung, H. Y., King, T. E. et al. A predominantly neolithic origin for European paternal lineages. *PLoS Biol*, **8** (1): e1000285 (2010)

[503] Capelli, C., Brisighelli, F., Scarnicci, F., Arredi, B., Caglia, A. et al. Y chromosome genetic variation in the Italian peninsula is clinal and supports an admixture model for the Mesolithic-Neolithic encounter. *Mol Phylogenet Evol*, **44** (1): 228–39 (2007)

[504] Chikhi, L., Destro-Bisol, G., Bertorelle, G., Pascali, V., and Barbujani, G. Clines of nuclear DNA markers suggest a largely neolithic ancestry of the European gene pool. *Proc Natl Acad Sci U S A*, **95** (15): 9053–8 (1998)

[505] Malmström, H., Gilbert, M. T., Thomas, M. G., Brandström, M., Stora, J. et al. Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. *Curr Biol*, **19** (20): 1758–62 (2009)

[506] Melchior, L., Lynnerup, N., Siegismund, H. R., Kivisild, T., and Dissing, J. Genetic diversity among ancient Nordic populations. *PLoS ONE*, **5** (7) (2010)

[507] Bramanti, B., Thomas, M. G., Haak, W., Unterlaender, M., Jores, P. et al. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science*, **326** (5949): 137–40 (2009)

[508] Haak, W., Forster, P., Bramanti, B., Matsumura, S., Brandt, G. et al. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science*, **310** (5750): 1016–8 (2005)

[509] Haak, W., Balanovsky, O., Sanchez, J. J., Koshel, S., Zaporozhchenko, V. et al. Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities. *PLoS Biol*, **8** (11): e1000536 (2010)

[510] Sampietro, M. L., Caramelli, D., Lao, O., Calafell, F., Comas, D. et al. The genetics of the pre-Roman Iberian Peninsula: a mtDNA study of ancient Iberians. *Ann Hum Genet*, **69** (Pt 5): 535–48 (2005)

[511] Hervella, M., Izagirre, N., Alonso, S., Fregel, R., Alonso, A. et al. Ancient DNA from hunter-gatherer and farmer groups from northern Spain supports a random dispersion model for the neolithic expansion into Europe. *PLoS ONE*, **7** (4) (2012)

[512] Watkin, I. ABO blood groups and racial characteristics in rural Wales. *Heredity*, **10** (2): 161–193 (1956)

[513] Falsetti, A. B. and Sokal, R. R. Genetic structure of human populations in the British Isles. *Ann Hum Biol*, **20** (3): 215–29 (1993)

[514] McEvoy, B., Richards, M., Forster, P., and Bradley, D. G. The Longue Durée of genetic ancestry: multiple genetic marker systems and Celtic origins on the Atlantic facade of Europe. *Am J Hum Genet*, **75** (4): 693–702 (2004)

[515] Wilson, J. F., Weiss, D. A., Richards, M., Thomas, M. G., Bradman, N. et al. Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc Natl Acad Sci U S A*, **98** (9): 5078–83 (2001)

[516] Capelli, C., Redhead, N., Abernethy, J. K., Gratrix, F., Wilson, J. F. et al. A Y chromosome census of the British Isles. *Curr Biol*, **13** (11): 979–84 (2003)

[517] $_{C}ase-Control_{C}onsortium, T.Genome - wideassociationstudyof14,000casesofsevencommon$
$661 - -78(2007)$

[518] O'Dushlaine, C. T., Morris, D., Moskvina, V., Kirov, G., Consortium, I. S. et al. Population structure and genome-wide patterns of variation in Ireland and Britain. *Eur J Hum Genet* (2010)

[519] Weale, M. E., Weiss, D. A., Jager, R. F., Bradman, N., and Thomas, M. G. Y chromosome evidence for Anglo-Saxon mass migration. *Mol Biol Evol*, **19** (7): 1008–21 (2002)

[520] Thomas, M. G., Stumpf, M. P., and Härke, H. Evidence for an apartheid-like social structure in early Anglo-Saxon England. *Proc Biol Sci*, **273** (1601): 2651–7 (2006)

[521] Thomas, M. G., Stumpf, M. P., and Härke, H. Integration versus apartheid in post-Roman Britain: a response to Pattison. *Proc Biol Sci*, **275** (1650): 2419–21 (2008)

[522] Winney, B., Boumertit, A., Day, T., Davison, D., Echeta, C. et al. People of the British Isles: preliminary analysis of genotypes and surnames in a UK control population. *Eur J Hum Genet*, **20**: 203–210 (2012)

[523] Grignani, P., Peloso, G., Achilli, A., Turchi, C., Tagliabracci, A. et al. Subtyping mtDNA haplogroup H by SNaPshot minisequencing and its application in forensic individual identification. *Int J Legal Med*, **120** (3): 151–6 (2006)

[524] Grignani, P., Turchi, C., Achilli, A., Peloso, G., Alu, M. et al. Multiplex mtDNA coding region SNP assays for molecular dissection of haplogroups U/K and J/T. *Forensic Sci Int Genet*, **4** (1): 21–5 (2009)

[525] Köhnemann, S., Sibbing, U., Pfeiffer, H., and Hohoff, C. A rapid mtDNA assay of 22 SNPs in one multiplex reaction increases the power of forensic testing in European Caucasians. *Int J Legal Med*, **122** (6): 517–23 (2008)

[526] Quintans, B., Alvarez-Iglesias, V., Salas, A., Phillips, C., Lareu, M. V. et al. Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. *Forensic Sci Int*, **140** (2-3): 251–7 (2004)

[527] Herrnstadt, C., Elson, J. L., Fahy, E., Preston, G., Turnbull, D. M. et al. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet*, **70** (5): 1152–71 (2002)

[528] Dubut, V., Chollet, L., Murail, P., Cartault, F., Beraud-Colomb, E. et al. mtDNA polymorphisms in five French groups: importance of regional sampling. *Eur J Hum Genet*, **12** (4): 293–300 (2004)

[529] Achilli, A., Olivieri, A., Pala, M., Metspalu, E., Fornarino, S. et al. Mitochondrial DNA variation of modern Tuscans supports the near eastern origin of Etruscans. *Am J Hum Genet*, **80** (4): 759–68 (2007)

[530] Richard, C., Pennarun, E., Kivisild, T., Tambets, K., Tolk, H. V. et al. An mtDNA perspective of French genetic variation. *Ann Hum Biol*, **34** (1): 68–79 (2007)

[531] Benn, M., Schwartz, M., Nordestgaard, B. G., and Tybjaerg-Hansen, A. Mitochondrial haplogroups: ischemic cardiovascular disease, other diseases, mortality, and longevity in the general population. *Circulation*, **117** (19): 2492–501 (2008)

[532] Simoons, F. Primary adult lactose intolerance and the milking habit: a problem in biological and cultural interrelations. I. Review of the medical research. *Am J Dig Dis*, **14** (12): 1819–36 (1969)

[533] Enattah, N. S., Sahi, T., Savilahti, E., Terwilliger, J. D., Peltonen, L. et al. Identification of a variant associated with adult-type hypolactasia. *Nat Genet*, **30** (2): 233–7 (2002)

[534] Duffy, D. L., Box, N. F., Chen, W., Palmer, J. S., Montgomery, G. W. et al. Interactive effects of MC1R and OCA2 on melanoma risk phenotypes. *Hum Mol Genet*, **13** (4): 447–61 (2004)

[535] Branicki, W., Brudnik, U., Draus-Barini, J., Kupiec, T., and Wojas-Pelc, A. Association of the SLC45A2 gene with physiological human hair colour variation. *J Hum Genet*, **53** (11-12): 966–71 (2008)

[536] Casto, A. and Feldman, M. Genome-wide association study SNPs in the human genome diversity project populations: does selection affect unlinked SNPs with shared trait associations? *PLoS Genet*, **7** (1) (2011)

[537] Hancock, A., Witonsky, D., Alkorta-Aranburu, G., Beall, C., Gebremedhin, A. et al. Adaptations to Climate-Mediated Selective Pressure in Humans. *PLoS Genet*, **7** (4) (2011)

[538] Quillen, E. and Shriver, M. Unpacking human evolution to find the genetic determinants of human skin pigmentation. *J Invest Dermatol*, **131** (E1): E5–7 (2011)

[539] Quillen, E. E., Bauchet, M., Bigham, A. W., Delgado-Burbano, M. E., Faust, F. X. et al. OPRM1 and EGFR contribute to skin pigmentation differences between Indigenous Americans and Europeans. *Hum Genet*, **131** (7): 1073–80 (2012)

[540] Candille, S. I., Absher, D. M., Beleza, S., Bauchet, M., McEvoy, B. et al. Genome-wide association studies of quantitatively measured skin, hair, and eye pigmentation in four European populations. *PLoS ONEt*, **7** (10) (2012)

[541] Sturm, R. and Duffy, D. Human pigmentation genes under environmental selection. *Genome Biol*, **13** (9) (2012)

[542] Paik, S. H., Kim, H. J., Son, H. Y., Lee, S., Im, S. W. et al. Gene mapping study for constitutive skin color in an isolated Mongolian population. *Exp Mol Med*, **44** (3): 241–9 (2012)

[543] Beleza, S., Santos, A. M., McEvoy, B., Alves, I., Martinho, C. et al. The timing of pigmentation lightening in Europeans. *Mol Biol Evol*, **30** (1): 24–35 (2013)

[544] Reissman, M. and Ludwig, A. Pleiotropic effects of coat colour-associated mutations in humans, mice and other mammals. *Semin Cell Dev Biol* (2013)

[545] Winter, R. M. What's in a face? *Nat Genet*, **12** (2): 124–9 (1996)

[546] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria (2011). ISBN 3-900051-07-0

[547] Weir, B. and Cockerham, C. Estimating F-Statistic for the Analysis of Population Structure. *Evolution*, **38** (6): 1358–1370 (1984)

[548] Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C. et al. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet*, **39** (1): 31–40 (2007)

[549] Passarino, G., Cavalleri, G. L., Lin, A. A., Cavalli-Sforza, L. L., Borresen-Dale, A. L. et al. Different genetic components in the Norwegian population revealed by the analysis of mtDNA and Y chromosome polymorphisms. *Eur J Hum Genet*, **10** (9): 521–9 (2002)

[550] Parsons, F. The Colour Index of the British Isles. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, **50**: 159–182 (1920)

[551] Goodacre, S., Helgason, A., Nicholson, J., Southam, L., Ferguson, L. et al. Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity*, **95** (2): 129–35 (2005)

[552] www.uniprot.org. Uniprot website

[553] Gloria-Bottini, F., Magrini, A., Cozzoli, E., Neri, A., Pietroiusti, A. et al. ABH secretor genetic polymorphism: evidence of intrauterine selection. *Eur J Obstet Gynecol Reprod Biol*, **154** (1): 20–3 (2011)

[554] Fumagalli, M., Cagliani, R., Pozzoli, U., Riva, S., Comi, G. P. et al. Widespread balancing selection and pathogen-driven selection at blood group antigen genes. *Genome Res*, **19** (2): 199–212 (2009)

[555] Ferrer-Admetlla, A., Sikora, M., Laayouni, H., Esteve, A., Roubinet, F. et al. A natural history of FUT2 polymorphism in humans. *Mol Biol Evol*, **26** (9): 1993–2003 (2009)

[556] McGovern, D. P., Jones, M. R., Taylor, K. D., Marciante, K., Yan, X. et al. Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. *Hum Mol Genet*, **19** (17): 3468–76 (2010)

[557] Tanwar, V. S., Chand, M. P., Kumar, J., Garg, G., Seth, S. et al. Common variant in FUT2 gene is associated with levels of vitamin B(12) in Indian population. *Gene*, **515** (1): 224–8 (2012)

[558] Seixas, S., Suriano, G., Carvalho, F., Seruca, R., Rocha, J. et al. Sequence diversity at the proximal 14q32.1 SERPIN subcluster: evidence for natural selection favoring the pseudogenization of SERPINA2. *Mol Biol Evol*, **24** (2): 587–98 (2007)

[559] Lucien, N., Sidoux-Walter, F., Roudier, N., Ripoche, P., Huet, M. et al. Antigenic and functional properties of the human red blood cell urea transporter hUT-B1. *J Biol Chem*, **277** (37): 34101–8 (2002)

[560] Valdes, A. M., Thomson, G., and Barcellos, L. F. Genetic variation within the HLA class III influences T1D susceptibility conferred by high-risk HLA haplotypes. *Genes Immun*, **11** (3): 209–18 (2010)

[561] Ingram, C. and Swallow, D. Population Genetics of Lactase Persistence and Lactose Intolerance. In *Encyclopedia of Life Sciences*. John Wiley Sons, Ltd., Chichester (2007)

[562] Sherratt, A. Plough and pastoralism: aspects of the secondary products revolution. In I. Hodder, G. Isaac, and N. Hammond, editors, *Pattern of the Past: Studies in Honour of David Clark*. Cambridge University Press, Cambridge (1981)

[563] Craig, O. The development of dairying in Europe: potential evidence from food residues on ceramics. *Documenta Praehistorica*, **29**: 97–107 (2001)

[564] Copley, M., Berstan, R., Dudd, S., Aillaud, S., Mukherjee, A. et al. Processing of milk products in pottery vessels through British prehistory. *Antiquity*, **79**: 895–908 (2005)

[565] Jablonski, N. Human skin pigmentation, migration, and disease susceptibility (2011)

[566] Djian-Zaouche, J., Campagne, C., Reyes-Gomez, E., Gadin-Czerw, S., Bernex, F. et al. Pax3(GFP), a new reporter for the melanocyte lineage, highlights novel aspects of PAX3 expression in the skin. *Pigment Cell Melanoma Res*, **25** (5): 545–554 (2012)

[567] Cannon, D. M., Walshe, M., Dempster, E., Collier, D. A., Marshall, N. et al. The association of white matter volume in psychotic disorders with genotypic variation in NRG1, MOG and CNP: a voxel-based analysis in affected individuals and their unaffected relatives. *Transl Psychiatry*, **2** (2012)

[568] Luzon-Toro, B., Torroglosa, A., Nunez-Torres, R., Enguix-Riego, M. V., Fernandez, R. M. et al. Comprehensive analysis of NRG1 common and rare variants in Hirschsprung patients. *PLoS One*, **7** (5) (2012)

[569] Choi, W., Kolbe, L., and Hearing, V. J. Characterization of the bioactive motif of neuregulin-1, a fibroblast-derived paracrine factor that regulates the constitutive color and the function of melanocytes in human skin. *Pigment Cell Melanoma Res*, **25** (4): 477–81 (2012)

[570] Jacobs, L. C., Wollstein, A., Lao, O., Hofman, A., Klaver, C. C. et al. Comprehensive candidate gene study highlights UGT1A and BNC2 as new genes determining continuous skin color variation in Europeans. *Hum Genet*, **132** (2): 147–58 (2013)

[571] Lang, M. R., Patterson, L. B., Gordon, T. N., Johnson, S. L., and Parichy, D. M. Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility. *PLoS Genet*, **5** (11) (2009)

[572] Calboli, F. C., Tozzi, F., Galwey, N. W., Antoniades, A., Mooser, V. et al. A genome-wide association study of neuroticism in a population-based sample. *PLoS One*, **5** (7) (2010)

[573] Gillespie, R. *Colonial Ulster*. Cork University Press, Cork (1985)

[574] Gillespie, R. *Seventeenth-century Ireland: Making Ireland Modern.* Gill & Macmillan, Dublin (2006)

[575] Connolly, S. *Contested Island – Ireland 1460-1630.* Oxford University Press, Oxford (2007)

[576] Nakayama, K., Soemantri, A., Jin, F., Dashnyam, B., Ohtsuka, R. et al. Identification of novel functional variants of the melanocortin 1 receptor gene originated from Asians. *Hum Genet*, **119** (3): 322–30 (2006)

[577] Yuasa, I., Umetsu, K., Harihara, S., Kido, A., Miyoshi, A. et al. Distribution of two Asian-related coding SNPs in the MC1R and OCA2 genes. *Biochem Genet*, **45** (7-8): 535–42 (2007)

[578] Teo, Y. Y., Sim, X., Ong, R. T., Tan, A. K., Chen, J. et al. Singapore Genome Variation Project: A haplotype map of three South-East Asian populations. *Genome Res* (2009)

[579] Masui, S., Nakatome, M., and Matoba, R. Variants of the melanocortin 1 receptor gene (MC1R) and P gene as indicators of the population origin of an individual. *Int J Legal Med*, **123** (3): 205–11 (2009)

[580] Peng, S., Lu, X. M., Luo, H. R., Xiang-Yu, J. G., and Zhang, Y. P. Melanocortin-1 receptor gene variants in four Chinese ethnic populations. *Cell Res*, **11** (1): 81–4 (2001)

[581] Mengel-From, J., Borsting, C., Sanchez, J., Eiberg, H., and Morling, N. Determination of cis/trans phase of variations in the MC1R gene with allelespecific PCR and single base extension. *Electrophoresis*, **29**: 4780–4787 (2009)

[582] Matichard, E., Verpillat, P., Meziani, R., Gerard, B., Descamps, V. et al. Melanocortin 1 receptor (MC1R) gene variants may increase the risk of melanoma in France independently of clinical risk factors and UV exposure. *J Med Genet*, **41** (2): e13 (2004)

[583] Chaudru, V., Laud, K., Avril, M. F., Miniere, A., Chompret, A. et al. Melanocortin-1 receptor (MC1R) gene variants and dysplastic nevi modify penetrance of CDKN2A mutations in French melanoma-prone pedigrees. *Cancer Epidemiol Biomarkers Prev*, **14** (10): 2384–90 (2005)

[584] Scherer, D., Nagore, E., Bermejo, J. L., Figl, A., Botella-Estrada, R. et al. Melanocortin receptor 1 variants and melanoma risk: a study of 2 European populations. *Int J Cancer*, **125** (8): 1868–75 (2009)

[585] Gerstenblith, M. R., Goldstein, A. M., Fargnoli, M. C., Peris, K., and Landi, M. T. Comprehensive evaluation of allele frequency differences of MC1R variants across populations. *Hum Mutat*, **28** (5): 495–505 (2007)

[586] Stratigos, A. J., Dimisianos, G., Nikolaou, V., Poulou, M., Sypsa, V. et al. Melanocortin receptor-1 gene polymorphisms and the risk of cutaneous melanoma in a low-risk southern European population. *J Invest Dermatol*, **126** (8): 1842–9 (2006)

[587] Pastorino, L., Cusano, R., Bruno, W., Lantieri, F., Origone, P. et al. Novel MC1R variants in Ligurian melanoma patients and controls. *Hum Mutat*, **24** (1): 103 (2004)

[588] Fargnoli, M. C., Chimenti, S., Keller, G., Hofler, H., and Peris, K. Identification of four novel melanocortin 1 receptor (MC1R) gene variants in a Mediterranean population. *Hum Mutat*, **21** (6): 655 (2003)

[589] Landi, M. T., Kanetsky, P. A., Tsang, S., Gold, B., Munroe, D. et al. MC1R, ASIP, and DNA repair in sporadic and familial melanoma in a Mediterranean population. *J Natl Cancer Inst*, **97** (13): 998–1007 (2005)

[590] Motokawa, T., Kato, T., Hongo, M., Ito, M., Takimoto, H. et al. Characteristic MC1R polymorphism in the Japanese population. *J Dermatol Sci*, **41** (2): 143–5 (2006)

[591] Na, G. Y., Lee, K. H., Kim, M. K., Lee, S. J., Kim, D. W. et al. Polymorphisms in the melanocortin-1 receptor (MC1R) and agouti signaling protein (ASIP) genes in Korean vitiligo patients. *Pigment Cell Res*, **16** (4): 383–7 (2003)

[592] Fernandez, L., Milne, R., Bravo, J., Lopez, J., Aviles, J. et al. MC1R: three novel variants identified in a malignant melanoma association study in the Spanish population. *Carcinogenesis*, **28** (8): 1659–64 (2007)

[593] Höiom, V., Tuominen, R., Kaller, M., Linden, D., Ahmadian, A. et al. MC1R variation and melanoma risk in the Swedish population in relation to clinical and pathological parameters. *Pigment Cell Melanoma Res*, **22** (2): 196–204 (2009)

[594] Galeeva, A. R., Khusnutdinova, E. K., Slominskii, P. A., and Limborskaia, S. A. [Distribution of the 32 bp deletion in the CCR5 chemokine receptor gene in populations of the Volga-Ural region]. *Genetika*, **34** (8): 1160–2 (1998)

[595] Lucotte, G. and Mercier, G. Distribution of the CCR5 gene 32-bp deletion in Europe. *J Acquir Immune Defic Syndr Hum Retrovirol*, **19** (2): 174–7 (1998)

[596] Libert, F., Cochaux, P., Beckman, G., Samson, M., Aksenova, M. et al. The deltaCCR5 mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. *Hum Mol Genet*, **7** (3): 399–406 (1998)

[597] Martinson, J. J., Chapman, N. H., Rees, D. C., Liu, Y. T., and Clegg, J. B. Global distribution of the CCR5 gene 32-basepair deletion. *Nat Genet*, **16** (1): 100–3 (1997)

[598] Struyf, F., Thoelen, I., Charlier, N., Keyaerts, E., Van der Donck, I. et al. Prevalence of CCR5 and CCR2 HIV-coreceptor gene polymorphisms in Belgium. *Hum Hered*, **50** (5): 304–7 (2000)

[599] Slominskii, P. A., Shadrina, M. I., Spitsyn, V. A., Mikulich, V. A., Khusnutdinova, E. K. et al. [A simple and rapid method for determining a 32-bp deletion in the gene for the chemokine receptor CCR5]. *Genetika*, **33** (11): 1596–8 (1997)

[600] Xiao, J., Hu, F., Xu, H., Su, B., Jiang, Y. et al. Provincial distribution of three HIV-1 resistant polymorphisms (CCR5-Delta32, CCR2-64I, and SDF1-3' A) in China. *Sci China C Life Sci*, **43** (1): 16–20 (2000)

[601] Zhang, C., Fu, S., Xue, Y., Wang, Q., Huang, X. et al. Distribution of the CCR5 gene 32-basepair deletion in 11 Chinese populations. *Anthropol Anz*, **60** (3): 267–71 (2002)

[602] Ristic, S., Starcevic Cizmarevic, N., Brajenovic-Milic, B., Crnic-Martinovic, M., and Kapovic, M. Frequency of CCR5 gene 32-basepair deletion in Croatian normal population. *Croat Med J*, **46** (4): 693–4 (2005)

[603] Christodoulou, C., Poullikas, M., Neumann, A. U., and Kostrikis, L. G. Low frequency of CCR5delta32 allele among Greeks in Cyprus. *AIDS Res Hum Retroviruses*, **13** (16): 1373–4 (1997)

[604] Magierowska, M., Lepage, V., Boubnova, L., Carcassi, C., de Juan, D. et al. Distribution of the CCR5 gene 32 base pair deletion and SDF1-3'A variant in healthy individuals from different populations. *Immunogenetics*, **48** (6): 417–9 (1998)

[605] Shahbazi, M., Ebadi, H., Fathi, D., Roshandel, D., Mahamadhoseeni, M. et al. CCR5-delta 32 allele is associated with the risk of developing multiple sclerosis in the Iranian population. *Cell Mol Neurobiol*, **29** (8): 1205–9 (2009)

[606] Voevodin, A., Samilchuk, E., and Dashti, S. A Survey for 32 Nucleotide Deletion in the CCR-5 Chemokine Receptor Gene Conferring Resistance to Human Immunodeficiency Virus Type 1 in Different Ethnic Groups and in Chimpanzees. *J Med Virol*, **55** (1998)

[607] Zamarchi, R., Indraccolo, S., Minuzzo, S., Coppola, V., Gringeri, A. et al. Frequency of a mutated CCR-5 allele (delta32) among Italian healthy donors and individuals at risk of parenteral HIV infection. *AIDS Res Hum Retroviruses*, **15** (4): 337–44 (1999)

[608] Battiloro, E., Andreoni, M., Parisi, S. G., Mura, M. S., Sotgiu, G. et al. Distribution of the CCR5 delta32 allele in Italian HIV type 1-infected and normal individuals. *AIDS Res Hum Retroviruses*, **16** (2): 181–2 (2000)

[609] Karam, W., Jurjus, R., Khoury, N., Khansa, H., Assad, C. et al. Frequency of the CCR5-delta 32 chemokine receptor gene mutation in the Lebanese population. *East Mediterr Health J*, **10** (4-5): 671–5 (2004)

[610] Sidoti, A., D'Angelo, R., Rinaldi, C., De Luca, G., Pino, F. et al. Distribution of the mutated delta 32 allele of the CCR5 gene in a Sicilian population. *Int J Immunogenet*, **32** (3): 193–8 (2005)

[611] Alvarez, V., Lopez-Larrea, C., and Coto, E. Mutational analysis of the CCR5 and CXCR4 genes (HIV-1 co-receptors) in resistance to HIV-1 infection and AIDS development among intravenous drug users. *Hum Genet*, **102** (4): 483–6 (1998)

[612] Ruiz, A., Royo, J. L., Rubio, A., Borrego, S., Leal, M. et al. Spectrofluorimetric analysis of CCR5-delta 32 allele using real-time polymerase chain reaction: prevalence in southern Spanish HIV(+) patients and noninfected population. *AIDS Res Hum Retroviruses*, **17** (2): 191–3 (2001)

[613] Aseev, M. V., Shawi, A., Dean, M., and Baranov, V. S. [Population features of gene mutation frequency of the CKR-5 receptor, determining sensitivity to the AIDS virus]. *Genetika*, **33** (12): 1724–6 (1997)

[614] Harakalova, M., van Harssel, J. J., Terhal, P. A., van Lieshout, S., Duran, K. et al. Dominant missense mutations in ABCC9 cause Cantu syndrome. *Nat Genet*, **44** (7) (2012)

[615] Hutchinson, D. O., Charlton, A., Laing, N. G., Ilkovski, B., and North, K. N. Autosomal dominant nemaline myopathy with intranuclear rods due to mutation of the skeletal muscle ACTA1 gene: clinical and pathological variability within a kindred. *Neuromuscul Disord*, **16** (2): 113–21 (2006)

[616] van Bon, B. W., Koolen, D. A., Borgatti, R., Magee, A., Garcia-Minaur, S. et al. Clinical and molecular characteristics of 1qter microdeletion syndrome: delineating a critical region for corpus callosum agenesis/hypogenesis. *J Med Genet*, **45** (6): 346–54 (2008)

[617] Simovich, M. J., Bland, S. D., Peiffer, D. A., Gunderson, K. L., Cheung, S. W. et al. Delineation of the proximal 3q microdeletion syndrome. *Am J Med Genet A*, **146A** (13): 1729–35 (2008)

[618] Twigg, S. R., Versnel, S. L., Nurnberg, G., Lees, M. M., Bhat, M. et al. Frontorhiny, a distinctive presentation of frontonasal dysplasia caused by recessive mutations in the ALX3 homeobox gene. *Am J Hum Genet*, **84** (5): 698–705 (2009)

[619] Wuyts, W., Waeber, G., Meinecke, P., Schuler, H., Goecke, T. O. et al. Proximal 11p deletion syndrome (P11pDS): additional evaluation of the clinical and molecular aspects. *Eur J Hum Genet*, **12** (5): 400–6 (2004)

[620] Hoischen, A., van Bon, B. W., Rodriguez-Santiago, B., Gilissen, C., Vissers, L. E. et al. De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. *Nat Genet*, **43** (8): 729–31 (2011)

[621] Abidi, F. E., Cardoso, C., Lossi, A. M., Lowry, R. B., Depetris, D. et al. Mutation in the 5' alternatively spliced region of the XNP/ATR-X gene causes Chudley-Lowry syndrome. *Eur J Hum Genet*, **13** (2): 176–83 (2005)

[622] Faiyaz-Ul-Haque, M., Zaidi, S. H., Al-Ali, M., Al-Mureikhi, M. S., Kennedy, S. et al. A novel missense mutation in the galactosyltransferase-I (B4GALT7) gene in a family exhibiting facioskeletal anomalies and Ehlers-Danlos syndrome resembling the progeroid type. *Am J Med Genet A*, **128A** (1): 39–45 (2004)

[623] Fan, Z., Yamaza, T., Lee, J. S., Yu, J., Wang, S. et al. BCOR regulates mesenchymal stem cell function by epigenetic mechanisms. *Nat Cell Biol*, **11** (8): 1002–9 (2009)

[624] Allanson, J. E., Anneren, G., Aoki, Y., Armour, C. M., Bondeson, M. L. et al. Cardio-facio-cutaneous syndrome: does genotype predict phenotype? *Am J Med Genet C Semin Med Genet*, **157** (2): 129–35 (2011)

[625] Koolen, D. A., Sharp, A. J., Hurst, J. A., Firth, H. V., Knight, S. J. et al. Clinical and molecular delineation of the 17q21.31 microdeletion syndrome. *J Med Genet* (2008)

[626] Gillis, J., Burashnikov, E., Antzelevitch, C., Blaser, S., Gross, G. et al. Long QT, syndactyly, joint contractures, stroke and novel CACNA1C mutation: Expanding the spectrum of Timothy syndrome. *Am J Med Genet A* (2011)

[627] Moog, U., Kutsche, K., Kortum, F., Chilian, B., Bierhals, T. et al. Phenotypic spectrum associated with CASK loss-of-function mutations. *J Med Genet*, **48** (11): 741–51 (2011)

[628] Bonaldi, A., Mazzeu, J. F., Costa, S. S., Honjo, R. S., Bertola, D. R. et al. Microduplication of the ICR2 domain at chromosome 11p15 and familial Silver-Russell syndrome. *Am J Med Genet A*, **155A** (101): 2479–83 (2011)

[629] Weksberg, R., Shuman, C., and Smith, A. C. Beckwith-Wiedemann syndrome. *Am J Med Genet C Semin Med Genet*, **137C** (1): 12–23 (2005)

[630] Vissers, L. E., van Ravenswaaij, C. M., Admiraal, R., Hurst, J. A., de Vries, B. B. et al. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet*, **36** (9): 955–7 (2004)

[631] Fernandez, T., Morgan, T., Davis, N., Klin, A., Morris, A. et al. Disruption of contactin 4 (CNTN4) results in developmental delay and other features of 3p deletion syndrome. *Am J Hum Genet*, **74** (6): 1286–93 (2004)

[632] Van Camp, G., Snoeckx, R. L., Hilgert, N., van den Ende, J., Fukuoka, H. et al. A new autosomal recessive form of Stickler syndrome is caused by a mutation in the COL9A1 gene. *Am J Hum Genet*, **79** (3): 449–57 (2006)

[633] Pace, J. M., Chitayat, D., Atkinson, M., Wilcox, W. R., Schwarze, U. et al. A single amino acid substitution (D1441Y) in the carboxyl-terminal propeptide of the proalpha1(I) chain of type I collagen results in a lethal variant of osteogenesis imperfecta with features of dense bone diseases. *J Med Genet*, **39** (1): 23–9 (2002)

[634] Varela, M. C., Krepischi-Santos, A. C., Paz, J. A., Knijnenburg, J., Szuhai, K. et al. A 17q21.31 microdeletion encompassing the MAPT gene in a mentally impaired patient. *Cytogenet Genome Res*, **114** (1): 89–92 (2006)

[635] Cardoso, C., Leventer, R. J., Ward, H. L., Toyo-Oka, K., Chung, J. et al. Refinement of a 400-kb critical region allows genotypic differentiation between isolated lissencephaly, Miller-Dieker syndrome, and other phenotypes secondary to deletions of 17p13.3. *Am J Hum Genet*, **72** (4): 918–30 (2003)

