MARI JÄRVE

Different genetic perspectives on human
history in Europe and the Caucasus:
the stories told by uniparental
and autosomal markers

Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu, Estonia

The dissertation was accepted for the commencement of the degree of Doctor of Philosophy (in Molecular Biology) on April 26th, 2012, by the Council of the Institute of Molecular and Cell Biology, University of Tartu.

Supervisors: Professor Richard Villems, Dr.Sci., M.D., Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu and the Estonian Biocentre, Tartu, Estonia; Siiri Rootsi, Ph.D., Estonian Biocentre, Tartu Estonia.

Opponent: Professor Antti Sajantila, Ph.D., M.D., Hjelt Institute, Department of Forensic Medicine, Helsinki University, Helsinki, Finland.

Commencement: Room No 217, 23 Riia St, Tartu, on June 11th, 2012, at 13:00.

European Union
European Social Fund

Investing in your future

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ORIGINAL PUBLICATIONS

I. Myres NM, Rootsi S, Lin AA, **Järve M**, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Pshenichnov A, Yunusbayev B, Balanovsky O, Balanovska E, Rudan P, Baldovic M, Herrera RJ, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. (2011). A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur J Hum Genet* 19, 95–101.

II. Underhill PA, Myres NM, Rootsi S, Metspalu M, Zhivotovsky LA, King RJ, Lin AA, Chow CET, Semino O, Battaglia V, Kutuev I, **Järve M**, Chaubey G, Ayub Q, Mohyuddin A, Mehdi SQ, Sengupta S, Rogaev EI, Khusnutdinova EK, Pshenichnov A, Balanovsky O, Balanovska E, Jeran N, Havas Augustin D, Baldovic M, Herrera RJ, Thangaraj K, Singh V, Singh L, Majumder P, Rudan P, Primorac D, Villems R, Kivisild T. (2010). Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *Eur J Hum Genet* 18, 479–484.

III. Yunusbayev B*, Metspalu M*, **Järve M***, Kutuev I, Rootsi S, Metspalu E, Behar DM, Varendi K, Sahakyan H, Khusainova R, Yepiskoposyan L, Khusnutdinova EK, Underhill PA, Kivisild T, Villems R. (2012). The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol* 29, 359–365.
    *equal contribution

IV. **Järve M**, Zhivotovsky LA, Rootsi S, Help H, Rogaev EI, Khusnutdinova EK, Kivisild T, Sanchez JJ. (2009). Decreased Rate of Evolution in Y Chromosome STR Loci of Increased Size of the Repeat Unit. *PLoS ONE* 4, e7276.

My contributions to the listed articles referred to in the present thesis are as follows:

Ref. I – performed experiments and statistical analysis of the data, participated in the writing of the paper.

Ref. II – performed experiments.

Ref. III – performed experiments and statistical analysis of the data, wrote the paper with the contribution of other co-authors.

Ref. IV – had a role in the design of the study, performed experiments and statistical analysis of the data, wrote the paper with the contribution of other co-authors.

ABBREVIATIONS

AMH	anatomically modern human(s)
bp	base pair(s)
GWA	genome-wide association
HVS	hypervariable segment of mtDNA
kb	thousand (kilo-) base pairs
ky(a)	thousand (kilo-) years (ago)
LD	linkage disequilibrium
LGM	Last Glacial Maximum
Mb	million (Mega-) base pairs
MRCA	most recent common ancestor
MSY	male-specific region of the Y chromosome
mtDNA	mitochondrial DNA
Mya	million (Mega-) years ago
N_e	effective population size
NRY	non-recombining region of the Y chromosome
RFLP	restriction fragment length polymorphism
SNP	single nucleotide polymorphism
STR	short tandem repeat
TMRCA	time to the most recent common ancestor
UEP	unique event polymorphism
YCC	Y Chromosome Consortium
Y-STR	Y chromosome short tandem repeat

1. INTRODUCTION

Genetic studies of human demographic history have a long tradition, one that began with studies of blood type variation and other protein polymorphisms in humans and is now largely upheld by research into the uniparentally inherited mitochondrial DNA (mtDNA) and Y chromosome. However, the field of human population genetics is currently undergoing a major phase of broadening, both in its scope and in the technologies involved. Whole genome single nucleotide polymorphism (SNP) genotyping and ancient DNA analyses have expanded the horizons of scientific enquiry. Large numbers of high quality full genome sequences are still just slightly out of reach for wide-scale population genetic studies, but the genotyping of hundreds of thousands of SNPs all over the genome has already become a routine.

Especially considering the recent developments in the field, the population genetic approach has become a powerful way of studying human history. Even a new term – archaeogenetics – has been coined by Colin Renfrew (Renfrew and Boyle 2000). The formation of the extant populations of Europe has long been a focus of attention, with numerous debates, such as those on the mode and timing of the peopling of Europe and the cultural *versus* demic diffusion of the Neolithic, still not settled despite extensive study both in the fields of archaeology and genetics. Another area of interest is the Caucasus, a region linking the Near/Middle East and the East European Plain between the Black and Caspian Seas and exhibiting high ethnic as well as linguistic diversity. So far, the role of the Caucasus in human dispersals in Eurasia, including the peopling of East Europe, has remained obscure. Indeed, one may wonder: does it link or separate the Near/Middle East and East Europe?

The present dissertation aims, firstly, to offer a brief overview of the current knowledge about genetic research into human demographic history, in particular concerning Europe and the Caucasus and, secondly, to add new insights to the understanding of the genetic structuring of western Eurasia. Our novel results about the phylogeography of Y-chromosomal, mitochondrial DNA and autosomal variation enhance the current understanding of the genetic heritage of the Caucasus populations as well as enable comparisons between the Caucasians and their neighbours, revealing the role of the Caucasus in ancient human migrations in Eurasia, including the peopling of East Europe. Furthermore, the spatial and temporal distribution of two Y chromosome haplogroups widespread in Europe and elsewhere is analysed and discussed in terms of potential Late Pleistocene and Holocene demographic history of West Eurasian populations.

2. LITERATURE OVERVIEW

2.1. The characteristics and applications of different genetic markers in studies of human ancestry

2.1.1. Classical markers

The first attempts to study human variation with genetic markers made use of polymorphisms of various blood proteins to analyse gene variation by proxy. The beginning of the 20th century saw the discovery of major blood type systems: ABO [Landsteiner (1901), cited from Cavalli-Sforza et al. (1994)] and Rh [Levine and Stetson (1939); Landsteiner and Wiener (1940), cited from Cavalli-Sforza et al. (1994)], which were among the first examples of clear-cut genetic variation not influenced by the environment and possessing a rather simple mode of hereditary transmission (ABO more so than Rh). Cavalli-Sforza and colleagues used allele frequency data for these and other blood groups as well as various hemoglobins, immunoglobulins, human leukocyte antigen histo-compatibility types and other blood proteins to compare human populations (Cavalli-Sforza et al. 1994). Analysis of the classical marker data revealed patterns such as decreased genetic variation in non-Africans, associated with the out-of-Africa migration, and a genetic cline from the Middle East towards Northwest Europe, suggested to be associated with the Neolithic agricultural dispersal (Cavalli-Sforza et al. 1994). However, the classical markers could only be a logical stepping-stone to direct DNA analysis. The latter can target different aspects of heritage, but here we discuss those directly relevant to us: variation in haploid uniparental – patrilineal Y-chromosomal and matrilineal mitochondrial DNA – markers, and the so-called high-density whole genome variation in the autosomally inherited part of the human genome, briefly mentioning also complete re-sequencing of the latter.

2.1.2. Uniparental markers

The human Y chromosome and mtDNA allow the tracing of paternal and maternal lineages, respectively, due to their mode of inheritance that differs from that of most of the genome. The Y chromosome [about 64 megabases (Mb) of DNA, of which ~41 Mb is heterochromatic and ~23 Mb euchromatic (Skaletsky et al. 2003)] is present only in males and therefore inherited from father to son. Mitochondrial DNA [a circular double stranded DNA molecule of 16,568 base pairs (bp) (Andrews et al. 1999)] is inherited from mother to offspring due to its location outside the nucleus, in the mitochondria, since only the mitochondria of the egg are transferred to the next generation, the small amount of cytoplasm from the sperm being lost in fertilization. Neither mtDNA nor, for the most part, the Y chromosome recombine in each generation, which

results in haplotypes being inherited intact, changing only through mutations accumulating with time.

There is one exception to the lack of recombination in case of the Y chromosome. Over the course of its evolution, the Y chromosome has lost its original homology with the X chromosome through a series of large-scale inversions that suppressed X–Y crossing over and resulted in the deletion of large portions of the Y chromosome (Lahn and Page 1999). However, short pseudoautosomal regions remain at the ends of the Y chromosome that are homologous with the X chromosome, enabling the pairing of the sex chromosomes during meiosis, and can, in principle, recombine [reviewed in Skaletsky et al. (2003)]. The portion of the Y chromosome that has no homology with other chromosomes (about 60 Mb) is referred to as the non-recombining region of the Y chromosome (NRY) or the male-specific region of the Y chromosome (MSY). Also, the argument has been made that gene conversion, which is a form of recombination, occurs on the Y chromosome (Rozen et al. 2003; Skaletsky et al. 2003). However, even acknowledging small exceptions, 95% of the Y chromosome does not recombine (in the classical sense, meaning crossing over between chromosome homologues), and in practice, the NRY can be considered as a single locus.

The mutation rates of both mtDNA and the Y chromosome are higher than that of the autosomal genome, which is estimated to be 10^{-8} per base pair per generation (The International HapMap Consortium 2005). Owing to the lack of a proofreading mechanism in mitochondria, mtDNA evolves more rapidly than nuclear DNA, as was observed already by Brown et al. (1979). Within mtDNA, the SNP mutation rate of the hypervariable segment (HVS) is higher than that of the coding region, the HVS-I substitution rate estimated as 4.48×10^{-6} per site per 25 years, the length of a generation (Forster et al. 1996), while the coding region substitution rate is estimated as 3.15×10^{-7} per site per 25 years (Mishmar et al. 2003) and the synonymous substitution rate in mitochondrial protein-coding genes as somewhat higher at 8.75×10^{-7} per site per 25 years (Kivisild et al. 2006). In case of the Y chromosome, the SNP mutation rate of 3.0×10^{-8} per base pair per generation (Xue et al. 2009) is several folds higher than that of autosomal chromosomes due to the fact that Y chromosomes pass only through the male germline, where a larger number of cell divisions occurs and mutation rate increases with time through continuing divisions of spermatogenic stem cells (Jobling and Tyler-Smith 2003). However, the difference between the Y-chromosomal and autosomal mutation rates is not large enough to be given serious consideration, especially given that mutation rate varies within auto-somes as well – not only the obvious difference between the coding and non-coding parts of the genome, but also deterministic variation within the non-coding portion of the genome (Smith et al. 2002). The Y chromosome, slow to evolve compared to mtDNA, conveniently contains genetic markers evolving at different rates, at the orders of magnitude of 10^{-8} per generation for SNPs and 10^{-3} per generation for short tandem repeats (STRs) (for details, see the section

"Estimating TMRCA in Y chromosome studies"), their combined use allowing for the examination of various time depths (de Knijff 2000). While the faster accumulating mutations generate more intra-species variation in mtDNA than in the nuclear genome, mtDNA markers have also been considered to be uninformative about earlier periods of human demographic history due to saturation. However, such an opinion is overly stringent as long as one studies just *Homo sapiens* and not the related species, unless one uses only HVS, which is not the practice in contemporary research on human ancestry.

The effective population sizes (N_e) of the Y chromosome and mtDNA are one-quarter of that of autosomes and one-third of that of the X chromosome, since both are haploid. This feature results in lower sequence diversity in the pool of human Y chromosomes compared to the rest of the nuclear genome, provided the same mutation processes act on all chromosomes (Thomson et al. 2000), and in a higher susceptibility of both the Y chromosome and mtDNA to genetic drift and founder effect (Jobling and Tyler-Smith 2003; Underhill and Kivisild 2007). Drift expedites the genetic differentiation of populations, but also causes the rapid change of haplotype frequencies through time, especially evident in case of some marked founder effects of the Y chromosome (Zerjal et al. 2003; Xue et al. 2005).

Since patrilocality, the custom of the wife moving near to the birthplace of her husband after marriage rather than *vice versa* (Burton et al. 1996), is practiced by approximately 70% of modern societies, most men live closer to their birthplaces than do women, which further heightens the local differentiation of Y chromosomes (Oota et al. 2001; Seielstad et al. 1998). Logically, this would result in a difference between the geographic distributions of Y chromosome and mtDNA lineages, since in patrilocal societies, the female-transmitted mtDNA would exhibit reduced geographic clustering, as was indeed shown in early studies of European (Seielstad et al. 1998) and island Southeast Asian populations (Kayser et al. 2001). However, it was later found based on re-sequencing a portion of the coding region of mtDNA and several separate regions of the NRY in the subjects that although the higher migration rate among females may be important at the local scale, it does not influence global-scale patterns of human population structure (Wilder et al. 2004).

Long-distance migrations have historically also been characteristic of the male sex, which produces an effect opposite to patrilocality, as can be seen in the case of the European expansion into the Americas and Oceania over the last 500 years, which has resulted in a strong introgression of European Y chromosomes with the retention of indigenous mtDNA lineages in Polynesia (Hurles et al. 1998), Greenland (Bosch et al. 2003) and the Americas (Carvajal-Carmona et al. 2000; Carvalho-Silva et al. 2001; Hammer et al. 2006).

In summary, the Y chromosome and mtDNA are highly suitable for tracing human genetic ancestry mainly due to their respective paternal and maternal mode of inheritance and lack of recombination. Very importantly, they allow for the detection of gender-specific gene flows, virtually impossible to infer

from the patterns of variation of autosomes (except for some information from the X chromosome), that are highly informative in the reconstruction of the demographic history of our species.

2.1.3. Whole genome markers

Despite their useful attributes, the Y chromosome and mtDNA provide only two specialised perspectives on the reconstruction of human genetic history. The whole human genome contains by far more information (mtDNA being ~5 orders of magnitude and the Y chromosome ~50 times smaller than the autosomal haploid genome), and lately that information has begun to be available for large numbers of samples, not just a few reference genomes. Due to recombination, the genome of each individual is an assemblage of DNA segments of different histories. These segments are combined in each generation by chance, ancestry, recombination and natural selection. Each segment can provide independent information about ancestry, however, the ancestry of different segments is not entirely independent, as humans are a subdivided population and segment ancestry is constrained by population history (Colonna et al. 2011). The particular combination of alleles in each of these segments along a chromosome is termed a haplotype, the correlation of alleles (the existence of the segments) is known as linkage disequilibrium (LD). Most of the recombination events occur in hotspots (Jeffreys et al. 2001), giving LD a block-like structure in the human genome (The International HapMap Consortium 2005). Haplotype blocks are inherited, being eroded only slowly by recombination and mutation (each occurring at an average rate of about 10^{-8} per base pair per generation) and thus giving information about population ancestry (The International HapMap Consortium 2005). Knowledge of the genetic structure of the populations studied is crucial not only for demographic ancestry research, but equally for genome-wide association (GWA) studies oriented towards the hunt for 'disease genes', since population stratification – allele frequency differences due to systematic ancestry differences – can and often has caused false signals in GWA studies (Price et al. 2006). This is equally valid for research with the wider goal of understanding the genetic changes caused by adaptation, as well as quantitative traits in general.

Genome-wide information can be obtained by either genotyping or, ultimately, re-sequencing human genomes. Determining the allelic state of hundreds of thousands or even millions of positions (mostly SNPs) across the genome that have prior evidence of variability has become routine – i.e. affordable enough to be performed on thousands of samples. The International HapMap Project began by typing one million genome-wide SNPs in 270 individuals of European, East Asian or West African descent (The International HapMap Consortium 2005), but the HapMap 3 dataset includes both SNPs and copy number polymorphisms, common and rare alleles, from 11 global populations, 1184 individuals in total (The International HapMap 3 Consortium

2010). The latest HapMap release (#28) includes data from the four original populations (Utah Americans of European descent, Han Chinese, Japanese and Bantu Yorubans), genotyped for about 4 million non-redundant SNPs, and from the seven additional populations, genotyped for about 1.5 million non-redundant SNPs (http://hapmap.ncbi.nlm.nih.gov/). Large numbers of the characterised SNPs are included in the commercial chips of companies such as Illumina (http://www.illumina.com) and Affymetrix (http://www.affymetrix.com), and used for GWA [among others, The Wellcome Trust Case Control Consortium (2007); Easton et al. (2007); Franke et al. (2010)] as well as population genetic studies [among others, Li et al. (2008)].

Sequencing is developing at a rapid pace, with next generation sequencing technologies [for a review, see Metzker (2010)] making whole genome re-sequencing projects ever more affordable and common. Sequencing avoids ascertainment bias and captures new as well as known variants. However, the next generation technologies have much higher error rates than the traditional Sanger methodology, for example, the Pacific Biosciences real-time sequencing platform having read accuracy of only 83% (Metzker 2010), and the market leader Illumina's Genome Analyzer and HiSeq platforms also producing various kinds of errors (Minoche et al. 2011). Thus, detecting and minimizing low read quality by appropriate data filtering criteria, as well as multiple coverage of the same DNA fragments, is essential. The latter is especially efficient in lowering the error rate, with companies such as Complete Genomics (http://www.completegenomics.com) claiming that a 50-fold coverage reduces the error rate to about 1–2 per 10,000 base pairs. Despite these complications, it is hard to overestimate the potential of large-scale re-sequencing, and several studies of immediate interest for human ancestry research have already been published, generating population-scale whole genome sequence data (The 1000 Genomes Project Consortium 2010) and developing statistical approaches to analyse them [e.g., Gronau et al. (2011); Li and Durbin (2011)].