[636] Knappskog, P. M., Majewski, J., Livneh, A., Nilsen, P. T., Bringsli, J. S. et al. Cold-induced sweating syndrome is caused by mutations in the CRLF1 gene. *Am J Hum Genet*, **72** (2): 375–83 (2003)

[637] Irving, M., Hanson, H., Turnpenny, P., Brewer, C., Ogilvie, C. M. et al. Deletion of the distal long arm of chromosome 10; is there a characteristic phenotype? A report of 15 de novo and familial cases. *Am J Med Genet A*, **123A** (2): 153–63 (2003)

[638] Ho, N., Punturieri, A., Wilkin, D., Szabo, J., Johnson, M. et al. Mutations of CTSK result in pycnodysostosis via a reduction in cathepsin K protein. *J Bone Miner Res*, **14** (10): 1649–53 (1999)

[639] Cella, W., de Vasconcellos, J. P., de Melo, M. B., Kneipp, B., Costa, F. F. et al. Structural assessment of PITX2, FOXC1, CYP1B1, and GJA1 genes in patients with Axenfeld-Rieger syndrome with developmental glaucoma. *Invest Ophthalmol Vis Sci*, **47** (5): 1803–9 (2006)

[640] Witsch-Baumgartner, M., Schwentner, I., Gruber, M., Benlian, P., Bertranpetit, J. et al. Age and origin of major Smith-Lemli-Opitz syndrome (SLOS) mutations in European populations. *J Med Genet*, **45** (4): 200–9 (2008)

[641] Ng, S. B., Buckingham, K. J., Lee, C., Bigham, A. W., Tabor, H. K. et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet*, **42** (1) (2010)

[642] Bartsch, O., Rasi, S., Delicado, A., Dyack, S., Neumann, L. M. et al. Evidence for a new contiguous gene syndrome, the chromosome 16p13.3 deletion syndrome alias severe Rubinstein-Taybi syndrome. *Hum Genet*, **120** (2): 179–86 (2006)

[643] Monreal, A. W., Zonana, J., and Ferguson, B. Identification of a new splice form of the EDA1 gene permits detection of nearly all X-linked hypohidrotic ectodermal dysplasia mutations. *Am J Hum Genet*, **63** (2): 380–9 (1998)

[644] Bondurand, N., Dastot-Le Moal, F., Stanchina, L., Collot, N., Baral, V. et al. Deletions at the SOX10 gene locus cause Waardenburg syndrome types 2 and 4. *Am J Hum Genet*, **81** (6): 1169–85 (2007)

[645] Wieland, I., Jakubiczka, S., Muschke, P., Cohen, M., Thiele, H. et al. Mutations of the ephrin-B1 gene cause craniofrontonasal syndrome. *Am J Hum Genet*, **74** (6): 1209–15 (2004)

[646] Kleefstra, T., Smidt, M., Banning, M. J., Oudakker, A. R., Van Esch, H. et al. Disruption of the gene Euchromatin Histone Methyl Transferase1 (Eu-HMTase1) is associated with the 9q34 subtelomeric deletion syndrome. *J Med Genet*, **42** (4): 299–306 (2005)

[647] Nowaczyk, M. J., Carter, M. T., Xu, J., Huggins, M., Raca, G. et al. Paternal deletion 6q24.3: a new congenital anomaly syndrome associated with intrauterine growth failure, early developmental delay and characteristic facial appearance. *Am J Med Genet A*, **146** (3): 354–60 (2008)

[648] Schule, B., Oviedo, A., Johnston, K., Pai, S., and Francke, U. Inactivating mutations in ESCO2 cause SC phocomelia and Roberts syndrome: no phenotype-genotype correlation. *Am J Hum Genet*, **77** (6): 1117–28 (2005)

[649] Tompson, S. W., Ruiz-Perez, V. L., Blair, H. J., Barton, S., Navarro, V. et al. Sequencing EVC and EVC2 identifies mutations in two-thirds of Ellis van Creveld syndrome patients. *Hum Genet*, **120** (5): 663–70 (2007)

[650] Ludecke, H. J., Schaper, J., Meinecke, P., Momeni, P., Gross, S. et al. Genotypic and phenotypic spectrum in tricho-rhino-phalangeal syndrome types I and III. *Am J Hum Genet*, **68** (1): 81–91 (2001)

[651] Tatton-Brown, K., Hanks, S., Ruark, E., Zachariou, A., Duarte Sdel, V. et al. Germline mutations in the oncogene EZH2 cause Weaver syndrome and increased human height. *Oncotarget*, **2** (12): 1127–33 (2011)

[652] Abzhanov, A., Cordero, D. R., Sen, J., Tabin, C. J., and Helms, J. A. Cross-regulatory interactions between Fgf8 and Shh in the avian frontonasal prominence. *Congenit Anom (Kyoto)*, **47** (4): 136–48 (2007)

[653] Abzhanov, A., Protas, M., Grant, B. R., Grant, P. R., and Tabin, C. J. Bmp4 and morphological variation of beaks in Darwin's finches. *Science*, **305** (5689): 1462–5 (2004)

[654] Zhang, D., Herring, J. A., Swaney, S. S., McClendon, T. B., Gao, X. et al. Mutations responsible for Larsen syndrome cluster in the FLNB protein. *J Med Genet*, **43** (5): e24 (2006)

[655] Papa, F. T., Mencarelli, M. A., Caselli, R., Katzaki, E., Sampieri, K. et al. A 3 Mb deletion in 14q12 causes severe mental retardation, mild facial dysmorphisms and Rett-like features. *Am J Med Genet A*, **146A** (15): 1994–8 (2008)

[656] Crisponi, L., Deiana, M., Loi, A., Chiappe, F., Uda, M. et al. The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. *Nat Genet*, **27** (2): 159–66 (2001)

[657] Slavotinek, A. M., Baranzini, S. E., Schanze, D., Labelle-Dumais, C., Short, K. M. et al. Manitoba-oculo-tricho-anal (MOTA) syndrome is caused by mutations in FREM1. *J Med Genet* (2011)

[658] Tzahor, E., Kempf, H., Mootoosamy, R. C., Poon, A. C., Abzhanov, A. et al. Antagonists of Wnt and BMP signaling promote the formation of vertebrate head muscle. *Genes Dev*, **17** (24): 3087–99 (2003)

[659] Archer, H. L., Gupta, S., Enoch, S., Thompson, P., Rowbottom, A. et al. Distinct phenotype associated with a cryptic subtelomeric deletion of 19p13.3-pter. *Am J Med Genet A*, **136** (1): 38–44 (2005)

[660] Cacheux, V., Dastot-Le Moal, F., Kaariainen, H., Bondurand, N., Rintala, R. et al. Loss-of-function mutations in *SIP1* Smad interacting protein result in a syndromic Hirschsprung disease. *Hum Mol Genet*, **10** (14): 1503–10 (2001)

[661] Pilia, G., Hughes-Benzie, R. M., MacKenzie, A., Baybayan, P., Chen, E. Y. et al. Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. *Nat Genet*, **12** (3): 241–7 (1996)

[662] Joyce, C. A., Sharp, A., Walker, J. M., Bullman, H., and Temple, I. K. Duplication of 7p12.1-p13, including GRB10 and IGFBP1, in a mother and daughter with features of Silver-Russell syndrome. *Hum Genet*, **105** (3): 273–80 (1999)

[663] Estep, A. L., Tidyman, W. E., Teitell, M. A., Cotter, P. D., and Rauen, K. A. HRAS mutations in Costello syndrome: detection of constitutional activating mutations in codon 12 and 13 and loss of wild-type allele in malignancy. *Am J Med Genet A*, **140** (1): 8–16 (2006)

[664] Zucchero, T. M., Cooper, M. E., Maher, B. S., Daack-Hirsch, S., Nepomuceno, B. et al. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *N Engl J Med*, **351** (8): 769–80 (2004)

[665] Samejima, H., Torii, C., Kosaki, R., Kurosawa, K., Yoshihashi, H. et al. Screening for Alagille syndrome mutations in the JAG1 and NOTCH2 genes using denaturing high-performance liquid chromatography. *Genet Test*, **11** (3): 216–27 (2007)

[666] Clayton-Smith, J., O'Sullivan, J., Daly, S., Bhaskar, S., Day, R. et al. Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am J Hum Genet*, **89** (5): 675–81 (2011)

[667] Caselli, R., Mencarelli, M. A., Papa, F. T., Ariani, F., Longo, I. et al. Delineation of the phenotype associated with 7q36.1q36.2 deletion: long QT syndrome, renal hypoplasia and mental retardation. *Am J Med Genet A*, **146A** (9): 1195–9 (2008)

[668] Modoni, A., Bianchi, M. L., Vitulano, N., Pagliarani, S., Perna, F. et al. Lack of any cardiac involvement in a patient with Andersen-Tawil syndrome associated with the c.574AtoG mutation in KCNJ2. *Cardiology*, **120** (4): 200–3 (2011)

[669] Kratz, C. P., Schubbert, S., Bollag, G., Niemeyer, C. M., Shannon, K. M. et al. Germline mutations in components of the Ras signaling pathway in Noonan syndrome and related disorders. *Cell Cycle*, **5** (15): 1607–11 (2006)

[670] Schlickum, S., Moghekar, A., Simpson, J. C., Steglich, C., O'Brien, R. J. et al. LETM1, a gene deleted in Wolf-Hirschhorn syndrome, encodes an evolutionarily conserved mitochondrial protein. *Genomics*, **83** (2): 254–61 (2004)

[671] Ng, S. B., Bigham, A. W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J. et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet*, **42** (9): 790–3 (2010)

[672] Schoumans, J., Wincent, J., Barbaro, M., Djureinovic, T., Maguire, P. et al. Comprehensive mutational analysis of a cohort of Swedish Cornelia de Lange syndrome patients. *Eur J Hum Genet*, **15** (2): 143–9 (2007)

[673] Yue, P. and Moult, J. Identification and analysis of deleterious human SNPs. *J Mol Biol*, **356** (5): 1263–74 (2006)

[674] Meins, M., Hagh, J. K., Gerresheim, F., Einhoff, E., Olschewski, H. et al. Novel case of dup(3q) syndrome due to a de novo interstitial duplication 3q24-q26.31 with minimal overlap to the dup(3q) critical region. *Am J Med Genet A*, **132A** (1): 84–9 (2005)

[675] van den Ende, J. J., Mattelaer, P., Declau, F., Vanhoenacker, F., Claes, J. et al. The facio-audio-symphalangism syndrome in a four generation family with a nonsense mutation in the NOG-gene. *Clin Dysmorphol*, **14** (2): 73–80 (2005)

[676] Zahir, F. R., Baross, A., Delaney, A. D., Eydoux, P., Fernandes, N. D. et al. A patient with vertebral, cognitive and behavioural abnormalities and a de novo deletion of NRXN1alpha. *J Med Genet*, **45** (4): 239–43 (2008)

[677] Tatton-Brown, K., Douglas, J., Coleman, K., Baujat, G., Cole, T. R. et al. Genotype-phenotype associations in Sotos syndrome: an analysis of 266 individuals with NSD1 aberrations. *Am J Hum Genet*, **77** (2): 193–204 (2005)

[678] Tassabehji, M., Read, A. P., Newton, V. E., Patton, M., Gruss, P. et al. Mutations in the PAX3 gene causing Waardenburg syndrome type 1 and type 2. *Nat Genet*, **3** (1): 26–30 (1993)

[679] Kasnauskiene, J., Ciuladaite, Z., Preiksaitiene, E., Matuleviciene, A., Alexandrou, A. et al. A single gene deletion on 4q28.3: PCDH18 - A new candidate gene for intellectual disability? *Eur J Med Genet*, **55** (4): 274–7 (2012)

[680] Griffith, E., Walker, S., Martin, C. A., Vagnarelli, P., Stiff, T. et al. Mutations in pericentrin cause Seckel syndrome with defective ATR-dependent DNA damage signaling. *Nat Genet*, **40** (2): 232–6 (2008)

[681] Michot, C., Le Goff, C., Goldenberg, A., Abhyankar, A., Klein, C. et al. Exome sequencing identifies PDE4D mutations as another cause of acrodysostosis. *Am J Hum Genet*, **90** (4): 740–5 (2012)

[682] Lower, K. M., Turner, G., Kerr, B. A., Mathews, K. D., Shaw, M. A. et al. Mutations in PHF6 are associated with Borjeson-Forssman-Lehmann syndrome. *Nat Genet*, **32** (4): 661–5 (2002)

[683] Horn, D., Krawitz, P., Mannhardt, A., Korenke, G. C., and Meinecke, P. Hyperphosphatasia-mental retardation syndrome due to PIGV mutations: expanded clinical spectrum. *Am J Med Genet A*, **155A** (8): 1917–22 (2011)

[684] Roa, B. B., Garcia, C. A., Suter, U., Kulpa, D. A., Wise, C. A. et al. Charcot-Marie-Tooth disease type 1A. Association with a spontaneous point mutation in the PMP22 gene. *N Engl J Med*, **329** (2): 96–101 (1993)

[685] Tartaglia, M., Mehler, E. L., Goldberg, R., Zampino, G., Brunner, H. G. et al. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat Genet*, **29** (4): 465–8 (2001)

[686] Digilio, M. C., Pacileo, G., Sarkozy, A., Limongelli, G., Conti, E. et al. Familial aggregation of genetically heterogeneous hypertrophic cardiomyopathy: a boy with LEOPARD syndrome due to PTPN11 mutation and his nonsyndromic father lacking PTPN11 mutations. *Birth Defects Res A Clin Mol Teratol*, **70** (2): 95–8 (2004)

[687] Deardorff, M. A., Wilde, J. J., Albrecht, M., Dickinson, E., Tennstedt, S. et al. RAD21 Mutations Cause a Human Cohesinopathy. *Am J Hum Genet*, **90** (6) (2012)

[688] Slager, R. E., Newton, T. L., Vlangos, C. N., Finucane, B., and Elsea, S. H. Mutations in RAI1 associated with Smith-Magenis syndrome. *Nat Genet*, **33** (4): 466–8 (2003)

[689] Arron, J. R., Winslow, M. M., Polleri, A., Chang, C. P., Wu, H. et al. NFAT dysregulation by increased dosage of DSCR1 and DYRK1A on chromosome 21. *Nature*, **441** (7093): 595–600 (2006)

[690] van Bokhoven, H., Celli, J., Kayserili, H., van Beusekom, E., Balci, S. et al. Mutation of the gene encoding the ROR2 tyrosine kinase causes autosomal recessive Robinow syndrome. *Nat Genet*, **25** (4): 423–6 (2000)

[691] Delaunoy, J., Abidi, F., Zeniou, M., Jacquot, S., Merienne, K. et al. Mutations in the X-linked RSK2 gene (RPS6KA3) in patients with Coffin-Lowry syndrome. *Hum Mutat*, **17** (2): 103–16 (2001)

[692] Kohlhase, J., Taschner, P. E., Burfeind, P., Pasche, B., Newman, B. et al. Molecular analysis of SALL1 mutations in Townes-Brocks syndrome. *Am J Hum Genet*, **64** (2): 435–45 (1999)

[693] Hoischen, A., van Bon, B. W., Gilissen, C., Arts, P., van Lier, B. et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet*, **42** (6): 483–5 (2010)

[694] Vidal-Taboada, J. M., Bergonon, S., Scartezzini, P., Egeo, A., Nizetic, D. et al. High-resolution physical map and identification of potentially regulatory sequences of the human SH3BGR located in the Down syndrome chromosomal region. *Biochem Biophys Res Commun*, **241** (2): 321–6 (1997)

[695] Bonaglia, M. C., Giorda, R., Borgatti, R., Felisari, G., Gagliardi, C. et al. Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. *Am J Hum Genet*, **69** (2): 261–8 (2001)

[696] Colmenares, C., Heilstedt, H. A., Shaffer, L. G., Schwartz, S., Berk, M. et al. Loss of the SKI proto-oncogene in individuals affected with 1p36 deletion syndrome is predicted by strain-dependent defects in Ski-/- mice. *Nat Genet*, **30** (1): 106–9 (2002)

[697] Deardorff, M. A., Kaur, M., Yaeger, D., Rampuria, A., Korolev, S. et al. Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. *Am J Hum Genet*, **80** (3): 485–94 (2007)

[698] Rizzoti, K. and Lovell-Badge, R. SOX3 activity during pharyngeal segmentation is required for craniofacial morphogenesis. *Development*, **134**: 3437–48 (2007)

[699] Lee, Y. and Saint-Jeannet, J. Sox9 function in craniofacial development and disease. *Genesis*, **49** (4): 200–8 (2011)

[700] Vieira, A. R., Avila, J. R., Daack-Hirsch, S., Dragan, E., Felix, T. M. et al. Medical sequencing of candidate genes for nonsyndromic cleft lip and palate. *PLoS Genet*, **1** (6): e64 (2005)

[701] Goodnough, L. H., Brugmann, S. A., Hu, D., and Helms, J. A. Stage-dependent craniofacial defects resulting from Sprouty2 overexpression. *Dev Dyn*, **236** (7): 1918–28 (2007)

[702] Song, T., Li, G., Jing, G., Jiao, X., Shi, J. et al. SUMO1 polymorphisms are associated with non-syndromic cleft lip with or without cleft palate. *Biochem Biophys Res Commun* (2008)

[703] Parvari, R., Hershkovitz, E., Grossman, N., Gorodischer, R., Loeys, B. et al. Mutation of TBCE causes hypoparathyroidism-retardation-dysmorphism and autosomal recessive Kenny-Caffey syndrome. *Nat Genet*, **32** (3): 448–52 (2002)

[704] Yagi, H., Furutani, Y., Hamada, H., Sasaki, T., Asakawa, S. et al. Role of TBX1 in human del22q11.2 syndrome. *Lancet*, **362** (9393): 1366–73 (2003)

[705] Zweier, C., Peippo, M. M., Hoyer, J., Sousa, S., Bottani, A. et al. Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am J Hum Genet*, **80** (5): 994–1001 (2007)

[706] Dixon, J., Trainor, P., and Dixon, M. J. Treacher Collins syndrome. *Orthod Craniofac Res*, **10** (2): 88–95 (2007)

[707] Milunsky, J. M., Maher, T. A., Zhao, G., Roberts, A. E., Stalker, H. J. et al. TFAP2A mutations result in branchio-oculo-facial syndrome. *Am J Hum Genet*, **82** (5): 1171–7 (2008)

[708] Xin, B., Puffenberger, E. G., Turben, S., Tan, H., Zhou, A. et al. Homozygous frameshift mutation in TMCO1 causes a syndrome with craniofacial dysmorphism, skeletal anomalies, and mental retardation. *Proc Natl Acad Sci U S A* (2009)

[709] Baynam, G., Goldblatt, J., and Walpole, I. Deletion of 8p23.1 with features of Cornelia de Lange syndrome and congenital diaphragmatic hernia and a review of deletions of 8p23.1 to 8pter? A further locus for Cornelia de Lange syndrome. *Am J Med Genet A*, **146A** (12): 1565–70 (2008)

[710] Chan, I., McGrath, J. A., and Kivirikko, S. Rapp-Hodgkin syndrome and the tail of p63. *Clin Exp Dermatol*, **30** (2): 183–6 (2005)

[711] Tukel, T., Sosic, D., Al-Gazali, L. I., Erazo, M., Casasnovas, J. et al. Homozygous nonsense mutations in TWIST2 cause Setleis syndrome. *Am J Hum Genet*, **87** (2): 289–96 (2010)

[712] Kishino, T., Lalande, M., and Wagstaff, J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet*, **15** (1): 70–3 (1997)

[713] Kolehmainen, J., Black, G. C., Saarinen, A., Chandler, K., Clayton-Smith, J. et al. Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. *Am J Hum Genet*, **72** (6): 1359–69 (2003)

[714] Yu, C. E., Oshima, J., Fu, Y. H., Wijsman, E. M., Hisama, F. et al. Positional cloning of the Werner's syndrome gene. *Science*, **272** (5259): 258–62 (1996)

[715] de Greef, J. C., Wang, J., Balog, J., den Dunnen, J. T., Frants, R. R. et al. Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Am J Hum Genet*, **88** (6): 796–804 (2011)

[716] Zweier, C., Thiel, C. T., Dufke, A., Crow, Y. J., Meinecke, P. et al. Clinical and mutational spectrum of Mowat-Wilson syndrome. *Eur J Med Genet*, **48** (2): 97–111 (2005)

[717] Nagaya, M., Kato, J., Niimi, N., Tanaka, S., and Wakamatsu, N. Clinical features of a form of Hirschsprung's disease caused by a novel genetic abnormality. *J Pediatr Surg*, **37** (8): 1117–22 (2002)

[718] Brown, S. A., Warburton, D., Brown, L. Y., Yu, C. Y., Roeder, E. R. et al. Holoprosencephaly due to mutations in ZIC2, a homologue of Drosophila odd-paired. *Nat Genet*, **20** (2): 180–3 (1998)

[719] Lawson, D., Hellenthal, G., Myers, S., and Falush, D. Inference of population structure using dense haplotype data. *PLoS Genet*, **8** (1): e1002453 (2012)

[720] Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C. C., Patsopoulos, N. A. et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*, **476** (7359): 214–219 (2011)

Appendix A

Glossary

Acculturation – the adoption of some traits from one culture into another
Aceramic – a culture lacking in pottery
Adstratum – a cultural overlay onto a native background
Allele – one of several alternative forms of a DNA sequence at a specific locus
Anthroponym – personal name
Artefact – an object made or modified by people
Assemblage – artefacts found together, probably deposited at the same time
Autosome – any chromosome that is not a sex chromosome (i.e. not X or Y)
Balancing selection – the process whereby (usually) two alleles at a given locus are actively maintained, usually due to a heterozygote advantage
Cephalic index – the length to breadth ratio of a skull
Chromosome – a molecule of DNA, of which the human genome has 23 pairs
Cognate words – words in different related languages that developed from the same word in the ancestral language
DNA – see deoxyribonucleic acid
Deoxyribonucleic acid – polymer encoding genetic (hereditary) information
Diploid – the state in which a cell/organism has two copies of a/each chromosome
Effective population size – the size of an population in which all individuals have the same probability of procreating, usually much smaller than the actual population it models
Ethnonym – ethnic/tribal/group name
Excarnation – the practice of defleshing a corpse, either actively or by natural means
Feature – non-artefactual archaeology that is part of the landscape; such as cropmarks, postholes
Fixation – the point at which one allele at a previously polymorphic locus reaches 100% in the population/species
Gametogenesis – the development of gametes (egg and sperm cells)
Gene – a stretch of DNA that codes for a single product, an RNA molecule or protein
Genetic drift – the change in allele frequency over generations due to random sampling of alleles
Genome – the entirety of the genetic information carried by an organism
Genotype – the allele states at both copies of a given locus
Genotyping – determining the genotype at a given locus
Haploid – the state in which a cell/organism has a single copy of a/each chromosome

Haplotype – the combination of alleles at loci along a section of a chromosome
Heterozygote – possessing two different alleles at a locus
Homozygote – possessing two identical alleles at a locus
Hydronym – name of a water feature, often a river
Inhumation – burial in a grave
Keratinocyte – a type of skin cell
Linkage disequilibrium – the non-random association of alleles, usually along the same chromosome
Lithics – stone artefact, most usually blades
Locus – a specific location on a chromosome
Mandible – lower jaw bone
Marker – a polymorphic DNA or protein sequence
Maxilla – bones forming the upper jaw
Melanocyte – a melanin–producing cell
Mesenchyme – undifferentiated connective tissue
Mitochondrial DNA – the chromosome of the maternally inherited mitochondrion, distinct from nuclear DNA
Negative selection – the process whereby detrimental genetic variants are eliminated/reduced in frequency in a population
Nucleotide – units of the DNA polymer
Ovicaprid – sheep/goat
Palimpsest – originally referring to a manuscript from which the text has been scraped for reuse, it is used to represent any surface that has been subject to successive and at least partially eradicating influences
Phenotype – observable structures, behaviours and physical appearance
Polymorphism – usages include a) any sequence variant, regardless of frequnecy, b) a genetic variant present at a frequency of ¿1%
Positive selection – the process whereby an advantageous genetic variant increases in frequency in a population
SNP – see single nucleotide polymorphism
Selection – see positive/adaptive, negative/purifying, balancing
Sequencing – determining the DNA sequence of a given stretch of the genome
Single nucleotide polymorphism – the difference of a single nucleotide between two sequences at a locus (technically, a SNP must be at a polymorphic frequency – see *polymorphism* – but it usually used to mean a single nucleotide variant, regardless of said variant's frequency)
Skeuomorph – a redundant feature in a new medium that was functional in the original medium
Stratigraphy – the layers of archaeology found in at a site, which built up over time
Substrate – language of a region that has been overlaid by another language
Superstrate – a new language in a region that overlays an older one
Toponym – place name
Tyrosine – one of twenty amino acids used in the production of proteins
Uniparentally inherited marker – a marker that is passed down strictly through the male or the female lines
Y chromosome – the male–specific chromosome, the counterpart of the X chromosome default

Appendix B

Populations typed for *MC1R*

Table B.1: **Populations typed for *MC1R* variants** 'Population code' is the unique identifier given to the sample set. Population codes with alphanumeric suffixes (e.g. Chinese4a & Chinese4b) are the same sampling, for which a different number of its constituent individuals has been typed for a given marker. 'Geographic homeland' is the area in which the ethnic/national group in question has its largest populations, 'Sampling location' is the exact place where the individual samples were collected, when known, 'N' is the number of individuals in the sample, and 'Source' details where the information was obtained.

Population code	**Geographic homeland**	**Sampling location**	**N**	**Source**
Adjar1	Georgia		35	this study
Ainu1	Japan		10	[576]
Ami1	Taiwan		48	[576]
Arab1	Arabian peninsula	Bukhara, Uzbekistan	53	this study
Armenian1	Armenia		77	this study
Asian1	East Asia	Taiwan, S China, N China, Cambodia, Japan, Vietnam	60	[153]
Azeri1	Azerbaijan		47	this study
Bangladeshi1	Bangladesh	Dhaka	118	[577]
Bartang1	Tajikistan		31	this study
Black Lahu1	Thailand		25	[576]
British1a	Britain		902	[522]
British1b	Britain		896	[522]
British2	Britain	Birmingham	51	this study

Continued on the next page

– continued from the previous page

Population code	Geographic homeland	Sampling location	N	Source
Bugis1	Indonesia	Sulawesi	39	[576]
Buryat1	Southeast Russia/ Mongolia	East Mongolia	50	[576]
Buryat2	Southeast Russia/ Mongolia	East Mongolia	143	[577]
Chinese1	China	Huizhou	111	[577]
Chinese2	China	Shenyang	87	[577]
Chinese3	China	Wuxi	119	[577]
Chinese4a	China		95	[578]
Chinese4b	China		94	[578]
Chinese5	China		24	[579]
Chinese6	China	Beijing	35	[576]
Chinese7	China	Xian	28	[576]
Chinese8a	China	Beijing	84	HapMap
Chinese8b	China	Beijing	83	HapMap
Chinese9a	China	Denver, USA	85	HapMap
Chinese9b	China	Denver, USA	84	HapMap
Chinese10	China		38	this study
Dai1	China	Yunnan	33	[580]
Danish1	Denmark		378	[581]
Dayak1	Indonesia	Borneo	42	[576]
Dungan1	Kyrgyzstan, Kazakhstan, Russia	Kyrgyzstan	63	this study
Dutch1	The Netherlands	Leiden	385	[147]
English1	Britain		32	[249]
European1a	Northern Europe	Utah. USA	111	HapMap
European1b	Northern Europe	Utah. USA	109	HapMap
Ewenki1	China, Russia, Mongolia	Manchuria	45	[576]
Finnish1	Finland		15	[249]
French1	France	Paris	105	[582]
French2	France		172	[583]
French3	France	Rheims	98	[577]
German1	Germany	Mannheim	1038	[584]
German2	Germany	Munich	92	[577]
German3	Germany	Westphalia	199	[577]
Greek1	Greece	Central & southern	155	[585, 586]
Indian1	India	Andhra Pradesh, Bengal, Gujarat, Maharashtra, Punjab, Tamil Nadu, UttarPradesh	20	[153]

Continued on the next page

– continued from the previous page

Population code	**Geographic homeland**	**Sampling location**	**N**	**Source**
Indian2	India	New Dehli	107	[577]
Indian3	India	Southern	22	[249]
Indian4	India	UK	11	[249]
Indian5	India		83	[578]
Indian6a	India	Houston. USA	87	HapMap
Indian6b	India	Houston. USA	86	HapMap
Indian6c	India	Houston. USA	88	HapMap
Indonesian1	Indonesia	Surabaya	105	[577]
Inuit1	Canada	Igoolik	20	[249]
Iranian1	Iran	Samarkand	90	this study
Iranian2	Iran	Tehran	37	this study
Iranian3	Iran	Esfahan	25	this study
Iranian4	Iran	Shiraz	23	this study
Irish1	Ireland		61	[249]
Ishkashim1	Tajikistan. Afghanistan	Tajikistan	25	this study
Italian1	Italy	Liguria	207	[587]
Italian2	Italy	Central	100	[588]
Italian3	Italy	Southern Emilia-Romagna and Northern Marche	171	[589]
Italian4	Italy	Rome. Sardinia	17	[249]
Italian5a	Italy	Tuscany	88	HapMap
Italian5b	Italy	Tuscany	87	HapMap
Japanese1	Japan		25	[576]
Japanese2	Japan		200	[579]
Japanese3	Japan	Okinawa	87	[577]
Japanese4	Japan	Tottori	103	[577]
Japanese5	Japan		238	[590]
Japanese6	Japan		15	[249]
Japanese7a	Japan	Tokyo	85	HapMap
Japanese7b	Japan	Tokyo	44	HapMap
Japanese7c	Japan	Tokyo	84	HapMap
Japanese7d	Japan	Tokyo	86	HapMap
Javanese1	Indonesia	Java	30	[576]
Jewish1	Uzbekistan	Samarkand	10	this study
Kallar1	India		85	this study
Karakalpak1	Uzbekistan		95	this study
Karakalpak2	Uzbekistan	Uzbekistan	11	this study
Kazakh1	Kazakhstan, Uzbekistan	Kazakhstan, Uzbekistan	146	this study

Continued on the next page

– continued from the previous page

Population code	Geographic homeland	Sampling location	N	Source
Kazbegi1	Georgia		47	this study
Khalkha1	Mongolia	Ulan Baator	173	[577]
Khalkha2	Mongolia		40	[576]
Korean1a	Korea		111	[591]
Korean1b	Korea		74	[591]
Korean2	Korea		51	this study
Kurd1	Western Asia, Turkmenistan, Uzbekistan	Turkmenistan. Uzbekistan	37	this study
Kyrgyz1	Kyrgyzstan, Uzbekistan	Kyrgyzstan. Uzbekistan	67	this study
Lebanese1	Lebanon		50	this study
Lezgi1	Georgia, Russia	Georgia	42	this study
Lisu1	Thailand		29	[576]
Madurian1	Indonesia	Java	28	[576]
Malay1	Malaysia		51	[576]
Malay2	Malaysia		88	[578]
Manchu1	China	Manchuria	50	[576]
Mlabri1	Thailand		25	[576]
Mongolian1	Mongolia		65	[579]
Orkney1a	Britain	Orkney	98	[522]
Orkney1b	Britain	Orkney	97	[522]
Orkney2	Britain	Orkney	72	this study
Ossetian1	Georgia		25	this study
Palawan1	Philippines	Palawan	28	[576]
Romani1		Uzbekistan	25	this study
Russian1	Russia	Tashkent	93	this study
Saami1	Northern Fennoscandia		13	[249]
Sardinian1	Sardinia, Italy		14	[249]
Semang1	Malaysia	Malay peninsula	27	[576]
Seramese1	Indonesia	Seram	29	[576]
Shan1	Thailand		25	[576]
Shugnan1	Tajikistan, Afghanistan		44	this study
Sourashtra1	India		47	this study
Spanish1	Spain	Madrid	188	[592]
Spanish2	Spain	Valencia	544	[584]
Svan1	Georgia		30	this study
Swedish1	Sweden		663	[593]
Swedish2	Sweden		26	[249]

Continued on the next page

– continued from the previous page

Population code	Geographic homeland	Sampling location	N	Source
Tajik1	Tajikistan, Uzbekistan, Afghanistan	Samarkand	74	this study
Tajik2	Tajikistan, Uzbekistan, Afghanistan	Khojant	31	this study
Tajik3	Tajikistan, Uzbekistan, Afghanistan	Dushanbe	16	this study
Tatar1	Russia	Uzbekistan	40	this study
Tatar2	Russia, Central Asia	Uzbekistan	43	this study
Tibetan1	China, India	Qinghai province	20	[580]
Timorese1	Indonesia	Timor	45	[576]
Toraja1	Indonesia	Sulawesi	51	[576]
Turkish1	Turkey	West Germany	200	[577]
Turkmen1	Turkmenistan	Turkmenistan, Uzbekistan	66	this study
Tuva1	Southeast Russia		51	this study
Uighur1	China, Kyrgyzstan, Uzbekistan, Kazakhstan	Xinjiang province	35	[580]
Uighur2	China, Kyrgyzstan, Kazakhstan, Uzbekistan	Kazakhstan, Uzbekistan	56	this study
Uzbek1	Uzbekistan	Fergana Valley	74	this study
Uzbek2	Uzbekistan	Surkhandarya	68	this study
Uzbek3	Uzbekistan	Tashkent	55	this study
Uzbek4	Uzbekistan	Bukhara	88	this study
Uzbek5	Uzbekistan	Kashkadarya	29	this study
Uzbek6	Uzbekistan	Samarkand	77	this study
Uzbek7	Uzbekistan	Khorezm	95	this study
Vedda1	Sri Lanka		20	[576]
Wa1	China	Yunnan province	34	[580]
Yadhava1	India		129	this study
Yagnobi1	Tajikistan		31	this study
Yami1	Taiwan		48	[576]

Appendix C

Populations typed for *CCR5del32*

Table C.1: **Populations typed for *CCR5del32*** 'Population code' is the unique identifier given to the sample set. Population codes with alphanumeric suffixes (e.g. Chinese4a & Chinese4b) are the same sampling, for which a different number of its constituent individuals has been typed for a given marker. 'Geographic homeland' is the area in which the ethnic/national group in question has its largest populations, 'Sampling location' is the exact place where the individual samples were collected, when known, 'N' is the number of individuals in the sample, and 'Source' details where the information was obtained.