At present, the main methods used in the analysis of whole genome geno-type data, such as principal component analysis [adapted for whole genome genotyping by Patterson et al. (2006)] or *structure*-like analyses, first developed to find patterns in the variation of limited sets of repeat elements (Pritchard et al. 2000), are mostly descriptive, if increasingly sophisticated (Lawson et al. 2012). They do provide the means to find similarities and recognise patterns, yet the results cannot be simply interpreted as either deep shared ancestry or recent admixture between populations under study (Li et al. 2008). From the human ancestry research point of view, the main disadvantage of the rest of the genome compared to mtDNA and the NRY is recombination. Therefore, in order to use whole genome data to trace ancestry, one must first account for recombination. Already the first high-density whole genome genotyping study of global populations (Li et al. 2008) revealed that a nearly linear correlation exists between the length of shared LD blocks and the geographic distance of a population from an assumed ancestral population. There are several develop-

ments in such approaches [e.g., Browning and Browning (2011)], but they are often based on simplified assumptions, such as single-admixture scenarios, applicable efficiently only to historically recent, well documented demographic events, for example, the trans-Atlantic slave trade from Africa, or the colonisation of the Americas by people of largely European descent. The analytical sophistication of such inferences needs to be improved, whereas the shortage of computational power adds further obstacles to the desired progress.

In summary, the wealth of information in the entire human genome, analysed at a population scale, has already opened vast new avenues both for human ancestry research and disease studies. However, it has also become clear that much remains to be done in order to fully explore the avalanche of empirical data gathered already and to be gathered in the near future.

2.2. The archaeological view on the history of anatomically modern humans in Europe and the Caucasus

For much longer than genetic inquiries into human history have been made, our past has been studied with archaeological methods, which have yielded valuable insights into the history of Europe and the Caucasus. After significant improvements in radiocarbon dating methodology (Fairbanks et al. 2005; Hughen et al. 2004; Mellars 2006a; Ramsey et al. 2004; van der Plicht et al. 2004), dates provided by archaeology are usually more direct and reliable than those based on genetic data. A brief overview of industries well represented in the archaeological record, including several early ones dated with methods other than radiocarbon dating, can be found in Table 1.

Anatomically modern humans (AMH) are defined based on cranial morphology, and the oldest AMH fossil remains – two partial skulls, arm, leg, foot and pelvis bones from Omo Kibish, Ethiopia – have been dated to about 196 thousand years ago (kya) (McDougall et al. 2005; Stringer 2011). Another find of early AMH is also from Ethiopia – the Herto fossils dated to 154–160 kya (Clark et al. 2003; Stringer 2011).

The earliest AMH arrived in Europe at the transition from the Middle to the Upper Palaeolithic era, and their dispersal across Europe, as well as the Caucasus, has often been associated with the subsequent disappearance of the Neanderthals (Mellars 2004; Pinhasi et al. 2011). The arrival of AMH initiated the cultural change from late Neanderthal Mousterian or transitional industries to Early Aurignacian technocomplexes (Bailey et al. 2009), attested to by numerous sites (Nigst 2006; Szmidt et al. 2010). According to calibrated radiocarbon dates, the spread of the early AMH was relatively fast (Mellars 2006a), and it has recently been shown based on archaeological material that they spread even wider and faster than previously thought, reaching southern

England by 44.2–41.5 kya (Higham et al. 2011) and southern Italy by 45–43 kya (Benazzi et al. 2011).

Table 1. An overview of the main archaeological cultures of the genus *Homo* (*H.*), with a focus on Europe. Adapted from Wikipedia and Soares et al. (2010).

Geological epoch	Archaeo-logical era	Culture	Age	Region	Species
Pleistocene	Lower Palaeolithic	Oldowan	2.6–1.8 Mya	Africa	*H. habilis, H. ergaster, H. erectus, H. heidel-bergensis*
		Acheulean	1.7–0.1 Mya	Africa, West Asia, South Asia, Europe	*H. erectus, H. sapiens*
	Middle Palaeolithic	Mousterian	300–30 kya	North Africa, Near East, Europe	*H. neander-thalensis, H. sapiens*
	Upper Palaeolithic	Uluzzian	45–43 kya	South Europe	*H. sapiens*
		Aurignacian	45–30 kya	Near East, Europe, Caucasus	*H. sapiens*
		Châtel-perronian	41–35 kya	West Europe	*H. neander-thalensis, H.sapiens*
		Gravettian	33–20 kya	Europe	*H. sapiens*
		Solutrean	24–18 kya	West Europe	*H. sapiens*
		Epigravettian	20–10 kya	Europe	*H. sapiens*
		Magdalenian	17–12 kya	Europe	*H. sapiens*
Pleistocene / Holocene	Mesolithic	Natufian	12.5–9.5 kya	Near East	*H. sapiens*
Holocene	Upper Palaeolithic	Swiderian	~9.5 kya	East Europe	*H. sapiens*
	Mesolithic	Kunda	8–5 kya	Northeast Europe	*H. sapiens*
	Neolithic	Linear Pottery	7.5–6.9 kya	Europe	*H. sapiens*

The Aurignacian technologies reflect an apparently sudden shift to distinctly 'modern' cultural behaviour, represented by the first complex bone tools, numerous personal ornaments such as various beads, perforated marine shells, etc., and sophisticated abstract and figurative art, including the elaborate cave

paintings of the Chauvet cave in Southeast France (Clottes 2001; Conard and Bolus 2003; Lewis-Williams 2002; White 1993; White 1997). Such symbolic behaviour was conspicuously lacking from preceding Neanderthal Mousterian communities (d'Errico 2003; Mellars 1996), and it is now generally agreed that Aurignacian was a modern human, not a Neanderthal culture (Bailey et al. 2009; Mellars 2005; Trinkaus 2005). The cultural change affected a wide range of different aspects of behaviour, having potentially profound social and cognitive implications, and was fast and abrupt enough to have been named a cultural and technological 'revolution' (Mellars 2005). Since the present study focuses on Europe and the Caucasus, the beginnings of 'modern' human behaviour are not discussed, but it should be borne in mind that cultural innovations such as shell ornaments and the use of ochre have been dated to as early as about 100 kya in Africa and the Levant (d'Errico et al. 2005; Henshilwood et al. 2002; Hovers et al. 2003; Mellars 2006b).

Based on archaeological Aurignacian evidence, there were two main dispersal routes of AMH into Europe, the 'classic' route through Central Europe along the Danube valley and the Proto-Aurignacian route along the Mediterranean coast (Mellars 2004; Mellars 2005; Mellars 2011; Figure 1). A third, Uluzzian dispersal has also been proposed (Mellars 2011; Figure 1). The Uluzzian technology (Kozlowski 2007) was a transitional Middle/Upper Palaeolithic industry, until recently believed to have been the result of acculturation between indigenous Neanderthal populations in Italy and intrusive Proto-Aurignacian AMH (d'Errico 2003; d'Errico et al. 1998; Mellars 2005), but now implied to have been a fully modern human culture (Benazzi et al. 2011).

In case of the Caucasus, a mountainous region between the Black and Caspian Seas, linking the Near/Middle East and the East European Plain, the discovery of numerous probable *Homo erectus* skulls [dated to 1.77 million years ago (Mya)] from Dmanisi, Georgia, makes it the region of the earliest evidence of the dispersal of the genus *Homo* outside Africa (Lieberman 2007; Lordkipanidze et al. 2007), and AMH, carriers of the Aurignacian technology (Bar-Yosef et al. 2006; Pinhasi et al. 2008), appeared there at least 42 kya (Adler et al. 2008). The early Upper Palaeolithic sites in the Caucasus (Bar-Yosef et al. 2006; Pinhasi et al. 2008) and the adjacent East European Plain (Krause et al. 2010; Prat et al. 2011) are nearly contemporary, while the lower Danube valley Upper Palaeolithic temporal estimates (Mellars 2006a) appear to predate the northern Pontocaspian Upper Palaeolithic by a few thousands of years. Thus, the early Upper Palaeolithic presence of AMH is documented both south and north of the Caucasus, and based on archaeological (chronological) evidence alone, it is impossible to say which route – across the Caucasus or via Anatolia and the Balkans – was used in the pioneer phase of the peopling of East Europe.
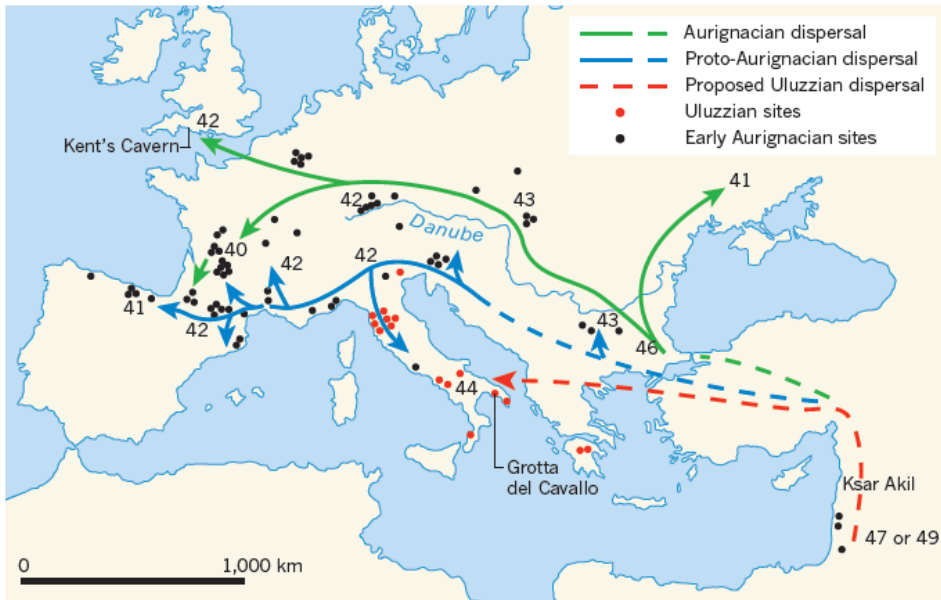
Figure 1. Principal dispersal routes of the earliest modern humans across Europe. From Mellars (2011).

The Last Glacial Maximum (LGM) in Europe (25–19.5 kya) resulted in the simplification and fragmentation of the Gravettian technology – a modern human industry spread across Europe before the LGM (Djindjian 2000; Perles 2000). During the LGM, people generally retreated southward with the climate change. However, the southward movement should not be taken as an exclusive trajectory, and in fact it seems that ambient temperature was less important than the availability of food, in some instances large game such as mammoths. The refugial areas human (and other faunal) populations became concentrated in included Southwest Europe, the Mediterranean coast, the Balkans, the Levant, and the East European Plain (Dolukhanov 1993; Gamble et al. 2005; Gamble et al. 2004), possibly also the eastern coast of the Black Sea (Adams and Faure 1997; Tarasov et al. 1999; Tarasov et al. 2000). The fragmentation of the human habitat in Europe resulted in regional industries such as the Solutrean 24–18 kya and the Epigravettian 20–10 kya [Soares et al. (2010) and references therein]. The Late Glacial Magdalenian technology 17–12 kya, on the other hand, is linked with the post-LGM population expansion [Soares et al. (2010) and references therein].

The Swiderian culture in East Europe marks the transition from the Palaeolithic to the Mesolithic (Dolukhanov 1997). It may have given rise to the Kunda culture and other 'post-Swiderian' cultures in the region (Dolukhanov 1997; Velichko et al. 2009). Chronologically preceding these cultures, the Mesolithic Near Easterners, the Natufian people (Bar-Yosef 1998; Brown et al.

2008; Sherratt 1997), were beginning to adopt cereal agriculture by 12 kya in the Levant. The Mesolithic, a rather arbitrary period, ended at different times in different places as agriculture reached them.

The Neolithic transition, meaning the adoption of agriculture, began in the Near East around 12–11 kya (Bar-Yosef 1998; Brown et al. 2008; Kuijt and Goring-Morris 2002; Sherratt 1997), with the domestication of plants preceding that of animals (Kuijt and Goring-Morris 2002). It has been argued that the fundamental change was a necessary economic response triggered by a crisis of sedentism generated by the brief cold period called the Younger Dryas (Bar-Yosef 1998). Interestingly, it appears that approximately the same two routes of the Aurignacian/Proto-Aurignacian dispersal, the northern along the Danube valley and the southern along the Mediterranean coast, were also taken by the earliest agricultural communities dispersing into Europe during the Neolithic transition 10–6 kya (Mellars 2004).

By about 8 kya, the Neolithic revolution had also reached the Caucasus, as evidenced by the presence of Neolithic cultural layers directly covering Mesolithic deposits in the Chokh (Dagestan, the Northeast Caucasus) and Darkveti (the South Caucasus) sites (Kushnareva 1997, pp. 154–158). The South Caucasus, particularly modern-day Armenia, has been suggested to be the region where the hexaploid bread wheat *Triticum aestivum* originated – about 10–8 kya (Feldman 2001) – through the spontaneous hybridisation of the tetraploid wild emmer *Triticum turgidum* and the diploid goatgrass *Aegilops tauschii*, and where it became a cultivated crop (Dubcovsky and Dvorak 2007; Dvorak et al. 1998). The South Caucasus has also been proposed as the area of grape domestication (McGovern 2003; Olmo 1995), a hypothesis that has recently found genetic support (Myles et al. 2011).

The Neolithic was followed by the Chalcolithic/Æneolithic or the Copper Age characterised by copper metallurgy that was well established in Southeast Europe by 7 kya and may have been invented independently in several locations rather than spread from a single source (Radivojevic et al. 2010). The technology of mining copper ore and copper smelting dispersed quickly, transforming into Bronze Age industries at different times in different regions. The Copper Age became widespread, evidence of such cultures has been found in West and Central Asian steppes, reaching Altay (Frachetti 2012), the Caucasus (Kushnareva 1997) and islands of the Mediterranean such as Cyprus (Peltenburg 1991) in addition to the Near East and Europe. The Copper Age also intensified the use of trade networks; the movement of people along the trade routes is difficult to differentiate today from large-scale migrations. Naturally, the Copper Age did not mark the end of major movements of people in Europe, which continued into the Bronze and Iron Ages and beyond, with examples such as several invasions of the British Isles, the Migration Period, the Slavic expansion, etc.

2.3. The formation of the European gene pool: a brief overview

2.3.1. The uniparental view

Several major demographic episodes over the last 50,000 years have very likely shaped major aspects of the variation of the gene pool of European populations: the pioneer colonisation of the Upper Palaeolithic, the Late Glacial recolonisation of the continent from refugia in southern Europe and the East European Plain after the Last Glacial Maximum (LGM), the postglacial recolonisation of deserted areas after the Younger Dryas cold period, the Neolithic diffusion from the Near East, and a multitude of lesser-scale migrations along continent-wide trade networks, which intensified starting with the Copper Age (Mellars 2004; Mellars 2006c; Soares et al. 2010). Less clear, but potentially equally important, is the role of the mid-Upper Palaeolithic Gravettian period that covered much of the continent, from southern France to the East European Plain, extending to the Ural Mountains.

All current genetic evidence supports a dispersal route of anatomically modern humans (AMH) from Africa through the Levantine corridor into Europe (Soares et al. 2010), a route represented in the archaeological record by the distribution of Aurignacian technologies (Mellars 2004; Mellars 2011). The most ancient mtDNA haplogroups in Europe are U5 and U8, which apparently originated locally, their ages proposed as around 32 and 44 ky, respectively (Behar et al. 2012), or alternatively even older, around 37 and 50 ky, respectively (Soares et al. 2010). The Y-chromosomal candidate for a signal of Upper Palaeolithic male ancestry is the mostly European-specific haplogroup I; its age of 24 ± 7.1 kya (Rootsi et al. 2004) can be tentatively linked to the spread of the largely pan-European Gravettian technology ~28–23 kya (Djindjian 2000; Perles 2000).

It has been suggested repeatedly that the main signal of the rise of variation in the modern European mtDNA and Y chromosome pools derives from the re-expansion from the refugia in the wake of the warming phase after the LGM 15 kya [e.g., Soares et al. (2010)]. Mitochondrial DNA haplogroups V (Torroni et al. 1998; Torroni et al. 2001), H1, H3 (Achilli et al. 2004; Loogväli et al. 2004; Pereira et al. 2005), U5b1b (Achilli et al. 2005; Tambets et al. 2004), and U5b3 (Pala et al. 2009) apparently originated in South Europe, and U4 and U5a (Malyarchuk et al. 2010; Malyarchuk et al. 2008) in the East European Plain, all expanding after the Ice Age. Several sub-clades of the Y chromosome haplogroup I have been proposed to have expanded from the Franco-Cantabrian and Balkan glacial refugia (Rootsi et al. 2004; Underhill et al. 2007), whereas the phylogeographic distribution of different sub-clades of the Y chromosome haplogroup R1 (discussed in more detail in the next section) has been suggested to reflect expansion from the Iberian Peninsula and from the present-day Ukraine (Semino et al. 2000).

The Neolithic transition or the spread of agriculture into Europe after its adoption in the Near East around 12–11 kya (Bar-Yosef 1998; Brown et al. 2008; Kuijt and Goring-Morris 2002; Sherratt 1997) has been researched extensively, using both archaeological and genetic data. The main debate in the field of genetics has centred on the models of cultural and demic diffusion, in other words, whether the spread of agriculture involved simply the adoption of new technologies or large-scale movement of people and population replacement (Cavalli-Sforza et al. 1994; Richards et al. 1996; Soares et al. 2010). The first couple of decades of pertinent genetic work, summarized by Cavalli-Sforza et al. (1994), made use of classical markers. The first principal component of classical marker data analysis was interpreted as reflecting the Neolithic dispersal, since it showed a cline from the Middle East towards Northwest Europe, and even though it only accounted for less than one third of the genetic variation of Europeans [summarized in Cavalli-Sforza et al. (1994)], these studies sparked a debate that has not ceased since.

The Y chromosome haplogroups associated with the spread of farming into Southeast Europe are E1b1b1 (Semino et al. 2000; 2004) and J, especially J2 (Di Giacomo et al. 2004; Semino et al. 2004; Semino et al. 2000). E1b1b1a and J2b* have been suggested to reflect the subsequent diffusion of people from the southern Balkans to the west (Semino et al. 2004). For mtDNA, haplogroups J and T1 have been proposed as candidates for the Neolithic dispersal, but founder analysis of European mtDNAs has suggested that the immigrant Neolithic component likely comprises less than one quarter of the mtDNA pool of modern Europeans (Richards et al. 2000). However, a new interpretation, based on an enlarged complete mitochondrial genome database, proposes that a substantial signal from mtDNA haplogroups J and T may in fact reflect dispersals from a Near Eastern refugium during the Late Glacial, 19–12 kya (Pala et al., accepted for publication).

Despite the designation of several Y chromosome and mtDNA haplogroups as 'Neolithic markers', it has been argued that the present-day mtDNA and Y chromosome pools in Europe show only limited Neolithic contribution from the Near East, which has been taken to suggest Late Glacial/postglacial origin for the majority of the lineages (Richards et al. 2000; Semino et al. 2000; Soares et al. 2010). On the other hand, several authors support the demic diffusion model, that is, substantial genetic input from the Near East during the Neolithic. For instance, a study of 840 men belonging to the Y chromosome haplogroup R1b that has high frequency in Europe concluded that this haplogroup reflects a recent genetic heritage uniformly introduced by Neolithic farmers from West Anatolia (Balaresque et al. 2010); however, these results have been challenged both on the basis of poor phylogenetic resolution (ref. I) and the dating based on 9 Y-STRs [Busby et al. (2012); for details, see the section "Estimating TMRCA in Y chromosome studies"]. It is also important to note that the mtDNA and Y chromosome results may differ if an original migration, of whatever proportions, is followed by subsequent influxes involving mostly men.

Direct ancient DNA evidence appears to be more in accord with the demic diffusion model. Ancient DNA from the largest Linear Pottery culture genetic dataset analysed to date (n = 42) reveals that the Neolithic samples share an affinity with the modern-day Near East and Anatolia, which supports a significant genetic input from this region to Europe during the Neolithic transition (Haak et al. 2010). The Linear Pottery culture populations have also been shown to have had a distribution of mitochondrial haplogroups clearly distinct from that of modern Europeans, suggesting that they left few descendants beyond the Neolithic and that major demographic events took place in Europe in later times (Haak et al. 2010; Haak et al. 2005). Another recent study found ancient mtDNA discontinuity between North and East European Late Palaeolithic/Mesolithic samples and Central European Neolithic samples and proposed that it implied large-scale Neolithic replacement in North and East Europe, although the authors emphasized that this observation does not resolve the question of the extent to which modern Europeans are descended from the Neolithic farmers, their hunter-gatherer forerunners, or later incoming groups (Bramanti et al. 2009). Ancient Y chromosome and autosomal data are more difficult to obtain due to the copy number of nuclear DNA in the cell that is orders of magnitude smaller than that of mtDNA, but two recent studies have succeeded in typing Y chromosome markers in ancient samples from France (~5 ky old) and Spain (~7 ky old) (Lacan et al. 2011a; Lacan et al. 2011b). The studies showed a high frequency of the Y chromosome haplogroup G2a, associated with the Neolithic diffusion (Battaglia et al. 2009; Behar et al. 2004), among the samples typed, but found that the haplotypes of the ancient G2a samples are rare among modern Europeans, concluding that the lineages were probably lost between the end of the Neolithic and today (Lacan et al. 2011a; Lacan et al. 2011b).

Thus, the debate of cultural versus demic diffusion during the Neolithic transition is far from being settled, with ancient DNA studies bringing new evidence to light. However, these studies, while having mostly overcome the issue of contamination, still understandably lack sample sizes sufficient to make large inferences. It would seem that if the Neolithic transition did involve a considerable degree of population replacement in Europe, these immigrants from the Near East make up a limited portion of modern Europeans. Apparently, later migrations have additionally influenced the genetic landscape of different sub-continental areas of Europe (Haak et al. 2010; Soares et al. 2010).

2.3.2. The whole genome view

So far, to the best of my knowledge, no dedicated European-centric population-scale re-sequencing studies have been published. A couple of recent studies using whole genome sequence data have investigated the historical relationships of major human groups (African, Asian and European) (Gronau et al. 2011) and the effective population size of humans in general back to several million years

ago (Li and Durbin 2011) – large and intriguing questions, but with no special focus on Europe. Both of these studies have estimated the time of the out-of-Africa split, one of the proposed dates being rather recent, ~50 kya (Gronau et al. 2011), the other slightly earlier, ~60–80 kya (Li and Durbin 2011). Interestingly, Li and Durbin (2011) also infer substantial gene flow between sub-Saharan Africans and Europeans/Asians until 20–40 kya, a pattern that is not apparent from uniparental marker data (Jobling and Tyler-Smith 2003; Torroni et al. 2006) or from high density whole genome genotyping results (Li et al. 2008).

Whole genome genotype data of Europeans mainly display a close corre-lation between geographic and genetic affiliation (Nelis et al. 2009; Novembre et al. 2008). *Structure*-like analyses show a clear European 'ancestry compo-nent' (The 1000 Genomes Project Consortium 2010; Behar et al. 2010; Li et al. 2008; Rasmussen et al. 2010). In general, there is a smooth transition in whole genome genotype data from the Near/Middle East to Europe, evident from both principal component and *structure*-like plots (Behar et al. 2010; Li et al. 2008; ref. III), and linearly decreasing haplotype heterozygosity from sub-Saharan Africa to the Middle East to Europe has also been detected (Li et al. 2008). Thus, whole genome genotype data support the model of a serial founder effect with origin in sub-Saharan Africa in the peopling of the world (Colonna et al. 2011; Li et al. 2008), with humans arriving in Europe through the Near/Middle East. However, since there are currently no reliable methods of sufficiently precise dating based on whole genome genotype data, more elaborate models of the formation of the European gene pool that would make use of these data remain a prospect for the future.

2.4. Two opposing sub-clades
of Y chromosome haplogroup R

In the early years of Y chromosome phylogenetic research, several unrelated and non-systematic nomenclatures for Y chromosome haplogroups emerged, making it difficult to navigate between results published by different sources. The various parallel nomenclatures of the Y chromosome phylogeny were unified in 2002 by the Y Chromosome Consortium (YCC), with a simple set of rules developed to unambiguously label the clades nested within the hierarchi-cal topology based on unique SNPs (YCC 2002). There are two complementary nomenclature systems: the first defining hierarchical sub-clades within each major haplogroup, denoted by capital letters, using an alphanumeric system (e.g., J2, J2a, J2a2, J2a2a, etc.), the shorter alternative naming haplogroups by the terminal mutation defining them (e.g., J-M92) (YCC 2002). The discovery of novel SNPs may have the effect of splitting or joining previous clades and thus alter the alphanumeric nomenclature system, whereas the mutation-based nomenclature is always unambiguous. Lineages representing interior nodes of

the tree, not defined by a derived character, are indicated by the symbol * due to their potentially paraphyletic nature (YCC 2002). The most recent published Y chromosome phylogeny is based on 599 SNPs and contains major clades A to T (Karafet et al. 2008); it has already been refined repeatedly (ISOGG Y-DNA Haplogroup Tree 2012).

In Europe, the Y chromosomes of about 50% of men belong to the relatively young (Karafet et al. 2008) haplogroup R (Jobling and Tyler-Smith 2003; Rosser et al. 2000; Semino et al. 2000), and essentially all of these European R affiliates belong to the sub-clade R1 defined by M173 (YCC 2002). This haplogroup has been the object of much study due to its high frequency among Europeans and supposed links to major demographic processes.

The two major sub-clades of haplogroup R1, R1a and R1b, show opposite clinal patterns in Europe – R1a is common in Northeast Europe, with frequency declining towards the Southwest, whereas R1b is most frequent in Southwest Europe and rare in the Northeast (Balanovsky et al. 2008; Peričić et al. 2005; Rosser et al. 2000; Semino et al. 2000; ref. I, II). While haplogroup R1a (named differently in various early papers) had a defining mutation already in the infancy stage of Y chromosome phylogenetic research – $SRY_{10831.2}$ (SRY_{1532}) in Rosser et al. (2000) and M17 in Semino et al. (2000) –, haplogroup R1b was long defined as M173(xM17) (Semino et al. 2000), 92R7(xSRY_{1532}) (Rosser et al. 2000) or the like, sometimes even after the mutation M269 defining the European-specific R1b1b2 clade had been discovered (Cruciani et al. 2002). The blind assignment of all M173-derived non-R1a Y chromosomes to R1b (again, named differently at the time) worked for European samples, as it has since been shown that in western Europe, these Y chromosomes do indeed belong almost exclusively to R1b1b2, but it was hardly good practice.

An early study proposed that the mutation M173 the haplogroups that later became known as R1a and R1b have in common is an ancient Eurasiatic marker brought by or arisen in the group of *Homo sapiens* that entered Europe and diffused from east to west about 40–35 kya, spreading the Aurignac culture (Semino et al. 2000). The same study interpreted the contrasting geographic distribution of two haplotypes within R1 – Eu19 defined by M17, a proxy for R1a, and Eu18, at the time defined as M173(xM17) and later found to corres-pond almost exclusively to R1b1b2 – as the result of re-colonisations from isolated population nuclei in the present-day Ukraine and the Iberian Peninsula, following the LGM 20–13 kya (Semino et al. 2000). This proposal was supported by the maximum variation of microsatellites linked to Eu19 and Eu18 being found in the Ukraine and the Iberian Peninsula, respectively (Semino et al. 2000).

However, simple opposing clines dominate the phylogeographic distribution of R1a and R1b only as long as one examines it in Europe. The first hint that R1b is not uniform even within Europe came from the *Taq*I haplotypes ht15 and ht35 associated with the complex restriction fragment length polymorphism (RFLP) 49a,f locus, ht15 found to be common in Iberia (Semino et al. 1996)

and ht35 distributed across Europe (Santachiara Benerecetti et al. 1993; Torroni et al. 1990). Later, additional markers such as M73 (Underhill et al. 2001) and V88 (Cruciani et al. 2010) were discovered, both defining sub-clades of R1b spread mostly outside Europe, the former in the Caucasus, Turkey, the Circum-Uralic and North Pakistan (ref. I) and the latter in *trans*-Saharan Africa, possibly reflecting a migration from Asia to Africa.

The comprehensive pattern of the spread of haplogroup R1a especially (ref. II), and R1b as well (ref. I), is much more complex than overall frequency clines in Europe seem to show, suggesting that simplistic interpretations tying these haplogroups to any single demographic process are likely to be insufficient. For instance, the authors of a recent study interpreted the phylo-geography of haplogroup R1b quite differently from previous research by analysing 9 Y-STR loci associated with 840 R1b Y chromosomes (Balaresque et al. 2010). They concluded that all such chromosomes in Europe reflect a recent genetic heritage that was uniformly introduced by Neolithic farmers migrating from West Anatolia. However, in this study, the samples were resolved only to the level of M269 (Cruciani et al. 2002) that defines the entire European-specific R1b1b2 clade, and any conclusions based on such shallow phylogenetic resolution are unlikely to capture the complexities of the spread of haplogroup R1b. Also, the age of ~6 ky for the R-M269 linage calculated based on Y-STRs (Balaresque et al. 2010) has been challenged by the claim that since Y-STR-based coalescence age estimation depends on the choice of STRs, existing data and methodology are insufficient to make credible estimates for the age of haplogroup R1b (Busby et al. 2012).

2.5. The Caucasus: general facts

"*Caucasian variety – I have taken the name of this variety from Mount Caucasus, both because its neighbourhood, and especially its southern slope, produces the most beautiful race of men, I mean the Georgian; and because all physiological reasons converge to this, that in that region, if anywhere, it seems we ought with the greatest probability to place the autochthones of mankind.*" The famous German anthropologist Johann Friedrich Blumenbach stated this in his "*De generis humani varietate native*" [Blumenbach (1795), quoted e.g. in Keith (1940)], naming the white race Caucasian, nowadays known as Cau-casoid. It is a humorous quirk of history that Blumenbach made his judgement on the 'most beautiful race of men' based on the skull of a Georgian woman, having never been to the Caucasus himself.

The erroneous view of the origin of humans in the Caucasus aside, this is certainly a region that merits interest from the point of view of human past. The Caucasus is among the more ethnically diverse areas of the world, being home to over 50 ethnic groups (Encyclopaedia Britannica). The linguistic diversity of the Caucasus is likewise remarkable: there are three language families

indigenous for the area – Abkhazian-Adyghe, Nakh-Dagestanian, and Kart-velian – and in addition, several Indo-European, Turkic, Mongolic and Semitic languages are spoken in the region (Comrie 2008; Figure 2). It is notable that the three autochthonous language families are very distantly related, the split between the Kartvelian family and the other two having occurred over 14 kya (The Global Lexicostatistical Database).

Figure 2. Larger ethno-linguistic groups in the Caucasus region in 2009. From Wikimedia Commons.

Geographically, the Caucasus lies between the East European Plain and the Near/Middle East, linking these regions between the Black and Caspian Seas (Figure 2). It is a mountainous area divided by the High Caucasus Mountain Range into the North and South Caucasus. The principal political entities (internationally recognised statehoods) in the region are the Russian Federation in the North Caucasus, and Georgia, Armenia and Azerbaijan, and partly Turkey and Iran, in the South Caucasus (Figure 2).

The earliest evidence of anatomically modern humans in the Caucasus dates to the early Upper Palaeolithic, at least 42 kya (Adler et al. 2008). It is unclear whether or not the region has been continuously occupied by AMH since then. Evidence for the LGM human refugium at the Black Sea coastal area of the Caucasus (Adams and Faure 1997; Tarasov et al. 1999; Tarasov et al. 2000) is as yet superficial; on the other hand, numerous middle/late Upper Palaeolithic sites from before and after the LGM support constant or nearly constant human settlement of the area (Pinhasi et al. 2008).

It has been remarked that since mountains present a less favourable environment than lowlands, the Caucasus has usually been in economic and political dependence of the peoples or political entities ruling the steppes, and that at least in historical times, settlers have only moved to the High Caucasus in order to avoid threats, such as the invading Mongols, and have spread out onto the plains again when opportunity arose (Лавров 1978). Nevertheless, a continuous inhabitation of the High Caucasus appears to also have existed throughout history.

The autochthonous languages of the Caucasus can be divided into one southern and two northern branches based on the areas inhabited by their speakers – the Kartvelian languages are spoken in the South Caucasus, the Abkhazian-Adyghe languages in the Northwest, and the Nakh-Dagestanian languages in the Northeast, a region known as Dagestan. Most of the ethnic groups speaking the indigenous languages are small, especially in the Russian North Caucasus, where only one population exceeds one million (the Nakh-Dagestanian-speaking Chechens), three more exceed half a million (the Nakh-Dagestanian-speaking Avars and Dargins and the Abkhazian-Adyghe-speaking Kabardins) and another two fall just under it (the Nakh-Dagestanian-speaking Lezgins and Ingush) (Russian census 2010). The Kartvelian-speaking Georgians number approximately 3.7 million in Georgia (Georgian census 2002), constituting the largest population to speak an indigenous Caucasian language.

The largest Indo-European-speaking ethnic group in the Caucasus is the Armenians with a population of approximately 3.1 million in Armenia (Armenian census 2001). The Armenians are presumed to have migrated to the South Caucasus from Anatolia in the early 1st millennium BC (Encyclopaedia Britannica). Another ancient Indo-European group is the Ossetians, commonly believed to represent the descendants of nomads successively known as Scythians, Sarmatians, and Alans, who retreated to the Caucasus from the East European Plain in the 4th century AD, having been displaced by the Huns

(Minahan 2000, p. 518). However, nowadays Slavic-speaking peoples, mainly Russians and Ukrainians, who have migrated to the North Caucasus fairly recently, account for more than one-third of the total population there (Encyclopaedia Britannica). Smaller Indo-European-speaking groups such as Kurds, Talysh, Mountain Jews, Greeks, and Roma (Gypsies) are distributed in various areas of the Caucasus (Encyclopaedia Britannica; Figure 2).

The Turkic-speaking peoples of the Caucasus are the Azerbaijani in the South Caucasus [population approximately 8.2 million in Azerbaijan (Azerbaijani census 2009)] and the Kipchak Turks in the North Caucasus. The largely Muslim Kipchak Turks consist of Kumyks, Nogays, Karachays, and Balkars. The Nogays have a special status among the Caucasian populations due to their recent, late 18th – early 19th century arrival from the Pontocaspian steppes (Kolga et al. 2001). They trace their ancestry from the Golden Horde nomads, a confederation of tribes of different ethnic descent (Kolga et al. 2001). Due to long geographic separation of the whole Nogay population, they speak three different dialects: Kara Nogay, Nogay Proper, and Ak Nogay, the last also known as Kuban Nogay since its speakers live by the River Kuban [in Turkic, 'kara' is black and 'ak' white – the Turkic peoples have commonly divided their tribes into black and white, 'black' meaning northern and 'white' western (Kolga et al. 2001)].

The only Mongolic-speaking people in the vicinity of the Caucasus region are the Kalmyks inhabiting a dry steppe area to the north of the Caucasus (Figure 2), and the only Semitic group in the Caucasus are the Assyrians, claiming to be the descendants of the Mesopotamian Assyrians and nowadays living mainly in the cities (Encyclopaedia Britannica).

2.6. Genetic studies of the Caucasus

Numerous genetic studies of the Caucasus region have been conducted. Based on 8 Alu insertion polymorphisms, an early paper found that the Caucasus populations exhibit high levels of between-population differentiation, with an average F_{ST} of 0.113 – a value which is almost as large as the F_{ST} of 0.157 for worldwide populations (Nasidze et al. 2001). High F_{ST} values for Dagestanian populations, based on autosomal STR allele frequency distributions, were also reported by (Bulayeva et al. 2006), whereas genetic diversity within the populations was low, both observations confirming the small population sizes and high levels of isolation among the ethnic groups of Dagestan. Analysing autosomal variation (15 STRs) among Armenians, a recent study provided evidence for the genetic differentiation of distinct communities from across Historical Armenia, proposing genetic influences from Turkey and the Balkans for the Armenian populations of Sasun and Lake Van, respectively – unsurprisingly, since these two populations inhabit areas west of Armenia, in present-day Turkey (Lowery et al. 2011).

High between-population differentiation evident from Y chromosome data, rather than mtDNA, has also been detected in the mountainous areas of the Caucasus, likely reflective of patrilocal societies and genetic drift in small isolated populations (Nasidze et al. 2004a; Weale et al. 2001). However, these results, like those of a later study that found reduced Y-chromosomal but not mtDNA or autosomal genetic diversity among highland Dagestanian populations due to genetic drift (Marchani et al. 2008), are based on rather limited phylogenetic resolution compared to the current state-of-the-art knowledge, and examine only a part of the wide range of populations in the Caucasus.