Population code	**Geographic homeland**	**Sampling location**	**N**	**Source**
Adjar1	Georgia	Georgia	35	this study
Albanian1	Albania		73	[107]
Arab1	Arabian peninsula	Bukhara	53	this study
Armenian1	Armenia		77	this study
Austrian1	Austria		36	[107]
Azeri1	Azerbaijan		47	this study
Azeri2	Azerbaijan		40	[107]
Bartang1	Tajikistan		31	this study
Bashkir1	Russia	Abselilovski	60	[594, 109]
Bashkir2	Russia	Ilishevski	50	[594, 109]
Bashkir3	Russia	Arkhangelski	46	[594, 109]
Bashkir4	Russia	Sterlibashevski	49	[109]

Continued on the next page

– continued from the previous page

Population code	**Geographic homeland**	**Sampling location**	**N**	**Source**
Basque1	Spain, France	Biarritz	111	[595]
Basque2	Spain, France	Bilbao	89	[596]
Basque3	Spain, France		29	[597]
Belgian1	Belgium	Brussels	704	[596]
Belgian2	Belgium	Leuven	310	[598]
Bengali1	India		25	[597]
British2	Britain	Birmingham	51	this study
British3	Britain		283	[597]
British4	Britain		422	[107]
Bulgarian1	Bulgaria		29	[107]
Burmese1	Burma		67	[597]
Byelorussian1	Byelorussia		80	[109, 599]
Catalan1	Spain	Perpignan	102	[595]
Catalan2	Spain	Spain	49	[597]
Chinese11	China		1046	[600]
Chinese12	China		715	[601]
Chinese13	China		40	[107]
Chuvash1	Russia	Morgaushevski	79	[109]
Corsican1	Corsica	Ajaccio	104	[595]
Croatian1	Croatia		303	[602]
Cypriot1	Cyprus		84	[597]
Cypriot2	Cyprus		1002	[603]
Czech1	Czech Republic		161	[107]
Daghestani1	Russia		110	[597]
Danish2	Denmark	Copenhagen	100	[596]
Danish3	Denmark	Copenhagen	239	[595]
Danish4	Denmark		24	[107]
Dungan1	Kyrgyzstan, Kazakhstan, Russia	Kyrgyzstan	63	this study
Estonian1	Estonia		158	[107]
Finnish2	Finland	Helsinki	98	[596]
Finnish3	Finland		195	[107]
French4	France	Lille	101	[595]
French5	France	Reims	276	[595]
French6	France	Nancy-Strasbourg	291	[595]
French7	France	Paris	294	[595]
French8	France	Brest	107	[595]
French9	France	Montpellier	326	[595]
French10	France	Montpellier	99	[596]
French11	France		230	[107]

Continued on the next page

– continued from the previous page

Population code	**Geographic homeland**	**Sampling location**	**N**	**Source**
Georgian1	Georgia		50	[107]
German4	Germany	Mulheim	99	[595]
German5	Germany		208	[107]
Greek2	Greece		63	[597]
Greek3	Greece	Athens	143	[604]
Greek4	Greece		160	[107]
Gujarati1	India		32	[597]
Hong Kong1	China		50	[597]
Hungarian1	Hungary	Budapest	99	[596]
Icelandic1	Iceland		102	[597]
Iranian1	Iran	Samarkand	96	this study
Iranian2	Iran	Tehran	37	this study
Iranian3	Iran	Esfahan	25	this study
Iranian4	Iran	Shiraz	23	this study
Iranian5	Iran		380	[605]
Iranian6	Iran		84	[606]
Irish2	Ireland		44	[597]
Irish3	Ireland		31	[107]
Ishkashim1	Tajikistan, Afghanistan		25	this study
Italian6	Italy		91	[597]
Italian7	Italy	Milan	98	[596]
Italian8	Italy	Padua	371	[607]
Italian9	Italy	Rome	548	[608]
Italian10	Italy		172	[107]
Jewish2		Ashkenazi	503	[107]
Karakalpak1.2	Uzbekistan	Uzbekistan	95	this study
Kazakh1	Kazakhstan, Uzbekistan	Kazakhstan, Uzbekistan	146	this study
Kazakh2	Kazakhstan, Uzbekistan		50	[107]
Kazbegi1	Georgia		47	this study
Komi Ziryanes1	Russia	Sisolski	50	[594, 109]
Korean2	Korea		51	this study
Korean3	Korea		50	[107]
Kurd1	Western Asia, Turkmenistan, Uzbekistan	Turkmenistan, Uzbekistan	37	this study
Kyrgyz1	Kyrgyzstan, Uzbekistan	Kyrgystan, Uzbekistan	67	this study

Continued on the next page

– continued from the previous page

Population code	**Geographic homeland**	**Sampling location**	**N**	**Source**
Lebanese1	Lebanon		50	this study
Lebanese2	Lebanon		209	[609]
Lebanese3	Lebanon		51	[107]
Lezgi1	Russia, Azerbaijan		42	this study
Lithuanian1	Lithuania	Vilnius	283	[596]
Mari1	Russia	Zvenigovski	47	[594, 109]
Moldovan1	Moldova	Kishinev	56	[109]
Mongolian2	Mongolia		59	[597]
Mordvinian1	Russia	Saransk	86	[596]
Mordvinian2	Russia	Staroshayginski	51	[594, 109]
Norwegian1	Norway	Oslo	100	[596]
Orkney2	Britain		72	this study
Ossetian1	Georgia		25	this study
Pakistani1	Pakistan		34	[597]
Polish1	Poland		30	[107]
Portguese1	Portugal	Lisbon	101	[596]
Portuguese2	Portugal	Porto	124	[595]
Punjabi1	India, Pakistan		34	[597]
Romani1		Uzbekistan	25	this study
Romani2		Bulgaria	47	[107]
Russian1	Russia	Tashkent	93	this study
Russian2	Russia	Moscow	50	[596]
Russian3	Russia	Ryazan	78	[109, 599]
Russian4	Russia	Moscow	176	[606]
Russian5	Russia		50	[107]
Saami2	Northern Fennoscandia	Sweden	120	[596]
Sardinian2	Sardinia, Italy		100	[596]
Sardinian3	Sardinia, Italy		264	[608]
Saudi-Arabian1	Arabian peninsula		241	[597]
Saudi-Arabian2	Arabian peninsula		100	[107]
Shugnan1	Tajikistan, Afghanistan		44	this study
Sicilian1	Sicily, Italy		901	[610]
Sindh1	Pakistan		29	[597]
Slovakian1	Slovakia		30	[107]
Slovenian1	Slovenia		110	[107]
Spanish3	Spain	Murcia	100	[596]
Spanish4	Spain	Oviedo	250	[611]
Spanish5	Spain	Barcelona	98	[595]

Continued on the next page

– continued from the previous page

Population code	**Geographic homeland**	**Sampling location**	**N**	**Source**
Spanish6	Spain	Seville	208	[612]
Spanish7	Spain		56	[107]
Sri Lanka1	Sri Lanka		37	[597]
Svan1	Georgia		30	this study
Swedish3	Sweden	Umeå	204	[596]
Swedish4	Sweden		131	[107]
Swiss1	Switzerland	Bern	64	[595]
Syrian1	Syria		106	[606]
Taiwan1	Taiwan		83	[597]
Tajik1	Tajikistan, Uzbekistan, Afghanistan	Samarkand	74	this study
Tajik2	Tajikistan, Uzbekistan, Afghanistan	Khojant	31	this study
Tajik3	Tajikistan, Uzbekistan, Afghanistan	Dushanbe	16	this study
Tatar1	Russia	Uzbekistan	40	this study
Tatar2	Russia, Central Asia	Uzbekistan	43	this study
Tatar3	Russia, Turkey, Central Asia	Russia	40	[596]
Tatar4	Russia, Turkey, Central Asia	Almetyevsk	48	[109]
Tatar5	Russia, Turkey, Central Asia	Elabuga	45	[109]
Tatar6	Russia, Turkey, Central Asia	Crimea	90	[109]
Tatar7	Russia, Turkey, Central Asia		50	[107]
Thai1	Thailand		101	[597]
Turkish2	Turkey	Ankara	104	[596]
Turkish3	Turkey		40	[107]
Turkmen1	Turkmenistan	Turkmenistan. Uzbekistan	66	this study
Tuva1	Russia		51	this study
Tuva2	Russia		50	[109, 613]
Tuva3	Russia		50	[107]
Udmurts1	Russia	Malo-Purginski	52	[594, 109]
Uighur2	China, Kyrgyzstan, Uzbekistan, Kazakhstan	Kazakhstan. Uzbekistan	56	this study

Continued on the next page

– continued from the previous page

Population code	Geographic homeland	Sampling location	N	Source
Uighur3	China, Kyrgyzstan, Uzbekistan, Kazakhstan	Russia	50	[109, 613]
Uighur4	China, Kyrgyzstan, Uzbekistan, Kazakhstan		45	[107]
Ukrainian1	Ukraine	Lvov	88	[109]
Ukrainian2	Ukraine	Kiev	83	[109]
Ukrainian3	Ukraine	Lugansk	86	[109]
Uzbek1	Uzbekistan	Fergana Valley	74	this study
Uzbek2	Uzbekistan	Surhandarya	68	this study
Uzbek3	Uzbekistan	Tashkent	55	this study
Uzbek4	Uzbekistan	Bukhara	88	this study
Uzbek5	Uzbekistan	Kashkadarya	29	this study
Uzbek6	Uzbekistan	Samarkand	77	this study
Uzbek7	Uzbekistan	Khorezm	95	this study
Uzbek8	Uzbekistan		40	[109, 613]
Uzbek9	Uzbekistan		29	[107]
Yemeni1	Yemen		34	[597]

Appendix D

CCR5del32 F$_{ST}$ tables

Table D.1: Significant values are in bold

	Adjar	Arab (Uzb.)	Armenian	Azeri	Dungan	Iranian (Uzb.)	Kazakh	Kazbegi	Korean	Kurd	Kyrgyz	Lezgi	Orcadian	Ossetian	Santi Romani
Adjar	0.00														
Arab (Uzb.)	0.00	0.00													
Armenian	0.00	0.00	0.00												
Azeri	0.01	0.02	0.01	0.00											
Dungan	0.00	0.00	0.00	0.00	0.00										
Iranian (Uzb.)	0.00	0.00	0.00	0.00	0.00	0.00									
Kazakh	0.00	0.00	0.00	0.00	0.00	0.00	0.00								
Kazbegi	0.04	**0.07**	**0.05**	0.00	0.04	**0.04**	**0.04**	0.00							
Korean	0.00	0.00	0.00	0.01	0.00	0.00	0.00	**0.04**	0.00						
Kurd	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00					
Kyrgyz	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00	0.00				
Lezgi	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00			
Orcadian	0.04	**0.05**	**0.04**	0.00	**0.03**	0.03	**0.03**	0.00	**0.03**	0.02	**0.02**	0.00	0.00		
Ossetian	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.02	0.00	
Santi Romani	0.00	0.00	0.00	0.03	0.00	0.01	0.01	0.06	0.00	0.01	0.01	0.02	0.05	0.00	0.00
Russian (Uzb.)	0.02	0.04	0.03	0.00	0.02	0.02	**0.02**	0.00	0.02	0.01	0.01	0.00	0.00	0.01	0.04
Svan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00
Turkmen	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.02	0.00	0.01
Tuva	0.00	0.00	0.00	0.01	0.00	0.00	0.00	**0.04**	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00
Uighur	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00	0.00	0.00	0.03	0.00	0.01
Yagnobi	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
Corsican	0.00	0.00	0.00	**0.03**	0.00	0.00	0.01	**0.09**	0.00	0.00	0.01	0.03	**0.07**	0.00	0.00
Croatian	0.02	**0.03**	**0.02**	0.00	0.01	0.01	**0.01**	0.00	0.02	0.01	0.01	0.00	0.00	0.01	0.03
Finnish	**0.08**	**0.10**	**0.10**	**0.04**	**0.09**	**0.09**	**0.10**	0.01	**0.09**	**0.07**	**0.08**	**0.05**	**0.01**	**0.07**	**0.09**
German	**0.04**	**0.06**	**0.05**	0.01	**0.04**	**0.04**	**0.04**	0.00	**0.04**	0.03	**0.03**	0.01	0.00	0.03	**0.06**
Hungarian	0.03	**0.04**	**0.03**	0.00	0.03	**0.03**	**0.03**	0.00	0.03	0.02	0.02	0.00	0.00	0.02	0.04
Icelandic	**0.07**	**0.09**	**0.09**	**0.03**	**0.07**	**0.08**	**0.09**	0.00	**0.08**	**0.06**	**0.07**	0.04	0.01	**0.06**	**0.09**
Irish	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.02
Lithuanian	**0.04**	**0.05**	**0.05**	0.01	**0.04**	**0.04**	**0.04**	0.00	**0.04**	**0.03**	**0.04**	**0.02**	0.00	**0.04**	**0.06**
Norwegian	0.04	**0.06**	**0.05**	0.01	**0.04**	**0.04**	**0.04**	0.00	**0.04**	0.03	**0.03**	0.01	0.00	0.03	**0.06**
Saami	**0.03**	**0.04**	**0.03**	0.00	0.02	**0.02**	**0.02**	0.00	0.03	0.02	0.02	0.00	0.00	0.02	0.04
Sicilian	0.00	0.01	0.00	0.00	0.00	0.00	0.00	**0.02**	0.00	0.00	0.00	0.00	0.02	0.00	0.01
Swedish	**0.06**	**0.08**	**0.07**	**0.03**	**0.06**	**0.06**	**0.07**	0.00	**0.06**	**0.05**	**0.06**	**0.03**	0.01	0.05	**0.08**
Swiss	**0.03**	**0.05**	**0.04**	0.00	0.03	0.03	**0.03**	0.00	**0.03**	0.02	0.02	0.00	0.00	0.02	0.05
Daghestani	0.02	**0.03**	0.02	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.03
Syrian	0.00	0.00	0.00	0.02	0.00	0.00	0.00	**0.08**	0.00	0.00	0.00	0.02	**0.06**	0.00	0.00
Turkish	0.01	**0.03**	**0.02**	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.03
Udmurts	**0.06**	**0.08**	**0.07**	0.01	**0.06**	**0.06**	**0.06**	0.00	**0.06**	0.04	**0.05**	0.02	0.00	0.04	**0.07**
Komi-Ziryanes	0.01	0.03	0.02	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.03
Mari	0.04	**0.07**	**0.05**	0.00	0.04	**0.04**	**0.04**	0.00	0.04	0.03	**0.03**	0.01	0.00	0.03	**0.06**
Mongolian	0.01	0.00	0.01	**0.05**	0.01	0.02	0.02	**0.10**	0.01	0.02	0.02	**0.05**	**0.08**	0.02	0.00
Taiwanese	0.01	0.00	0.01	**0.06**	0.02	0.02	0.02	**0.12**	0.02	0.03	0.03	**0.06**	**0.09**	0.03	0.00
Burmese	0.01	0.00	0.01	**0.06**	0.02	0.02	0.02	**0.11**	0.01	0.03	0.02	**0.05**	**0.08**	0.02	0.00
Thai	0.00	0.00	0.00	**0.05**	0.01	0.01	0.01	**0.11**	0.00	0.01	0.01	**0.04**	**0.08**	0.00	0.00
Pamir	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	**0.02**	0.00	0.01
British	**0.04**	**0.05**	**0.04**	0.01	**0.04**	**0.04**	**0.04**	0.00	**0.04**	**0.03**	**0.03**	0.02	0.00	0.03	**0.05**
Iranian	0.01	**0.02**	**0.01**	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.02
Indian	0.00	0.00	0.00	**0.05**	0.00	0.01	**0.01**	**0.15**	0.00	0.00	0.01	0.03	**0.12**	0.00	0.00
Pakistani	0.00	0.01	0.00	0.00	0.00	0.00	0.00	**0.02**	0.00	0.00	0.00	0.00	0.01	0.00	0.01
Lebanese	0.00	0.00	0.00	0.01	0.00	0.00	0.00	**0.06**	0.00	0.00	0.00	**0.00**	**0.05**	0.00	0.00
Tajik	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.02
Tatar	0.03	**0.05**	**0.03**	0.00	**0.03**	0.02	**0.02**	0.00	**0.03**	0.02	**0.02**	0.00	0.00	0.02	0.04
Uzbek	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.02
Bashkir	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.04**	0.00	0.00	0.00	0.00	**0.03**	0.00	0.01
Basque	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.02	0.00	0.01
Belgian	**0.03**	**0.04**	**0.03**	0.01	**0.03**	**0.03**	**0.03**	0.00	**0.03**	0.02	**0.02**	0.01	0.00	0.03	0.04
Catalan	0.02	0.03	0.02	0.00	0.02	0.01	0.01	0.00	0.02	0.01	0.01	0.00	0.00	0.01	0.03
Chinese	0.00	0.00	0.02	**0.17**	0.03	**0.04**	**0.05**	**0.39**	0.02	0.04	**0.05**	**0.14**	**0.33**	0.02	0.00
Cypriot	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.06**	0.00	0.00	0.00	0.00	**0.05**	0.00	0.01
Danish	**0.05**	**0.06**	**0.05**	0.02	**0.04**	**0.04**	**0.04**	0.00	**0.05**	**0.03**	**0.04**	0.02	0.00	0.04	**0.06**
French	**0.04**	**0.05**	**0.04**	0.01	**0.03**	**0.03**	**0.03**	0.00	**0.04**	**0.03**	**0.03**	0.01	0.00	0.03	**0.05**
Greek	0.00	0.01	0.00	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00	0.00	0.00	**0.03**	0.00	0.01
Italian	0.01	0.02	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.02
Mordvinian	**0.05**	**0.07**	**0.06**	0.02	**0.05**	**0.05**	**0.06**	0.00	**0.05**	**0.04**	**0.05**	0.02	0.00	**0.04**	**0.07**
Portuguese	0.01	**0.02**	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.02
Russian	**0.05**	**0.06**	**0.06**	0.02	**0.05**	**0.05**	**0.05**	0.00	**0.05**	**0.04**	**0.04**	0.02	0.00	**0.04**	**0.06**
Sardinian	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.04**	0.00	0.00	0.00	0.00	**0.03**	0.00	0.01
Saudi-Arabian	0.00	0.00	0.00	0.02	0.00	0.00	0.00	**0.09**	0.00	0.00	0.00	0.01	**0.07**	0.00	0.00
Yemeni	0.02	**0.03**	**0.02**	0.00	**0.01**	**0.01**	**0.01**	0.00	0.02	0.01	0.01	0.00	0.00	0.01	**0.03**
Spanish	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05

Table D.2:

	Russian (Uzb.)	Svan	Turkmen	Tuva	Uighur	Yagnobi	Corsican	Croatian	Finnish	German	Hungarian	Icelandic	Irish	Lithuanian	Norwegian
Adjar															
Arab (Uzb.)															
Armenian															
Azeri															
Dungan															
Iranian (Uzb.)															
Kazakh															
Kazbegi															
Korean															
Kurd															
Kyrgyz															
Lezgi															
Orcadian															
Ossetian															
Santi Romani															
Russian (Uzb.)	0.00														
Svan	0.02	0.00													
Turkmen	0.01	0.00	0.00												
Tuva	0.02	0.00	0.00	0.00											
Uighur	0.01	0.00	0.00	0.00	0.00										
Yagnobi	0.00	0.00	0.00	0.00	0.00	0.00									
Corsican	**0.05**	0.00	0.01	0.00	0.00	0.03	0.00								
Croatian	0.00	0.01	0.00	0.02	0.01	0.00	**0.03**	0.00							
Finnish	**0.03**	**0.08**	**0.07**	**0.09**	**0.08**	**0.04**	**0.13**	**0.04**	0.00						
German	0.00	0.04	0.03	**0.04**	**0.04**	0.01	**0.08**	0.01	0.01	0.00					
Hungarian	0.00	0.03	0.01	0.03	0.02	0.00	**0.06**	0.00	**0.02**	0.00	0.00				
Icelandic	**0.02**	**0.07**	**0.06**	**0.08**	**0.07**	**0.03**	**0.12**	**0.03**	0.00	0.00	0.01	0.00			
Irish	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	**0.05**	0.01	0.00	**0.04**	0.00		
Lithuanian	0.00	**0.04**	**0.03**	**0.04**	**0.04**	0.01	**0.06**	**0.01**	0.01	0.00	0.00	0.00	**0.02**	0.00	
Norwegian	0.00	**0.04**	0.02	**0.04**	**0.04**	0.01	**0.08**	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00
Saami	0.00	0.02	0.01	0.03	0.02	0.00	**0.05**	0.00	0.02	0.00	0.00	**0.02**	0.00	0.00	0.00
Sicilian	0.01	0.00	0.00	0.00	0.00	0.00	0.01	**0.01**	**0.11**	**0.03**	**0.01**	**0.09**	0.00	**0.04**	**0.03**
Swedish	**0.02**	**0.06**	**0.05**	**0.06**	**0.06**	0.03	**0.09**	**0.03**	0.00	0.00	0.01	0.00	**0.03**	0.00	0.00
Swiss	0.00	0.03	0.01	0.03	0.02	0.00	**0.07**	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00
Daghestani	0.00	0.01	0.00	0.01	0.01	0.00	**0.04**	0.00	**0.04**	0.01	0.00	**0.03**	0.00	0.01	0.01
Syrian	**0.04**	0.00	0.01	0.00	0.00	0.02	0.00	**0.03**	**0.12**	**0.07**	**0.05**	**0.11**	0.01	**0.06**	**0.07**
Turkish	0.00	0.01	0.00	0.01	0.01	0.00	**0.03**	0.00	**0.04**	0.01	0.00	**0.03**	0.00	0.01	0.01
Udmurts	0.00	0.05	0.04	**0.06**	**0.05**	0.01	**0.12**	0.01	0.00	0.00	0.00	0.00	0.02	0.00	0.00
Komi-Ziryanes	0.00	0.01	0.00	0.01	0.00	0.00	**0.04**	0.00	0.04	0.01	0.00	**0.03**	0.00	0.01	0.00
Mari	0.00	0.04	0.02	**0.04**	**0.03**	0.00	**0.09**	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00
Mongolian	**0.06**	0.01	0.03	0.01	0.02	0.06	0.00	**0.04**	**0.12**	**0.08**	**0.06**	**0.11**	0.04	**0.07**	**0.08**
Taiwanese	**0.07**	0.02	**0.04**	0.02	0.03	**0.08**	0.00	**0.04**	**0.14**	**0.09**	**0.07**	**0.13**	**0.06**	**0.07**	**0.09**
Burmese	**0.06**	0.01	**0.03**	0.01	0.02	0.06	0.00	**0.04**	**0.13**	**0.08**	**0.07**	**0.12**	**0.05**	**0.07**	**0.08**
Thai	**0.06**	0.00	0.02	0.00	0.01	**0.05**	0.00	**0.04**	**0.14**	**0.09**	**0.07**	**0.13**	0.04	**0.07**	**0.09**
Pamir	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	**0.08**	**0.03**	0.02	**0.07**	0.00	**0.03**	**0.03**
British	0.00	**0.04**	**0.03**	**0.04**	**0.03**	0.01	**0.06**	**0.01**	**0.01**	0.00	0.00	0.00	0.02	0.00	0.00
Iranian	0.00	0.01	0.00	0.01	0.01	0.00	**0.03**	0.00	**0.05**	0.01	0.00	**0.04**	0.00	**0.02**	0.01
Indian	**0.08**	0.00	0.02	0.00	0.00	0.04	0.00	**0.05**	**0.23**	**0.14**	**0.10**	**0.21**	0.03	**0.10**	**0.13**
Pakistani	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	**0.07**	**0.03**	0.01	**0.06**	0.00	**0.03**	**0.03**
Lebanese	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	**0.02**	**0.13**	**0.06**	**0.04**	**0.11**	0.00	**0.05**	**0.06**
Tajik	0.00	0.00	0.00	0.00	0.00	0.00	**0.02**	0.00	**0.07**	0.02	0.01	**0.06**	0.00	**0.02**	0.02
Tatar	0.00	0.03	0.01	**0.03**	0.02	0.00	**0.06**	0.00	**0.02**	0.00	0.00	0.01	0.00	0.00	0.00
Uzbek	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.00	**0.08**	**0.02**	0.01	**0.07**	0.00	**0.03**	**0.02**
Bashkir	0.02	0.00	0.00	0.00	0.00	0.00	0.01	**0.01**	**0.10**	**0.04**	**0.02**	**0.08**	0.00	**0.04**	**0.04**
Basque	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.08**	**0.03**	0.01	**0.07**	0.00	**0.03**	**0.03**
Belgian	0.00	0.03	**0.02**	**0.03**	**0.03**	0.01	**0.05**	**0.00**	**0.01**	0.00	0.00	0.01	0.01	0.00	0.00
Catalan	0.00	0.02	0.01	0.02	0.01	0.00	**0.04**	0.00	**0.03**	0.00	0.00	**0.03**	0.00	0.01	0.00
Chinese	**0.25**	0.01	**0.09**	0.02	**0.04**	**0.15**	0.00	**0.14**	**0.52**	**0.36**	**0.29**	**0.48**	**0.13**	**0.26**	**0.36**
Cypriot	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	**0.02**	**0.18**	**0.08**	**0.04**	**0.15**	0.00	**0.07**	**0.07**
Danish	0.01	**0.04**	**0.03**	**0.05**	**0.04**	0.02	**0.07**	**0.01**	0.00	0.00	0.00	0.00	0.02	0.00	0.00
French	0.00	0.03	**0.02**	**0.04**	**0.03**	0.01	**0.05**	**0.01**	**0.01**	0.00	0.00	0.00	0.01	0.00	0.00
Greek	0.01	0.00	0.00	0.00	0.00	0.00	0.01	**0.01**	**0.10**	**0.04**	**0.02**	**0.08**	0.00	**0.04**	**0.04**
Italian	0.00	0.01	0.00	0.01	0.00	0.00	**0.02**	0.00	**0.08**	**0.02**	0.01	**0.06**	0.00	**0.03**	**0.02**
Mordvinian	0.01	**0.05**	**0.04**	**0.05**	**0.05**	0.02	**0.09**	**0.01**	0.00	0.00	0.00	0.00	0.02	0.00	0.00
Portuguese	0.00	0.01	0.00	0.01	0.00	0.00	**0.02**	0.00	**0.06**	**0.01**	0.00	**0.05**	0.00	**0.02**	0.01
Russian	0.01	**0.05**	**0.04**	**0.05**	**0.05**	0.02	**0.07**	**0.02**	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00
Sardinian	**0.02**	0.00	0.00	0.00	0.00	0.00	0.01	**0.01**	**0.12**	**0.05**	**0.03**	**0.10**	0.00	**0.05**	**0.05**
Saudi-Arabian	**0.05**	0.00	0.00	0.00	0.00	0.01	0.00	**0.03**	**0.17**	**0.09**	**0.06**	**0.15**	0.01	**0.07**	**0.09**
Yemeni	0.00	0.02	0.01	0.02	0.01	0.00	**0.03**	0.00	**0.04**	0.00	0.00	**0.03**	0.00	**0.01**	0.00
Spanish	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00

Table D.3:

	Saami	Sicilian	Swedish	Swiss	Daghestani	Syrian	Turkish	Udmurts	Komi-Ziryanes	Mari	Mongolian	Taiwanese	Burmese	Thai	Pamir
Adjar															
Arab (Uzb.)															
Armenian															
Azeri															
Dungan															
Iranian (Uzb.)															
Kazakh															
Kazbegi															
Korean															
Kurd															
Kyrgyz															
Lezgi															
Orcadian															
Ossetian															
Santi Romani															
Russian (Uzb.)															
Svan															
Turkmen															
Tuva															
Uighur															
Yagnobi															
Corsican															
Croatian															
Finnish															
German															
Hungarian															
Icelandic															
Irish															
Lithuanian															
Norwegian															
Saami	0.00														
Sicilian	**0.01**	0.00													
Swedish	0.01	**0.07**	0.00												
Swiss	0.00	0.01	0.01	0.00											
Daghestani	0.00	0.00	**0.03**	0.00	0.00										
Syrian	**0.04**	0.01	**0.08**	**0.06**	**0.03**	0.00									
Turkish	0.00	0.00	**0.03**	0.00	0.00	**0.03**	0.00								
Udmurts	0.00	**0.05**	0.00	0.00	0.01	**0.10**	0.01	0.00							
Komi-Ziryanes	0.00	0.00	0.02	0.00	0.00	0.03	0.00	0.01	0.00						
Mari	0.00	**0.02**	0.00	0.00	0.00	**0.08**	0.00	0.00	0.00	0.00					
Mongolian	**0.06**	**0.02**	**0.09**	**0.08**	**0.04**	0.00	**0.04**	**0.11**	**0.06**	**0.10**	0.00				
Taiwanese	**0.07**	**0.02**	**0.10**	**0.09**	**0.05**	0.01	**0.05**	**0.14**	**0.07**	**0.12**	0.00	0.00			
Burmese	**0.06**	**0.02**	**0.09**	**0.08**	**0.05**	0.01	**0.05**	**0.12**	**0.06**	**0.11**	0.00	0.00	0.00		
Thai	**0.06**	**0.02**	**0.10**	**0.08**	**0.04**	0.00	**0.04**	**0.13**	**0.06**	**0.11**	0.00	0.00	0.00	0.00	
Pamir	0.02	0.00	**0.05**	0.02	0.00	0.00	0.00	**0.05**	0.00	0.03	0.02	**0.03**	**0.02**	0.02	0.00
British	0.00	**0.04**	0.00	0.00	0.01	**0.05**	0.01	0.00	0.01	0.00	**0.06**	**0.07**	**0.07**	**0.07**	**0.03**
Iranian	0.00	**0.00**	**0.04**	0.00	0.00	**0.02**	0.00	0.01	0.00	0.00	**0.03**	**0.03**	**0.03**	**0.03**	0.00
Indian	**0.09**	**0.02**	**0.15**	**0.11**	**0.06**	0.00	**0.05**	**0.19**	**0.06**	**0.15**	0.00	0.00	0.00	0.00	**0.01**
Pakistani	0.01	0.00	**0.05**	0.01	0.00	0.01	0.00	**0.04**	0.00	0.02	**0.03**	**0.03**	**0.03**	**0.02**	0.00
Lebanese	**0.03**	0.00	**0.09**	**0.04**	**0.01**	0.00	**0.01**	**0.09**	0.01	**0.06**	0.01	0.01	0.01	0.01	0.00
Tajik	0.01	0.00	**0.04**	0.01	0.00	0.01	0.00	**0.03**	0.00	0.02	0.03	**0.03**	**0.03**	**0.03**	0.00
Tatar	0.00	0.01	0.01	0.00	0.00	**0.05**	0.00	0.00	0.00	0.00	**0.07**	**0.08**	**0.07**	**0.07**	0.02
Uzbek	0.01	0.00	**0.06**	0.01	0.00	0.01	0.00	**0.03**	0.00	0.01	**0.02**	**0.03**	0.03	**0.02**	0.00
Bashkir	**0.02**	0.00	**0.07**	**0.03**	0.01	0.00	0.01	**0.06**	0.00	**0.04**	0.02	0.02	**0.02**	0.01	0.00
Basque	**0.01**	0.00	**0.06**	0.01	0.00	0.01	0.00	**0.04**	0.00	0.02	**0.02**	**0.03**	**0.02**	**0.02**	0.00
Belgian	0.00	**0.02**	**0.01**	0.00	0.01	**0.04**	0.01	0.00	0.00	0.00	**0.05**	**0.05**	**0.05**	**0.05**	**0.02**
Catalan	0.00	0.01	**0.02**	0.00	0.00	**0.03**	0.00	0.01	0.00	0.00	**0.05**	**0.05**	**0.05**	**0.05**	0.01
Chinese	**0.26**	**0.05**	**0.37**	**0.32**	**0.19**	0.01	**0.19**	**0.46**	**0.21**	**0.39**	0.00	0.00	0.00	0.00	**0.07**
Cypriot	**0.04**	**0.00**	**0.12**	**0.04**	**0.02**	0.00	**0.01**	**0.10**	0.01	**0.06**	0.01	**0.01**	0.01	**0.01**	0.00
Danish	0.00	**0.04**	0.00	0.00	0.01	**0.06**	**0.01**	0.00	0.01	0.00	**0.07**	**0.07**	**0.07**	**0.07**	**0.04**
French	0.00	**0.03**	**0.00**	0.00	0.01	**0.04**	0.01	0.00	0.01	0.00	**0.06**	**0.06**	**0.06**	**0.05**	**0.03**
Greek	0.02	0.00	**0.06**	0.02	0.00	0.01	0.00	**0.05**	0.00	0.03	0.02	**0.02**	**0.02**	**0.02**	0.00
Italian	0.00	0.00	**0.05**	0.00	0.00	**0.02**	0.00	**0.03**	0.00	0.01	**0.03**	**0.03**	**0.03**	**0.02**	0.00
Mordvinian	0.00	**0.06**	0.00	0.00	**0.02**	**0.08**	0.02	0.00	0.01	0.00	**0.09**	**0.10**	**0.09**	**0.10**	**0.04**
Portuguese	0.00	0.00	**0.04**	0.00	0.00	**0.02**	0.00	0.02	0.00	0.01	**0.03**	**0.04**	**0.03**	**0.03**	0.00
Russian	0.01	**0.05**	0.00	0.00	**0.02**	**0.07**	**0.02**	0.00	**0.02**	0.00	**0.08**	**0.08**	**0.08**	**0.08**	**0.04**
Sardinian	**0.02**	0.00	**0.08**	0.03	0.01	0.00	0.01	**0.07**	0.00	**0.04**	0.02	**0.02**	0.02	**0.01**	0.00
Saudi-Arabian	**0.05**	**0.01**	**0.11**	**0.07**	**0.03**	0.00	**0.03**	**0.13**	**0.03**	**0.09**	0.01	0.01	0.01	0.00	0.00
Yemeni	0.00	**0.01**	**0.03**	0.00	0.00	**0.03**	0.00	0.01	0.00	0.00	**0.04**	**0.04**	**0.04**	**0.04**	**0.01**
Spanish	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.13	0.18	0.15	0.08	0.00

Table D.4:

	British	Iranian	Indian	Pakistani	Lebanese	Tajik	Tatar	Uzbek	Bashkir	Basque	Belgian	Catalan	Chinese	Cypriot	Danish
Adjar															
Arab (Uzb.)															
Armenian															
Azeri															
Dungan															
Iranian (Uzb.)															
Kazakh															
Kazbegi															
Korean															
Kurd															
Kyrgyz															
Lezgi															
Orcadian															
Ossetian															
Santi Romani															
Russian (Uzb.)															
Svan															
Turkmen															
Tuva															
Uighur															
Yagnobi															
Corsican															
Croatian															
Finnish															
German															
Hungarian															
Icelandic															
Irish															
Lithuanian															
Norwegian															
Saami															
Sicilian															
Swedish															
Swiss															
Daghestani															
Syrian															
Turkish															
Udmurts															
Komi-Ziryanes															
Mari															
Mongolian															
Taiwanese															
Burmese															
Thai															
Pamir															
British	0.00														
Iranian	**0.01**	0.00													
Indian	**0.09**	**0.03**	0.00												
Pakistani	**0.03**	0.00	**0.02**	0.00											
Lebanese	**0.05**	**0.01**	0.00	0.00	0.00										
Tajik	**0.02**	0.00	**0.03**	0.00	0.00	0.00									
Tatar	0.00	0.00	**0.10**	0.01	**0.04**	0.01	0.00								
Uzbek	**0.03**	0.00	**0.02**	0.00	0.01	0.00	0.01	0.00							
Bashkir	**0.04**	0.01	0.01	0.00	0.00	0.00	**0.02**	0.00	0.00						
Basque	**0.03**	0.00	**0.02**	0.00	0.00	0.00	0.01	0.00	0.00	0.00					
Belgian	0.00	**0.01**	**0.05**	**0.02**	**0.03**	**0.02**	0.00	**0.02**	**0.03**	**0.02**	0.00				
Catalan	0.01	0.00	**0.06**	0.00	**0.02**	0.00	0.00	0.00	**0.01**	0.01	0.00	0.00			
Chinese	**0.23**	**0.09**	**0.00**	**0.09**	**0.03**	**0.11**	**0.29**	**0.07**	**0.05**	**0.08**	**0.12**	**0.20**	0.00		
Cypriot	**0.07**	**0.01**	**0.01**	0.00	0.00	0.00	**0.04**	**0.01**	0.00	0.00	**0.04**	**0.02**	**0.02**	0.00	
Danish	0.00	**0.02**	**0.10**	**0.03**	**0.06**	**0.03**	0.00	**0.03**	**0.04**	**0.03**	0.00	0.01	**0.25**	**0.08**	0.00
French	0.00	**0.01**	**0.06**	**0.02**	**0.04**	**0.02**	0.00	**0.02**	**0.03**	**0.02**	0.00	0.00	**0.10**	**0.04**	0.00
Greek	**0.04**	0.01	**0.01**	0.00	0.00	0.00	**0.02**	0.00	0.00	0.00	**0.02**	0.01	**0.06**	0.00	**0.04**
Italian	**0.02**	0.00	**0.02**	0.00	**0.01**	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	**0.06**	**0.01**	**0.03**
Mordvinian	0.00	**0.02**	**0.15**	**0.04**	**0.08**	**0.03**	0.00	**0.04**	**0.06**	**0.04**	0.00	0.01	**0.38**	**0.10**	0.00
Portuguese	**0.02**	0.00	**0.04**	0.00	**0.01**	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	**0.12**	**0.01**	**0.02**
Russian	0.00	**0.02**	**0.12**	**0.04**	**0.07**	**0.03**	0.00	**0.04**	**0.05**	**0.04**	0.00	**0.01**	**0.30**	**0.09**	0.00
Sardinian	**0.04**	**0.01**	**0.01**	0.00	0.00	0.00	**0.03**	0.00	0.00	0.00	**0.03**	**0.01**	**0.05**	0.00	**0.05**
Saudi-Arabian	**0.06**	**0.02**	0.00	0.01	0.00	0.01	**0.06**	**0.01**	0.00	0.01	**0.04**	**0.04**	**0.01**	0.00	**0.07**
Yemeni	**0.01**	0.00	**0.04**	0.01	**0.02**	0.00	0.00	**0.00**	**0.01**	0.01	**0.00**	0.00	**0.10**	**0.02**	**0.01**
Spanish	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00

Table D.5:

	French	Greek	Italian	Mordvinian	Portuguese	Russian	Sardinian	Saudi-Arabian	Yemeni	Spanish
Adjar										
Arab (Uzb.)										
Armenian										
Azeri										
Dungan										
Iranian (Uzb.)										
Kazakh										
Kazbegi										
Korean										
Kurd										
Kyrgyz										
Lezgi										
Orcadian										
Ossetian										
Santi Romani										
Russian (Uzb.)										
Svan										
Turkmen										
Tuva										
Uighur										
Yagnobi										
Corsican										
Croatian										
Finnish										
German										
Hungarian										
Icelandic										
Irish										
Lithuanian										
Norwegian										
Saami										
Sicilian										
Swedish										
Swiss										
Daghestani										
Syrian										
Turkish										
Udmurts										
Komi-Ziryanes										
Mari										
Mongolian										
Taiwanese										
Burmese										
Thai										
Pamir										
British										
Iranian										
Indian										
Pakistani										
Lebanese										
Tajik										
Tatar										
Uzbek										
Bashkir										
Basque										
Belgian										
Catalan										
Chinese										
Cypriot										
Danish										
French	0.00									
Greek	**0.03**	0.00								
Italian	**0.02**	0.00	0.00							
Mordvinian	0.00	**0.05**	**0.04**	0.00						
Portuguese	**0.01**	0.00	0.00	**0.03**	0.00					
Russian	0.00	**0.05**	**0.04**	0.00	**0.03**	0.00				
Sardinian	**0.03**	0.00	**0.00**	**0.07**	**0.00**	**0.06**	0.00			
Saudi-Arabian	**0.05**	0.00	**0.01**	**0.10**	**0.02**	**0.08**	0.00	0.00		
Yemeni	**0.01**	**0.01**	0.00	**0.01**	0.00	**0.02**	**0.01**	**0.03**	0.00	
Spanish	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Appendix E

Mitochondrial DNA F$_{ST}$ tables

Table E.1: **Mitochondrial DNA F$_{ST}$ 1:3** Significant values are in bold

	Orkney2	British2	Tatar2	Adjar1	Armenian1	Kazbegi1	Lebanese1	Azeri1	Iranian1	Iranian2	Iranian3	Iranian4	Kurd1	Tatar1
Orkney2	0.000													
British2	0.003	0.000												
Tatar2	**0.012**	0.008	0.000											
Adjar1	**0.028**	**0.025**	**0.042**	0.000										
Armenian1	**0.008**	**0.005**	**0.014**	**0.018**	0.000									
Kazbegi1	**0.014**	**0.008**	**0.020**	**0.021**	**0.007**	0.000								
Lebanese1	0.003	0.001	0.005	**0.029**	0.003	**0.010**	0.000							
Azeri1	**0.020**	**0.016**	**0.033**	**0.013**	**0.006**	**0.011**	**0.019**	0.000						
Iranian1	**0.016**	**0.014**	**0.030**	**0.016**	**0.005**	**0.010**	**0.014**	0.002	0.000					
Iranian2	**0.014**	**0.010**	**0.026**	**0.013**	0.003	0.007	**0.013**	0.000	0.000	0.000				
Iranian3	0.018	**0.018**	**0.032**	**0.016**	0.003	0.012	0.013	0.000	0.000	0.000	0.000			
Iranian4	0.003	0.001	0.013	0.006	0.000	0.007	0.003	**0.006**	**0.007**	0.002	0.008	0.000		
Kurd1	**0.018**	**0.013**	**0.029**	0.009	**0.010**	**0.014**	**0.017**	**0.008**	**0.011**	**0.008**	**0.011**	0.003	0.000	
Tatar1	**0.023**	**0.018**	**0.037**	**0.013**	**0.008**	**0.011**	**0.019**	0.000	0.001	**0.003**	0.001	0.006	**0.011**	0.000
Saami1	**0.054**	**0.050**	**0.065**	**0.053**	**0.042**	**0.044**	**0.053**	**0.038**	**0.041**	**0.039**	**0.041**	**0.045**	**0.048**	**0.041**
Russian1	**0.005**	0.000	0.011	**0.021**	0.003	**0.007**	0.001	**0.012**	**0.008**	**0.008**	0.011	0.001	**0.011**	0.008
Arab1	**0.022**	**0.017**	**0.030**	**0.015**	**0.007**	**0.011**	**0.019**	**0.001**	0.003	**0.001**	0.001	**0.009**	**0.012**	0.000
Uzbek5	0.005	0.000	0.008	0.013	0.002	0.007	0.002	**0.010**	**0.009**	0.005	**0.010**	0.000	0.010	**0.013**
Tadjik1_2	**0.018**	**0.013**	**0.029**	**0.017**	**0.005**	**0.009**	**0.015**	0.002	**0.004**	0.002	0.003	**0.007**	**0.011**	0.001
Tadjik3	**0.028**	**0.020**	**0.042**	**0.022**	**0.013**	0.015	0.020	**0.006**	0.004	0.003	**0.006**	**0.014**	0.017	0.000
Romani1	0.013	0.008	0.015	**0.048**	**0.021**	**0.027**	0.010	**0.040**	**0.034**	**0.031**	**0.035**	0.015	**0.033**	**0.043**
Lezgi1	0.005	0.004	0.007	**0.032**	0.006	**0.016**	0.000	**0.021**	**0.018**	**0.017**	0.018	0.005	**0.020**	**0.025**
Ossetian1	**0.022**	**0.021**	**0.037**	**0.014**	**0.011**	**0.013**	**0.024**	**0.004**	**0.006**	0.001	0.004	0.008	**0.015**	**0.007**
Svan1	**0.029**	**0.026**	**0.044**	0.012	**0.016**	**0.021**	**0.029**	**0.009**	**0.011**	**0.008**	0.010	**0.013**	**0.017**	**0.009**
Ishkashim1	**0.020**	**0.015**	**0.026**	**0.028**	**0.015**	**0.020**	**0.018**	**0.016**	**0.019**	**0.016**	**0.021**	0.011	**0.021**	**0.014**
Shugnan1	**0.016**	**0.011**	**0.025**	**0.017**	**0.008**	**0.013**	**0.012**	**0.009**	**0.009**	0.007	0.010	0.005	**0.013**	**0.009**
Bartang1	**0.051**	**0.048**	**0.066**	**0.047**	**0.037**	**0.043**	**0.051**	**0.028**	**0.033**	**0.031**	**0.033**	**0.040**	**0.042**	**0.035**
Uzbek3	**0.016**	**0.013**	**0.029**	**0.015**	**0.005**	**0.011**	**0.013**	0.002	0.002	0.000	0.001	0.003	**0.011**	0.000
Uzbek7	0.005	0.001	**0.012**	**0.019**	0.001	0.005	0.002	0.006	**0.005**	0.002	0.006	0.000	0.009	**0.006**
Uzbek4	**0.013**	**0.009**	**0.024**	**0.015**	0.004	0.005	**0.011**	**0.002**	0.002	0.000	0.002	0.003	**0.009**	0.002
Uzbek2	**0.011**	**0.007**	**0.019**	**0.013**	0.001	**0.006**	**0.007**	0.002	0.003	0.000	0.000	0.001	**0.009**	**0.005**
Uzbek6	**0.015**	**0.010**	**0.024**	**0.012**	**0.005**	**0.009**	**0.014**	0.001	0.003	0.001	0.003	0.004	**0.009**	0.004
Uzbek1	**0.010**	**0.006**	**0.019**	**0.013**	0.002	**0.007**	**0.008**	0.003	**0.003**	0.001	0.003	0.000	0.006	0.004
Yagnobi1	**0.035**	**0.030**	**0.048**	**0.035**	**0.024**	**0.025**	**0.033**	**0.022**	**0.020**	**0.016**	0.012	**0.024**	**0.030**	**0.022**
Uighur2	**0.015**	**0.010**	**0.024**	**0.015**	0.004	**0.007**	**0.012**	0.001	0.002	0.000	0.001	0.004	**0.009**	0.001
Kazakh1	**0.013**	**0.011**	**0.023**	**0.014**	**0.004**	**0.008**	**0.012**	0.002	**0.003**	0.000	0.000	0.004	**0.009**	**0.004**
Dungan1	**0.024**	**0.019**	**0.036**	**0.017**	**0.011**	**0.014**	**0.023**	0.002	**0.006**	0.004	0.004	**0.012**	**0.010**	**0.006**
Tuva1	**0.048**	**0.044**	**0.061**	**0.043**	**0.034**	**0.039**	**0.047**	**0.028**	**0.030**	**0.028**	**0.029**	**0.036**	**0.039**	**0.030**
Karakalpak1	**0.022**	**0.018**	**0.030**	**0.017**	**0.009**	**0.012**	**0.020**	0.002	**0.004**	0.002	**0.002**	**0.010**	**0.010**	0.003
Kyrgyz1	**0.025**	**0.020**	**0.036**	**0.018**	**0.011**	**0.013**	**0.023**	**0.003**	**0.007**	0.004	0.005	**0.011**	**0.012**	0.004
Korean2	**0.022**	**0.019**	**0.035**	**0.017**	**0.009**	**0.013**	**0.021**	0.002	**0.004**	0.002	0.002	**0.010**	**0.012**	**0.003**

Table E.2: **Mitochondrial DNA F_{ST} 2:3**

	Saami1	Russian1	Arab1	Uzbek5	Tadjik1_2	Tadjik3	Romani1	Lezgi1	Ossetian1	Svan1	Ishkashim1	Shugnan1	Bartang1	Uzbek3
Orkney2														
British2														
Tatar2														
Adjar1														
Armenian1														
Kazbegi1														
Lebanese1														
Azeri1														
Iranian1														
Iranian2														
Iranian3														
Iranian4														
Kurd1														
Tatar1														
Saami1	0.000													
Russian1	**0.040**	0.000												
Arab1	**0.040**	**0.012**	0.000											
Uzbek5	**0.046**	0.002	**0.011**	0.000										
Tadjik1_2	**0.040**	**0.009**	0.003	**0.008**	0.000									
Tadjik3	**0.048**	0.010	0.007	**0.017**	0.004	0.000								
Romani1	**0.073**	**0.014**	**0.041**	0.012	**0.034**	**0.049**	0.000							
Lezgi1	**0.057**	0.007	**0.024**	0.004	**0.019**	**0.031**	0.011	0.000						
Ossetian1	**0.044**	**0.016**	**0.004**	**0.012**	**0.007**	**0.010**	**0.044**	**0.026**	0.000					
Svan1	**0.049**	**0.021**	**0.009**	**0.016**	**0.012**	**0.016**	**0.050**	**0.032**	0.007	0.000				
Ishkashim1	**0.052**	**0.014**	**0.021**	0.014	**0.016**	0.023	**0.032**	**0.019**	**0.024**	**0.028**	0.000			
Shugnan1	**0.047**	**0.009**	**0.010**	0.008	**0.009**	0.006	**0.029**	**0.017**	**0.014**	**0.018**	0.007	0.000		
Bartang1	**0.071**	**0.043**	**0.033**	**0.042**	**0.031**	**0.032**	**0.073**	**0.055**	**0.036**	**0.042**	**0.052**	**0.022**	0.000	
Uzbek3	**0.040**	**0.007**	0.001	**0.009**	0.003	0.001	**0.034**	**0.019**	0.000	**0.012**	**0.016**	**0.008**	**0.033**	0.000
Uzbek7	**0.043**	0.000	**0.008**	0.000	**0.004**	0.008	**0.015**	0.005	**0.011**	**0.016**	**0.010**	0.005	**0.036**	0.004
Uzbek4	**0.040**	0.005	0.002	0.005	0.002	0.005	**0.029**	**0.016**	0.004	**0.011**	**0.014**	**0.007**	**0.033**	0.001
Uzbek2	**0.040**	0.004	**0.004**	0.001	**0.003**	**0.007**	**0.023**	**0.011**	**0.006**	**0.012**	**0.014**	**0.005**	**0.030**	0.000
Uzbek6	**0.041**	**0.007**	0.001	0.004	**0.004**	**0.010**	**0.029**	**0.017**	**0.007**	**0.011**	**0.016**	**0.009**	**0.032**	0.003
Uzbek1	**0.040**	**0.005**	0.003	0.000	**0.003**	**0.010**	**0.022**	**0.011**	**0.006**	**0.011**	**0.014**	**0.007**	**0.034**	**0.003**
Yagnobi1	**0.060**	**0.027**	**0.023**	**0.027**	**0.021**	**0.024**	**0.041**	**0.038**	**0.023**	**0.031**	**0.035**	**0.026**	**0.053**	**0.017**
Uighur2	**0.039**	**0.007**	0.000	0.004	0.002	0.004	**0.030**	**0.016**	0.003	**0.011**	**0.016**	**0.007**	**0.032**	0.000
Kazakh1	**0.039**	**0.007**	0.001	0.002	**0.003**	**0.008**	**0.028**	**0.015**	**0.004**	**0.011**	**0.016**	**0.008**	**0.032**	0.002
Dungan1	**0.042**	**0.015**	**0.005**	**0.012**	**0.007**	**0.010**	**0.043**	**0.027**	**0.007**	**0.013**	**0.023**	**0.014**	**0.035**	**0.006**
Tuva1	**0.066**	**0.040**	**0.023**	**0.038**	**0.031**	**0.036**	**0.068**	**0.051**	**0.031**	**0.038**	**0.048**	**0.038**	**0.060**	**0.029**
Karakalpak1	**0.041**	**0.012**	0.002	**0.009**	**0.004**	0.005	**0.040**	**0.024**	0.003	**0.011**	**0.020**	**0.011**	**0.033**	0.003
Kyrgyz1	**0.042**	**0.016**	0.000	**0.011**	**0.007**	**0.011**	**0.043**	**0.027**	0.003	**0.014**	**0.024**	**0.013**	**0.035**	0.003
Korean2	**0.041**	**0.014**	0.002	**0.011**	**0.004**	**0.008**	**0.041**	**0.025**	0.005	**0.010**	**0.019**	**0.011**	**0.033**	**0.004**

Table E.3: **Mitochondrial DNA F$_{ST}$ 3:3**

	Uzbek7	Uzbek4	Uzbek2	Uzbek6	Uzbek1	Yagnobi1	Uighur2	Kazakh1	Dungan1	Tuva1	Karakalpak1	Kyrgyz1	Korean2
Orkney2													
British2													
Tatar2													
Adjar1													
Armenian1													
Kazbegi1													
Lebanese1													
Azeri1													
Iranian1													
Iranian2													
Iranian3													
Iranian4													
Kurd1													
Tatar1													
Saami1													
Russian1													
Arab1													
Uzbek5													
Tadjik1_2													
Tadjik3													
Romani1													
Lezgi1													
Ossetian1													
Svan1													
Ishkashim1													
Shugnan1													
Bartang1													
Uzbek3													
Uzbek7	0.000												
Uzbek4	0.002	0.000											
Uzbek2	0.001	0.000	0.000										
Uzbek6	0.003	0.003	0.002	0.000									
Uzbek1	0.000	0.000	0.001	0.000	0.000								
Yagnobi1	**0.023**	**0.021**	**0.018**	**0.020**	**0.020**	0.000							
Uighur2	0.001	0.000	0.001	0.000	0.000	**0.021**	0.000						
Kazakh1	**0.003**	0.001	0.001	0.000	0.000	**0.020**	0.000	0.000					
Dungan1	**0.010**	**0.004**	**0.006**	**0.004**	**0.006**	**0.025**	**0.004**	**0.004**	0.000				
Tuva1	**0.034**	**0.027**	**0.030**	**0.029**	**0.030**	**0.049**	**0.027**	**0.027**	**0.030**	0.000			
Karakalpak1	**0.007**	0.002	**0.004**	**0.004**	**0.003**	**0.020**	0.001	0.000	**0.005**	**0.028**	0.000		
Kyrgyz1	**0.011**	**0.004**	**0.007**	**0.004**	**0.004**	**0.025**	0.002	0.002	**0.005**	**0.027**	0.002	0.000	
Korean2	**0.009**	0.003	**0.005**	**0.004**	**0.003**	**0.023**	**0.002**	0.002	**0.005**	**0.029**	0.001	0.004	0.000