Several early mtDNA studies encompassing Caucasian populations found a continuity of mtDNA lineages from the South Caucasus and Anatolia to Europe, although these haplogroups showed systematically much later signs of expansion in Europe than in the South Caucasus and Anatolia (Metspalu et al. 1999; Tambets et al. 2000). However, this early research used low phylogenetic resolution. A later higher resolution study focusing on the mtDNA haplogroup H common in both Europe and the Caucasus showed that irrespective of their common Upper Palaeolithic origin, the distribution of H sub-haplogroups differs significantly in Europe and in the Near East and the South Caucasus, implying limited post-LGM maternal gene flow between these regions (Roostalu et al. 2007). In contrast, the study found that the North Caucasian maternal gene pool has received an influx of H lineages from East Europe (Roostalu et al. 2007).

Two studies, the first exploring mtDNA diversity (Nasidze and Stoneking 2001) and the second Y chromosome data (Nasidze et al. 2003), showed that genetic relationships among Caucasian populations reflect geographic rather than linguistic relationships since the Indo-European-speaking Armenians and Turkic-speaking Azerbaijani were found to be genetically closer to their geographic neighbours speaking indigenous Caucasian languages than to their linguistic relatives. Hence, the authors postulated language replacement among Armenians and Azerbaijani. Furthermore, a number of studies have focused on single populations from the Caucasus region, such as the Armenians (Herrera et al. 2012; Lowery et al. 2011; Weale et al. 2001), the Ossetians (Nasidze et al. 2004b), the Nogays (Bermisheva et al. 2004), the Kurds (Nasidze et al. 2005a), and the Kalmyks (Nasidze et al. 2005b).

A recent high resolution Y chromosome study included large samples from 14 North Caucasian populations and focused again on the relationship of linguistic and genetic diversity in the Caucasus (Balanovsky et al. 2011), arriving, however, at a different conclusion than previous research. Similarly to several earlier studies, the authors suggest that Caucasian male lineages originate from the Near East, followed by high levels of isolation, differentiation and genetic drift in the mountainous terrain of the Caucasus. However, in the matter of the relationship of genetic and linguistic diversity, the authors present a decidedly different viewpoint. The study reports a strong correlation of Y chromosome haplogroup frequencies with language in the North Caucasus,

further showing a number of haplotype clusters within haplogroups to be specific to individual populations and languages. The authors also claim an unprecedented level of gene–language co-evolution in the Caucasus, backing it with a comparison of genetic and linguistic reconstructions that match in both topology and dates. This position disagrees with the previously held view that geography mainly influences genetic relationships in the Caucasus (Nasidze et al. 2003; Nasidze and Stoneking 2001). It must be noted, however, that the genetic and linguistic concordance was shown based mainly on populations speaking indigenous Caucasian languages (the only exception being the Indo-European-speaking Ossetians) and that the ethnic groups speaking these languages, several of them relatively small, are geographically sharply demarcated (see the section "The Caucasus: general facts"), making it challenging to differentiate between the correlation of genetic data with language and geography.

Another recent Y chromosome study proposed a scenario of the Armenian plateau having been repopulated after the LGM by agriculturalists from the Fertile Crescent (Herrera et al. 2012). Besides finding a high prevalence among the Armenian populations of Y chromosome haplogroups associated with the spread of agriculture from the Near East, the study reported a restricted genetic affinity of Armenians with Europeans (Herrera et al. 2012), in accordance with previous results (Nasidze et al. 2004a; Nasidze et al. 2003). However, although the authors insist that the STR variation time window linked to Y chromosome haplogroups studied, J2 and R1b, testifies to the spread of the Neolithic to Armenia, the majority of their time estimates fall into the Chalcolithic (Herrera et al. 2012). Furthermore, the same is largely true for time estimates for Syria, Anatolia, Iran, and Greece, where certainly any signal linked to Neolithisation should be expected not around 4–6 kya, but many thousands of years earlier. Hence, if anything, this expansion of Y chromosomes can be much more reliably attributed to later demographic processes, possibly linked to the mastering of metal industries that reached both the Northwest and the Southeast Caucasus already in the early Bronze Age, over 5 kya.

2.7. Estimating TMRCA in Y chromosome studies

Dating genetic lineages is a well-known difficulty, no less so for Y chromosome than for mtDNA studies (Jobling and Tyler-Smith 2003). Major haplogroups of the Y chromosome tree have been dated using SNP data, for example in the study of Karafet et al. (2008), where the method relied on a uniform probability distribution for the age of mutations in the ancestry of a lineage. The calibration point of the molecular clock was inferred using archaeological evidence – the age of the CT clade representing all non-African variation was fixed using the 70 kya date of the out-of-Africa migration supported by archaeological data (Karafet et al. 2008). Based on this calibration, the ages of 10 other major

clades were estimated, ranging from 68.9 (64.6–69.9) kya for the CF clade to 18.5 (12.5–25.7) kya for the R1 clade (Karafet et al. 2008). In general, earlier estimates of the time to the most recent common ancestor (TMRCA) for the Y chromosome have been recent – under 100 kya (Pritchard et al. 1999; Thomson et al. 2000) – compared to those made for mtDNA, the X chromosome and autosomes, but it has been argued that it would be premature to regard this as proof of a departure from neutrality in the evolution of the Y chromosome (Jobling and Tyler-Smith 2003). However, a recent estimate – based on 138 SNPs from 206-kb Y-chromosomal stretches re-sequenced in seven African males from the deep-rooting branches A and B of the Y chromosome tree – sets the most recent common ancestor of all present-day Y chromosomes at a much earlier date, 141.5 ± 15.6 kya, and geographically in North or Central Africa (Cruciani et al. 2011). While the precise geographic origin of anatomically modern humans within Africa is still subject to debate and will perhaps remain so (Batini and Jobling 2011), it seems clear that the root of Y chromosome phylogeny lies in Africa.

However, the Y chromosome base substitution rate, measured as 3.0×10^{-8} mutations/nucleotide/generation (Xue et al. 2009), is too slow, unless one re-sequences long stretches of the Y chromosomes under study, to allow for the dating of younger and minor haplogroups, which is of interest for many studies. Therefore, a common method of estimating the TMRCA in Y chromosome studies is the use of faster evolving Y chromosome short tandem repeats (Y-STRs). Y chromosome STRs, or microsatellites, consist of 1–6-bp units that are, on average, repeated 9.7 (nonpolymorphic loci) or 14.4 times (polymorphic loci) (Kayser et al. 2004).

Several factors complicate dating based on Y-STRs, including differences between individual STR loci, STR locus saturation, and the difference between the 'genealogical' and 'evolutionary' mutation rates. Studies of deep rooting pedigrees have yielded an average Y-STR mutation rate of 2.0×10^{-3} per generation (Heyer et al. 1997), which compares to the average rates of 2.5×10^{-3} (Goedbloed et al. 2009) and 2.1×10^{-3} (Ge et al. 2009) per generation observed in father/son pairs. These so-called 'genealogical' rates have turned out to be an order of magnitude higher than the 'evolutionary' rate estimate of 2.6×10^{-4} per generation for the same STR loci, obtained in a study based on counting the number of mutations on the branches of a haplotype network (Forster et al. 2000). This discrepancy, estimated as a 3.6× difference, might be explained by a large share of STR variation derived within a haplogroup being effectively removed by genetic drift, rendering mutation rate estimates based on evolutionary considerations lower than those based on pedigree studies (Zhivotovsky et al. 2006). However, a wide-scale study of father/son pairs obtained Y-STR 'genealogical' mutation rates (based on 186 Y-STRs) ranging from 3.78×10^{-4} to 7.44×10^{-2} per marker per generation, highlighting the importance of mutation rate variation between different Y-STR loci in addition to the 'genealogical'/'evolutionary' rate distinction (Ballantyne et al. 2010). Therefore,

irrespective of the fact that temporal estimates based on STR variation have been used in hundreds of research papers, it is by no means an ideal approach.

For the purpose of dating in population genetic studies, it is rather common (Cinnioglu et al. 2004; Derenko et al. 2006; King et al. 2008; Lappalainen et al. 2008; Peričić et al. 2005; Semino et al. 2004; Sengupta et al. 2006) to use the 'evolutionary' mutation rate of 6.9×10^{-4} per 25 years for tri- and tetranucleotide Y-STRs, and a method of estimating age as the average squared difference (ASD) in repeat count between all current chromosomes and the median haplotype, averaged over STR loci and divided by the mutation rate (Zhivotovsky et al. 2004). The mutation rate was first estimated using data on Y-STR variation in several populations with documented short-term histories (Zhivotovsky et al. 2004). It could be argued that this mutation rate is even too widely used, since despite the fact that the authors of the ASD method have cautioned against it (Zhivotovsky et al. 2006), several studies have employed the mutation rate of 6.9×10^{-4} per 25 years to estimate coalescence times of large haplogroups (Cinnioglu et al. 2004; King et al. 2008; Sengupta et al. 2006).

Other methods of STR-based dating of (Y chromosome) lineages include the ρ (rho) estimator (Forster et al. 1996; Saillard et al. 2000) and the BATWING software (Wilson et al. 2003). Both depend on STR networks of samples constructed by a program such as Network (Fluxus-Engineering, http://www.fluxus-engineering.com). The ρ statistic estimates the TMRCA of clusters of STR haplotypes identified in the networks. Examining a cluster that appears to have evolved within a specific population, the resulting age estimate will serve as the lower bound for the time that the population may have been isolated following a split. BATWING provides a mechanism to identify bounds, established by coalescent events, between which a unique event polymorphism (UEP) may have emerged. UEPs are typically identified by SNPs, but can be generalized to include 'virtual UEPs' that mark clusters of phylogenetically related STR haplotypes identified in a network. BATWING requires as input several prior conditions such as ancestral effective population size and a demographic model (stable population / expansion / etc.) that may often be difficult or impossible to estimate correctly.

Another software called BARCODE (Basic ASD Regression for Comparative Dating Estimates) was developed and compared with BATWING by Luca et al. (2005). BARCODE makes use of the concept of 'comparative dating' in which direct measures of STR variability for different UEP-defined lineages are compared to each other with the aim of evaluating the timing of each UEP's appearance in terms of proportions of the antiquity of the clade defined by them. The benefit of this method is that it only requires minimal assumptions and can be applied in the absence of knowledge on STR mutation rates; however, by definition, it cannot return estimates in absolute times. Nevertheless, it represents an independent aid in the reconstruction of the intervals separating the nodes of the UEP tree. The authors of BARCODE found that its results closely match those of BATWING as far as the relative antiquity of

clades is concerned (Luca et al. 2005). They also recommended the use of specific priors for each locus in a particular dataset in order to obtain more reliable dating results, at odds with their own use of unified parameters for BARCODE, since locus-specific STR mutation rates were unavailable (Luca et al. 2005). Indeed, the difficulty of estimating locus- and haplogroup-specific mutation rates or, in addition, ancestral population sizes and demographic models required for BATWING, is the reason the simple ASD method using a single prior, an averaged mutation rate (Zhivotovsky et al. 2004), remains commonly used, being free of biases potentially introduced by other priors or models.

The 'evolutionary' mutation rate, and even the use of Y-STRs for dating purposes in general, have also been heavily criticised. The main shortcoming of the 'evolutionary' rate is that it tends to overestimate dates. This is mostly due to men having different numbers of offspring, with the majority of men alive today descending from a relatively small percentage of forefathers (likely of higher social status). Therefore, a small number of more prolific forefathers have contributed the largest share of the accumulated Y-STR variance, which in its turn means that the variance is higher among the descendants of the prolific men than among the descendants of the less prolific men. Given this expansion of a few lineages at the expense of others and the fact that the more numerous descendants of the prolific forefathers are most likely to be sampled, we overestimate the time to the MRCA by mostly sampling men from lineages where Y-STR variance accumulates faster than the 'evolutionary' rate. The authors of the ASD method themselves have pointed out that an effective mutation rate higher than the 'evolutionary' rate would be observed in case of a sudden jump in the size of the haplogroup after its appearance or in case of an expanding population (Zhivotovsky et al. 2006), both of which are often reasonable assumptions for post-Neolithic human populations. The confidence intervals of Y-STR-based age estimates are wide, influenced by the chosen generation length, the inherent stochasticity of the mutation process, and the mutation rate, which in its turn is influenced by the demographic history of a particular haplogroup in a particular population (Zhivotovsky et al. 2006). In addition to these uncertainties, a recent study casts into doubt the use of Y-STRs for dating in general, claiming that age estimates depend explicitly on the choice of STRs, more specifically, their duration of linearity (the length of time into the past over which variance and the average coalescence time continue to be linearly related for a specific STR), and that commonly used STRs do not allow linearity to be assumed further into the past, rendering all Y-STR-based age estimates suspect (Busby et al. 2012).

However, despite the biases surrounding dating based on Y-STRs, at present this remains the only method readily available for estimating the TMRCA of younger lineages in Y chromosome population studies. The common set of 10 Y-STRs (DYS19, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS439, DYS461) is consistently used, and since these are the same

markers used in the original mutation rate study by Zhivotovsky et al. (2004) and 8 of them are tetra- and 2 trinucleotide repeats, the choice of the mutation rate of 6.9×10^{-4} per 25 years for tri- and tetranucleotide Y-STRs is justified. Considering this, any bias in the age estimates of the studies that use these Y-STRs, this mutation rate and the ASD method, is at least consistent, and since the tendency of the method to overestimate coalescence times is known, the dates obtained are regarded as the upper bounds of past demographic events, and with a degree of reservation. Until new and hopefully more reliable methods can be implemented, based for instance on SNPs from massive re-sequencing of a considerable portion (e.g., some tens of millions of bp) of the NRY combined with reliably dated calibration (derived, e.g., from ancient DNA studies), Y-STR variation analysis, despite its known imperfections, will continue to be used for dating. It is also likely that the new re-sequencing-based methods will offer, first and foremost, more precise estimates to date the major splits of the Y chromosome tree, whereas in case of the younger branches, the faster Y-STR clock may long remain the method of choice.

3. AIMS OF THE PRESENT STUDY

The current era of extensive expansion in the field of human population genetics is constantly opening up new possibilities, both in data acquisition and analysis. The developments that are already accessible enable us to approach several long-standing issues in human ancestry studies in a more complex manner. Previous research, reviewed in the literature overview, allows us to formulate several thus far indistinct yet major problems concerning human demographic history in Europe and the Caucasus that this study aims to explore.

Firstly, the present study investigated the insights into human demographic history in western Eurasia, including the peopling of Europe, offered by the spatial and temporal distribution of Y chromosome haplogroups R1a and R1b, both widespread in the region. The aim was a detailed phylogenetic and phylogeographic analysis of both haplogroups that would allow for a more fine-grained view of the formation of present-day West Eurasian populations than overall haplogroup frequencies.

The second goal of the present study was to examine the hitherto not comprehensively analysed region – the Caucasus –, encompassing all its sub-regions and linguistic groups, from a population genetic viewpoint, including Y-chromosomal, mtDNA and autosomal analyses. The intriguing questions related to the Caucasus are its role in human dispersals in Eurasia and the extent to which the high ethnic and linguistic diversity of the region is reflected in its genetic variation.

Lastly, the present study compared penta- and hexanucleotide Y-STRs to commonly used tri- and tetranucleotide loci with the aim of estimating their mutation rate and applicability in evolutionary studies.

4. SUBJECTS AND METHODS

The subjects of the present study, analysed in four separate publications, but often overlapping, are listed in the respective articles and/or their supplementary material. The DNA samples were obtained from unrelated volunteers after receiving informed consent in accordance with the guidelines of the ethical committees of the institutions involved.

The methods used in the present study, both experimental and computational, are described in detail in the respective publications and their supplementary material.

5. RESULTS AND DISCUSSION

The discussion of the results of this study presented here is a summary of four articles published in peer-reviewed journals and does not recapture their contents word for word; many details can be found in the articles themselves and their supplementary material.

The human demographic history of any region cannot be reliably inferred unless the data are placed in the context of those of the neighbouring regions. Extensive genetic studies of Europe have revealed the links of the gene pool of this continent to the gene pools of other regions, most importantly the Near/Middle East. Any comprehensive understanding of the formation of the current European male gene pool with its major components, the Y chromosome haplogroups R1a and R1b, clearly requires a broader view than that of only Europe itself.

However, there is one region bordering with Europe that has thus far not been studied in detail in terms of its relationship with the European gene pool – namely, the Caucasus. Occupying a junction between the Near/Middle East and East Europe and exhibiting high ethnic and linguistic diversity, the Caucasus is of interest both in its own right and in connection to a wider understanding of human dispersals in Eurasia.

Following are the results of the present study, based on DNA samples mainly from Europe and the Caucasus, with other populations from Africa to East Asia providing the context. While the results on Europe are based on Y chromosome data, the Caucasus is studied from the aspects of its autosomal and maternal as well as its paternal gene pool. New insights into human demographic history in Eurasia are discussed, as well as several issues concerning Y-STR-based dating of genetic lineages.

5.1. Major components of the European paternal gene pool – Y chromosome haplogroups R1a and R1b (ref. I and II)

The two major sub-clades of Y chromosome haplogroup R1, R1a and R1b, have been studied extensively (Balanovsky et al. 2008; Balaresque et al. 2010; Peričić et al. 2005; Rosser et al. 2000; Semino et al. 2000; ref. I, II) due to their high frequency among Europeans and therefore, *ipso facto*, links to major demographic processes.