Appendix F

MC1R variant frequency pie charts

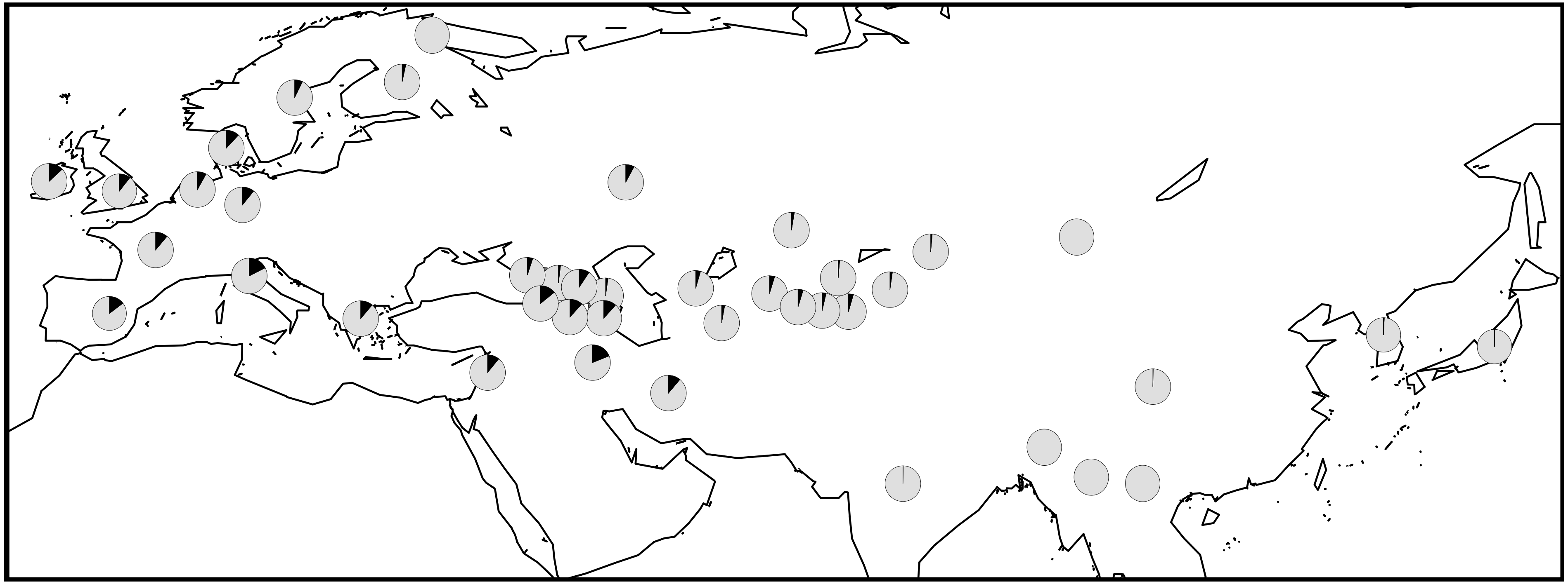


Figure F.1: **V60L frequencies**

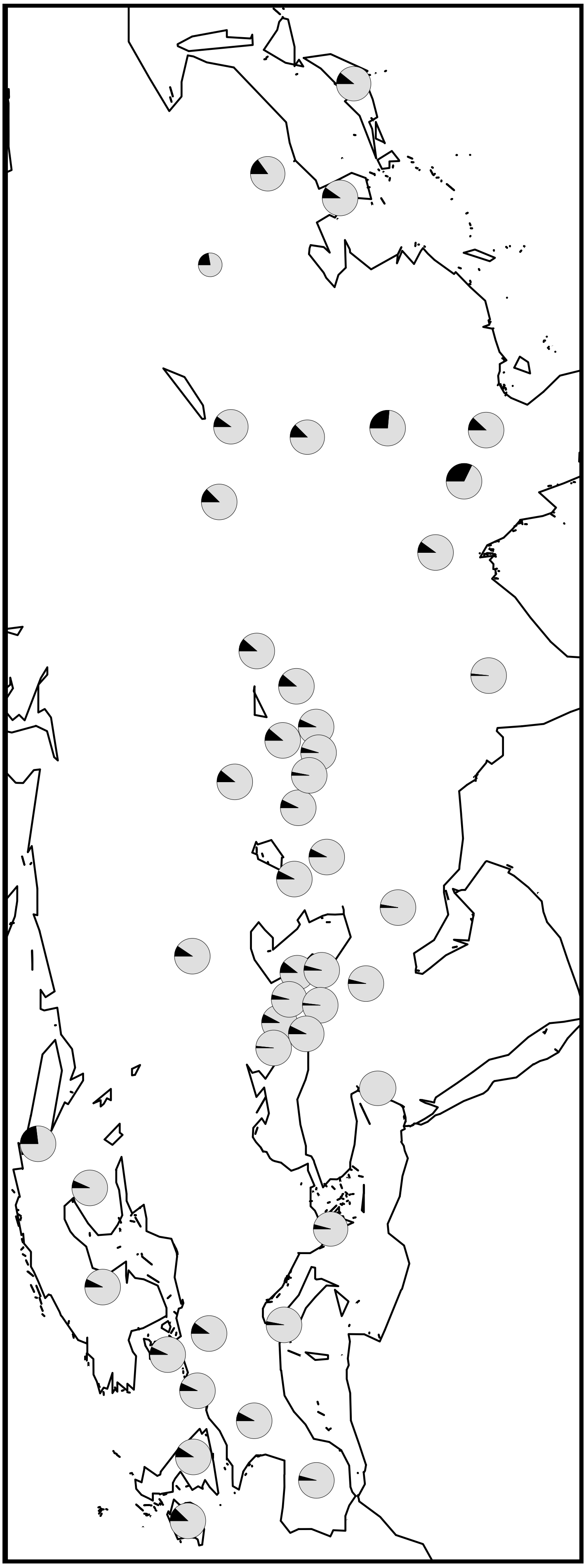


Figure F.2: **V92M frequencies**

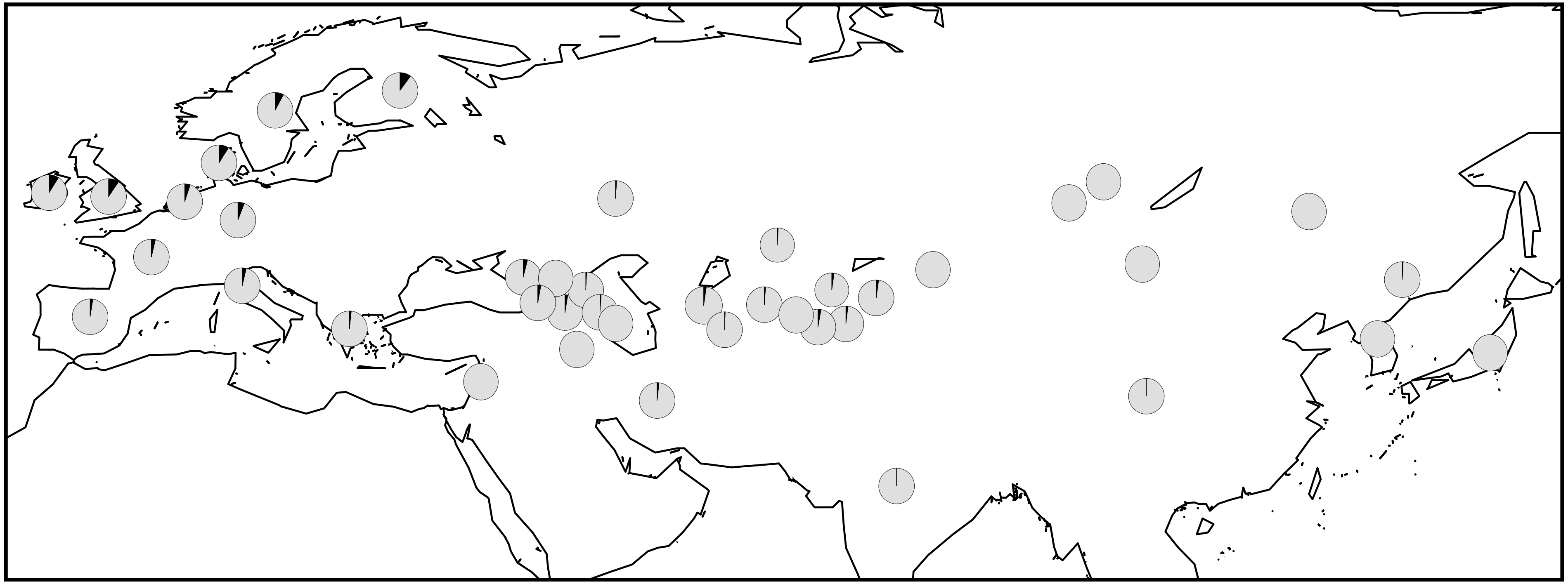


Figure F.3: **R151C frequencies**

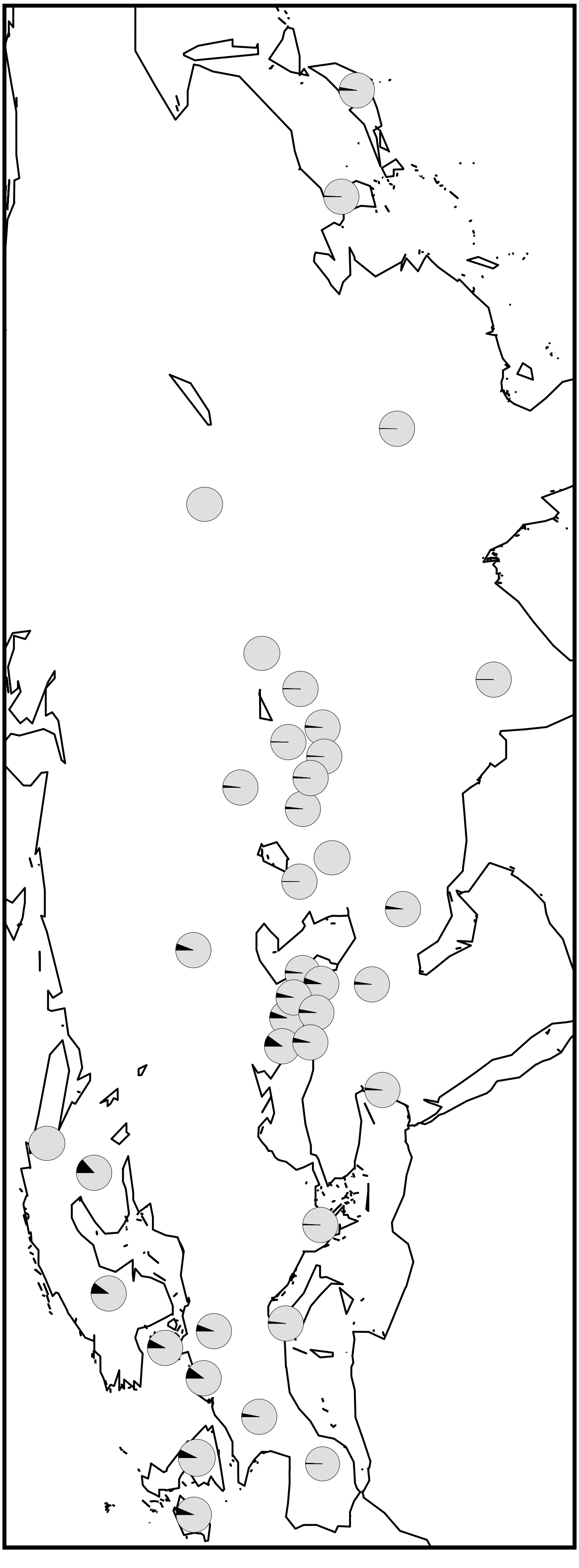


Figure F.4: **R160W frequencies**

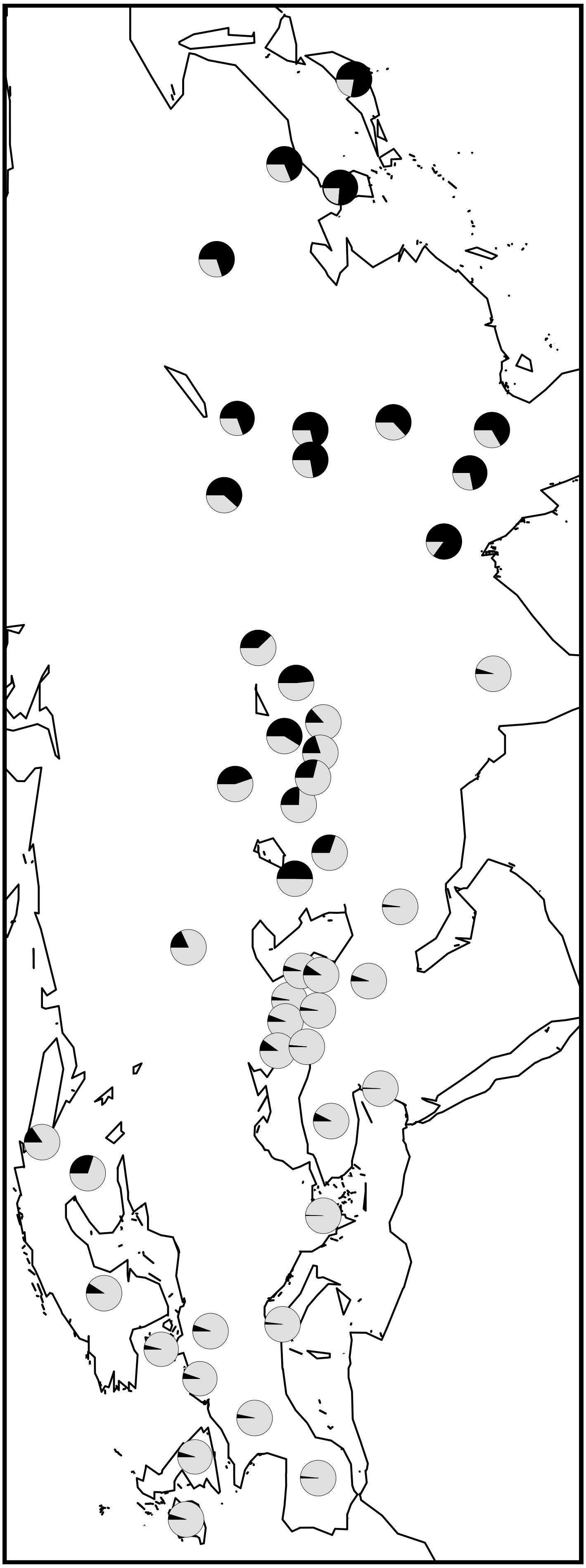


Figure F.5: **R163Q** frequencies

Appendix G

MC1R F$_{ST}$ tables

For the files on the following pages, all significant values are in bold font.

Table G.1: **V60L 1:3**

	Adjar	Arab (Uzb.)	Armenian	Azeri	British	Dungan	European (CEU)	Iranian (Uzb.)	Japanese	Jewish (Uzb.)	Kazakh	Kazbegi
Adjar	0.000											
Arab (Uzb.)	0.033	0.000										
Armenian	0.000	0.014	0.000									
Azeri	0.000	0.013	0.000	0.000								
British	0.000	**0.059**	0.006	0.004	0.000							
Dungan	**0.098**	0.006	**0.054**	**0.062**	**0.122**	0.000						
European (CEU)	0.017	0.000	0.004	0.003	**0.044**	**0.019**	0.000					
Iranian (Uzb.)	**0.060**	0.000	0.030	0.032	**0.092**	0.000	0.004	0.000				
Japanese	**0.181**	**0.043**	**0.099**	**0.124**	**0.193**	0.005	**0.050**	0.023	0.000			
Jewish (Uzb.)	0.000	0.035	0.000	0.000	0.000	0.153	0.008	0.074	**0.341**	0.000		
Kazakh	**0.121**	0.006	**0.066**	**0.075**	**0.157**	0.000	**0.022**	0.000	0.007	**0.166**	0.000	
Kazbegi	0.000	0.001	0.000	0.000	0.017	0.040	0.000	0.014	**0.096**	0.000	**0.047**	0.000
Korean	**0.127**	0.024	**0.072**	**0.086**	**0.144**	0.000	**0.035**	0.012	0.000	**0.233**	0.000	**0.064**
Kurd	0.000	**0.075**	0.012	0.008	0.000	**0.152**	**0.057**	**0.115**	**0.240**	0.000	**0.194**	0.025
Kyrgyz	**0.127**	0.018	**0.071**	**0.084**	**0.148**	0.000	**0.031**	0.008	0.000	**0.218**	0.000	**0.060**
Lebanese	0.000	0.009	0.000	0.000	0.008	**0.053**	0.000	0.026	**0.112**	0.000	**0.065**	0.000
Lezgi	**0.082**	0.003	**0.046**	**0.051**	**0.105**	0.000	0.016	0.000	0.005	**0.130**	0.000	0.033
Orcadian	0.038	0.000	0.017	0.017	**0.067**	0.005	0.000	0.000	**0.037**	0.041	0.006	0.003
Ossetian	**0.071**	0.001	0.041	0.045	**0.092**	0.000	0.013	0.000	0.000	0.113	0.000	0.029
Santi Romani	**0.110**	0.028	**0.069**	**0.078**	**0.125**	0.003	0.037	0.016	0.000	**0.209**	0.005	0.061
Russian (Uzb.)	0.029	0.000	0.011	0.010	**0.058**	0.011	0.000	0.000	**0.042**	0.026	0.013	0.000
Spanish	0.000	**0.022**	0.000	0.000	0.002	**0.056**	**0.012**	**0.037**	**0.086**	0.000	**0.065**	0.000
Svan	0.032	0.000	0.014	0.013	0.056	0.000	0.000	0.000	**0.044**	0.037	0.000	0.001
Turkmen	**0.083**	0.000	**0.045**	**0.050**	**0.110**	0.000	0.013	0.000	0.011	0.122	0.000	0.030
Tuva	**0.160**	0.046	**0.092**	**0.112**	**0.168**	0.013	**0.050**	0.028	0.000	**0.333**	0.012	**0.090**
Uighur	0.109	0.012	**0.061**	**0.071**	**0.130**	0.000	0.025	0.003	0.000	**0.182**	0.000	**0.049**
Yagnobi	0.034	0.000	0.015	0.015	0.059	0.000	0.000	0.000	0.041	0.042	0.000	0.002
Pamir	**0.061**	0.000	**0.030**	**0.032**	**0.094**	0.000	0.004	0.000	0.023	0.074	0.000	0.014
Karakalpak	**0.058**	0.000	**0.028**	**0.030**	**0.091**	0.001	0.003	0.000	**0.025**	0.068	0.002	0.012
Tajik	**0.084**	0.000	**0.044**	**0.048**	**0.118**	0.000	0.010	0.000	**0.015**	0.108	0.000	0.026
Tatar	0.013	0.000	0.002	0.001	0.038	0.021	0.000	0.004	**0.057**	0.005	0.025	0.000
Uzbek	**0.074**	0.000	**0.038**	**0.039**	**0.125**	0.002	0.005	0.000	**0.019**	0.076	0.003	0.016
South Asian	**0.412**	**0.127**	**0.233**	**0.302**	**0.414**	0.031	**0.121**	**0.070**	0.000	**0.626**	**0.027**	**0.248**
Iranian	0.000	0.011	0.000	0.000	0.010	**0.048**	0.002	0.026	**0.090**	0.000	**0.058**	0.000
Chinese	**0.298**	**0.078**	**0.162**	**0.209**	**0.307**	0.014	**0.081**	**0.042**	0.000	**0.493**	**0.015**	**0.166**
Italian	0.000	0.036	0.003	0.000	0.000	**0.070**	**0.025**	**0.052**	**0.099**	0.000	**0.080**	0.009

Table G.2: **V60L 2:3**

	Korean	Kurd	Kyrgyz	Lebanese	Lezgi	Orcadian	Ossetian	Santi Romani	Russian (Uzb.)	Spanish	Svan	Turkmen
Adjar												
Arab (Uzb.)												
Armenian												
Azeri												
British												
Dungan												
European (CEU)												
Iranian (Uzb.)												
Japanese												
Jewish (Uzb.)												
Kazakh												
Kazbegi												
Korean	0.000											
Kurd	**0.177**	0.000										
Kyrgyz	0.000	**0.183**	0.000									
Lebanese	**0.077**	0.014	**0.074**	0.000								
Lezgi		**0.129**	0.000	0.045	0.000							
Orcadian	0.021	**0.084**	0.016	0.012	0.003	0.000						
Ossetian	0.000	**0.111**	0.000	0.039	0.000	0.001	0.000					
Santi Romani	0.000	**0.149**	0.000	0.071	0.003	0.025	0.000	0.000				
Russian (Uzb.)	0.027	**0.074**	0.023	0.006	0.008	0.000	0.006	0.030	0.000			
Spanish	**0.070**	0.005	**0.068**	0.000	**0.050**	0.025	0.047	**0.070**	**0.020**	0.000		
Svan	0.020	**0.070**	0.012	0.009	0.000	0.000	0.000	0.028	0.000	0.022	0.000	
Turkmen	0.001	**0.137**	0.000	**0.043**	0.000	0.000	0.000	0.007	0.006	**0.049**	0.000	0.000
Tuva	0.000	**0.208**	0.005	**0.102**	0.016	**0.040**	0.015	0.000	**0.043**	**0.083**	0.054	0.018
Uighur	0.000	**0.161**	0.000	**0.062**	0.000	0.011	0.000	0.000	0.017	**0.061**	0.005	0.000
Yagnobi	0.018	**0.073**	0.011	0.011	0.000	0.000	0.000	0.026	0.000	0.024	0.000	0.000
Pamir	0.012	**0.117**	0.008	0.026	0.000	0.000	0.000	0.016	0.000	**0.037**	0.000	0.000
Karakalpak	0.013	**0.114**	0.009	0.024	0.000	0.000	0.000	0.018	0.000	**0.036**	0.000	0.000
Tajik	0.006	**0.147**	0.003	**0.040**	0.000	0.000	0.000	0.011	0.003	**0.048**	0.000	0.000
Tatar	**0.039**	**0.050**	0.035	0.000	0.017	0.000	0.015	0.040	0.000	0.010	0.000	0.015
Uzbek	0.012	**0.151**	0.009	**0.031**	0.000	0.000	0.000	0.016	0.000	**0.050**	0.000	0.000
South Asian	0.001	**0.498**	0.011	**0.276**	0.035	**0.105**	0.025	0.000	**0.110**	**0.165**	**0.141**	**0.046**
Iranian	**0.066**	0.016	**0.064**	0.000	0.041	0.014	0.037	0.063	0.008	0.000	0.011	**0.040**
Chinese	0.000	**0.378**	0.002	**0.189**	0.015	**0.065**	0.007	0.000	**0.071**	**0.123**	**0.082**	0.024
Italian	**0.084**	0.000	**0.082**	0.003	**0.065**	**0.039**	**0.062**	**0.084**	**0.034**	0.000	**0.036**	**0.064**

Table G.3: **V60L 3:3**

	Tuva	Uighur	Yagnobi	Pamir	Karakalpak	Tajik	Tatar	Uzbek	South Asian	Iranian	Chinese	Italian
Adjar												
Arab (Uzb.)												
Armenian												
Azeri												
British												
Dungan												
European (CEU)												
Iranian (Uzb.)												
Japanese												
Jewish (Uzb.)												
Kazakh												
Kazbegi												
Korean												
Kurd												
Kyrgyz												
Lebanese												
Lezgi												
Orcadian												
Ossetian												
Santi Romani												
Russian (Uzb.)												
Spanish												
Svan												
Turkmen												
Tuva	0.000											
Uighur	0.008	0.000										
Yagnobi	0.051	0.004	0.000									
Pamir	0.027	0.004	0.000	0.000								
Karakalpak	**0.028**	0.005	0.000	0.000	0.000							
Tajik	0.020	0.000	0.000	0.000	0.000	0.000						
Tatar	**0.057**	0.028	0.000	0.004	0.003	0.011	0.000					
Uzbek	0.022	0.006	0.000	0.000	0.000	0.000	0.005	0.000				
South Asian	0.000	0.017	**0.132**	**0.067**	**0.069**	**0.047**	**0.146**	**0.034**	0.000			
Iranian	**0.084**	**0.055**	0.012	**0.026**	**0.024**	**0.038**	0.001	**0.033**	**0.211**	0.000		
Chinese	0.000	0.005	**0.077**	**0.041**	**0.043**	**0.028**	**0.097**	**0.026**	0.000	**0.146**	0.000	
Italian	**0.096**	**0.076**	**0.038**	**0.053**	**0.051**	**0.064**	0.022	**0.072**	**0.170**	0.005	**0.132**	0.000

Table G.4: **V92M 1:3**

	Adjar	Arab (Uzb.)	Armenian	Azeri	British	Dai	Dungan	European (CEU)	Iranian (Uzb.)	Japanese	Jewish (Uzb.)	Kazakh	Kazbegi
Adjar	0.000												
Arab (Uzb.)	0.000	0.000											
Armenian	0.000	0.000	0.000										
Azeri	0.001	0.000	0.001	0.000									
British	0.021	0.008	**0.024**	0.000	0.000								
Dai	**0.280**	**0.287**	**0.356**	**0.241**	**0.191**	0.000							
Dungan	**0.054**	**0.041**	**0.064**	0.022	0.002	**0.123**	0.000						
European (CEU)	0.018	0.008	**0.020**	0.000	0.000	**0.235**	0.006	0.000					
Iranian (Uzb.)	0.000	0.000	0.000	0.000	0.006	**0.325**	**0.042**	0.007	0.000				
Japanese	0.042	**0.030**	**0.049**	0.014	0.000	**0.151**	0.000	0.002	0.030	0.000			
Jewish (Uzb.)	0.000	0.000	0.000	0.000	0.000	**0.142**	0.000	0.000	0.000	0.000	0.000		
Kazakh	**0.045**	**0.034**	**0.050**	0.019	0.003	**0.146**	0.000	0.007	**0.034**	0.000	0.000	0.000	
Kazbegi	0.001	0.000	0.001	0.000	0.000	**0.241**	0.022	0.000	0.000	0.014	0.000	0.019	0.000
Kurd	0.000	0.000	0.000	0.000	0.000	**0.228**	0.021	0.000	0.000	0.014	0.000	0.019	0.000
Kyrgyz	**0.053**	**0.041**	**0.064**	0.022	0.002	**0.124**	0.000	0.007	**0.042**	0.000	0.000	0.000	0.022
Lebanese	0.005	0.018	0.008	0.034	**0.059**	**0.362**	**0.093**	**0.043**	0.019	**0.074**	0.111	**0.070**	0.034
Lezgi	**0.056**	0.041	**0.068**	0.019	0.000	**0.118**	0.000	0.003	0.041	0.000	0.000	0.000	0.019
Malay	**0.212**	**0.213**	**0.253**	**0.185**	**0.151**	0.000	**0.103**	**0.183**	**0.237**	**0.123**	**0.126**	**0.123**	**0.185**
Orcadian	0.020	0.009	0.023	0.000	0.000	**0.207**	0.003	0.000	0.008	0.000	0.000	0.004	0.000
Ossetian	0.036	0.016	0.041	0.000	0.000	**0.139**	0.000	0.000	0.013	0.000	0.000	0.000	0.000
Santi Romani	0.000	0.000	0.000	0.000	0.009	**0.239**	0.038	0.008	0.000	0.029	0.000	0.033	0.000
Russian (Uzb.)	**0.050**	**0.039**	**0.059**	0.022	0.004	**0.130**	0.000	0.008	**0.040**	0.000	0.000	0.000	0.022
Spanish	0.000	0.000	0.000	0.000	0.008	**0.394**	**0.051**	0.009	0.000	**0.036**	0.000	**0.041**	0.000
Svan	0.000	0.000	0.000	0.000	0.015	**0.261**	**0.047**	0.013	0.000	0.036	0.000	**0.040**	0.000
Tibetan	0.063	0.038	0.074	0.011	0.000	**0.106**	0.000	0.000	0.035	0.000	0.000	0.000	0.011
Turkmen	0.025	0.013	0.029	0.000	0.000	**0.189**	0.000	0.000	0.012	0.000	0.000	0.001	0.000
Tuva	**0.071**	**0.058**	**0.087**	**0.035**	0.010	**0.095**	0.000	0.015	**0.061**	0.000	0.000	0.000	0.035
Wa	**0.070**	0.054	**0.087**	0.028	0.003	**0.098**	0.000	0.008	**0.055**	0.000	0.000	0.000	0.028
Yagnobi	0.000	0.000	0.000	0.000	0.000	**0.232**	0.027	0.001	0.000	0.019	0.000	0.024	0.000
Pamir	0.020	0.009	0.022	0.000	0.000	**0.223**	0.004	0.000	0.008	0.000	0.000	0.005	0.000
Karakalpak	0.023	0.013	0.026	0.001	0.000	**0.212**	0.002	0.000	0.012	0.000	0.000	0.003	0.001
Tajik	0.006	0.000	0.007	0.000	0.000	**0.296**	0.022	0.000	0.000	**0.014**	0.000	**0.020**	0.000
Tatar	**0.039**	**0.028**	**0.046**	0.012	0.000	**0.157**	0.000	0.000	**0.027**	0.000	0.000	0.000	0.012
Uzbek	0.019	0.010	**0.019**	0.001	0.000	**0.276**	0.006	0.000	0.009	0.002	0.000	0.007	0.001
Iranian	0.000	0.000	0.000	0.000	0.010	**0.333**	**0.047**	0.010	0.000	**0.035**	0.000	**0.038**	0.000
Korean	**0.062**	0.049	**0.076**	0.027	0.004	**0.108**	0.000	0.009	**0.050**	0.000	0.000	0.000	0.027
Italian	0.002	0.000	0.002	0.000	0.003	**0.390**	**0.042**	0.005	0.000	**0.029**	0.000	**0.035**	0.000
Uighur	**0.053**	**0.041**	**0.061**	0.024	0.005	**0.125**	0.000	0.009	**0.042**	0.000	0.000	0.000	0.024
Chinese	**0.149**	**0.141**	**0.160**	**0.123**	**0.097**	0.000	**0.061**	**0.110**	**0.147**	**0.073**	**0.090**	**0.070**	**0.123**
South Asian	0.000	0.000	0.000	0.008	**0.046**	**0.579**	**0.127**	**0.039**	0.002	**0.096**	0.000	**0.095**	0.008

Table G.5: **V92M 2:3**

	Kurd	Kyrgyz	Lebanese	Lezgi	Malay	Orcadian	Ossetian	Santi Romani	Russian (Uzb.)	Spanish	Svan	Tibetan	Turkmen
Adjar													
Arab (Uzb.)													
Armenian													
Azeri													
British													
Dai													
Dungan													
European (CEU)													
Iranian (Uzb.)													
Japanese													
Jewish (Uzb.)													
Kazakh													
Kazbegi													
Kurd	0.000												
Kyrgyz	0.021	0.000											
Lebanese	0.036	**0.091**	0.000										
Lezgi	0.019	0.000	**0.106**	0.000									
Malay	**0.177**	**0.103**	**0.257**	**0.099**	0.000								
Orcadian	0.000	0.004	**0.051**	0.000	**0.162**	0.000							
Ossetian	0.000	0.000	**0.102**	0.000	**0.117**	0.000	0.000						
Santi Romani	0.000	0.039	0.015	0.038	**0.190**	0.009	0.018	0.000					
Russian (Uzb.)	0.021	0.000	**0.082**	0.000	**0.108**	0.005	0.000	0.037	0.000				
Spanish	0.000	**0.051**	0.016	**0.049**	**0.295**	0.010	0.015	0.000	**0.049**	0.000			
Svan	0.000	0.047	0.009	0.048	**0.202**	0.015	0.027	0.000	0.045	0.000	0.000		
Tibetan	0.011	0.000	**0.148**	0.000	**0.092**	0.000	0.000	0.038	0.000	0.039	0.051	0.000	
Turkmen	0.000	0.000	**0.059**	0.000	**0.149**	0.000	0.000	0.013	0.001	0.014	0.020	0.000	0.000
Tuva	0.033	0.000	**0.118**	0.000	**0.081**	0.011	0.000	0.052	0.000	**0.074**	**0.062**	0.000	0.006
Wa	0.027	0.000	**0.131**	0.000	**0.084**	0.004	0.000	0.049	0.000	**0.065**	0.060	0.000	0.000
Yagnobi	0.000	0.028	0.029	0.026	**0.182**	0.001	0.004	0.000	0.027	0.000	0.000	0.020	0.004
Pamir	0.000	0.005	**0.046**	0.001	**0.174**	0.000	0.000	0.010	0.006	0.010	0.015	0.000	0.000
Karakalpak	0.001	0.002	**0.050**	0.000	**0.167**	0.000	0.000	0.013	0.003	0.014	0.019	0.000	0.000
Tajik	0.000	**0.023**	**0.029**	0.019	**0.223**	0.000	0.000	0.000	0.023	0.000	0.003	0.009	0.000
Tatar	0.012	0.000	**0.071**	0.000	**0.127**	0.000	0.000	0.027	0.000	**0.033**	0.034	0.000	0.000
Uzbek	0.001	0.006	0.036	0.002	**0.239**	0.000	0.000	0.011	0.008	**0.011**	0.015	0.000	0.000
Iranian	0.000	**0.048**	0.016	**0.048**	**0.241**	0.011	0.019	0.000	**0.045**	0.000	0.000	0.044	0.016
Korean	0.026	0.000	**0.108**	0.000	**0.091**	0.006	0.000	0.045	0.000	**0.061**	**0.054**	0.000	0.002
Italian	0.000	**0.043**	0.019	**0.038**	**0.304**	0.005	0.007	0.000	**0.042**	0.000	0.000	0.027	0.009
Uighur	0.023	0.000	**0.085**	0.000	**0.105**	0.006	0.000	0.039	0.000	**0.051**	0.046	0.000	0.002
Chinese	**0.121**	**0.061**	**0.173**	**0.061**	0.001	**0.101**	**0.078**	**0.137**	**0.063**	**0.169**	**0.144**	**0.059**	**0.094**
South Asian	0.005	**0.126**	0.005	**0.130**	**0.454**	**0.046**	**0.072**	0.000	**0.117**	0.004	0.000	**0.124**	**0.057**

Table G.6: **V92M 3:3**

	Tuva	Wa	Yagnobi	Pamir	Karakalpak	Tajik	Tatar	Uzbek	Iranian	Korean	Italian	Uighur	Chinese	South Asian
Adjar														
Arab (Uzb.)														
Armenian														
Azeri														
British														
Dai														
Dungan														
European (CEU)														
Iranian (Uzb.)														
Japanese														
Jewish (Uzb.)														
Kazakh														
Kazbegi														
Kurd														
Kyrgyz														
Lebanese														
Lezgi														
Malay														
Orcadian														
Ossetian														
Santi Romani														
Russian (Uzb.)														
Spanish														
Svan														
Tibetan														
Turkmen														
Tuva	0.000													
Wa	0.000	0.000												
Yagnobi	0.041	0.035	0.000											
Pamir	0.013	0.005	0.002	0.000										
Karakalpak	0.009	0.002	0.005	0.000	0.000									
Tajik	**0.037**	0.028	0.000	0.000	0.001	0.000								
Tatar	0.000	0.000	0.017	0.000	0.000	0.012	0.000							
Uzbek	0.015	0.006	0.004	0.000	0.000	0.002	0.000	0.000						
Iranian	**0.067**	**0.063**	0.000	0.011	0.015	0.000	**0.032**	0.012	0.000					
Korean	0.000	0.000	0.033	0.007	0.004	**0.028**	0.000	0.008	**0.057**	0.000				
Italian	**0.063**	**0.052**	0.000	0.006	0.009	0.000	**0.026**	**0.008**	0.000	**0.051**	0.000			
Uighur	0.000	0.000	0.029	0.007	0.004	0.025	0.000	**0.009**	**0.047**	0.000	**0.044**	0.000		
Chinese	**0.047**	**0.051**	**0.127**	**0.106**	**0.101**	**0.134**	**0.076**	**0.141**	**0.150**	**0.054**	**0.173**	**0.061**	0.000	
South Asian	**0.170**	**0.161**	0.000	**0.044**	**0.051**	0.017	**0.090**	**0.031**	0.000	**0.148**	**0.008**	**0.122**	**0.254**	0.000

Table G.7: **R151C 1:3**

	Adjar	Arab (Uzb.)	Armenian	Azeri	Dungan	European (CEU)	Iranian (Uzb.)	Japanese	Jewish (Uzb.)	Kazakh	Kazbegi
Adjar	0.000										
Arab (Uzb.)	0.000	0.000									
Armenian	0.000	0.000	0.000								
Azeri	0.021	0.000	0.013	0.000							
Dungan	0.000	0.000	0.000	0.012	0.000						
European (CEU)	**0.038**	**0.071**	**0.053**	**0.083**	**0.053**	0.000					
Iranian (Uzb.)	0.000	0.000	0.000	0.000	0.000	**0.083**	0.000				
Japanese	0.020	0.000	0.012	0.000	0.011	**0.081**	0.000	0.000			
Jewish (Uzb.)	0.000	0.000	0.000	0.000	0.000	0.045	0.000	0.000	0.000		
Kazakh	0.003	0.000	0.003	0.000	0.001	**0.103**	0.000	0.000	0.000	0.000	
Kazbegi	0.000	0.000	0.000	0.000	0.000	**0.066**	0.000	0.000	0.000	0.000	0.000
Korean	0.023	0.000	0.014	0.000	0.013	**0.085**	0.001	0.000	0.000	0.000	0.001
Kurd	0.016	0.000	0.009	0.000	0.008	**0.076**	0.000	0.000	0.000	0.000	0.000
Kyrgyz	0.000	0.000	0.000	0.010	0.000	**0.056**	0.000	0.009	0.000	0.000	0.000
Lebanese	0.023	0.000	0.013	0.000	0.013	**0.085**	0.001	0.000	0.000	0.000	0.001
Lezgi	0.000	0.000	0.000	0.001	0.000	**0.062**	0.000	0.001	0.000	0.000	0.000
Ossetian	0.000	0.010	0.000	0.042	0.000	0.022	0.012	0.039	0.000	0.019	0.006
Santi Romani	0.008	0.000	0.004	0.000	0.003	**0.067**	0.000	0.000	0.000	0.000	0.000
Russian (Uzb.)	0.000	0.004	0.000	0.017	0.000	**0.048**	0.005	0.016	0.000	0.009	0.002
Spanish	0.000	0.000	0.000	0.009	0.000	**0.077**	0.000	0.008	0.000	0.002	0.000
Svan	0.011	0.000	0.007	0.000	0.005	**0.071**	0.000	0.000	0.000	0.000	0.000
Turkmen	0.004	0.000	0.003	0.000	0.001	**0.080**	0.000	0.000	0.000	0.000	0.000
Tuva	0.023	0.000	0.014	0.000	0.013	**0.085**	0.001	0.000	0.000	0.000	0.001
Uighur	0.026	0.001	0.015	0.000	0.014	**0.088**	0.002	0.000	0.000	0.001	0.002
Yagnobi	0.012	0.000	0.007	0.000	0.006	**0.072**	0.000	0.000	0.000	0.000	0.000
Pamir	0.000	0.000	0.000	0.007	0.000	**0.069**	0.000	0.006	0.000	0.000	0.000
Karakalpak	0.000	0.000	0.000	0.010	0.000	**0.064**	0.000	0.009	0.000	0.002	0.000
Tajik	0.000	0.002	0.000	0.013	0.000	**0.057**	0.003	0.012	0.000	0.006	0.000
Tatar	0.000	0.000	0.000	0.001	0.000	**0.078**	0.000	0.001	0.000	0.000	0.000
Uzbek	0.000	0.000	0.002	0.002	0.000	**0.158**	0.000	0.001	0.000	0.000	0.000
South Asian	0.053	0.000	**0.033**	0.000	**0.031**	**0.198**	0.004	0.000	0.000	0.003	0.002
Iranian	0.000	0.000	0.000	0.006	0.000	**0.069**	0.000	0.005	0.000	0.000	0.000
Italian	0.000	0.008	0.000	0.018	0.000	**0.050**	0.010	0.018	0.000	**0.013**	0.007
British	0.017	**0.037**	**0.024**	**0.046**	**0.025**	0.002	**0.038**	**0.045**	0.026	**0.042**	**0.035**
Chinese	0.022	0.000	0.014	0.000	0.012	**0.111**	0.000	0.000	0.000	0.000	0.000
Orcadian	0.034	**0.061**	**0.046**	**0.072**	**0.046**	0.000	**0.069**	**0.071**	0.042	**0.084**	**0.058**

Table G.8: **R151C 2:3**

	Korean	Kurd	Kyrgyz	Lebanese	Lezgi	Ossetian	Santi Romani	Russian (Uzb.)	Spanish	Svan	Turkmen	Tuva
Adjar												
Arab (Uzb.)												
Armenian												
Azeri												
Dungan												
European (CEU)												
Iranian (Uzb.)												
Japanese												
Jewish (Uzb.)												
Kazakh												
Kazbegi												
Korean	0.000											
Kurd	0.000	0.000										
Kyrgyz	0.011	0.007	0.000									
Lebanese	0.000	0.000	0.011	0.000								
Lezgi	0.002	0.000	0.000	0.002	0.000							
Ossetian	0.045	0.033	0.000	0.045	0.002	0.000						
Santi Romani	0.000	0.000	0.002	0.000	0.000	0.020	0.000					
Russian (Uzb.)	0.018	0.013	0.000	0.017	0.000	0.000	0.008	0.000				
Spanish	0.009	0.006	0.000	0.009	0.000	0.000	0.002	0.000	0.000			
Svan	0.000	0.000	0.004	0.000	0.000	0.026	0.000	0.011	0.004	0.000		
Turkmen	0.000	0.000	0.000	0.000	0.000	0.018	0.000	0.008	0.002	0.000	0.000	
Tuva	0.000	0.000	0.011	0.000	0.002	0.045	0.000	0.018	0.009	0.000	0.000	0.000
Uighur	0.000	0.000	0.013	0.000	0.003	0.050	0.000	0.019	0.010	0.000	0.000	0.000
Yagnobi	0.000	0.000	0.005	0.000	0.000	0.027	0.000	0.011	0.005	0.000	0.000	0.000
Pamir	0.008	0.005	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.008
Karakalpak	0.010	0.007	0.000	0.010	0.000	0.000	0.002	0.000	0.000	0.005	0.001	0.010
Tajik	0.014	0.010	0.000	0.014	0.000	0.000	0.006	0.000	0.000	0.008	0.005	0.014
Tatar	0.002	0.000	0.000	0.002	0.000	0.008	0.000	0.003	0.000	0.000	0.000	0.002
Uzbek	0.002	0.000	0.000	0.002	0.000	0.014	0.000	0.008	0.002	0.000	0.000	0.002
South Asian	0.000	0.000	0.027	0.000	0.004	**0.107**	0.000	**0.043**	**0.020**	0.000	0.000	0.000
Iranian	0.006	0.003	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.006
Italian	0.019	0.016	0.000	0.019	0.005	0.000	0.012	0.000	0.002	0.014	0.012	0.019
British	**0.046**	**0.044**	**0.026**	**0.046**	**0.033**	0.007	**0.041**	**0.020**	**0.029**	**0.042**	**0.040**	**0.046**
Chinese	0.000	0.000	0.010	0.000	0.000	0.050	0.000	0.020	0.009	0.000	0.000	0.000
Orcadian	**0.074**	**0.068**	**0.048**	**0.074**	**0.054**	0.020	**0.061**	**0.041**	**0.064**	**0.064**	**0.068**	**0.074**

Table G.9: **R151C 3:3**

	Uighur	Yagnobi	Pamir	Karakalpak	Tajik	Tatar	Uzbek	South Asian	Iranian	Italian	British	Chinese	Orcadian
Adjar													
Arab (Uzb.)													
Armenian													
Azeri													
Dungan													
European (CEU)													
Iranian (Uzb.)													
Japanese													
Jewish (Uzb.)													
Kazakh													
Kazbegi													
Korean													
Kurd													
Kyrgyz													
Lebanese													
Lezgi													
Ossetian													
Santi Romani													
Russian (Uzb.)													
Spanish													
Svan													
Turkmen													
Tuva													
Uighur	0.000												
Yagnobi	0.000	0.000											
Pamir	0.009	0.003	0.000										
Karakalpak	0.011	0.005	0.000	0.000									
Tajik	0.015	0.008	0.000	0.000	0.000								
Tatar	0.003	0.000	0.000	0.000	0.002	0.000							
Uzbek	0.003	0.000	0.000	0.001	0.005	0.000	0.000						
South Asian	0.000	0.000	**0.019**	**0.024**	**0.032**	0.005	0.005	0.000					
Iranian	0.007	0.001	0.000	0.000	0.000	0.000	0.000	**0.015**	0.000				
Italian	0.020	0.014	0.003	0.001	0.000	0.008	**0.015**	**0.036**	0.004	0.000			
British	**0.047**	**0.043**	**0.030**	**0.027**	**0.023**	**0.037**	**0.049**	**0.060**	**0.032**	**0.017**	0.000		
Chinese	0.000	0.000	0.007	0.010	**0.015**	0.000	0.001	0.000	0.005	**0.020**	**0.048**	0.000	
Orcadian	**0.076**	**0.064**	**0.058**	**0.054**	**0.048**	**0.066**	**0.127**	**0.153**	**0.058**	**0.042**	0.002	**0.091**	0.000

Table G.10: **R160W 1:3**

	Adjar	Arab (Uzb.)	Armenian	Azeri	Dungan	European (CEU)	Iranian (Uzb.)	Japanese	Jewish (Uzb.)	Kazakh	Kazbegi	Korean
Adjar	0.000											
Arab (Uzb.)	0.013	0.000										
Armenian	0.000	0.004	0.000									
Azeri	0.000	0.023	0.000	0.000								
Dungan	0.019	0.000	0.007	0.029	0.000							
European (CEU)	0.025	**0.072**	**0.045**	0.019	**0.079**	0.000						
Iranian (Uzb.)	0.000	0.008	0.000	0.000	0.012	**0.039**	0.000					
Japanese	0.013	0.000	0.004	0.024	0.000	**0.080**	0.009	0.000				
Jewish (Uzb.)	0.000	0.000	0.000	0.003	0.000	0.046	0.000	0.000	0.000			
Kazakh	0.001	0.000	0.000	0.012	0.000	**0.079**	0.002	0.000	0.000	0.000		
Kazbegi	0.000	0.012	0.000	0.000	0.017	**0.028**	0.000	0.013	0.000	0.003	0.000	
Korean	0.012	0.000	0.003	0.021	0.000	**0.070**	0.008	0.000	0.000	0.000	0.011	0.000
Kurd	0.000	0.000	0.000	0.000	0.002	**0.041**	0.000	0.000	0.000	0.000	0.000	0.000
Kyrgyz	0.021	0.000	0.008	0.031	0.000	**0.082**	0.013	0.000	0.000	0.000	0.019	0.000
Lebanese	0.000	0.000	0.000	0.005	0.000	**0.055**	0.000	0.000	0.000	0.000	0.000	0.000
Lezgi	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000	0.000	0.000	0.000	0.000
Ossetian	0.000	0.036	0.000	0.000	0.046	0.008	0.000	0.038	0.008	0.018	0.000	0.034
Santi Romani	0.020	0.000	0.009	0.026	0.000	**0.068**	0.013	0.000	0.000	0.001	0.018	0.000
Russian (Uzb.)	0.004	**0.045**	**0.019**	0.000	**0.051**	0.002	0.013	**0.050**	0.023	**0.043**	0.006	**0.044**
Spanish	0.053	0.000	**0.023**	**0.068**	0.000	**0.135**	**0.031**	0.000	0.000	0.007	0.047	0.000
Svan	0.010	**0.084**	0.034	0.003	**0.099**	0.000	0.025	**0.097**	0.037	**0.081**	0.013	**0.081**
Turkmen	0.050	0.002	0.023	**0.055**	0.000	**0.095**	0.028	0.004	0.000	0.010	**0.042**	0.003
Tuva	0.041	0.000	0.019	**0.046**	0.000	**0.086**	0.023	0.002	0.000	0.007	**0.034**	0.000
Uighur	0.044	0.001	0.021	**0.049**	0.000	**0.089**	0.025	0.003	0.000	0.008	0.037	0.001
Yagnobi	0.000	0.000	0.000	0.005	0.000	**0.051**	0.000	0.000	0.000	0.000	0.000	0.000
Pamir	0.001	0.000	0.000	0.011	0.000	**0.070**	0.001	0.000	0.000	0.000	0.002	0.000
Karakalpak	0.041	0.000	0.018	**0.052**	0.000	**0.104**	0.024	0.000	0.000	0.005	0.036	0.000
Tajik	0.015	0.000	0.005	0.028	0.000	**0.092**	0.010	0.000	0.000	0.000	0.015	0.000
Tatar	0.000	0.025	0.002	0.000	**0.030**	0.017	0.000	0.027	0.007	**0.018**	0.000	0.024
Uzbek	0.005	0.000	0.000	0.020	0.000	**0.128**	0.005	0.000	0.000	0.000	0.007	0.000
South Asian	**0.082**	0.000	**0.038**	**0.104**	0.000	**0.192**	**0.048**	0.001	0.000	**0.012**	**0.073**	0.000
Iranian	0.000	0.002	0.000	0.000	0.005	**0.051**	0.000	0.002	0.000	0.000	0.000	0.001
Chinese	**0.059**	0.000	**0.026**	**0.075**	0.000	**0.143**	**0.034**	0.000	0.000	0.008	**0.052**	0.000
Italian	0.000	0.003	0.000	0.001	0.005	**0.071**	0.000	0.003	0.000	0.000	0.000	0.002
British	0.000	**0.024**	0.008	0.000	**0.027**	0.016	0.005	**0.024**	0.013	**0.019**	0.001	**0.024**
Orcadian	0.006	**0.041**	0.019	0.002	**0.046**	0.003	0.014	**0.044**	0.024	**0.039**	0.008	**0.040**

Table G.11: **R160W 2:3**

	Kurd	Kyrgyz	Lebanese	Lezgi	Ossetian	Santi Romani	Russian (Uzb.)	Spanish	Svan	Turkmen	Tuva	Uighur
Adjar												
Arab (Uzb.)												
Armenian												
Azeri												
Dungan												
European (CEU)												
Iranian (Uzb.)												
Japanese												
Jewish (Uzb.)												
Kazakh												
Kazbegi												
Korean												
Kurd	0.000											
Kyrgyz	0.003	0.000										
Lebanese	0.000	0.000	0.000									
Lezgi	0.000	0.000	0.000	0.000								
Ossetian	0.000	0.050	0.010	0.002	0.000							
Santi Romani	0.006	0.000	0.000	0.003	0.041	0.000						
Russian (Uzb.)	0.017	**0.053**	**0.029**	0.022	0.000	**0.044**	0.000					
Spanish	0.018	0.000	0.007	0.012	**0.108**	0.000	**0.096**	0.000				
Svan	0.032	**0.104**	**0.052**	**0.040**	0.000	**0.076**	0.000	**0.202**	0.000			
Turkmen	0.027	0.000	0.014	0.021	0.092	0.000	**0.068**	0.000	**0.141**	0.000		
Tuva	0.021	0.000	0.010	0.016	0.075	0.000	**0.061**	0.000	**0.119**	0.000	0.000	
Uighur	0.023	0.000	0.012	0.017	**0.081**	0.000	**0.063**	0.000	**0.126**	0.000	0.000	0.000
Yagnobi	0.000	0.000	0.000	0.000	0.010	0.000	0.027	0.000	0.047	0.013	0.008	0.010
Pamir	0.000	0.000	0.000	0.000	0.017	0.000	**0.038**	0.007	**0.072**	0.010	0.008	0.009
Karakalpak	0.015	0.000	0.005	0.010	**0.084**	0.000	**0.073**	0.000	**0.153**	0.000	0.000	0.000
Tajik	0.000	0.000	0.000	0.000	0.043	0.000	**0.057**	0.000	**0.113**	0.004	0.002	0.003
Tatar	0.002	**0.032**	0.011	0.006	0.000	0.028	0.000	**0.063**	0.001	**0.048**	**0.042**	**0.044**
Uzbek	0.000	0.000	0.000	0.000	0.028	0.000	**0.069**	0.005	**0.115**	0.007	0.006	0.006
South Asian	0.032	0.000	0.014	0.023	**0.159**	0.000	**0.141**	0.000	**0.288**	0.000	0.000	0.000
Iranian	0.000	0.006	0.000	0.000	0.000	0.006	**0.023**	0.018	**0.042**	0.020	0.016	0.017
Chinese	0.022	0.000	0.009	0.015	**0.118**	0.000	**0.103**	0.000	**0.218**	0.000	0.000	0.000
Italian	0.000	0.006	0.000	0.000	0.002	0.006	**0.031**	**0.014**	**0.053**	0.015	**0.013**	0.014
British	0.008	**0.028**	0.015	0.011	0.000	0.028	0.000	**0.035**	0.000	**0.035**	**0.034**	**0.034**
Orcadian	0.018	**0.047**	0.028	0.022	0.000	0.042	0.000	**0.076**	0.000	**0.059**	**0.054**	**0.055**

Table G.12: **R160W 3:3**

	Yagnobi	Pamir	Karakalpak	Tajik	Tatar	Uzbek	South Asian	Iranian	Chinese	Italian	British	Orcadian
Adjar												
Arab (Uzb.)												
Armenian												
Azeri												
Dungan												
European (CEU)												
Iranian (Uzb.)												
Japanese												
Jewish (Uzb.)												
Kazakh												
Kazbegi												
Korean												
Kurd												
Kyrgyz												
Lebanese												
Lezgi												
Ossetian												
Santi Romani												
Russian (Uzb.)												
Spanish												
Svan												
Turkmen												
Tuva												
Uighur												
Yagnobi	0.000											
Pamir	0.000	0.000										
Karakalpak	0.000	0.005	0.000									
Tajik	0.000	0.000	0.000	0.000								
Tatar	0.010	0.016	**0.048**	**0.031**	0.000							
Uzbek	0.000	0.000	0.004	0.000	**0.029**	0.000						
South Asian	0.005	**0.013**	0.000	0.002	**0.095**	**0.008**	0.000					
Iranian	0.000	0.000	0.014	0.003	0.005	0.000	**0.030**	0.000				
Chinese	0.001	0.008	0.000	0.000	**0.068**	0.005	0.000	**0.021**	0.000			
Italian	0.000	0.000	0.012	0.004	0.008	0.001	**0.021**	0.000	**0.016**	0.000		
British	0.015	**0.018**	**0.033**	**0.026**	0.000	**0.025**	**0.042**	0.010	**0.037**	**0.012**	0.000	
Orcadian	0.027	**0.035**	**0.061**	**0.049**	0.001	**0.062**	**0.107**	**0.023**	**0.080**	**0.030**	0.001	0.000

Table G.13: **R163Q 1:4**

	Adjar	Arab (Uzb.)	Armenian	Azeri	British	Buryat	Dai	Dungan	European (CEU)	French	Indonesian	Iranian (Uzb.)
Adjar	0.000											
Arab (Uzb.)	**0.116**	0.000										
Armenian	0.000	**0.114**	0.000									
Azeri	**0.044**	0.019	0.029	0.000								
British	0.000	**0.086**	0.000	0.015	0.000							
Buryat	**0.566**	**0.399**	**0.591**	**0.492**	**0.558**	0.000						
Dai	**0.706**	**0.467**	**0.745**	**0.595**	**0.698**	0.000	0.000					
Dungan	**0.393**	**0.179**	**0.436**	**0.289**	**0.384**	**0.083**	**0.103**	0.000				
European (CEU)	**0.041**	0.020	**0.030**	0.000	0.019	**0.518**	**0.624**	**0.324**	0.000			
French	0.000	**0.129**	0.000	**0.035**	0.000	**0.611**	**0.769**	**0.467**	**0.035**	0.000		
Indonesian	**0.400**	**0.211**	**0.431**	**0.310**	**0.390**	**0.055**	**0.071**	0.000	**0.341**	**0.456**	0.000	
Iranian (Uzb.)	0.025	**0.042**	0.013	0.000	0.005	**0.542**	**0.663**	**0.358**	0.000	0.017	**0.369**	0.000
Jewish (Uzb.)	0.000	0.032	0.000	0.000	0.000	**0.497**	**0.573**	**0.280**	0.000	0.000	**0.312**	0.000
Kazakh	**0.313**	**0.136**	**0.331**	**0.225**	**0.298**	**0.111**	**0.132**	0.000	**0.246**	**0.351**	0.007	**0.273**
Kazbegi	0.000	**0.096**	0.000	0.023	0.000	**0.562**	**0.704**	**0.390**	**0.025**	0.000	**0.395**	0.010
Khalkha	**0.588**	**0.429**	**0.608**	**0.517**	**0.579**	0.000	0.000	**0.107**	**0.540**	**0.625**	**0.074**	**0.562**
Kurd	0.010	**0.057**	0.000	0.000	0.000	**0.528**	**0.648**	**0.337**	0.005	0.000	**0.353**	0.000
Kyrgyz	**0.497**	**0.290**	**0.542**	**0.401**	**0.492**	0.018	0.029	0.015	**0.440**	**0.571**	0.003	**0.472**
Lebanese	0.000	**0.143**	0.003	**0.063**	0.008	**0.590**	**0.749**	**0.434**	**0.054**	0.002	**0.430**	0.037
Lezgi	0.004	**0.068**	0.000	0.006	0.000	**0.540**	**0.668**	**0.356**	0.010	0.000	**0.367**	0.000
Malay	**0.409**	**0.213**	**0.445**	**0.316**	**0.400**	**0.055**	**0.072**	0.000	**0.350**	**0.472**	0.000	**0.379**
Orcadian	0.015	0.059	0.003	0.000	0.000	**0.549**	**0.680**	**0.371**	0.004	0.006	**0.379**	0.000
Ossetian	0.015	**0.043**	0.000	0.000	0.000	**0.509**	**0.612**	**0.307**	0.000	0.000	**0.329**	0.000
Santi Romani	**0.127**	0.000	**0.114**	0.004	**0.081**	**0.408**	**0.476**	**0.184**	0.005	**0.131**	**0.217**	0.026
Russian (Uzb.)	**0.035**	0.028	0.022	0.000	0.012	**0.525**	**0.638**	**0.335**	0.000	**0.027**	**0.350**	0.000
Spanish	0.000	**0.240**	0.005	**0.103**	0.012	**0.698**	**0.861**	**0.603**	**0.083**	0.004	**0.570**	**0.057**
Svan	0.055	0.011	0.034	0.000	0.018	**0.474**	**0.566**	**0.262**	0.000	**0.042**	**0.289**	0.000
Tibetan	**0.861**	**0.608**	**0.866**	**0.738**	**0.837**	**0.045**	0.022	**0.221**	**0.740**	**0.878**	**0.177**	**0.782**
Turkish	0.021	**0.054**	0.011	0.000	0.004	**0.595**	**0.707**	**0.424**	0.001	0.013	**0.432**	0.000
Turkmen	**0.217**	0.032	**0.237**	**0.111**	**0.197**	**0.259**	**0.295**	**0.059**	**0.123**	**0.261**	**0.088**	**0.157**
Tuva	**0.550**	**0.329**	**0.599**	**0.447**	**0.545**	0.006	0.014	0.027	**0.488**	**0.629**	0.010	**0.523**
Wa	**0.636**	**0.389**	**0.684**	**0.520**	**0.629**	0.000	0.000	**0.051**	**0.556**	**0.712**	0.028	**0.597**
Yagnobi	**0.256**	0.023	**0.282**	**0.112**	**0.227**	**0.270**	**0.310**	**0.063**	**0.123**	**0.312**	**0.093**	**0.167**
Pamir	**0.066**	0.000	**0.057**	0.000	**0.041**	**0.470**	**0.558**	**0.261**	0.000	**0.064**	**0.287**	0.012
Karakalpak	**0.381**	**0.192**	**0.410**	**0.290**	**0.370**	**0.068**	**0.086**	0.000	**0.320**	**0.435**	0.000	**0.348**
Tajik	**0.113**	0.000	**0.111**	**0.032**	**0.090**	**0.385**	**0.444**	**0.166**	**0.034**	**0.122**	**0.200**	**0.055**
Tatar	**0.104**	0.000	**0.101**	0.020	**0.079**	**0.406**	**0.474**	**0.188**	0.021	**0.113**	**0.219**	**0.041**
Uzbek	**0.135**	0.011	**0.127**	**0.062**	**0.114**	**0.324**	**0.362**	**0.111**	**0.063**	**0.133**	**0.148**	**0.082**
Iranian	0.000	**0.140**	0.000	**0.046**	0.000	**0.610**	**0.773**	**0.467**	**0.043**	0.000	**0.455**	0.024
Korean	**0.661**	**0.500**	**0.684**	**0.590**	**0.654**	0.010	0.000	**0.159**	**0.612**	**0.702**	**0.118**	**0.636**
Italian	0.000	**0.271**	0.006	**0.115**	0.013	**0.731**	**0.881**	**0.648**	**0.094**	0.005	**0.613**	**0.064**
Uighur	**0.270**	**0.083**	**0.293**	**0.172**	**0.254**	**0.177**	**0.204**	0.016	**0.192**	**0.317**	**0.036**	**0.223**
South Asian	0.007	**0.124**	0.000	0.014	0.000	**0.725**	**0.804**	**0.585**	**0.020**	0.002	**0.594**	0.004
Chinese	**0.459**	**0.307**	**0.457**	**0.388**	**0.443**	0.006	0.012	**0.039**	**0.396**	**0.466**	**0.021**	**0.417**
German	0.004	**0.129**	0.000	0.020	0.000	**0.679**	**0.798**	**0.547**	**0.025**	0.000	**0.539**	0.007
Japanese	**0.638**	**0.498**	**0.648**	**0.576**	**0.628**	**0.012**	0.000	**0.170**	**0.590**	**0.660**	**0.128**	**0.609**

Table G.14: **R163Q 2:4**

	Jewish (Uzb.)	Kazakh	Kazbegi	Khalkha	Kurd	Kyrgyz	Lebanese	Lezgi	Malay	Orcadian	Ossetian	Santi Romani
Adjar												
Arab (Uzb.)												
Armenian												
Azeri												
British												
Buryat												
Dai												
Dungan												
European (CEU)												
French												
Indonesian												
Iranian (Uzb.)												
Jewish (Uzb.)	0.000											
Kazakh	**0.237**	0.000										
Kazbegi	0.000	**0.304**	0.000									
Khalkha	**0.526**	**0.136**	**0.583**	0.000								
Kurd	0.000	**0.267**	0.000	**0.552**	0.000							
Kyrgyz	**0.388**	**0.034**	**0.496**	**0.031**	**0.446**	0.000						
Lebanese	0.018	**0.336**	0.002	**0.608**	0.023	**0.536**	0.000					
Lezgi	0.000	**0.279**	0.000	**0.563**	0.000	**0.464**	0.016	0.000				
Malay	**0.313**	0.006	**0.405**	**0.075**	**0.360**	0.002	**0.443**	**0.376**	0.000			
Orcadian	0.000	**0.284**	0.001	**0.569**	0.000	**0.482**	0.026	0.000	**0.389**	0.000		
Ossetian	0.000	**0.248**	0.000	**0.535**	0.000	**0.415**	0.032	0.000	**0.333**	0.000	0.000	
Santi Romani	0.018	**0.144**	**0.096**	**0.439**	**0.046**	**0.293**	**0.167**	0.059	**0.218**	0.045	0.030	0.000
Russian (Uzb.)	0.000	**0.255**	0.018	**0.547**	0.000	**0.450**	**0.047**	0.005	**0.358**	0.000	0.000	0.012
Spanish	0.018	**0.449**	0.004	**0.703**	**0.033**	**0.693**	0.000	0.023	**0.593**	**0.040**	0.046	**0.275**
Svan	0.000	**0.209**	0.028	**0.502**	0.000	**0.374**	**0.080**	0.007	**0.292**	0.000	0.000	0.000
Tibetan	**0.752**	**0.248**	**0.846**	0.030	**0.799**	**0.122**	**0.889**	**0.814**	**0.179**	**0.806**	**0.775**	**0.636**
Turkish	0.000	**0.324**	0.008	**0.610**	0.000	**0.538**	**0.029**	0.000	**0.444**	0.000	0.000	0.034
Turkmen	0.121	**0.038**	**0.205**	**0.291**	**0.159**	**0.147**	**0.250**	**0.174**	**0.088**	**0.174**	**0.137**	0.038
Tuva	**0.426**	**0.049**	**0.550**	0.017	**0.495**	0.000	**0.594**	**0.515**	0.010	**0.534**	**0.460**	**0.331**
Wa	**0.491**	**0.077**	**0.635**	0.000	**0.574**	0.000	**0.684**	**0.596**	0.029	**0.613**	**0.534**	**0.393**
Yagnobi	**0.119**	**0.040**	**0.239**	**0.303**	**0.174**	**0.153**	**0.307**	**0.195**	**0.092**	**0.192**	**0.145**	0.029
Pamir	0.005	**0.199**	**0.049**	**0.495**	0.023	**0.378**	**0.081**	0.030	**0.292**	0.022	0.014	0.000
Karakalpak	**0.293**	0.002	**0.375**	**0.089**	**0.333**	0.008	**0.410**	**0.347**	0.000	**0.358**	**0.310**	**0.198**
Tajik	0.046	**0.124**	**0.098**	**0.414**	**0.067**	**0.279**	**0.131**	**0.076**	**0.201**	**0.069**	0.055	0.000
Tatar	0.032	**0.141**	**0.087**	**0.434**	**0.054**	**0.301**	**0.125**	**0.064**	**0.221**	**0.055**	**0.042**	0.000
Uzbek	0.081	**0.080**	**0.121**	**0.351**	**0.096**	**0.216**	**0.145**	**0.104**	**0.147**	**0.095**	**0.087**	0.014
Iranian	0.000	**0.351**	0.000	**0.624**	0.005	**0.570**	0.000	0.001	**0.471**	0.013	0.009	**0.151**
Korean	**0.596**	**0.186**	**0.657**	0.003	**0.625**	**0.066**	**0.683**	**0.636**	**0.120**	**0.644**	**0.607**	**0.512**
Italian	0.019	**0.490**	0.004	**0.734**	**0.036**	**0.732**	0.000	0.025	**0.637**	**0.045**	**0.049**	**0.302**
Uighur	**0.180**	0.005	**0.261**	**0.207**	**0.218**	**0.079**	**0.299**	**0.232**	**0.035**	**0.237**	**0.196**	**0.090**
South Asian	0.000	**0.482**	0.000	**0.733**	0.000	**0.685**	0.013	0.000	**0.605**	0.000	0.000	**0.095**
Chinese	**0.414**	**0.063**	**0.448**	**0.014**	**0.424**	0.000	**0.469**	**0.432**	**0.021**	**0.427**	**0.414**	**0.320**
German	0.000	**0.422**	0.000	**0.687**	0.000	**0.649**	0.010	0.000	**0.556**	0.000	0.000	**0.109**
Japanese	**0.593**	**0.200**	**0.632**	0.005	**0.608**	**0.073**	**0.652**	**0.616**	**0.130**	**0.617**	**0.596**	**0.511**

Table G.15: **R163Q 3:4**

	Russian (Uzb.)	Spanish	Svan	Tibetan	Turkish	Turkmen	Tuva	Wa	Yagnobi	Pamir	Karakalpak
Adjar											
Arab (Uzb.)											
Armenian											
Azeri											
British											
Buryat											
Dai											
Dungan											
European (CEU)											
French											
Indonesian											
Iranian (Uzb.)											
Jewish (Uzb.)											
Kazakh											
Kazbegi											
Khalkha											
Kurd											
Kyrgyz											
Lebanese											
Lezgi											
Malay											
Orcadian											
Ossetian											
Santi Romani											
Russian (Uzb.)	0.000										
Spanish	**0.074**	0.000									
Svan	0.000	**0.129**	0.000								
Tibetan	**0.757**	**0.935**	**0.723**	0.000							
Turkish	0.000	**0.041**	0.000	**0.800**	0.000						
Turkmen	**0.136**	**0.391**	0.095	**0.430**	**0.195**	0.000					
Tuva	**0.499**	**0.749**	**0.416**	0.101	**0.586**	**0.175**	0.000				
Wa	**0.570**	**0.822**	**0.487**	0.068	**0.650**	**0.221**	0.000	0.000			
Yagnobi	**0.139**	**0.483**	**0.094**	**0.461**	**0.202**	0.000	**0.182**	**0.231**	0.000		
Pamir	0.004	**0.128**	0.000	**0.680**	**0.017**	**0.078**	**0.422**	**0.486**	**0.071**	0.000	
Karakalpak	**0.328**	**0.549**	**0.269**	**0.195**	**0.410**	**0.074**	0.018	0.039	**0.078**	**0.266**	0.000
Tajik	**0.042**	**0.198**	0.024	**0.568**	**0.069**	0.022	**0.315**	**0.371**	0.012	0.011	**0.180**
Tatar	**0.028**	**0.201**	0.013	**0.604**	**0.053**	**0.034**	**0.339**	**0.399**	0.025	0.002	**0.199**
Uzbek	**0.070**	**0.172**	**0.055**	**0.478**	**0.095**	0.001	**0.246**	**0.293**	0.000	**0.037**	**0.130**
Iranian	0.035	0.000	0.057	**0.887**	0.019	**0.266**	**0.629**	**0.715**	**0.324**	**0.073**	**0.434**
Korean	**0.620**	**0.779**	**0.574**	0.006	**0.680**	**0.357**	**0.047**	0.019	**0.374**	**0.566**	**0.135**
Italian	**0.084**	0.000	**0.142**	**0.942**	**0.045**	**0.436**	**0.784**	**0.847**	**0.526**	**0.146**	**0.593**
Uighur	**0.203**	**0.437**	**0.154**	**0.330**	**0.273**	0.006	**0.101**	**0.138**	0.006	**0.142**	**0.026**
South Asian	0.012	**0.016**	0.015	**0.862**	0.004	**0.332**	**0.721**	**0.765**	**0.336**	**0.055**	**0.573**
Chinese	**0.403**	**0.511**	**0.379**	**0.084**	**0.443**	**0.185**	0.000	0.000	**0.194**	**0.360**	**0.030**
German	0.017	**0.013**	0.023	**0.870**	0.006	**0.308**	**0.695**	**0.754**	**0.336**	**0.059**	**0.518**
Japanese	**0.596**	**0.719**	**0.566**	0.006	**0.644**	**0.367**	**0.052**	0.022	**0.383**	**0.552**	**0.146**