The phylogeographic spread of haplogroup R1a (Figure 3; ref. II) continues to puzzle researchers, and the source area of haplogroup R1a dispersals remains to be determined. With the aim of discovering informative SNP markers that would allow the separation of the regional sub-clusters of R1a, more than 2000 Y chromosomes belonging to haplogroup R1a were analysed (ref. II). A largely European-specific sub-clade of the haplogroup, defined by the SNP M458, was identified, marking the onset of geographically informative sub-divisions of

R1a (Figure 3; ref. II). The high frequency (maximum around 30% in Poland) and relatively low diversity of R1a-M458 in East Europe suggest a founder effect. According to the coalescence time estimate of the clade the founder effect falls into the early Holocene period, 7.9 ± 2.6 kya, although based solely on the M458 carriers in Poland, where the frequency of the SNP is highest, it occurred earlier, 10.7 ± 4.1 kya. These dates fall into a juncture period between the Mesolithic and early Neolithic in Europe, not allowing for a straightforward connection of the founder effect to any particular culture. The fact that the spread of R1a-M458 is limited to Europe excludes the possibility of any significant patrilineal gene flow from Europe to Asia, at least since the mid-Holocene.
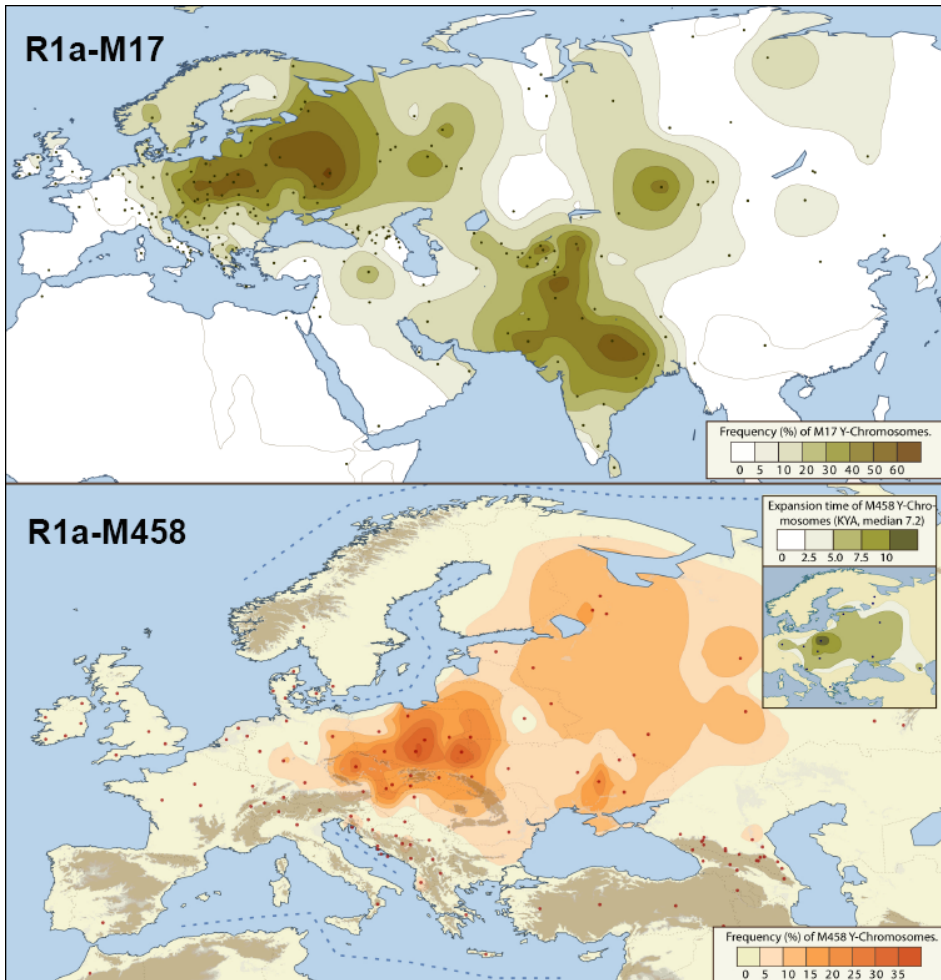

Figure 3. The phylogeographic distribution of the Y chromosome haplogroup R1a1 defined by the marker M17, and its sub-clade defined by the marker M458. Modified from ref. II.

The phylogeographic distribution of haplogroup R1b is more 'traditional' – its main European sub-clade defined by the SNP M269 exhibits a clear west-to-east cline and does not extend far outside Europe at notable frequencies (Figure 1d in ref. I). Haplogroup R1b probably originated in West Asia, its initial differentiation there followed by a rapid spread of the M269 sub-clade to Europe, where it has reached very high frequencies among West Europeans (over 80% in the British Isles and almost 100% in western Ireland) (Figure 1d in ref. I). However, the level of resolution of M269 does not allow for a detailed examination of the past demographic processes linked to this major western European haplogroup, and therefore 2043 M269-derived men from West Asia and Europe were analysed to discover SNP markers that would define geographically informative sub-divisions of R1b (ref. I).

It was found that the M412 SNP largely separates the majority of Central and West European R1b lineages from those observed among eastern populations, representing a major founder effect in West Europe (Figure 4; Figure 1f in ref. I), estimated to have occurred 8.8 ± 1.7 kya, which is considerably more recent than a possible post-LGM re-expansion from the Iberian refugium proposed by Semino et al. (2000), especially taking into account that the methodology used generally overestimates coalescence times (discussed in more detail in the literature overview section "Estimating TMRCA in Y chromosome studies").

The R1b-M412 assemblage is split into several SNP-defined sub-clades with different regional distributions in Europe (Figure 1g–o in ref. I). The frequency pattern and expansion time estimate of a major R1b-M412 sub-clade, defined by the SNP S116, closely approximate the spread of the Linear Pottery Neolithic culture in Europe (Figures 1j and 2 in ref. I). Three sub-haplogroups of the S116 assemblage exhibit further geographic localisation, with the S116*(xU152,M529) branch being most frequent in Iberia (Figure 1k in ref. I), the U152 clade in Switzerland, Italy, France and western Poland (Figure 1l in ref. I), and the M529 branch having the highest occurrence in the British Isles (Figure 1m in ref. I). While the whole S116 assemblage is generally distributed west of the Rhine River, its sister-clade U106 is more frequent east of the Rhine, being spread around the North Sea (Figure 1i in ref. I). While one or more branches within the M412 assemblage are likely associated with Neolithic events, more complex pre-Neolithic scenarios remain possible for the R1b lineage outside the M412 clade, defined by the SNP L23 and spread in East Europe, the Circum-Uralic region, the Near East, the Caucasus and Pakistan (Figure 4; Figure 1e in ref. I).

An analysis of 9 Y-STR loci associated with 840 R1b chromosomes resolved just to the level of M269 concluded that all such chromosomes in Europe reflect a recent genetic heritage that was uniformly introduced by farmers migrating from West Anatolia (Balaresque et al. 2010). However, a look at haplogroup R1b at a deeper phylogenetic and phylogeographic resolution level

offers a more complex picture of the post-glacial formation and expansion of populations in Europe (ref. I).
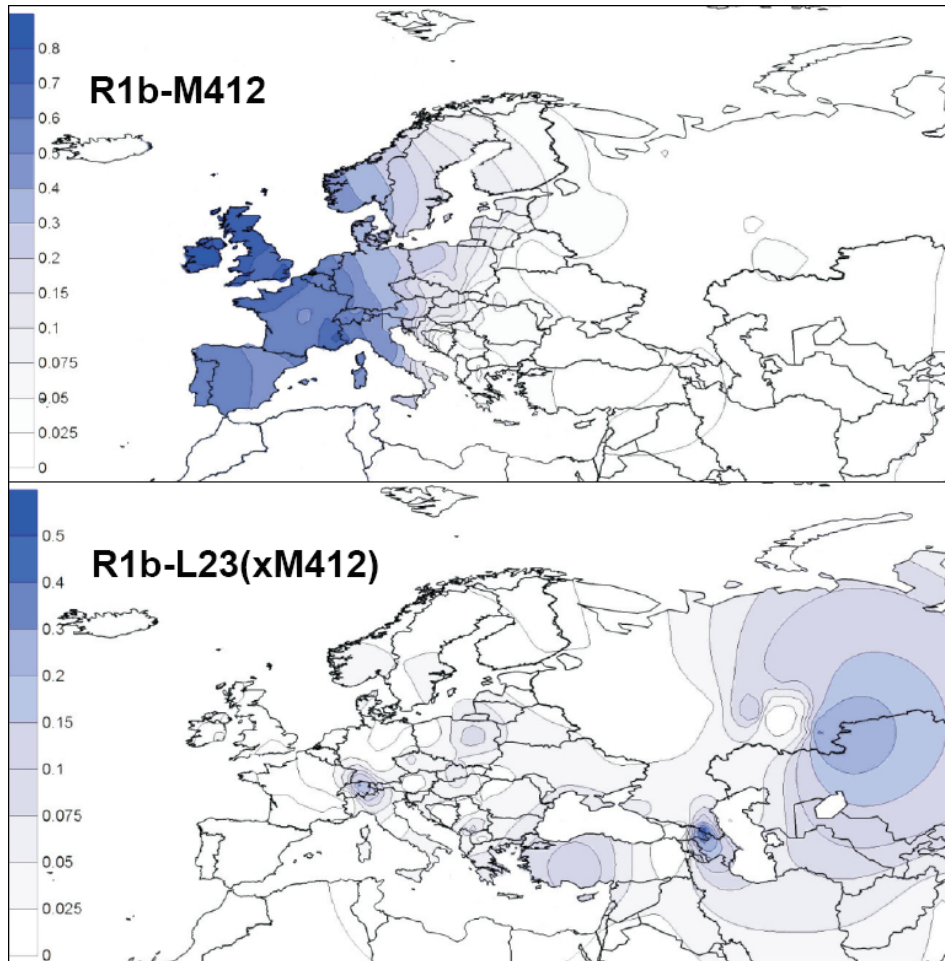
Figure 4. The phylogeographic distribution of two sub-clades of the Y chromosome haplogroup R1b, defined by the markers M412 and L23. Modified from ref. I.

Novel SNP markers M458 and M412 define major European-specific sub-clades, likely associated with Neolithic expansions, in Y chromosome haplo-groups R1a and R1b, respectively, providing a refined picture of the phylo-geographic distribution of these two haplogroups that comprise about half of the European male gene pool.

5.2. The Caucasus and its role in human dispersals in Eurasia (ref. III)

The geographic position of the Caucasus between the Near/Middle East and East Europe and its high ethnic and linguistic diversity raise the questions of how the region was peopled and which later prehistoric and historic migrations to and from it may have occurred, and to which extent the present-day remarkable linguistic diversity is reflected in the genetic variation of the Caucasian populations. The results of the present study, the first synthesis of autosomal, Y chromosome and mtDNA variation in populations from all major sub-regions and linguistic phyla of the Caucasus, attempt to answer these questions.

The populations of the Caucasus, irrespective of their linguistic differences, show the lowest autosomal F_{ST} distances to one another, followed closely by their distance to the populations of the Near/Middle East, Turks in particular (Figure 1B in ref. III). Autosomal, mtDNA and generally also Y chromosome variation in the Caucasus follows geography rather than language, the only exception being the Nakh and Dagestanian language speakers who differ considerably both from each other and the rest of the Caucasian populations due to the unique structure of Y chromosome haplogroup frequencies in Dagestan, but this distinction is not supported by autosomal or mtDNA data (Figure 1C and Supplementary Figures S7 and S8 in ref. III). This evidence of a surprising homogeneity and geography-based clustering in the genetic data of a linguistically diverse region disagrees, to some extent, with the view of gene–language co-evolution in the Caucasus (Balanovsky et al. 2011), suggesting rather that the basic aspects of the genetic structuring of the Caucasian populations have formed before its present-day linguistic diversity, including the language families autochthonous to the region, arose. On the other hand, the sharp contrasts in the variation of paternal lineages detected between the North Caucasus populations both by Balanovsky et al. (2011) and in the present study most likely reflect strong founder event(s) and random genetic drift in general, probably combined with social practices in mating.

However, there is a noticeable lapse in the concordance of genetic data and geography between the North Caucasus and the East European Plain – several analytical methods reveal a sharp border, evident from both autosomal and Y chromosome data and contrasting the continuous transition from the Near/Middle East to the Caucasus. Indeed, while a range of low genetic distances can be observed starting from the populations of the Levant and Syria and proceeding along the western coast of the Black Sea: from Turks to Bulgarians, to Romanians, to Ukrainians, to Russians, ending with Mordvins on the East European Plain, it does not extend through the Caucasus (Figure 1B in ref. III). Modelling genetic distances based on geographic distances, multiple regression analysis revealed that the best fit of the observed autosomal F_{ST} distances between populations to the model was achieved with the assumption of a

Caucasus barrier (Supplementary Table 2 in ref. III). The genetic discontinuity is observed between the populations of the North Caucasus and the East European Plain, not at the High Caucasus Mountain Range. It is possible that a putative post-LGM northward progression of people was halted due to pre-existing habitation of the East European Plain by relatively numerous hunter-gatherer tribes autochthonous to the steppe belt. It may have been additionally inhibited by the Khvalynian transgression, a connection between the Black and Caspian Seas dated 12–14 kya (Badyukova 2007; Svitoch 2009) which may have served as a natural barrier.

The combined autosomal and gender-specific genetic variation of the Caucasian populations testifies to their predominantly Near/Middle Eastern descent, and reveals much greater genetic uniformity and, therefore, deep shared genetic ancestry, between the populations than might be expected from a region of deep linguistic and cultural diversity. The genetic discontinuity between the North Caucasus and the East European Plain, contrasting with continuity through Anatolia and the Balkans, suggests that the movement of people from the Near/Middle East to Europe took place around the western flank of the Black Sea and not through the Caucasus.

5.3. The application of Y chromosome STR loci of increased repeat unit size in population genetic studies (ref. IV)

Compared to di-, tri-, and tetranucleotide repeats, Y chromosome STRs with longer repeat units occur more rarely and are less commonly used in population genetic and evolutionary studies. Previous estimates of 'evolutionary' mutation rate have set it about twice as high for dinucleotide STR loci compared to tri- and tetranucleotide repeats (Zhivotovsky et al. 2000; Zhivotovsky et al. 2004). It is intuitively obvious that the mutation rate of Y chromosome penta- or hexanucleotide STRs should be lower than that of STR loci with smaller repeat unit sizes, since replication slippage, the mechanism of repeat count changes of STRs, is less likely to occur in case of longer repeats. In order to estimate the scale of genetic variation of penta- and hexanucleotide Y-STRs and to compare their mutation rate to that of Y-STRs of smaller repeat unit size, this study analysed 1 tri-, 14 tetra-, 7 penta-, and 2 hexanucleotide Y-STR loci in 148 samples from diverse human populations, representing all the major Y chromosome haplogroups of the world (ref. IV). The study included too few tri- and hexanucleotide markers to make any definitive statements about them, but they were grouped together with tetra- and pentanucleotide markers, respectively, due to similar behaviour.

Both the average repeat variance and the average diversity of penta- and hexanucleotide markers were approximately two times lower than those of tri- and tetranucleotide STRs, an observation that was consistent through the use of

different data sets and haplogroups within the data of this study (Table 2 in ref. IV). Another variable – repeat count – also seemed to have an effect on marker variation, especially on repeat variance (higher repeat variance corresponding to higher repeat count), confirming previous results (Kayser et al. 2004) and making it impossible to state definitively whether STR marker variation, and hence mutation rate, depends on repeat unit size or repeat count (or both).

Perhaps more important than the mutation rate, which is known to vary widely among Y-STRs in any case (discussed in more detail in the literature overview section "Estimating TMRCA in Y chromosome studies"), are other features of penta/hexa markers that underline their applicability in population genetic and evolutionary studies. The penta- and hexanucleotide Y-STRs are more clock-like in their behaviour, their mean variance in better concordance with the age estimates of younger versus older clades than that of tri- and tetranucleotide markers (Table 4 in ref. IV). Penta/hexa STRs also surpass tri/tetra loci in the ability to distinguish Y chromosome haplogroups and, in some cases, subdivisions within haplogroups, without SNP data (Figures 1 and 2 in ref. IV). These characteristics, together with the slower locus saturation, apparently result from the lower rate of evolution of Y-STRs with longer repeat units, and even if their exact mutation rate and the age estimates based on it are subject to debate as is the case with all Y-STR markers, these features make Y-STRs of increased repeat unit size well applicable in evolutionary studies, perhaps even more so than their counterparts with shorter repeat units.

The new information on penta- and hexanucleotide Y-STRs can now be applied in wider population genetic research; however, time is required to type markers with longer repeat units in enough global populations to gather a sufficient amount of comparable data.

6. CONCLUSIONS

- The mutation M458, representing a founder effect around the Mesolithic/ Neolithic transition in Central and East Europe, separates most of the European R1a carriers from the spread of this Y chromosome haplogroup that spans from South Asia to Siberia and Europe.
- In West Europe, another major Holocene era founder effect is denoted by the mutation M412 within the Y chromosome haplogroup R1b, and the spatial and temporal pattern of a sub-clade within R1b-M412 is in close correlation with the spread of the Linear Pottery Neolithic culture.
- Populations of the Caucasus are autosomally much more uniform than might be expected from their diverse linguistic and ethnic backgrounds. Conversely, the variation of Caucasian Y chromosome lineages exhibits sharp differences between populations and sub-regions, likely due to founder effects and genetic drift in (patrilineally) isolated populations.
- The Caucasus has not served as a corridor for the movement of people from the Near/Middle East to East Europe or to Europe in general. Instead, there is a genetic continuity from the Near Eastern to East European populations along the western coast of the Black Sea, suggesting a predominant trajectory of the flow of genes and humans.
- Y chromosome STRs with longer repeat units have a lower rate of evolution, in some cases making them better suited for population genetic studies than their counterparts with shorter repeat units.

REFERENCES

Achilli A, Rengo C, Battaglia V, Pala M, Olivieri A, Fornarino S, Magri C, Scozzari R, Babudri N, Santachiara-Benerecetti AS et al. 2005. Saami and Berbers – An unexpected mitochondrial DNA link. American Journal of Human Genetics 76: 883–886.

Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V et al. 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. American Journal of Human Genetics 75: 910–918.

Adams JM, and Faure H. 1997. Preliminary vegetation maps of the world since the last glacial maximum: An aid to archaeological understanding. J Archaeol Sci 24: 623–647.

Adler DS, Bar-Yosef O, Belfer-Cohen A, Tushabramishvili N, Boaretto E, Mercier N, Valladas H, and Rink WJ. 2008. Dating the demise: neandertal extinction and the establishment of modern humans in the southern Caucasus. J Hum Evol 55: 817–833.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, and Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23: 147.

Armenian census 2001, http: //docs.armstat.am/census/pdfs/51.pdf

Azerbaijani census 2009, http: //en.wikipedia.org/wiki/Azerbaijani_people#cite_note-12

Badyukova EN. 2007. Age of Khvalynian transgressions in the Caspian Sea region. Oceanology 47: 400–405.

Bailey SE, Weaver TD, and Hublin JJ. 2009. Who made the Aurignacian and other early Upper Paleolithic industries? J Hum Evol 57: 11–26.

Balanovsky O, Dibirova K, Dybo A, Mudrak O, Frolova S, Pocheshkhova E, Haber M, Platt D, Schurr T, Haak W et al. 2011. Parallel Evolution of Genes and Languages in the Caucasus Region. Mol Biol Evol 28: 2905–2920.

Balanovsky O, Rootsi S, Pshenichnov A, Kivisild T, Churnosov M, Evseeva I, Pocheshkhova E, Boldyreva M, Yankovsky N, Balanovska E et al. 2008. Two sources of the Russian patrilineal heritage in their Eurasian context. Am J Hum Genet 82: 236–250.

Balaresque P, Bowden GR, Adams SM, Leung HY, King TE, Rosser ZH, Goodwin J, Moisan JP, Richard C, Millward A et al. 2010. A Predominantly Neolithic Origin for European Paternal Lineages. PLoS Biology 8: e1000285.

Ballantyne KN, Goedbloed M, Fang RX, Schaap O, Lao O, Wollstein A, Choi Y, van Duijn K, Vermeulen M, Brauer S et al. 2010. Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications. American Journal of Human Genetics 87: 341–353.

Bar-Yosef O. 1998. The Natufian culture in the Levant, threshold to the origins of agriculture. Evol Anthropol 6: 159–177.

Bar-Yosef O, Belfer-Cohen A, and Adler DS. 2006. The implications of the Middle-Upper Paleolithic chronological boundary in the Caucasus to Eurasian prehistory. Anthropologie XLIV: 49–60.

Batini C, and Jobling MA. 2011. The jigsaw puzzle of our African ancestry: unsolved, or unsolvable? Genome Biol 12: 118.

Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, Myres NM, King RJ, Rootsi S, Marjanovic D, Primorac D et al. 2009. Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. Eur J Hum Genet 17: 820–830.

Behar DM, Garrigan D, Kaplan ME, Mobasher Z, Rosengarten D, Karafet TM, Quintana-Murci L, Ostrer H, Skorecki K, and Hammer MF. 2004. Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations. Hum Genet 114: 354–365.

Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli E-L, Silva NM, Kivisild T, Torroni A, and Villems R. 2012. A "Copernican" Reassessment of the Human Mitochondrial DNA Tree from its Root. Am J Hum Genet 90: 675–684.

Behar DM, Yunusbayev B, Metspalu M, Metspalu E, Rosset S, Parik J, Rootsi S, Chaubey G, Kutuev I, Yudkovsky G et al. 2010. The genome-wide structure of the Jewish people. Nature 466: 238–242.

Benazzi S, Douka K, Fornai C, Bauer CC, Kullmer O, Svoboda J, Pap I, Mallegni F, Bayle P, Coquerelle M et al. 2011. Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. Nature 479: 525–528.

Bermisheva MA, Kutuev IA, Korshunova T, Dubova NA, Villems R, and Khus-nutdinova EK. 2004. Phylogeografic analysis of mitochondrial DNA in the Nogays: the high level of mixture of maternal lineages from Eastern and Western Eurasia. Mol Biol (Mosk) 38: 617–624.

Bosch E, Calafell F, Rosser ZH, Norby S, Lynnerup N, Hurles ME, and Jobling MA. 2003. High level of male-biased Scandinavian admixture in Greenlandic Inuit shown by Y-chromosomal analysis. Human Genetics 112: 353–363.

Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P, Tambets K, Antanaitis-Jacobs I, Haidle MN, Jankauskas R, Kind CJ et al. 2009. Genetic Discontinuity Between Local Hunter-Gatherers and Central Europe's First Farmers. Science 326: 137–140.

Brown TA, Jones MK, Powell W, and Allaby RG. 2008. The complex origins of domesticated crops in the Fertile Crescent. Trends Ecol Evol 24: 103–109.

Brown WM, George M, Jr., and Wilson AC. 1979. Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A 76: 1967–1971.

Browning BL, and Browning SR. 2011. A fast, powerful method for detecting identity by descent. Am J Hum Genet 88: 173–182.

Bulayeva KB, Jorde L, Watkins S, Ostler C, Pavlova TA, Bulayev OA, Tofanelli S, Paoli G, and Harpending H. 2006. Ethnogenomic diversity of Caucasus, Daghestan. Am J Hum Biol 18: 610–620.

Burton ML, Moore CC, Whiting JWM, and Romney AK. 1996. Regions based on social structure. Curr Anthropol 37: 87–123.

Busby GB, Brisighelli F, Sanchez-Diz P, Ramos-Luis E, Martinez-Cadenas C, Thomas MG, Bradley DG, Gusmao L, Winney B, Bodmer W et al. 2012. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. Proc Biol Sci 279: 884–892.

Carvajal-Carmona LG, Soto ID, Pineda N, Ortiz-Barrientos D, Duque C, Ospina-Duque J, McCarthy M, Montoya P, Alvarez VM, Bedoya G et al. 2000. Strong Amerind/ white sex bias and a possible sephardic contribution among the founders of a population in northwest Colombia. American Journal of Human Genetics 67: 1287–1295.

Carvalho-Silva DR, Santos FR, Rocha J, and Pena SDJ. 2001. The phylogeography of Brazilian Y-chromosome lineages. American Journal of Human Genetics 68: 281–286.

Cavalli-Sforza L, Menozzi P, and Piazza A. 1994. The History and Geography of Human Genes. Princeton: Princeton University Press.

Cinnioglu C, King R, Kivisild T, Kalfoglu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K et al. 2004. Excavating Y-chromosome haplotype strata in Anatolia. Hum Genet 114: 127–148.

Clark JD, Beyene Y, WoldeGabriel G, Hart WK, Renne PR, Gilbert H, Defleur A, Suwa G, Katoh S, Ludwig KR et al. 2003. Stratigraphic, chronological and behavioural contexts of Pleistocene Homo sapiens from Middle Awash, Ethiopia. Nature 423: 747–752.

Clottes J, editor. 2001. La Grotte Chauvet: l'art des origines. Paris: Seuil.

Colonna V, Pagani L, Xue Y, and Tyler-Smith C. 2011. A world in a grain of sand: human history from genetic data. Genome Biol 12: 234.

Comrie B. 2008. Linguistic Diversity in the Caucasus. Annu Rev Anthropol 37: 131–143.

Conard NJ, and Bolus M. 2003. Radiocarbon dating the appearance of modern humans and timing of cultural innovations in Europe: new results and new challenges. J Hum Evol 44: 331–371.

Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V et al. 2002. A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. Am J Hum Genet 70: 1197–1214.

Cruciani F, Trombetta B, Massaia A, Destro-Bisol G, Sellitto D, and Scozzari R. 2011. A revised root for the human Y chromosomal phylogenetic tree: the origin of patrilineal diversity in Africa. Am J Hum Genet 88: 814–818.

Cruciani F, Trombetta B, Sellitto D, Massaia A, Destro-Bisol G, Watson E, Beraud Colomb E, Dugoujon JM, Moral P, and Scozzari R. 2010. Human Y chromosome haplogroup R-V88: a paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. Eur J Hum Genet 18: 800–807.

d'Errico F. 2003. The invisible frontier. A multiple species model for the origin of behavioral modernity. Evol Anthropol 12: 188–202.

d'Errico F, Henshilwood C, Vanhaeren M, and van Niekerk K. 2005. Nassarius kraussianus shell beads from Blombos Cave: evidence for symbolic behaviour in the Middle Stone Age. J Hum Evol 48: 3–24.

d'Errico F, Zilhao J, Julien M, Baffier D, and Pelegrin J. 1998. Neanderthal acculturation in western Europe? A critical review of the evidence and its interpretation. Curr Anthropol 39: S1–S44.

de Knijff P. 2000. Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome. Am J Hum Genet 67: 1055–1061.

Derenko M, Malyarchuk B, Denisova GA, Wozniak M, Dambueva I, Dorzhu C, Luzina F, Miscicka-Sliwka D, and Zakharov I. 2006. Contrasting patterns of Y-chromosome variation in south Siberian populations from Baikal and Altai-Sayan regions. Human Genetics 118: 591–604.

Di Giacomo F, Luca F, Popa LO, Akar N, Anagnou N, Banyko J, Brdicka R, Barbujani G, Papola F, Ciavarella G et al. 2004. Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. Human Genetics 115: 357–371.

Djindjian F. 2000. The Mid Upper Palaeolithic (30,000 to 20,000 BP) in France. In: Roebroeks W, Mussi M, Svoboda J, and Fennema K, editors. Hunters of the golden age: the Mid Upper Palaeolithic of Eurasia 30,000–20,000 BP. Leiden: Analecta Praehistorica Leidensia, University of Leiden. p 313–324.

Dolukhanov PM. 1993. Foraging and farming groups in north-eastern and north-western Europe: identity and interaction. In: Chapman J, and Dolukhanov P, editors. Cultural Transformations and Interactions in Eastern Europe. Aldershot: Avebury.

Dolukhanov PM. 1997. The Pleistocene-Holocene transition in northern Eurasia: Environmental changes and human adaptations. Quatern Int 41–2: 181–191.

Dubcovsky J, and Dvorak J. 2007. Genome plasticity a key factor in the success of polyploid wheat under domestication. Science 318: 393–393.

Dvorak J, Luo MC, Yang ZL, and Zhang HB. 1998. The structure of the Aegilops tauschii genepool and the evolution of hexaploid wheat. Theor Appl Genet 97: 657–670.

Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R et al. 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447: 1087–1093.

Encyclopaedia Britannica, http: //www.britannica.com

Fairbanks RG, Mortlock RA, Chiu T-C, Cao L, Kaplan A, Guilderson TP, Fairbanks TW, Bloom AL, Grootes PM, and Nadeau M-J. 2005. Radiocarbon calibration curve spanning 0 to 50,000 years BP based on paired 230Th/ 234U/ 238U and 14C dates on pristine corals. Quaternary Sci Rev 24: 1781–1796.

Feldman MF. 2001. Origin of cultivated wheat. In: Bonjean AP, and Angus WJ, editors. The World Wheat Book: A History of Wheat Breeding. Paris: Lavoisier Publishing. p 1–56.

Forster P, Harding R, Torroni A, and Bandelt HJ. 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. Am J Hum Genet 59: 935–945.

Forster P, Rohl A, Lunnemann P, Brinkmann C, Zerjal T, Tyler-Smith C, and Brinkmann B. 2000. A short tandem repeat-based phylogeny for the human Y chromosome. Am J Hum Genet 67: 182–196.

Frachetti MD. 2012. Multiregional Emergence of Mobile Pastoralism and Nonuniform Institutional Complexity across Eurasia. Curr Anthropol 53: 2–38.

Franke A, Balschun T, Sina C, Ellinghaus D, Hasler R, Mayr G, Albrecht M, Wittig M, Buchert E, Nikolaus S et al. 2010. Genome-wide association study for ulcerative colitis identifies risk loci at 7q22 and 22q13 (IL17REL). Nat Genet 42: 292–294.

Gamble C, Davies W, Pettitt P, Hazelwood L, and Richards M. 2005. The archaeological and genetic foundations of the European population during the late glacial: Implications for 'agricultural thinking'. Camb Archaeol J 15: 193–223.

Gamble C, Davies W, Pettitt P, and Richards M. 2004. Climate change and evolving human diversity in Europe during the last glacial. Philos T Roy Soc B 359: 243–253.

Ge JY, Budowle B, Aranda XG, Planz JV, Eisenberg AJ, and Chakraborty R. 2009. Mutation rates at Y chromosome short tandem repeats in Texas populations. Forensic Sci Int-Gen 3: 179–184.

Georgian census 2002, http: //www.ecmicaucasus.org/upload/stats/Census 2002.pdf

Goedbloed M, Vermeulen M, Fang RN, Lembring M, Wollstein A, Ballantyne K, Lao O, Brauer S, Kruger C, Roewer L et al. 2009. Comprehensive mutation analysis of 17 Y-chromosomal short tandem repeat polymorphisms included in the AmpFlSTR Yfiler PCR amplification kit. Int J Legal Med 123: 471–482.

Gronau I, Hubisz MJ, Gulko B, Danko CG, and Siepel A. 2011. Bayesian inference of ancient human demography from individual genome sequences. Nat Genet 43: 1031–1034.

Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, Der Sarkissian CS, Brandt G, Schwarz C, Nicklisch N et al. 2010. Ancient DNA from European early neolithic farmers reveals their near eastern affinities. PLoS Biol 8: e1000536.

Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tanzer M, Villems R, Renfrew C, Gronenborn D, Alt KW et al. 2005. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. Science 310: 1016–1018.

Hammer MF, Chamberlain VF, Kearney VF, Stover D, Zhang G, Karafet T, Walsh B, and Redd AJ. 2006. Population structure of Y chromosome SNP haplogroups in the United States and forensic implications for constructing Y chromosome STR databases. Forensic Science International 164: 45–55.

Henshilwood CS, d'Errico F, Yates R, Jacobs Z, Tribolo C, Duller GAT, Mercier N, Sealy JC, Valladas H, Watts I et al. 2002. Emergence of modern human behavior: Middle Stone Age engravings from South Africa. Science 295: 1278–1280.

Herrera KJ, Lowery RK, Hadden L, Calderon S, Chiou C, Yepiskoposyan L, Regueiro M, Underhill PA, and Herrera RJ. 2012. Neolithic patrilineal signals indicate that the Armenian plateau was repopulated by agriculturalists. Eur J Hum Genet 20: 313–320.

Heyer E, Puymirat J, Dieltjes P, Bakker E, and deKnijff P. 1997. Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Human Molecular Genetics 6: 799–803.

Higham T, Compton T, Stringer C, Jacobi R, Shapiro B, Trinkaus E, Chandler B, Groning F, Collins C, Hillson S et al. 2011. The earliest evidence for anatomically modern humans in northwestern Europe. Nature 479: 521–524.

Hovers E, Ilani S, Bar-Yosef O, and Vandermeersch B. 2003. An early case of color symbolism – Ochre use by modern humans in Qafzeh cave. Curr Anthropol 44: 491–522.

Hughen K, Lehman S, Southon J, Overpeck J, Marchal O, Herring C, and Turnbull J. 2004. 14C Activity and Global Carbon Cycle Changes over the Past 50,000 Years. Science 303: 202–207.

Hurles ME, Irven C, Nicholson J, Taylor PG, Santos FR, Loughlin J, Jobling MA, and Sykes BC. 1998. European Y-chromosomal lineages in Polynesians: a contrast to the population structure revealed by mtDNA. Am J Hum Genet 63: 1793–1806.

ISOGG Y-DNA Haplogroup Tree 2012, http: //www.isogg.org/tree/

Jeffreys AJ, Kauppi L, and Neumann R. 2001. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat Genet 29: 217–222.

Jobling MA, and Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4: 598–612.

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, and Hammer MF. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res 18: 830–838.

Kayser M, Brauer S, Weiss G, Schiefenhovel W, Underhill PA, and Stoneking M. 2001. Independent histories of human Y chromosomes from Melanesia and Australia. Am J Hum Genet 68: 173–190.

Kayser M, Kittler R, Erler A, Hedman M, Lee AC, Mohyuddin A, Mehdi SQ, Rosser Z, Stoneking M, Jobling MA et al. 2004. A comprehensive survey of human Y-chromosomal microsatellites. Am J Hum Genet 74: 1183–1197.

Keith A. 1940. Blumenbach's Centenary. Man 40: 82–85.

King RJ, Ozcan SS, Carter T, Kalfoglu E, Atasoy S, Triantaphyllidis C, Kouvatsi A, Lin AA, Chow CET, Zhivotovsky LA et al. 2008. Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. Ann Hum Genet 72: 205–214.

Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K, Passarino G, Underhill PA, Scharfe C, Torroni A et al. 2006. The role of selection in the evolution of human mitochondrial genomes. Genetics 172: 373–387.

Kolga M, Tõnurist I, Vaba L, and Viikberg J. 2001. The Red Book of the Peoples of the Russian Empire. Tallinn: NGO Red Book.

Kozlowski JK. 2007. The archaeological and behavioural records in central and eastern Europe. The significance of blade technologies in the period 50–35 kyr BP for the middle Palaeolithic-upper Palaeolithic transition in central and eastern Europe. In: Mellars P, Boyle K, Bar-Yosef O, and Stringer C, editors. Rethinking the Human Revolution. Cambridge: McDonald Institute for Archaeological Research. p 317–328.

Krause J, Briggs AW, Kircher M, Maricic T, Zwyns N, Derevianko A, and Paabo S. 2010. A Complete mtDNA Genome of an Early Modern Human from Kostenki, Russia. Curr Biol 20: 231–236.

Kuijt I, and Goring-Morris N. 2002. Foraging, farming, and social complexity in the Pre-Pottery Neolithic of the Southern Levant: A review and synthesis. J World Prehist 16: 361–440.

Kushnareva KK. 1997. The southern Caucasus in prehistory: stages of cultural and socioeconomic development from the eighth to the second millennium B.C. Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology.

Lacan M, Keyser C, Ricaut FX, Brucato N, Duranthon F, Guilaine J, Crubezy E, and Ludes B. 2011a. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. Proc Natl Acad Sci U S A 108(24): 9788–9791.

Lacan M, Keyser C, Ricaut FX, Brucato N, Tarrus J, Bosch A, Guilaine J, Crubezy E, and Ludes B. 2011b. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. Proc Natl Acad Sci U S A 108: 18255–18259.

Lahn BT, and Page DC. 1999. Four evolutionary strata on the human Y chromosome. Science 286: 964–967.

Lappalainen T, Laitinen V, Salmela E, Andersen P, Huoponen K, Savontaus ML, and Lahermo P. 2008. Migration waves to the Baltic Sea region. Annals of Human Genetics 72: 337–348.

Лавров ЛИ. 1978. Историко-этнографические очерки Кавказа. Ленинград: Наука.

Lawson D, Hellenthal G, Myers S, and Falush D. 2012. Inference of population structure using dense genotype data. PloS Genet unpublished.

Lewis-Williams D. 2002. The mind in the cave. London: Thames & Hudson.

Li H, and Durbin R. 2011. Inference of human population history from individual whole-genome sequences. Nature 475: 493–496.

Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL et al. 2008. Worldwide human

relationships inferred from genome-wide patterns of variation. Science 319: 1100–1104.

Lieberman DE. 2007. Palaeoanthropology: homing in on early Homo. Nature 449: 291–292.