Table G.16: **R163Q 4:4**

	Tajik	Tatar	Uzbek	Iranian	Korean	Italian	Uighur	South Asian	Chinese	German	Japanese
Adjar											
Arab (Uzb.)											
Armenian											
Azeri											
British											
Buryat											
Dai											
Dungan											
European (CEU)											
French											
Indonesian											
Iranian (Uzb.)											
Jewish (Uzb.)											
Kazakh											
Kazbegi											
Khalkha											
Kurd											
Kyrgyz											
Lebanese											
Lezgi											
Malay											
Orcadian											
Ossetian											
Santi Romani											
Russian (Uzb.)											
Spanish											
Svan											
Tibetan											
Turkish											
Turkmen											
Tuva											
Wa											
Yagnobi											
Pamir											
Karakalpak											
Tajik	0.000										
Tatar	0.000	0.000									
Uzbek	**0.005**	**0.012**	0.000								
Iranian	**0.129**	**0.122**	**0.138**	0.000							
Korean	**0.480**	**0.504**	**0.406**	**0.702**	0.000						
Italian	**0.225**	**0.229**	**0.185**	0.000	**0.806**	0.000					
Uighur	**0.070**	**0.087**	**0.033**	**0.319**	**0.266**	**0.482**	0.000				
South Asian	**0.146**	**0.121**	**0.162**	0.005	**0.786**	**0.017**	**0.429**	0.000			
Chinese	**0.291**	**0.310**	**0.244**	**0.469**	**0.037**	**0.530**	**0.116**	**0.548**	0.000		
German	**0.136**	**0.120**	**0.142**	0.002	**0.755**	**0.015**	**0.385**	0.000	**0.498**	0.000	
Japanese	**0.479**	**0.500**	**0.411**	**0.661**	0.000	**0.742**	**0.280**	**0.747**	**0.041**	**0.706**	0.000

Appendix H

Facial feature candidate genes

Table H.1: **Facial features candidate genes**

Gene	**Chromosome**	**Relevance**	**Reference**
ABCC9	12	Cantu syndrome	[614]
ACTA1	1	Myopathy w/ tall, thin face in affected	[615]
ADSS	1	Minimal area for 1qter microdeletion syndrome	[616]
AKT3	1	Minimal area for 1qter microdeletion syndrome	[616]
ALCAM	3	Proximal 3q microdeletion syndrome	[617]
ALDH1A2	15	Signals important for craniofacial development	[99]
ALX3	1	Causative of frontorhiny	[618]
ALX4	11	P11pDS candidates	[619]
ARSE	X	Chondrodysplasia punctata	[545]
ASXL1	20	Bohring-Opitz syndrome	[620]
ATRX	X	Mutated in ChLS	[621]
B4GALT7	5	Mutated in Ehlers-Danlos syndrome-like & facioskeletal anomalies	[622]
BCOR	X	Mutated in OCFD	[623]
BLM	15	Bloom syndrome	[545]
BMP2	20	Signals important for craniofacial development	[99]

Continued on the next page

–continued from the previous page

Gene	Chromosome	Relevance	Reference
BMP4	14	Signals important for craniofacial development	[99]
BMP7	20	Signals important for craniofacial development	[99]
BRAF	7	Cardio-facio-cutaneous syndrome	[624]
C17orf69	17	Candidates for 17q21.31 microdeletion syndrome	[625]
C1orf100	1	Minimal area for 1qter microdeletion syndrome	[616]
C1orf101	1	Minimal area for 1qter microdeletion syndrome	[616]
CACNA1C	12	Timothy syndrome spectrum	[626]
CASK	X	Syndrome with distinctive facial features	[627]
CBLB	3	Proximal 3q microdeletion syndrome	[617]
CD300A	17	Silver-Russell syndrome	[628]
CDKN1C	11	Beckwith-Wiedemann syndrome	[629]
CHD7	8	CHARGE syndrome	[630]
CNTN4	3	3p deletion syndrome	[631]
COL11A1	1	Stickler syndrome; ophthalmic, orofacial, articular, and auditory manifestations	[632]
COL11A2	6	Stickler syndrome; ophthalmic, orofacial, articular, and auditory manifestations	[632]
COL1A1	17	Osteogenesis imperfecta w/ facial dysmorphia	[633]
COL2A1	12	Stickler syndrome; ophthalmic, orofacial, articular, and auditory manifestations	[632]
COL9A1	6	Stickler syndrome; ophthalmic, orofacial, articular, and auditory manifestations	[632]
CREBBP	16	Rubinstein-Taybi syndrome	[545]
CRHR1	17	Like Anglemann syndrome, candidates for 17q21.31 microdeletion syndrome	[634, 625]
CRK	17	Miller-Dieker syndrome, candidate genes for clinical features excepting standard lissencephaly	[635]
CRLF1	19	CISS w/ facial dysmorphia	[636]
CTBP2	10	10q deletion candidates for non-behavioral/mental retardation phenotype	[637]
CTNNB1	3	Signals important for craniofacial development	[99]

Continued on the next page

–continued from the previous page

Gene	**Chromosome**	**Relevance**	**Reference**
CTSK	1	Pycnodysostosis	[638]
CYP1B1	2	Axenfeld-Rieger syndrome/Peters' anomaly	[639]
DHCR7	11	Smith-Lemli-Opitz syndrome	[545, 640]
DHODH	16	Miller syndrome	[641]
DLX1	2	Signals important for craniofacial development	[99]
DLX2	2	Signals important for craniofacial development	[99]
DLX3	17	Signals important for craniofacial development	[99]
DLX5	7	Signals important for craniofacial development	[99]
DLX6	7	Signals important for craniofacial development	[99]
DNASE1	16	Severe RSTS	[642]
DNMT3B	20	Immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome	[615]
EDA	X	Anhydrotic ectodermal dysplasia	[643]
EDN1	6	Signals important for craniofacial development	[99]
EDN3	20	Mutated in syndromes that can display facial abnormalities	[644]
EDNRA	4	Signals important for craniofacial development	[99]
EDNRB	13	Mutated in syndromes that can display facial abnormalities	[644]
EFNB1	X	Craniofrontonasal Syndrome	[645]
EHMT1	9	9q34 subtelomeric deletion syndrome	[646]
ELN	7	Williams syndrome	[545]
ENPP1	6	Affects head height and face breadth	[94]
EPM2A	6	In the 6q24.3 region, deleted in patients showing dysmorphic features	[647]
ERCC3	2	Cockayne syndrome	[545]
ERCC5	13	Cockayne syndrome	[545]
ERCC6	10	Cockayne syndrome, type II	[545]
ERCC8	5	Cockayne syndrome, type I	[545]
ESCO2	8	SC phocomelia	[648]
EVC	4	Ellis van Creveld syndrome	[649]
EVC2	4	Ellis van Creveld syndrome	[649]
EXT1	8	Tricho-rhino-phalangeal syndrome (cf. Lagner-Giedion syndrome)	[650]
EXT2	11	P11pDS candidates	[619]
EXH2	7	Weaver syndrome	[651]

Continued on the next page

–continued from the previous page

Gene	Chromosome	Relevance	Reference
FGD1	X	Aarskog syndrome	[545]
FGF4	11	Signals important for craniofacial development	[99]
FGF8	10	Influences pattern and growth of upper face in model vertebrates	[652, 653]
FGF9	13	Signals important for craniofacial development	[99]
FGFR1	8	Pfeiffer syndrome	[545]
FGFR2	10	10q deletion candidates for non-behavioral/mental retardation phenotype, Pfeiffer syndrome, Crouzon syndrome (craniofacial dysostosis), Apert syndrome, Jackson-Weiss syndrome	[637, 545]
FGFR3	4	Signals important for craniofacial development	[99]
FLNB	3	Larsen syndrome	[654]
FMR1	X	Fragile X-linked mental retardation	[545]
FOXC1	6	Axenfeld-Rieger syndrome/Peters' anomaly	[639]
FOXG1	14	Mild facial dysmorphia, 14q12del	[655]
FOXL2	3	Blepharophimosis-Ptosis-Epicanthus inversus	[656]
FREM1	9	Manitoba-oculo-tricho-anal syndrome	[657]
FRZB	2	Possible promoters of vertebrate head muscle dev.	[658]
FSTL3	19	Syndrome w/ distinctive facial features	[659]
GEMIN2	14	Hirschsprung disease	[660]
GHR	5	Affects mandibular height	[92]
GJA1	6	Axenfeld-Rieger syndrome/Peters' anomaly	[639]
GLI1	12	Altered expression in frontonasal ectoderm and mesenchyme w/ Shh perturbation in model organisms	[97]
GLI2	2	Signals important for craniofacial development	[99]
GLI3	7	Greig cephalopolysyndactyly	[545]
GPC3	X	Simpson-Golabi-Behmel syndrome (SGBS)	[661]
GRB10	7	Candidates for Silver-Russell syndrome	[662]
GREM2	1	Possible promoters of vertebrate head muscle dev.	[658]
GSC	14	Signals important for craniofacial development	[99]
H19	11	Beckwith-Wiedemann syndrome	[629]

Continued on the next page

–continued from the previous page

Gene	Chromosome	Relevance	Reference
HRAS	11	Costello syndrome	[663]
IGF2	11	Beckwith-Wiedemann syndrome	[629]
IGFBP1	7	Candidates for Silver-Russell syndrome	[662]
IMP5	17	Candidates for 17q21.31 microdeletion syndrome	[625]
INPP5K	17	Miller-Dieker syndrome, candidate genes for clinical features excepting standard lissencephaly	[635]
IRF6	1	van der Woude syndrome	[664]
JAG1	20	Alagille syndrome	[665]
KAT6B	10	Say-Barber-Biesecker variant of Ohdo syndrome	[666]
KCNH2	7	Unusual facial appearance	[667]
KCNJ2	17	Andersen-Tawali syndrome	[668]
KCNQ1	11	Beckwith-Wiedemann syndrome	[629]
KCNQ1OT1	11	Beckwith-Wiedemann syndrome	[629]
KIAA1267	17	Candidates for 17q21.31 microdeletion syndrome	[625]
KRAS	12	Noonan/CFC syndrome, w/ facial anomalies	[669]
LETM1	4	Mutated in Wolf-Hirschorn syndrome	[670]
LHX8	1	Signals important for craniofacial development	[99]
MAP2K1	15	Cardio-facio-cutaneous syndrome	[624]
MAP2K2	19	Cardio-facio-cutaneous syndrome	[624]
MAPK8IP1	11	P11pDS candidates	[619]
MAPT	17	Like Anglemann syndrome, candidates for 17q21.31 microdeletion syndrome	[634, 625]
MITF	3	Mutated in syndromes that can display facial abnormalities, Waardenburg syndrome type 2	[545, 644]
MLL2	12	Kabuki syndrome	[671]
MSX1	4	Signals important for craniofacial development	[99]
MSX2	5	Boston craniosynostosis	[545]
MYBPC3	11	P11pDS candidates	[619]
MYO1C	17	Miller-Dieker syndrome, candidate genes for clinical features excepting standard lissencephaly	[635]
NIPBL	5	Cornelia de Lange syndrome	[672]
NKAIN2	6	Syndrome w/ distinctive facial features	[673]
NLGN1	3	Implicated as causal in Dup(3q) syndrome	[674]
NOG	17	Possible promoters of vertebrate head muscle development, mutated in facio-audio-symphalangism syndrome	[658, 675]

Continued on the next page

–continued from the previous page

Gene	Chromosome	Relevance	Reference
NOTCH2	1	Alagille syndrome	[665]
NRXN1	2	Unusual facial appearance	[676]
NSD1	5	Sotos syndrome	[677]
PAX3	2	Waardenburg syndrome type 1, mutated in syndromes that can display facial abnormalities	[678, 545]
PAX7	1	Signals important for craniofacial development	[99]
PAX9	14	Signals important for craniofacial development	[99]
PCDH18	4	Dysmorphic features with mutation	[679]
PCNT	21	Seckel syndrome, related to ATR-dependent DNA damage signalling	[680]
PDE4D	5	Acrodysostosis	[681]
PEX16	11	P11pDS candidates	[619]
PHF6	X	Borjeson-Fossman-Lehrmann syndrome	[682]
PIGV	1	Syndrome with distinctive facial features	[683]
PITPNA	17	Miller-Dieker syndrome, candidate genes for clinical features excepting standard lissencephaly	[635]
PITX1	5	Signals important for craniofacial development	[99]
PITX2	4	Axenfeld-Rieger syndrome/Peters' anomaly	[639]
PLD5	1	Minimal area for 1qter microdeletion syndrome	[616]
PMP22	17	Charcot-Marie-Tooth 1A	[684]
PPPDE1	1	minimal area for 1qter microdeletion syndrome	[616]
PRDX1	1	Signals important for craniofacial development	[99]
PRKAR1A	17	Acrodysostosis	[681]
PRKD1	14	Mild facial dysmorphia, 14q12del	[655]
PRPF8	17	Miller-Dieker syndrome, candidate genes for clinical features excepting standard lissencephaly	[635]
PTCH1	9	Gorlin syndrome, altered expression in frontonasal ectoderm and mesenchyme w/ Shh perturbation in model organisms	[97]
PTPN11	12	Noonan syndrome, Leopard syndrome	[685, 686]
RAD21	8	Cohesinopathy, similar to CdLS	[687]
RAI1	17	Smith-Magenis syndrome	[688]

Continued on the next page

–continued from the previous page

Gene	**Chromosome**	**Relevance**	**Reference**
RARA	17	Signals important for craniofacial development	[99]
RCAN1	21	May be related to facial features of Down syndrome (murine model)	[689]
RET	10	Multiple endocrine adenomatosis type 2B	[545]
RILP	17	Miller-Dieker syndrome, candidate genes for clinical features excepting standard lissencephaly	[635]
ROR2	9	Robinow syndrome	[690]
RPS6KA3	X	Coffin-Lowry syndrome	[691]
SALL1	16	Goldenhar syndrome-like phenotype	[692]
SDCCAG8	1	Minimal area for 1qter microdeletion syndrome	[616]
SETBP1	18	Schinzel-Giedion syndrome	[693]
SH3BGR	21	In a Down syndrome facial feature region	[694]
SHANK3	22	22q13.3 del syndrome	[695]
SHH	7	Influences pattern and growth of upper face in model vertebrates	[652, 653]
SHOC2	10	Cardio-facio-cutaneous syndrome	[624]
SKI	1	May contribute to 1p36 del syndrome	[696]
SMC1A	X	Cornelia de Lange syndrome	[672]
SMC3	10	Cornelia de Lange syndrome	[697]
SMYD3	1	Minimal area for 1qter microdeletion syndrome	[616]
SNAI2	8	Mutated in syndromes that can display facial abnormalities	[644]
SOX10	22	Mutated in syndromes that can display facial abnormalities	[644]
SOX2	3	Required for craniofacial morphogenesis	[698]
SOX3	X	Required for craniofacial morphogenesis	[698]
SOX9	17	Camponelic dysplasia	[699]
SPRY2	13	Assoc. w/cleft palate, exp.shown to be involved in frontonasal development	[700, 701]
STH	17	Candidates for 17q21.31 microdeletion syndrome	[625]
STK11	19	Syndrome w/ distinctive facial features	[659]
SUMO1	2	Haploinsufficiency results in cleft lip/palate	[702]
TBCE	1	Kenny-Caffey syndrome	[703]
TBX1	22	22q11DS	[704]
TBX22	X	Signals important for craniofacial development	[99]

Continued on the next page

–continued from the previous page

Gene	Chromosome	Relevance	Reference
TCF4	18	Pitt-Hopkins syndrome	[705]
TCOF1	5	Treacher Collins syndrome	[545, 706]
TFAP2A	6	Branchio-oculo-facial syndrome	[707]
TGFB2	1	Signals important for craniofacial development	[99]
TGFB3	14	Signals important for craniofacial development	[99]
TMCO1	1	Syndrome with craniofacial dysmorphia	[708]
TNKS	8	Cornelia de Lange syndrome	[709]
TP63	3	Rapp-Hodgkin syndrome	[710]
TRPS1	8	Tricho-rhino-phalangeal syndrome (cf. Lagner-Giedion syndrome)	[650]
TWIST1	7	Saethre-Chotzen syndrome	[91]
TWIST2	2	Setleis syndrome	[711]
UBE3A	15	Angelman syndrome	[712]
VPS13B	8	Cohen syndrome, causative	[713]
WNT5A	3	Signals important for craniofacial development	[99]
WRN	8	Werner syndrome	[714]
YWHAE	17	Miller-Dieker syndrome, candidate genes for clinical features excepting standard lissencephaly	[635]
ZBTB24	6	Immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome	[715]
ZEB2	2	Mowat-Wilson syndrome, Hirschsprung syndrome	[716, 717]
ZIC2	13	Holoprosencephaly	[718]
ZNF238	1	Minimal area for 1qter microdeletion syndrome	[616]

Appendix I

Frequency gradients maps of Britain for mitochondrial haplogroups

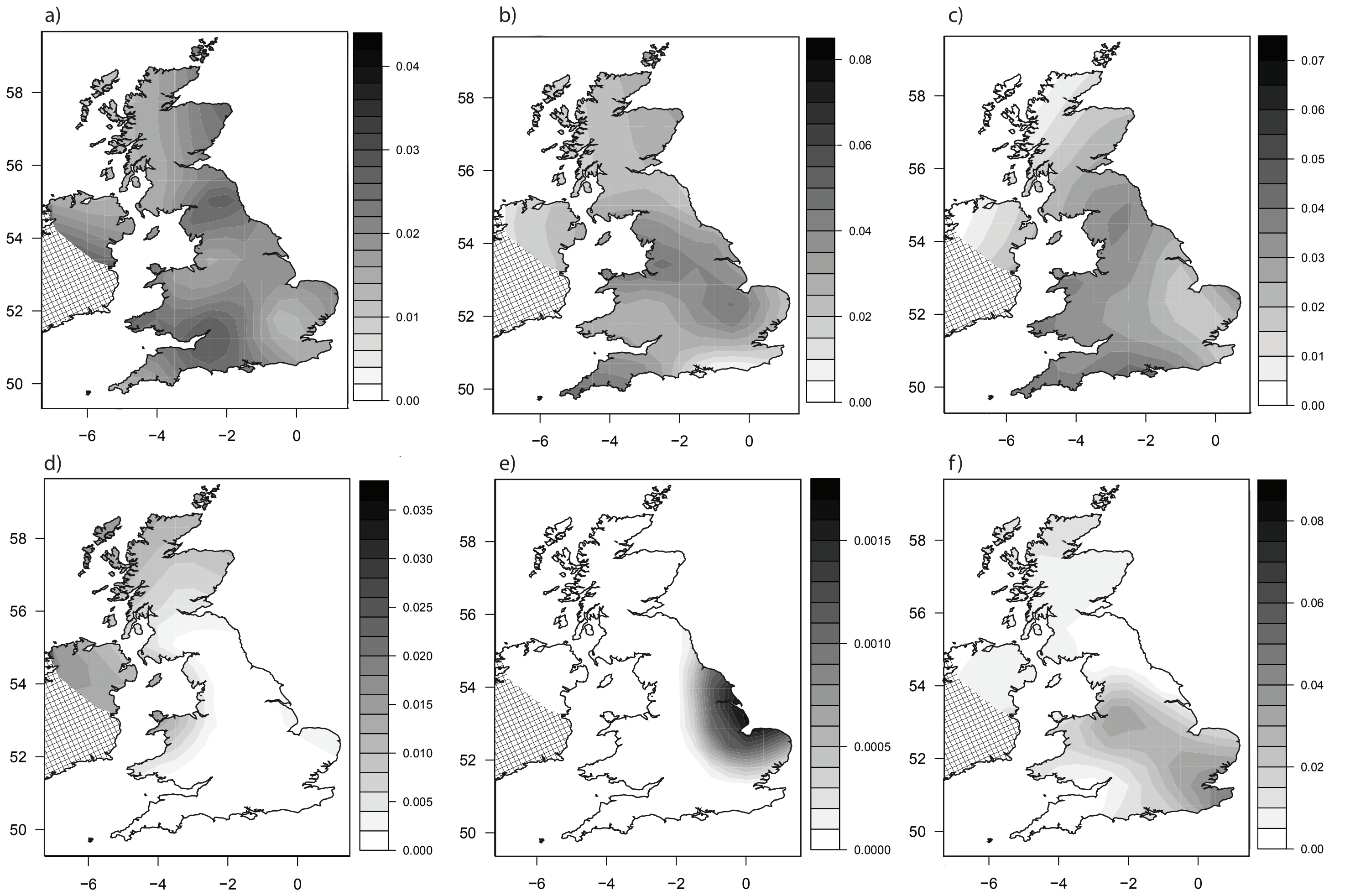


Figure I.1: **Mitochondrial haplogroup frequency distributions** a) H2, b) H5, c) H6a, d) H8, e) H9, and f) H10. Note that the frequency scale varies.

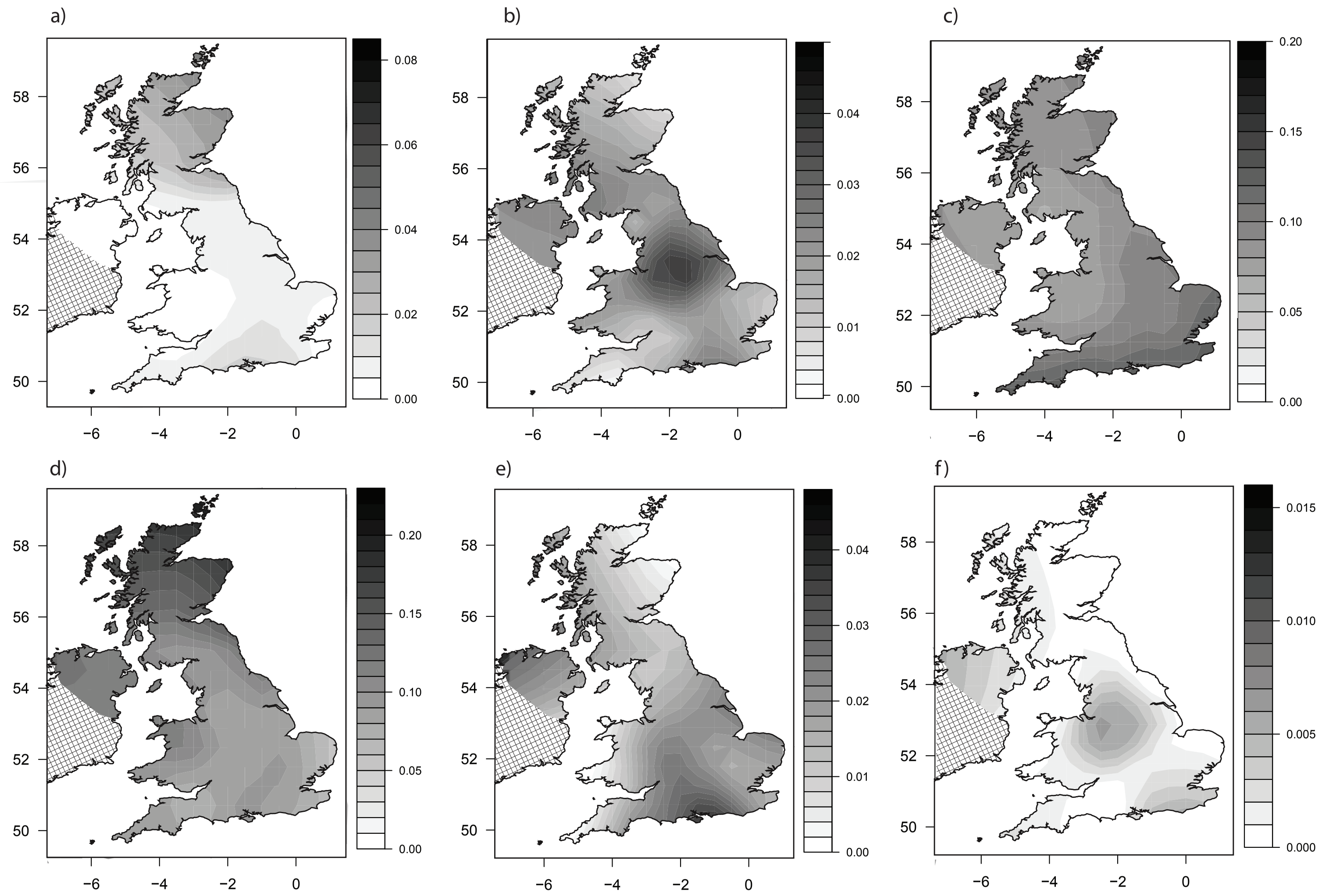


Figure I.2: **Mitochondrial haplogroup frequency distributions** a) H13, b) HV*(xH,V), c) H*(x1-5,6a,7a,8-10,13), d) J1, e) J2, and f) JT*(xJ,T). Note that the frequency scale varies.

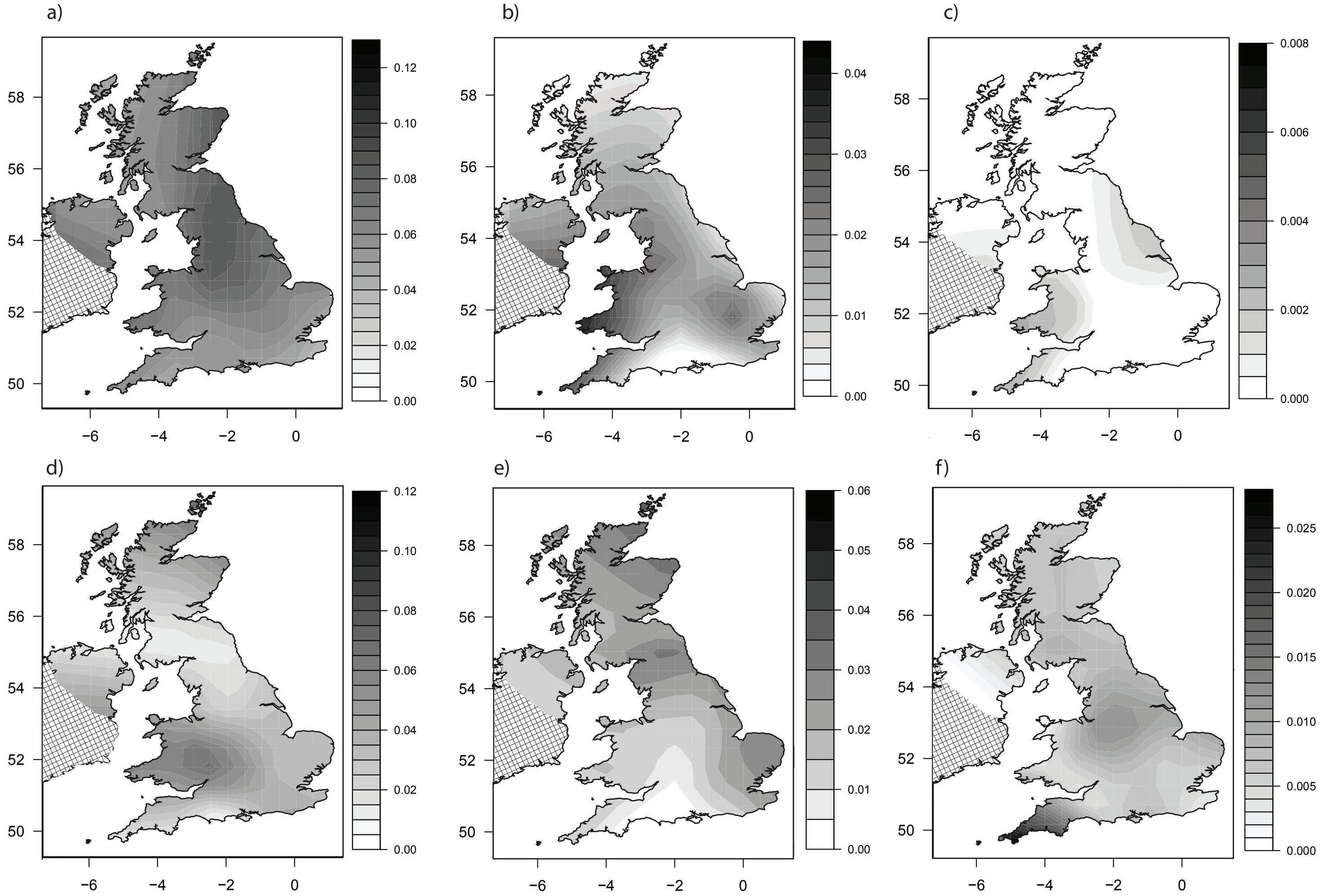


Figure I.3: **Mitochondrial haplogroup frequency distributions** a) K1, b) K2, c) K*(x1,2), d) N1e/I, e) T1, and f) Other and unresolved. Note that the frequency scale varies.

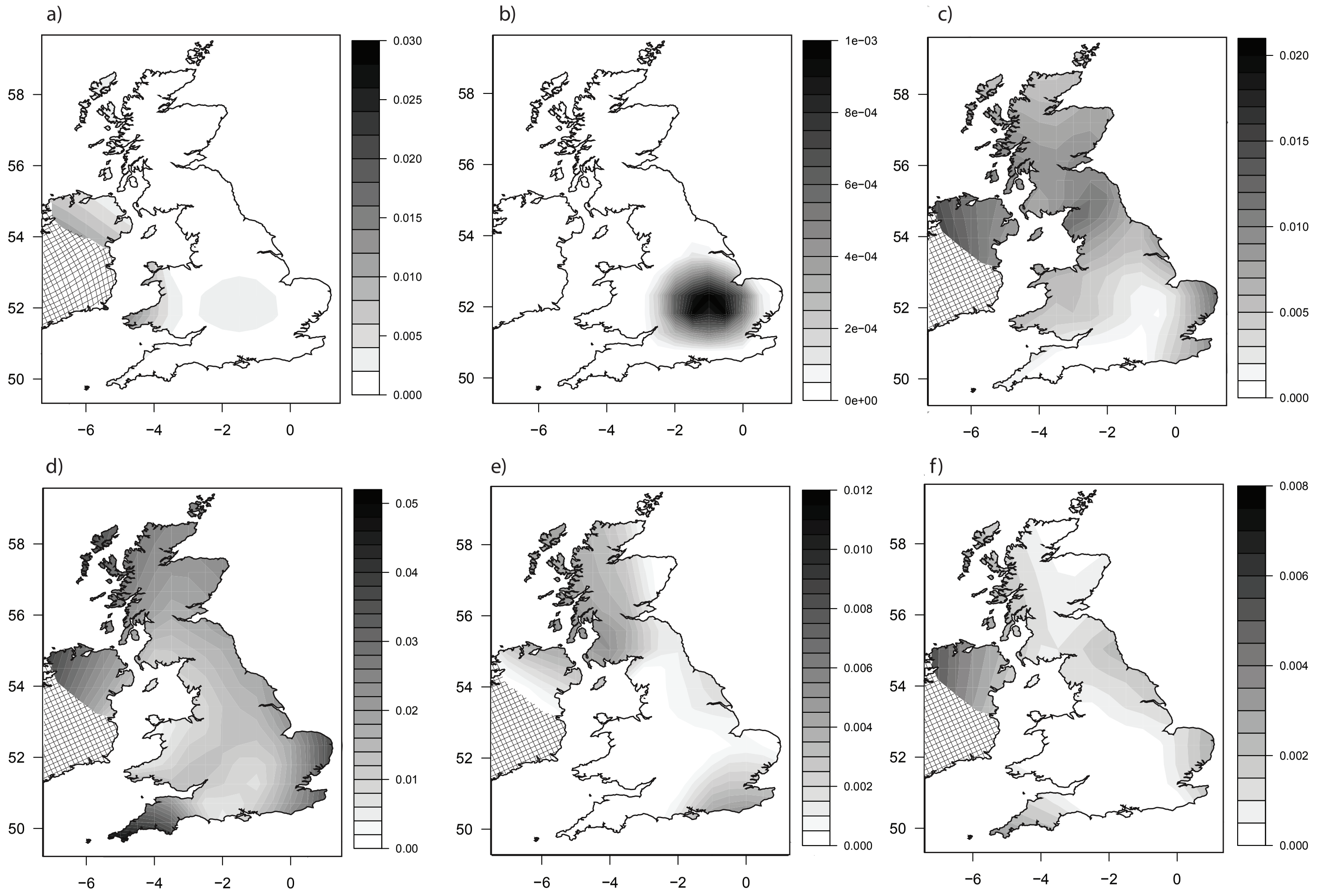


Figure I.4: **Mitochondrial haplogroup frequency distributions** a) T*(x1,2), b) U1, c) U3, d) U4, e) U5*(xa,b), and f) U6. Note that the frequency scale varies.