Loogväli EL, Roostalu U, Malyarchuk BA, Derenko MV, Kivisild T, Metspalu E, Tambets K, Reidla M, Tolk HV, Parik J et al. 2004. Disuniting uniformity: A pied cladistic canvas of mtDNA haplogroup H in Eurasia. Molecular Biology and Evolution 21: 2012–2021.

Lordkipanidze D, Jashashvili T, Vekua A, Ponce de Leon MS, Zollikofer CP, Rightmire GP, Pontzer H, Ferring R, Oms O, Tappen M et al. 2007. Postcranial evidence from early Homo from Dmanisi, Georgia. Nature 449: 305–310.

Lowery RK, Herrera KJ, Barrett DA, Rodriguez R, Hadden LRM, Harutyunyan A, Margaryan A, Yepiskoposyan L, and Herrera RJ. 2011. Regionalized Autosomal STR Profiles among Armenian Groups Suggest Disparate Genetic Influences. Am J Phys Anthropol 146: 171–178.

Luca F, Basile M, Di Giacomo F, and Novelletto A. 2005. Independent methods for evolutionary genetic dating provide insights into Y-chromosomal STR mutation rates confirming data from direct father-son transmissions. Hum Genet 118: 153–165.

Malyarchuk B, Derenko M, Grzybowski T, Perkova M, Rogalla U, Vanecek T, and Tsybovsky I. 2010. The Peopling of Europe from the Mitochondrial Haplogroup U5 Perspective. PLoS One 5: e10285.

Malyarchuk B, Grzybowski T, Derenko M, Perkova M, Vanecek T, Lazur J, Gomolcak P, and Tsybovsky I. 2008. Mitochondrial DNA phylogeny in eastern and western Slavs. Molecular Biology and Evolution 25: 1651–1658.

Marchani EE, Watkins WS, Bulayeva K, Harpending HC, and Jorde LB. 2008. Culture creates genetic structure in the Caucasus: autosomal, mitochondrial, and Y-chromosomal variation in Daghestan. Bmc Genet 9: 47.

McDougall I, Brown FH, and Fleagle JG. 2005. Stratigraphic placement and age of modern humans from Kibish, Ethiopia. Nature 433: 733–736.

McGovern PE. 2003. Ancient Wine: The Search for the Origins of Viniculture. Princeton: Princeton University Press.

Mellars P. 1996. The Neanderthal legacy: an archaeological perspective from western Europe. Princeton: Princeton University Press.

Mellars P. 2004. Neanderthals and the modern human colonization of Europe. Nature 432: 461–465.

Mellars P. 2005. The impossible coincidence. A single-species model for the origins of modern human behavior in Europe. Evol Anthropol 14: 12–27.

Mellars P. 2006a. A new radiocarbon revolution and the dispersal of modern humans in Eurasia. Nature 439: 931–935.

Mellars P. 2006b. Why did modern human populations disperse from Africa ca. 60,000 years ago? A new model. P Natl Acad Sci USA 103: 13560–13560.

Mellars P. 2006c. Archeology and the dispersal of modern humans in Europe: Deconstructing the "Aurignacian". Evol Anthropol 15: 167–182.

Mellars P. 2011. PALAEOANTHROPOLOGY The earliest modern humans in Europe. Nature 479: 483–485.

Metspalu E, Kivisild T, Kaldma K, Parik J, Reidla M, Tambets K, and Villems R. 1999. The Trans-Caucasus and the expansion of the Caucasoid-specific human mito-chrondrial DNA. In: Papiha D, and Chakraborty K, editors. Genomic Diversity:

Applications in Human Population Genetics. New York: Academic / Plenum Publishers. p 121–133.

Metzker ML. 2010. Sequencing technologies – the next generation. Nat Rev Genet 11: 31–46.

Minahan JB. 2000. One Europe, many nations: a historical dictionary of European national groups: Greenwood Publishing Group.

Minoche AE, Dohm JC, and Himmelbauer H. 2011. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. Genome Biol 12: R112.

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, Easley K, Chen E, Brown MD et al. 2003. Natural selection shaped regional mtDNA variation in humans. Proc Natl Acad Sci U S A 100: 171–176.

Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK, Prins B, Reynolds A, Chia JM, Ware D et al. 2011. Genetic structure and domestication history of the grape. P Natl Acad Sci USA 108: 3530–3535.

Nasidze I, Risch GM, Robichaux M, Sherry ST, Batzer MA, and Stoneking M. 2001. Alu insertion polymorphisms and the genetic structure of human populations from the Caucasus. Eur J Hum Genet 9: 267–272.

Nasidze I, and Stoneking M. 2001. Mitochondrial DNA variation and language replacements in the Caucasus. P Roy Soc Lond B Bio 268: 1197–1206.

Nasidze I, Sarkisian T, Kerimov A, and Stoneking M. 2003. Testing hypotheses of language replacement in the Caucasus: evidence from the Y-chromosome. Hum Genet 112: 255–261.

Nasidze I, Ling EY, Quinque D, Dupanloup I, Cordaux R, Rychkov S, Naumova O, Zhukova O, Sarraf-Zadegan N, Naderi GA et al. 2004a. Mitochondrial DNA and Y-chromosome variation in the Caucasus. Ann Hum Genet 68: 205–221.

Nasidze I, Quinque D, Dupanloup I, Rychkov S, Naumova O, Zhukova O, and Stoneking M. 2004b. Genetic evidence concerning the origins of South and North Ossetians. Ann Hum Genet 68: 588–599.

Nasidze I, Quinque D, Ozturk M, Bendukidze N, and Stoneking M. 2005a. MtDNA and Y-chromosome variation in Kurdish groups. Ann Hum Genet 69: 401–412.

Nasidze I, Quinque D, Dupanloup I, Cordaux R, Kokshunova L, and Stoneking M. 2005b. Genetic evidence for the Mongolian ancestry of Kalmyks. Am J Phys Anthropol 128: 846–854.

Nelis M, Esko T, Magi R, Zimprich F, Zimprich A, Toncheva D, Karachanak S, Piskackova T, Balascak I, Peltonen L et al. 2009. Genetic Structure of Europeans: A View from the North-East. PLoS One 4: e5472.

Nigst PR. 2006. In: Conard NJ, editor. When Neanderthals and Modern Humans Met. Tübingen: Kerns. p 269–304.

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR et al. 2008. Genes mirror geography within Europe. Nature 456: 98–U95.

Olmo H. 1995. Grapes. In: Smartt J, and Simmonds N, editors. Evolution of Crop Plants. 2nd ed. New York: Longman. p 485–490.

Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, and Stoneking M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. Nat Genet 29: 20–21.

Pala M, Achilli A, Olivieri A, Kashani BH, Perego UA, Sanna D, Metspalu E, Tambets K, Tamm E, Accetturo M et al. 2009. Mitochondrial Haplogroup U5b3: A Distant

Echo of the Epipaleolithic in Italy and the Legacy of the Early Sardinians. American Journal of Human Genetics 84: 814–821.

Pala M, Olivieri A, Achilli A, Accetturo M, Metspalu E, Reidla M, Tamm E, Karmin M, Kashani BH, Perego UA et al. 2012. Mitochondrial DNA signals of Late Glacial re-colonisation of Europe from Near Eastern refugia. Am J Hum Genet, accepted for publication.

Patterson N, Price AL, and Reich D. 2006. Population structure and eigenanalysis. PLoS Genet 2: 2074–2093.

Peltenburg E. 1991. Kissonerga-Mosphilia, a Major Chalcolithic Site in Cyprus. B Am Sch Oriental Re 282/283: 17–35.

Pereira L, Richards M, Goios A, Alonso A, Albarran C, Garcia O, Behar DM, Golge M, Hatina J, Al-Gazali L et al. 2005. High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. Genome Research 15: 19–24.

Peričić M, Lauc LB, Klaric IM, Janicijevic B, and Rudan P. 2005. Review of croatian genetic heritage as revealed by mitochondrial DNA and Y chromosomal lineages. Croat Med J 46: 502–513.

Perles C. 2000. Greece, 30,000–20,000 BP. In: Roebroeks W, Mussi M, Svoboda J, and Fennema K, editors. Hunters of the golden age: the Mid Upper Palaeolithic of Eurasia 30,000–20,000 BP. Leiden: Analecta Praehistorica Leidensia, University of Leiden. p 375–398.

Pinhasi R, Gasparian B, Wilkinson K, Bailey R, Bar-Oz G, Bruch A, Chataigner C, Hoffmann D, Hovsepyan R, Nahapetyan S et al. 2008. Hovk 1 and the Middle and Upper Paleolithic of Armenia: a preliminary framework. J Hum Evol 55: 803–816.

Pinhasi R, Higham TFG, Golovanova LV, and Doronichev VB. 2011. Revised age of late Neanderthal occupation and the end of the Middle Paleolithic in the northern Caucasus. P Natl Acad Sci USA 108: 8611–8616.

Prat S, Pean SC, Crepin L, Drucker DG, Puaud SJ, Valladas H, Laznickova-Galetova M, van der Plicht J, and Yanevich A. 2011. The Oldest Anatomically Modern Humans from Far Southeast Europe: Direct Dating, Culture and Behavior. PLoS One 6: e20834.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, and Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904–909.

Pritchard JK, Seielstad MT, Perez-Lezaun A, and Feldman MW. 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. Mol Biol Evol 16: 1791–1798.

Pritchard JK, Stephens M, and Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Radivojevic M, Rehren T, Pernicka E, Sljivar D, Brauns M, and Boric D. 2010. On the origins of extractive metallurgy: new evidence from Europe. J Archaeol Sci 37: 2775–2787.

Ramsey CB, Higham T, Bowles A, and Hedges R. 2004. Improvements to the pretreat-ment of bone at Oxford. Radiocarbon 46: 155–163.

Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R et al. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463: 757–762.

Renfrew C, and Boyle K, editors. 2000. Archaeogenetics: DNA and the population prehistory of Europe. Cambridge: McDonald Institute for Archaeological Research.

Richards M, CorteReal H, Forster P, Macaulay V, WilkinsonHerbots H, Demaine A, Papiha S, Hedges R, Bandelt HJ, and Sykes B. 1996. Paleolithic and neolithic lineages in the European mitochondrial gene pool. American Journal of Human Genetics 59: 747–747.

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T et al. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. Am J Hum Genet 67: 1251–1276.

Roostalu U, Kutuev I, Loogväli EL, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, and Villems R. 2007. Origin and expansion of haplo-group H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. Mol Biol Evol 24: 436–448.

Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barac L, Peričić M, Balanovsky O et al. 2004. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. American Journal of Human Genetics 75: 128–137.

Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G et al. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet 67: 1526–1543.

Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, and Page DC. 2003. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature 423: 873–876.

Russian census 2010, http: //en.wikipedia.org/wiki/Russians#cite_note-gks-1

Saillard J, Forster P, Lynnerup N, Bandelt HJ, and Norby S. 2000. mtDNA variation among Greenland Eskimos: The edge of the Beringian expansion. American Journal of Human Genetics 67: 718–726.

Santachiara Benerecetti AS, Semino O, Passarino G, Torroni A, Brdicka R, Fellous M, and Modiano G. 1993. The common, Near-Eastern origin of Ashkenazi and Sep-hardi Jews supported by Y-chromosome similarity. Ann Hum Genet 57: 55–64.

Seielstad MT, Minch E, and Cavalli-Sforza LL. 1998. Genetic evidence for a higher female migration rate in humans. Nat Genet 20: 278–280.

Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V, Maccioni L, Triantaphyllidis C, Shen P, Oefner PJ et al. 2004. Origin, diffusion, and differen-tiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. Am J Hum Genet 74: 1023–1034.

Semino O, Passarino G, Brega A, Fellous M, and SantachiaraBenerecetti AS. 1996. A view of the neolithic demic diffusion in Europe through two Y chromosome-specific markers. American Journal of Human Genetics 59: 964–968.

Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S et al. 2000. The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science 290: 1155–1159.

Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, Chow CE, Lin AA, Mitra M, Sil SK, Ramesh A et al. 2006. Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. Am J Hum Genet 78: 202–221.

Sherratt A. 1997. Climatic cycles and behavioural revolutions: The emergence of modern humans and the beginning of farming. Antiquity 71: 271–287.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423: 825–837.

Smith NG, Webster MT, and Ellegren H. 2002. Deterministic mutation rate variation in the human genome. Genome Res 12: 1350–1356.

Soares P, Achilli A, Semino O, Davies W, Macaulays V, Bandelt HJ, Torroni A, and Richards MB. 2010. The Archaeogenetics of Europe. Curr Biol 20: R174–R183.

Szmidt CC, Brou L, and Jaccottey L. 2010. Direct radiocarbon (AMS) dating of split-based points from the (Proto)Aurignacian of Trou de la Mere Clochette, Northeastern France. Implications for the characterization of the Aurignacian and the timing of technical innovations in Europe. J Archaeol Sci 37: 3320–3337.

Stringer C. 2011. The Origin of Our Species. London: Allen Lane.

Svitoch AA. 2009. Khvalynian transgression of the Caspian Sea was not a result of water overflow from the Siberian Proglacial lakes, nor a prototype of the Noachian flood. Quatern Int 197: 115–125.

Zerjal T, Xue Y, Bertorelle G, Wells RS, Bao W, Zhu S, Qamar R, Ayub Q, Mohyuddin A, Fu S et al. 2003. The genetic legacy of the Mongols. Am J Hum Genet 72: 717–721.

Zhivotovsky LA, Bennett L, Bowcock AM, and Feldman MW. 2000. Human population expansion and microsatellite variation. Molecular Biology and Evolution 17: 757–767.

Zhivotovsky LA, Underhill PA, Cinnioglu C, Kayser M, Morar B, Kivisild T, Scozzari R, Cruciani F, Destro-Bisol G, Spedini G et al. 2004. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. Am J Hum Genet 74: 50–61.

Zhivotovsky LA, Underhill PA, and Feldman MW. 2006. Difference between evolutionarily effective and germ line mutation rate due to stochastically varying haplogroup size. Molecular Biology and Evolution 23: 2268–2270.

Tambets K, Kivisild T, Metspalu E, Parik J, Kaldma K, Laos S, Tolk H-V, Gölge M, Demirtas H, Geberhiwot T et al. 2000. The Topology of the Maternal Lineages of the Anatolian and Trans-Caucasus Populations and the Peopling of Europe: Some Preliminary Considerations. In: Renfrew C, and Boyle K, editors. Archaeogenetics: DNA and the population prehistory in Europe. Cambridge: Cambridge Univ. Press. p 219–235.

Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogväli EL, Tolk HV, Reidla M, Metspalu E, Pliss L et al. 2004. The western and eastern roots of the Saami − The story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. American Journal of Human Genetics 74: 661–682.

Tarasov PE, Peyron O, Guiot J, Brewer S, Volkova VS, Bezusko LG, Dorofeyuk NI, Kvavadze EV, Osipova IM, and Panova NK. 1999. Last Glacial Maximum climate of the former Soviet Union and Mongolia reconstructed from pollen and plant macrofossil data. Clim Dynam 15: 227–240.

Tarasov PE, Volkova VS, Webb III T, Guiot J, Andreev AA, Bezusko LG, Bezusko TV, Bykova GV, Dorofeyuk NI, Kvavadze EV et al. 2000. Last glacial maximum biomes reconstructed from pollen and plant macrofossil data from northern Eurasia. J Biogeogr 27: 609–620.

The 1000 Genomes Project Consortium 2010. A map of human genome variation from population-scale sequencing. Nature 467: 1061–1073.

The Global Lexicostatistical Database, http: //starling.rinet.ru/new100/main.htm

The International HapMap 3 Consortium. 2010. Integrating common and rare genetic variation in diverse human populations. Nature 467: 52–58.

The International HapMap Consortium 2005. A haplotype map of the human genome. Nature 437: 1299–1320.

The Wellcome Trust Case Control Consortium 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.

The Y Chromosome Consortium 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12: 339–348.

Thomson R, Pritchard JK, Shen P, Oefner PJ, and Feldman MW. 2000. Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. Proc Natl Acad Sci U S A 97: 7360–7365.

Torroni A, Achilli A, Macaulay V, Richards M, and Bandelt HJ. 2006. Harvesting the fruit of the human mtDNA tree. Trends Genet 22: 339–345.

Torroni A, Bandelt HJ, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus ML, Bonne-Tamir B et al. 1998. mtDNA analysis reveals a major late paleolithic population expansion from southwestern to northeastern Europe. American Journal of Human Genetics 62: 1137–1152.

Torroni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, Martinez-Cabrera V, Villems R, Kivisild T, Metspalu E et al. 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. American Journal of Human Genetics 69: 844–852.

Torroni A, Semino O, Scozzari R, Sirugo G, Spedini G, Abbas N, Fellous M, and Santachiara Benerecetti AS. 1990. Y chromosome DNA polymorphisms in human populations: differences between Caucasoids and Africans detected by 49a and 49f probes. Ann Hum Genet 54: 287–296.

Trinkaus E. 2005. Early modern humans. Annu Rev Anthropol 34: 207–230.

Underhill PA, and Kivisild T. 2007. Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. Annu Rev Genet 41: 539–564.

Underhill PA, Myres NM, Rootsi S, Chow C-ET, Lin AA, Otillar RP, King R, Zhivotovsky LA, Balanovsky O, Pshenichnov A et al. 2007. New Phylogenetic Relationships for Y-chromosome Haplogroup I: Reappraising its Phylogeography and Prehistory. In: Mellars P, Boyle K, Bar-Yosef O, and Stringer C, editors. Rethinking the Human Revolution: New Behavioural and Biological Perspectives on the Origin and Dispersal of Modern Humans (McDonald Institute Monographs) Cambridge: McDonald Institute for Archaeological Research. p 33–42.

Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ, and Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplo-types and the origins of modern human populations. Ann Hum Genet 65: 43–62.

van der Plicht J, Beck JW, Bard E, Baillie MGL, Blackwell PG, Buck CE, Friedrich M, Guilderson TP, Hughen KA, Kromer B et al. 2004. NotCal04; comparison/calibration 14C records 26–50 cal kyr BP. Radiocarbon 46: 1225–1238.

Weale ME, Yepiskoposyan L, Jager RF, Hovhannisyan N, Khudoyan A, Burbage-Hall O, Bradman N, and Thomas MG. 2001. Armenian Y chromosome haplotypes reveal strong regional structure within a single ethno-national group. Human Genetics 109: 659–674.