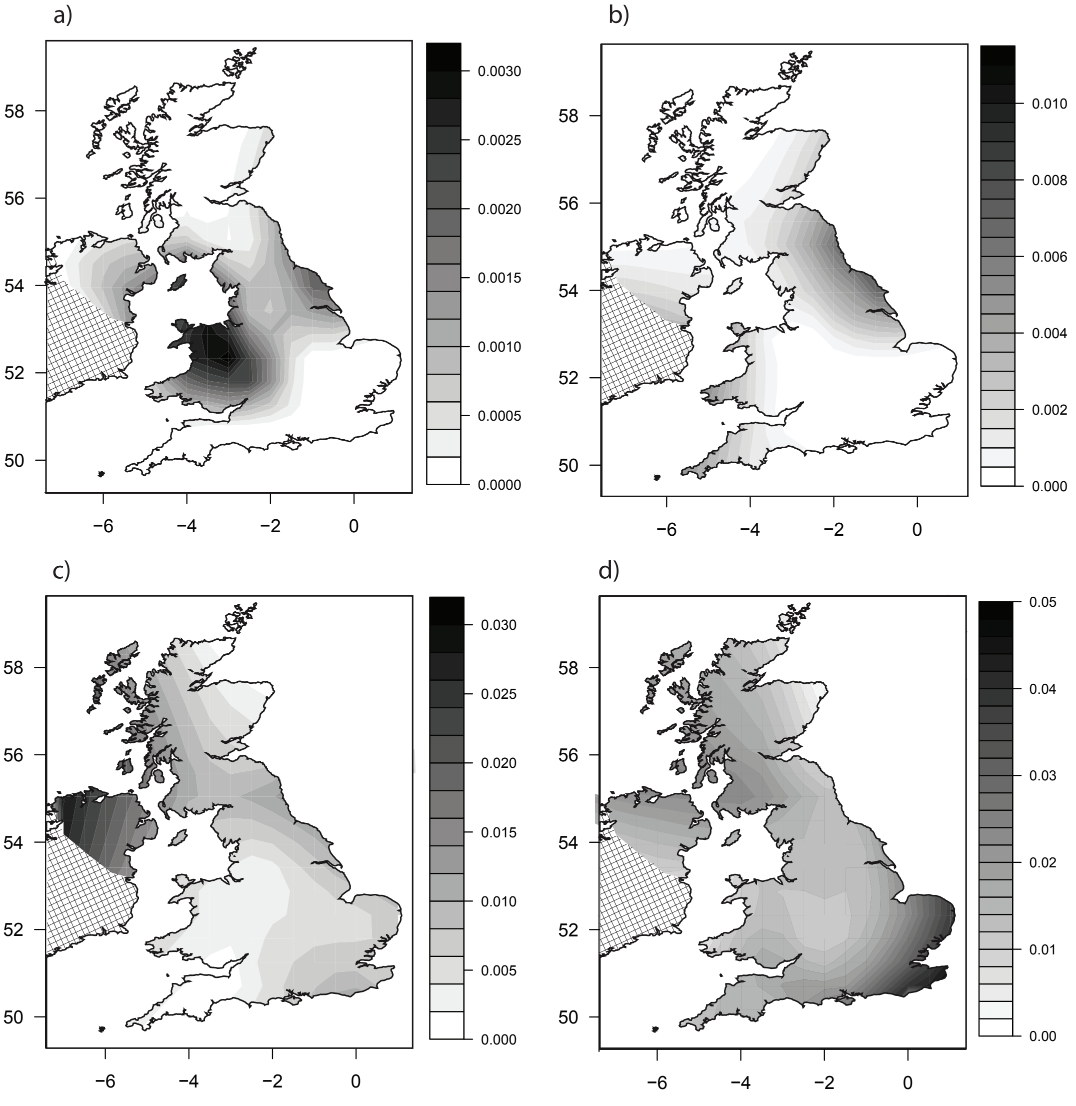


Figure I.5: **Mitochondrial haplogroup frequency distributions** a) U7, b) U8*(xK), c) U*(x1,3-8), d) W. Note that the frequency scale varies.

Appendix J

Regional mitochondrial haplogroup frequencies in Norway

Table J.1: N= the number of samples in each cohort

	H*	**H1**	**H13**	**H2**	**H3**	**H4**	**H5**	**H6a**	**HV***	**J1**	**J2**	**N=**
Nordfjord	2	8	0	3	1	0	2	0	0	5	2	37
Ottadalen	5	8	1	1	2	0	1	1	0	1	1	25
Setesdalen	0	1	0	2	0	1	10	0	0	2	0	22

	K2	**other**	**T1**	**T2**	**U4**	**U5***	**U5a**	**U5b**	**U8***	**V**	**W**	**N=**
Nordfjord	1	0	0	4	0	0	6	2	0	0	1	37
Ottadalen	0	0	0	0	0	0	1	0	1	2	0	25
Setesdalen	0	3	1	0	1	1	0	0	0	0	0	22

Appendix K

fineSTRUCTURE Analysis

K.1 fineSTRUCTURE Analysis

This analysis was performed by G. Hellenthal, and is provisional for a forthcoming paper from the PoBI project and its collaborators. The output provided to the author of this thesis was numerical data defining clusters, their individual constituents, cluster branching order, and copying proportions by individual.

The Illumina 1 million SNP data was analyzed with a modified variant of the STRUCTURE program, fineSTRUCTURE, which takes linkage into account, and thereby can pick out subtler substructure [719]. Each PoBI individual was also expressed in terms of chromosomal sections' similarity to those found in other European populations, *copying proportions*. European populations were represented by those used in Sawcer *et al.* [720].

Output used from this analysis was the clustering of PoBI individuals based on genetic profile alone (not geographic information), the sequential binary splitting of these clusters from the PoBI population as a whole, and the copying proportions for each PoBI individual from non-British European populations.

K.2 fineSTRUCTURE Results - Provisional

The fineSTRUCTURE analysis identified 72 genetic clusters in Britain, mostly tightly linked to the geographic provenance. By constraining the the program to increasing numbers from one and with increments of one, and approximation of a binary branching tree was acquired (see Figure K.1). In so far as this procedure indicates relative similarity of the clusters, Orkney is the most different from the rest of Britain, followed by Wales, northern Britain (and Ulster), Cornwall, the Welsh border areas of England, Devon, and a cluster comprising about half of northern Lowland Britain.

These clusters, at chosen points of the branching sequence, were used as a basis for comparing classical markers, lactose intolerance, pigmentation and facial feature candidate genes (see Table K.1).

Table K.1: **Cluster-based grouping**

Cluster	**Major constituents**
COR	89% of Cornish samples
DEV	84% of Devon samples
ENG	Majority of all Lowland British samples
NENG	86% of Cumbrian, and 62% of northeast English samples
ORK	99% of Orcadian samples
SCO	98% of eastern Scottish, 100% of western Scottish, 100% of northern Irish, 9% of northeast English, and 6% of Cumbrian samples
WAL	100% of south Welsh, 99% of north Welsh, and 15% of Herefordshire samples
WEM	98% of Forest of Dean, 78% of Herefordshire, 33% of Worcestershire samples 30% of Cheshire, 19% of Lancashire, and 17% of Gloucestershire samples

Figure K.1: **Sequential cluster branching** Panels a-e) represent a) the starting situation of all individuals in a single cluster, b-e) successive binary branching. For all panels, the regions shaded in red represent those that, in the given branching step, are to a large extent – sometimes wholly – removed from the residual population. These regions are shaded a different colour in the next step. For each step, the percentage of the regional population(s) belonging to the relevant cluster is given. The top right number in each panel gives the branching step.

The copying proportions from European populations - indicating levels of genetic similarity - showed only six groups that contributed significantly to makeup of the British population. By nation, these were Belgium, Denmark, France, Germany, Ireland and Norway. The contributions from Spain, Italy, Sweden, Poland, Finland each amounted to a maximum of 0.1-3%. Within Britain, the distribution of copying proprtions from the major contributors varies considerably (see Figure K.2). Similarities to Belgians, Danes and Germans all show the same pattern, peaking at between 13-25% in the south and fading towards the north and west, to as low as ca. 5% in northwest Scotland. Similarities to the French were more uniform, but with a peak of 35% in western Wales and the Cornwall, and elsewhere largely between 15-20%, with 25-26% in Devon and the Forest of Dean. Similarities to the Irish was greatest in Ulster and western Scotland, up to 57%, diminishing to around low 20% in Lowland Britain. Norwegian copying proportions of up to 19% are seen in Orkney, but the effect tails off very rapidly and lies at around 3-5% everywhere else.[1]

The European samples themselves were subject to fineSTRUCTURE clustering, and the copying proportions from major contributors to Britain based on these clusters was assessed (see Figure K.3). The overall patterns where the same as for the nation-based groupings above. The French contribution was largely made up of a cluster consisting of people from the northwest of France and the Paris region (the northeast of France was unfortunately not represented in the French cohort). The German contribution was accounted for by both a largely southern and a largely northern cluster, of broadly equal magnitude, but with the southern group having a slightly larger input with a maximum of 18% versus 10%. The Norwegian contribution came from three different clusters, and all have their frequency peak in Orkney. Minor contributions, 2-3.5%, are provided by clusters consisting primarily of Norwegians from the southeast and north-central parts of the country. The major contribution, at around 8% in Orkney, is dervied from western Norwegians. For Belgium and Denmark there were no sub-clusters, and

---

[1] Note that, as we have no samples from northern Scotland, higher Norwegian influence seen there may well be an artefact of the regression method used to create the map.

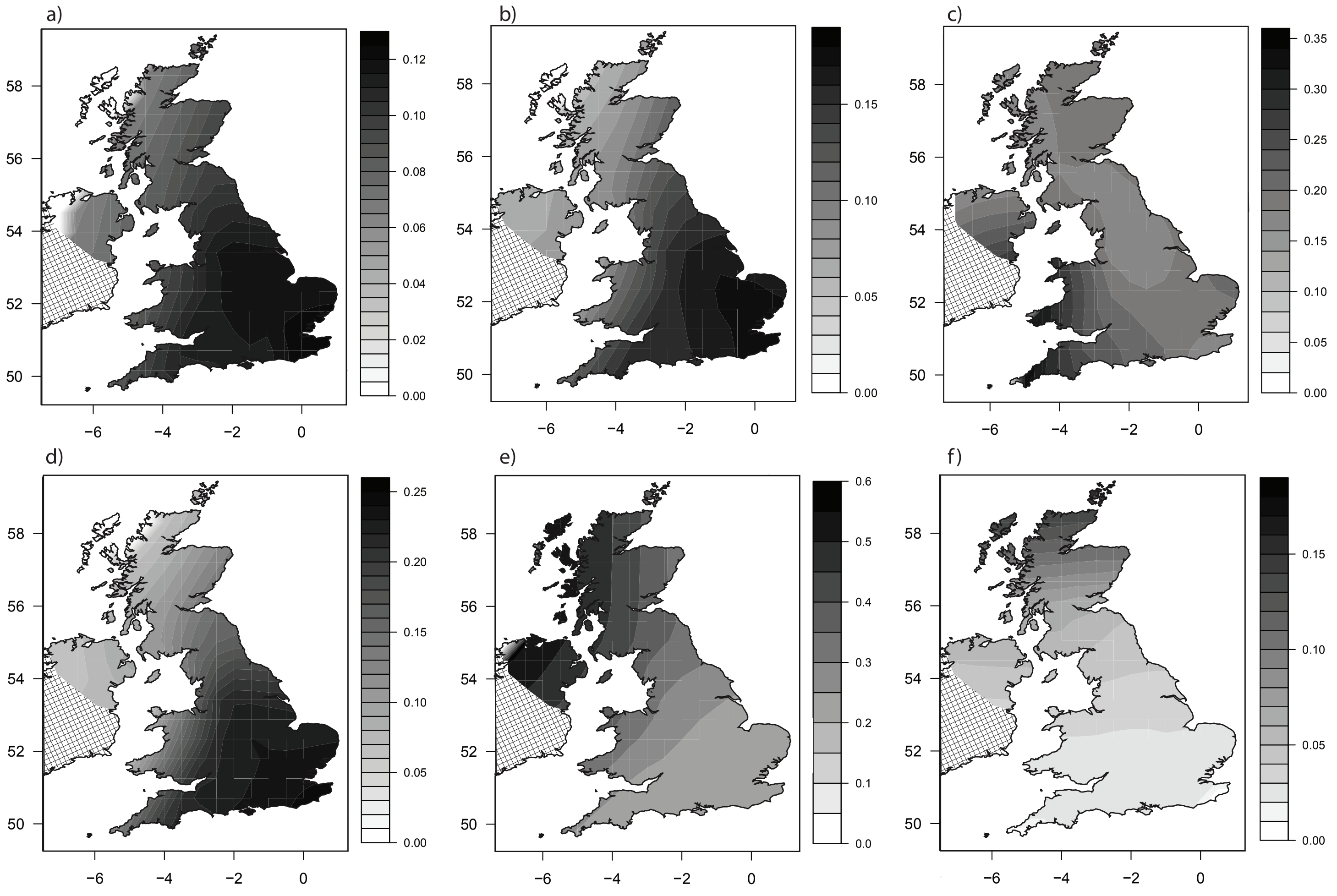


Figure K.2: **Major European chromosome copying proportions** Maps detailing the copying proportions in Britain from a) Belgium, b) Denmark, c) France, d) Germany, e) Ireland, and f) Norway. Note that the frequency scale varies.

the nation/cluster contributions were therefore the same.

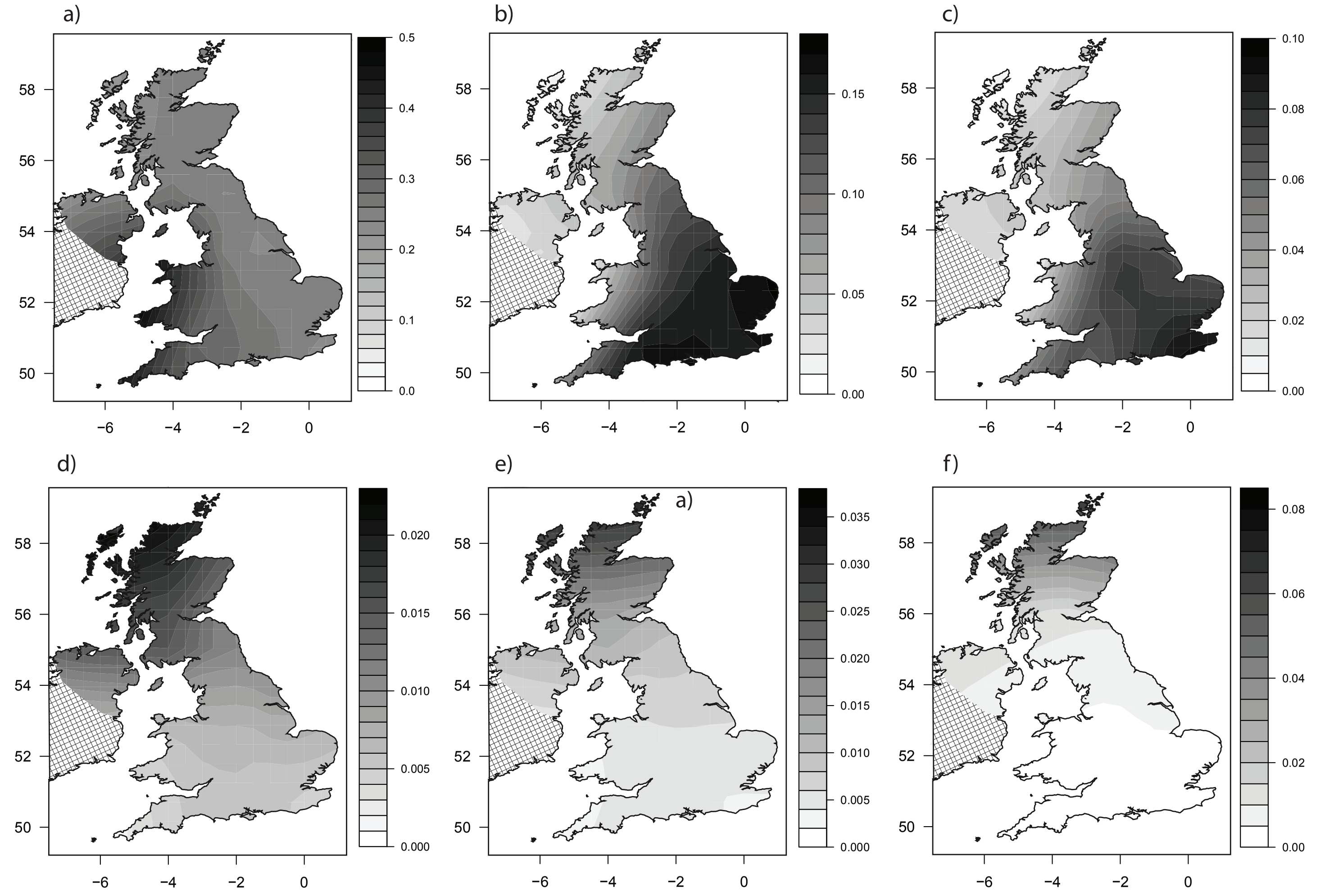


Figure K.3: **Major chromosome copying proportions by European cluster** Maps detailing the copying proportions in Britain from a) Mostly northwest France/Paris region, b) Mostly southern Germany, c) Mostly northern Germany, d) Mostly southeast Norway, e) Mostly north-central coastal Norway, and f) Western Norway. Note that the frequency scale varies.

Appendix L

Facial Candidate Genes SNPs on the Illumina 1M-Duo Array

Table L.1: NS-C – non-synonymous coding, S-C – synonymous coding, Near 5'/3' – within 2000bp of transcribed region.

Gene	**SNPs**	**NS-C**	**S-C**	**Intron**	**5'-UTR**	**3'-UTR**	**Near 5'**	**Near 3'**
ABCC9	102	2	5	90	0	1	2	2
ACTA1	2	0	0	1	0	0	1	0
ADSS	21	1	0	19	0	1	0	0
AKT3	156	0	0	154	0	0	0	2
ALCAM	72	4	4	60	1	2	1	0
ALDH1A2	128	1	2	114	0	10	0	1

Continued on the next page

– continued from the previous page

Gene	**SNPs**	**NS-C**	**S-C**	**Intron**	**5’-UTR**	**3’-UTR**	**Near 5’**	**Near 3’**
ALX3	8	0	0	6	0	0	1	1
ALX4	37	1	2	26	0	8	0	0
ARSE	7	0	0	7	0	0	0	0
ASXL1	40	2	0	32	0	2	3	1
ATRX	3	0	0	3	0	0	0	0
B4GALT7	9	0	1	5	0	0	1	2
BCOR	46	1	2	42	0	0	1	0
BLM	61	7	7	46	0	0	1	0
BMP2	10	1	0	3	1	2	3	0
BMP4	3	1	0	1	0	0	1	0
BMP7	50	0	0	46	0	3	0	1
BRAF	83	0	0	81	0	0	2	0
C17orf69	4	0	4	0	0	0	0	0
C1orf100	28	0	0	28	0	0	0	0
C1orf101	51	0	2	47	0	0	2	0
CACNA1C	317	1	3	310	0	1	2	0
CASK	194	2	1	178	0	5	5	3
CBLB	71	1	4	64	0	1	0	1
CBP	63	1	3	56	1	2	0	0
CD300A	14	0	0	11	0	2	1	0
CDKN1C	7	0	1	0	1	1	3	1
CHD7	60	0	0	59	0	0	1	0
CNTN4	634	0	4	617	3	3	6	1

Continued on the next page

– continued from the previous page

Gene	**SNPs**	**NS-C**	**S-C**	**Intron**	**5'-UTR**	**3'-UTR**	**Near 5'**	**Near 3'**
COL11A1	111	4	3	103	0	1	0	0
COL11A2	52	5	4	32	1	4	4	2
COL1A1	17	3	1	10	0	3	0	0
COL2A1	34	3	4	25	0	0	1	1
COL9A1	86	3	3	75	0	1	3	1
CRHR1	22	1	1	19	0	1	0	0
CRK	15	0	0	14	0	0	0	1
CRLF1	15	1	2	10	0	0	2	0
CTBP2	115	4	2	104	1	0	4	0
CTNNB1	27	1	5	17	0	3	1	0
CTSK	8	0	1	3	0	2	1	1
CYP1B1	27	6	1	5	2	11	0	2
DHCR7	11	1	2	6	0	2	0	0
DHODH	10	1	0	8	0	0	0	1
DLX1	6	1	0	1	0	3	0	1
DLX2	1	0	0	0	0	1	0	0
DLX3	5	0	1	1	1	2	0	0
DLX5	5	0	0	2	0	0	1	2
DLX6	3	0	0	0	0	2	1	0
DNASE1	16	7	3	0	1	0	2	3
DNMT3B	27	0	2	19	0	2	4	0
EDA	91	0	0	86	2	1	2	0
EDN1	14	2	1	6	1	1	1	2

Continued on the next page

– continued from the previous page

Gene	SNPs	NS-C	S-C	Intron	5'-UTR	3'-UTR	Near 5'	Near 3'
EDN3	26	2	0	11	1	10	2	0
EDNRA	46	0	3	31	1	10	1	0
EDNRB	40	5	2	30	1	1	0	1
EFNB1	8	2	1	3	0	1	1	0
EHMT1	57	1	4	50	0	0	2	0
ELN	22	1	1	19	0	1	0	0
ENPP1	60	2	3	44	0	9	2	0
EPM2A	0	0	0	0	0	0	0	0
ERCC3	24	2	2	17	0	2	1	0
ERCC5	42	12	3	20	2	3	2	0
ERCC6	57	12	5	29	4	3	3	1
ERCC8	32	0	0	26	0	4	2	0
ESCO2	15	1	0	6	2	0	6	0
EVC	165	9	6	141	0	6	3	0
EVC2	76	2	3	67	0	1	3	0
EXT1	133	0	1	132	0	0	0	0
EXT2	44	1	2	37	1	0	2	1
EZH2	19	0	0	19	0	0	0	0
FGD1	13	0	1	9	0	2	0	1
FGF4	2	0	0	1	0	1	0	0
FGF8	5	0	0	4	0	0	1	0
FGF9	17	0	1	10	0	4	2	0
FGFR1	41	0	2	31	0	6	2	0

Continued on the next page

– continued from the previous page

Gene	**SNPs**	**NS-C**	**S-C**	**Intron**	**5'-UTR**	**3'-UTR**	**Near 5'**	**Near 3'**
FGFR2	63	1	2	53	1	2	3	1
FGFR3	26	2	7	6	0	8	2	1
FLNB	85	4	3	72	0	6	0	0
FMR1	19	0	1	15	0	2	1	0
FOXC1	4	1	2	0	0	1	0	0
FOXG1	2	0	0	0	0	1	1	0
FOXL2	3	0	0	0	0	1	2	0
FREM1	151	5	8	137	1	0	0	0
FRZB	18	0	0	15	0	2	1	0
FSTL3	7	0	0	1	0	2	3	1
GHR	87	6	2	74	0	0	3	2
GJA1	4	0	0	2	0	0	2	0
GLI1	11	3	1	5	0	0	1	1
GLI2	78	3	1	71	0	0	3	0
GLI3	130	3	0	121	0	3	1	2
GPC3	202	2	2	180	0	0	9	9
GRB10	98	0	3	90	0	2	2	1
GSC	2	0	0	1	0	0	0	1
H19	98	2	1	23	0	0	3	1
HRAS	3	0	0	2	0	0	1	0
IGF2	54	5	1	25	15	1	0	7
IGFBP1	6	2	1	0	0	1	2	0
IMP5	14	8	1	0	0	2	3	0

Continued on the next page

– continued from the previous page

Gene	SNPs	NS-C	S-C	Intron	5’-UTR	3’-UTR	Near 5’	Near 3’
INPP5K	11	0	0	5	1	2	2	1
IRF6	11	1	1	7	1	0	1	0
JAG1	27	2	4	18	0	2	1	0
KAT6B	46	0	0	46	0	0	0	0
KCNH2	18	2	2	10	0	2	2	0
KCNJ2	7	0	1	2	0	2	1	1
KCNQ1	202	4	3	190	0	2	3	0
KCNQ1OT1	24	0	1	23	0	0	0	0
KIAA1267	44	4	3	33	0	4	0	0
KRAS	21	0	0	16	0	4	0	1
LETM1	19	1	1	16	0	1	0	0
LHX8	19	0	1	15	0	0	2	1
MAP2K1	25	0	0	24	0	0	1	0
MAP2K2	17	0	0	16	0	0	0	1
MAPK8IP1	6	1	0	2	0	0	2	1
MAPT	50	2	3	37	2	2	2	2
MITF	30	0	0	24	0	2	4	0
MLL2	22	2	4	9	0	2	4	1
MSX1	3	0	0	0	0	1	2	0
MSX2	6	1	0	2	0	2	1	0
MYBPC3	23	11	3	8	0	1	0	0
MYO1C	16	1	4	7	0	0	2	2
NIPBL	82	3	2	77	0	0	0	0

Continued on the next page

– continued from the previous page

Gene	**SNPs**	**NS-C**	**S-C**	**Intron**	**5’-UTR**	**3’-UTR**	**Near 5’**	**Near 3’**
NKAIN2	496	0	0	492	1	1	1	1
NLGN1	332	0	2	326	3	1	0	0
NOG	2	0	0	0	0	2	0	0
NOTCH2	76	4	2	59	0	9	1	1
NRXN1	469	0	1	466	0	1	1	0
NSD1	72	4	3	58	0	6	1	0
PAX3	56	0	0	49	0	2	4	1
PAX7	65	0	3	59	1	0	2	0
PAX9	21	0	0	15	0	4	2	0
PCDH18	8	0	1	4	0	1	1	1
PCNT	52	11	4	35	1	0	1	0
PDE4D	612	0	0	606	0	0	6	0
PEX16	5	1	0	1	1	1	0	1
PHF6	24	0	0	20	0	1	2	1
PIGV	7	0	1	4	0	1	1	0
PITPNA	19	0	0	16	0	3	1	0
PITX1	4	0	1	3	0	0	0	0
PITX2	14	1	0	8	0	1	3	1
PLD5	210	0	0	206	1	2	1	0
PMP22	25	0	0	20	0	3	2	0
PPPDE1	15	0	0	11	0	2	2	0
PRDX1	7	0	0	6	1	0	0	0
PRKAR1A	10	0	0	4	1	3	2	0

Continued on the next page

– continued from the previous page

Gene	**SNPs**	**NS-C**	**S-C**	**Intron**	**5'-UTR**	**3'-UTR**	**Near 5'**	**Near 3'**
PRKD1	105	0	1	103	0	1	0	0
PRPF8	22	0	4	15	0	1	2	0
PTCH1	38	0	5	30	0	1	2	0
PTPN11	40	0	0	39	0	0	1	0
RAD21	35	0	1	29	0	1	4	0
RAI1	56	2	5	45	0	2	0	2
RARA	0	0	0	0	0	0	0	0
RCAN1	54	0	0	50	0	1	3	0
RET	26	4	2	15	0	4	1	0
RILP	5	1	0	3	0	1	0	0
ROR2	101	0	3	96	0	1	1	0
RPS6KA3	56	0	1	51	0	4	0	0
SALL1	3	0	0	1	0	1	1	0
SDCCAG8	79	1	1	74	0	0	2	1
SETBP1	148	2	2	143	0	1	0	0
SH3BGR	30	2	0	24	0	2	2	0
SHANK3	23	1	0	18	0	1	0	3
SHH	8	0	1	2	1	0	3	1
SHOC2	25	0	0	21	0	1	3	0
SKI	32	0	0	30	0	2	0	0
SMC1A	19	1	1	13	1	1	2	0
SMC3	15	0	0	14	0	0	1	0
SMYD3	291	0	1	288	0	1	1	0

Continued on the next page

– continued from the previous page

Gene	SNPs	NS-C	S-C	Intron	5'-UTR	3'-UTR	Near 5'	Near 3'
SNAI2	5	2	0	2	0	0	0	1
SOX10	7	0	1	4	0	2	0	0
SOX2	1	0	0	0	0	0	1	0
SOX3	1	0	0	0	0	0	1	0
SOX9	6	0	1	0	1	2	2	0
SPRY2	8	2	0	1	1	2	0	2
STH	2	0	0	0	0	0	2	0
STK11	12	0	0	12	0	0	0	0
SUMO1	14	0	0	12	0	1	1	0
TBCE	25	2	0	20	0	2	0	1
TBX1	13	1	1	8	0	1	1	1
TBX22	6	0	0	5	0	0	1	0
TCF4	182	0	0	174	0	3	1	0
TCOF1	22	2	1	17	0	0	1	1
TFAP2A	7	0	1	3	1	1	1	0
TGFB2	51	1	1	45	1	2	1	0
TGFB3	51	1	2	42	1	5	0	0
TMCO1	15	0	0	10	0	5	0	0
TNKS	74	0	0	62	0	7	5	0
TP63	138	0	0	132	0	1	2	3
TRPS1	80	2	1	71	1	5	0	0
TWIST1	3	0	0	1	0	2	0	0
TWIST2	1	0	0	1	0	0	0	0

Continued on the next page

– continued from the previous page

Gene	**SNPs**	**NS-C**	**S-C**	**Intron**	**5'-UTR**	**3'-UTR**	**Near 5'**	**Near 3'**
UBE3A	28	0	0	22	0	1	5	0
VPS13B	404	3	2	393	0	2	1	3
WNT5A	9	0	0	7	0	1	0	1
WRN	80	14	4	54	3	3	2	0
YWHAE	21	0	0	19	0	2	0	0
ZBTB24	16	1	0	8	0	5	0	2
ZEB2	50	0	1	45	0	1	2	1
ZIC2	3	0	0	2	0	1	0	0
ZNF238	6	1	1	0	0	1	2	1