Velichko AA, Kurenkova EI, and Dolukhanov PM. 2009. Human socio-economic adaptation to environment in Late Palaeolithic, Mesolithic and Neolithic Eastern Europe. Quatern Int 203: 1–9.

White R. 1993. Technological and social dimensions of "Aurignacian age" body ornaments across Europe. In: Knecht H, Pike-Tay A, and White R, editors. Before Lascaux: the complex record of the early Upper Paleolithic. Boca Raton: CRC Press. p 277–300.

White R. 1997. Substantial acts: from materials to meaning in Upper Paleolithic representation. In: Conkey MW, Soffer O, Stratmann D, and Jablonski NG, editors. Beyond art: Pleistocene image and symbol. San Francisco: California Academy of Sciences. p 93–121.

Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, and Hammer MF. 2004. Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. Nature Genetics 36: 1238–1238.

Wilson IJ, Weale ME, and Balding DJ. 2003. Inferences from DNA data: population histories, evolutionary processes and forensic match probabilities. J R Stat Soc a Stat 166: 155–188.

Xue Y, Zerjal T, Bao W, Zhu S, Lim SK, Shu Q, Xu J, Du R, Fu S, Li P et al. 2005. Recent spread of a Y-chromosomal lineage in northern China and Mongolia. Am J Hum Genet 77: 1112–1116.

Xue Y, Wang Q, Long Q, Ng BL, Swerdlow H, Burton J, Skuce C, Taylor R, Abdellah Z, Zhao Y et al. 2009. Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. Curr Biol 19: 1453–1457.

SUMMARY IN ESTONIAN

Erinevad geneetilised perspektiivid inimajaloole Euroopas ja Kaukaasias: lood, mida räägivad haploidsed ja autosomaalsed markerid

Inimese demograafilise ajaloo uurimiseks kasutatakse laialdaselt geneetilisi andmeid. Juba paarkümmend aastat on läbi viidud emaliinis päranduva mito-kondriaalse DNA (mtDNA) ja isaliinis päranduva Y-kromosoomi põhiseid uuringuid, viimasel ajal on lisandunud ülegenoomsete markerite tüpiseerimine. Lähitulevikus on oodata hulgaliselt – populatsioonide tasandil – genoomide täisjärjestusi, aga juba praegu määratakse paljudel proovidel rutiinselt sadu tuhandeid punktmutatsioone (*single nucleotide polymorphism*, SNP) üle kogu genoomi.

Euroopa praeguse geneetilise struktuuri kujunemine on laiaulatuslik ja intri-geeriv küsimus, mida on pikka aega uuritud, sh nii selle maailmajao algset asus-tamist anatoomiliselt moodsa inimese poolt, viimase jääaja järgset taasasusta-mist refuugiumitest kui ka neoliitilist revolutsiooni, st põllumajanduse levimist Lähis-Idast. Seni pole jõutud üksmeelele, kas üleminek põllumajandusele toimus peamiselt kultuuriliselt või hõlmas ka rahvastiku ulatuslikku vahetumist. Käesolevas töös analüüsiti Lääne-Euraasia, sh Euroopa geneetilist struktuuri Y-kromosoomi andmete põhjal, käsitledes haplogruppe R1a ja R1b, mis on Euroopas sagedased (umbes 50% Euroopa meestest kuulub neid ühendavasse haplogruppi R), kuid vastandliku levikuga – R1a on kõrge sagedusega Ida- ja R1b Lääne-Euroopas. R1a levik ei piirdu aga Euroopaga, ulatudes India ja Siberini.

Kaukaasia, mägine piirkond Musta ja Kaspia mere vahel, mis on silmatorkav kõrge etnilise, kultuurilise ja keelelise mitmekesisuse poolest, pakub kõigepealt huvi seoses küsimusega, kuivõrd peegeldub see mitmekesisus piirkonna popu-latsioonide geneetilises struktuuris. Samuti on intrigeeriv Kaukaasia roll ini-mese rännetes Euraasias, sh Euroopa asustamises. Käesolevas töös kasutati Kaukaasia populatsioonide uurimisel esmakordselt nii mtDNA, Y-kromosoomi kui ka autosomaalsete SNP markerite andmeid.

Y-kromosoomi haplogruppide dateerimiseks kasutatakse lühikesi tandeem-seid kordusjärjestusi (*short tandem repeats*, STR). Laialdaselt on kasutusel kolme- ja neljanukleotiidsete kordustega STR markerid, samas kui pikemate kordusühikutega markereid pole põhjalikult uuritud.

Käesoleva töö esimeseks eesmärgiks oli esiteks analüüsida detailselt Y-kromosoomi haplogruppide R1a ja R1b fülogeneesi ja fülogeograafiat, saamaks täpsemat ettekujutust Lääne-Euraasia populatsioonide isaliinide geneetilise varieeruvuse mustri kujunemisest võrreldes seni teadaolevaga. Teiseks eesmärgiks uurida kõiki Kaukaasia alampiirkondi ja keelerühmi esinda-vaid populatsioone ning võrrelda neid naaberalade populatsioonidega, et hinnata, kuivõrd kajastub Kaukaasia suur etniline ja lingvistiline varieeruvus

geneetilistes andmetes ning milline on olnud selle piirkonna roll inimese rännetes Euraasias. Töö kolmandaks eesmärgiks oli võrrelda viie- ja kuuenukleotiidsete kordustega Y-kromosoomi STR markereid sagedamini kasutatavate kolme- ja neljanukleotiidsete markeritega, et hinnata nende mutatsioonikiirust ja sobivust populatsioonigeneetilisteks uuringuteks.

Käesoleva töö tulemuste põhjal formuleeriti alljärgnevad põhilised järeldused:

- Mutatsioon M458 Y-kromosoomi haplogrupis R1a esindab mesoliitikumi ja neoliitikumi ülemineku aegset asutajaefekti Kesk- ja Ida-Euroopas. Haplogrupp R1a on levinud Lõuna-Aasiast Siberi ja Euroopani, M458 on üks esimesi markereid, mis on informatiivne selle geograafilise alajaotuse suhtes, eristades enamiku haplogruppi kuuluvatest eurooplastest.
- Lääne-Euroopas märgib olulist holotseeni perioodi asutajaefekti mutatsioon M412 Y-kromosoomi haplogrupis R1b. Ühe R1b-M412 alamklaadi ajaline ja sagedusmuster korreleerub hästi neoliitilise paelkeraamika kultuuri levikuga.
- Kaukaasia populatsioonid on autosomaalselt palju ühtsemad, kui nende mitmekesise etnilise ja lingvistilise tausta põhjal eeldada võiks. Samas esineb Kaukaasia isaliinide varieeruvuses järske erinevusi populatsioonide ja alampiirkondade vahel, mille põhjuseks võib pidada asutajaefekte ja geneetilist triivi (isaliinis) isoleeritud populatsioonides. Geneetiliselt on kaukaaslased kõige lähedasemad Lähis-Ida populatsioonidele, mis viitab Kaukaasia asustamisele Lähis-Idast.
- Põhja-kaukaaslaste ja geograafiliselt nende lähedal elavate Ida-Euroopa populatsioonide vahel on nii Y-kromosoomi kui ka autosomaalsete andmete põhjal oluline geneetiline erinevus. Sellest võib järeldada, et Kaukaasia ei ole olnud koridoriks inimeste liikumisel Lähis-Idast Ida-Euroopasse ega Euroopasse üldiselt. Seevastu on märgatav sujuv geneetiline üleminek Lähis-Idast Ida-Euroopasse piki Musta mere läänekallast, mis viitab peamisele geenivoolu trajektoorile.
- Pikemate kordusühikutega Y-kromosoomi STR markerid evolutsioneeruvad aeglasemalt, mis muudab need mõningatel juhtudel populatsioonigeneetilisteks uuringuteks sobivamaks kui lühemate kordusühikutega markerid.

ACKNOWLEDGEMENTS

I am thankful to my supervisors, Prof. Richard Villems and Dr. Siiri Rootsi, for being both challenging and supportive in the right measure throughout my doctoral studies, for giving valuable constructive criticism to my work, and for widening my horizons of thought.

I would like to thank Jüri Parik, my first supervisor, who laid the basis of my methodological skills, and Prof. Toomas Kivisild, who helped and encouraged me when I encountered the first of the difficulties that inevitably rise in science, and was kind enough to welcome me to the University of Cambridge for a term. Thanks are also owed to Tiiu Rootslane, without whose timely help in bureaucratic manoeuvring I might not have made it into the PhD programme to begin with, at least not the year I had planned.

Science is a process that requires a fostering environment. Such an environment has been created at the Department of Evolutionary Biology, with a synergy of open-minded professional people from Estonia and abroad, and a general attitude that encourages independent thinking. The constant flow of researchers from a dozen collaborating laboratories around the world creates an international atmosphere that provides ample opportunities for the exchange of experience and ideas. This working environment has helped me complete this thesis, and I would like to thank the people who have contributed to it: Ene, Maere, Monika, Erika, Erwan, Gyaneshwer, Chandana, Alena, Hovhannes, Anne-Mai, Liisa, Erik, Anu and Helle-Viivi. Special thanks go to Eva-Liis, who is much better at statistics than I will ever be, and has willingly devoted her time to advising me, to Mait, who has saved my skin several times amidst the dangers of autosomal analyses, and to Sergey, who provided me with Russian texts on the subject of the Caucasus. I am also very grateful to Tuuli, Georgi, Viljo, Mari and Ille for expert technical assistance.

I am indebted to all my international collaborators and co-authors, especially Bayazit Yunusbayev, Lev A. Zhivotovsky, Juan J. Sanchez, Ildus Kutuev, Doron M. Behar, and Peter A. Underhill, for their constructive suggestions in preparing the manuscripts of the articles presented in this thesis. The valuable comments on this thesis made by Prof. Maido Remm and Dr. Kristiina Tambets are gratefully acknowledged.

anonymous donors of DNA samples without whose contribution none of this work would have been possible, and all the people who assisted in collecting the samples.

I thank my *alma mater*, the University of Tartu, for giving me an excellent education, and the Estonian Women Students' Society for the many things it has taught me, from speaking and organisational skills to a general academic and yet free way of thinking. The good friends I have made at the Society and elsewhere at the university during my years in Tartu inspire me, each in their own way, and have my lifelong admiration and respect.

My parents saw potential in me and encouraged it, and my mother especially stimulated my curiosity for genetics, being a direct link to a family tradition of science in this field. My mother has been a role model as a scientist since my childhood, and she has unfailingly given me good advice on every matter in science I have questioned her about. I owe much to the support of my parents and my brother, and I thank them for it.

PUBLICATIONS

CURRICULUM VITAE

Personal Details

Name	Mari Järve
Date of Birth	14th September 1983
Address for	Department of Evolutionary Biology
Correspondence	Institute of Molecular and Cell Biology
	University of Tartu and the Estonian Biocentre
	23 Riia St, Tartu
	Estonia, 51010
Phone (Office)	+372 7 375 054
E-mail	mari.jarve@ut.ee

Academic Background

Since 2008	PhD student (Molecular Biology), Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu, Estonia.
Since 2011	MSc student (Secondary School Science Teacher), University of Tartu, Estonia.
2011	Wellcome Trust Advanced Course "Computational Molecular Evolution", Hinxton, UK.
2010	One term of studies at the Leverhulme Centre for Human Evolutionary Studies, University of Cambridge, UK.
2008	MSc thesis "Refining the phylogeny and phylogeography of the human Y chromosome haplogroup R1b1b" defended at the Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu, Estonia. Supervisors Dr. Siiri Rootsi and Dr. Ene Metspalu.
2005–2008	MSc studies (Gene Technology), Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu, Estonia (*cum laude*).
2005	BSc thesis "Analysis of the *ABC* transporter gene ABCA4 mutation G1961E in several North-East African and Near Eastern populations" defended at the Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu, Estonia. Supervisor Jüri Parik, MSc.
2002–2005	BSc studies (Gene Technology), Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu, Estonia (*cum laude*).

Distinctions and Awards

2007   Artur Lind scholarship
2007   Rotalia Foundation scholarship
2004   Hansapank scholarship

Publications

Rootsi S, Myres NM, Lin AA, **Järve M**, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Varendi K, Sahakyan H, Behar DM, Khusainova R, Balanovsky O, Balanovska E, Rudan P, Yepiskoposyan L, Bahmanimehr A, Farjadian S, Kushniarevich A, Herrera RJ, Grugni V, Battaglia V, Nici C, Crobu F, Karachanak S, Kashani BH, Houshmand M, Sanati MH, Toncheva D, Lisa A, Semino O, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. *Eur J Hum Genet*, accepted for publication.

Yunusbayev B*, Metspalu M*, **Järve M***, Kutuev I, Rootsi S, Metspalu E, Behar DM, Varendi K, Sahakyan H, Khusainova R, Yepiskoposyan L, Khusnutdinova EK, Underhill PA, Kivisild T, Villems R. (2012). The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol* 29, 359–365.
*equal contribution

Scholes C, Siddle K, Ducourneau A, Crivellaro F, **Järve M**, Rootsi S, Bellatti M, Tabbada K, Mormina M, Reidla M, Villems R, Kivisild T, Mirazón Lahr M, Migliano AB. (2011). Genetic Diversity and Evidence for Population Admixture in Batak Negritos from Palawan. *Am J Phys Anthropol* 146, 62–72.

Myres NM, Rootsi S, Lin AA, **Järve M**, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Pshenichnov A, Yunusbayev B, Balanovsky O, Balanovska E, Rudan P, Baldovic M, Herrera RJ, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. (2011). A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur J Hum Genet* 19, 95–101.

Underhill PA, Myres NM, Rootsi S, Metspalu M, Zhivotovsky LA, King RJ, Lin AA, Chow CET, Semino O, Battaglia V, Kutuev I, **Järve M**, Chaubey G, Ayub Q, Mohyuddin A, Mehdi SQ, Sengupta S, Rogaev EI, Khusnutdinova EK, Pshenichnov A, Balanovsky O, Balanovska E, Jeran N, Havas Augustin D, Baldovic M, Herrera RJ, Thangaraj K, Singh V, Singh L, Majumder P, Rudan P, Primorac D, Villems R, Kivisild T. (2010). Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *Eur J Hum Genet* 18, 479–484.

Järve M, Zhivotovsky LA, Rootsi S, Help H, Rogaev EI, Khusnutdinova EK, Kivisild T, Sanchez JJ. (2009). Decreased Rate of Evolution in Y Chromosome STR Loci of Increased Size of the Repeat Unit. *PLoS ONE* 4, e7276.

ELULOOKIRJELDUS

Isikuandmed

Nimi	Mari Järve
Sünniaeg	14. september 1983
Aadress	Evolutsioonilise bioloogia õppetool
kirjavahetuseks	Molekulaar- ja Rakubioloogia Instituut
	Tartu Ülikool ja Eesti Biokeskus
	Riia 23, 51010 Tartu
	Eesti
Telefon tööl	+372 7 375 054
E-post	mari.jarve@ut.ee

Akadeemiline elukäik

Alates 2008	Doktoriõpe (molekulaarbioloogia) Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituudi evolutsioonilise bioloogia õppetoolis.
Alates 2011	Magistriõpe (gümnaasiumi loodusteaduste õpetaja, ekstern) Tartu Ülikoolis.
2011	Wellcome Trusti kursus *Computational Molecular Evolution* Hinxtonis Suurbritannias.
2010	Semester õpinguid Cambridge'i Ülikoolis keskuses *Leverhulme Centre for Human Evolutionary Studies*.
2008	Magistritöö „Inimese Y kromosoomi haplogrupi R1b1b fülogeneesi ja fülogeograafia täpsustamine", kaitstud Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituudi evolutsioonilise bioloogia õppetooli juures. Juhendajad dr Siiri Rootsi ja dr Ene Metspalu.
2005–2008	Magistriõpe (geenitehnoloogia) Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituudi evolutsioonilise bioloogia õppetoolis (*cum laude*).
2005	Bakalaureusetöö „*ABC* transporteri geeni ABCA4 mutatsiooni G1961E analüüs mõningates Kirde-Aafrika ja Lähis-Ida populatsioonides", kaitstud Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituudi evolutsioonilise bioloogia õppetooli juures. Juhendaja Jüri Parik, MSc.
2002–2005	Bakalaureuseõpe (geenitehnoloogia) Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituudi evolutsioonilise bioloogia õppetoolis (*cum laude*).

Tunnustused ja auhinnad

2007   Artur Linnu stipendium
2007   Rotalia Fondi stipendium
2004   Hansapanga stipendium

Avaldatud teadusartiklid

Rootsi S, Myres NM, Lin AA, **Järve M**, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Varendi K, Sahakyan H, Behar DM, Khusainova R, Balanovsky O, Balanovska E, Rudan P, Yepiskoposyan L, Bahmanimehr A, Farjadian S, Kushniarevich A, Herrera RJ, Grugni V, Battaglia V, Nici C, Crobu F, Karachanak S, Kashani BH, Houshmand M, Sanati MH, Toncheva D, Lisa A, Semino O, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. *Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus*. Eur J Hum Genet, avaldamiseks vastu võetud.

Yunusbayev B*, Metspalu M*, **Järve M***, Kutuev I, Rootsi S, Metspalu E, Behar DM, Varendi K, Sahakyan H, Khusainova R, Yepiskoposyan L, Khusnutdinova EK, Underhill PA, Kivisild T, Villems R. (2012). *The Caucasus as an asymmetric semipermeable barrier to ancient human migrations*. Mol Biol Evol 29, 359–365.

*võrdne panus

Scholes C, Siddle K, Ducourneau A, Crivellaro F, **Järve M**, Rootsi S, Bellatti M, Tabbada K, Mormina M, Reidla M, Villems R, Kivisild T, Mirazón Lahr M, Migliano AB. (2011). *Genetic Diversity and Evidence for Population Admixture in Batak Negritos from Palawan*. Am J Phys Anthropol 146, 62–72.

Myres NM, Rootsi S, Lin AA, **Järve M**, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Pshenichnov A, Yunusbayev B, Balanovsky O, Balanovska E, Rudan P, Baldovic M, Herrera RJ, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. (2011). *A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe*. Eur J Hum Genet 19, 95–101.

Underhill PA, Myres NM, Rootsi S, Metspalu M, Zhivotovsky LA, King RJ, Lin AA, Chow CET, Semino O, Battaglia V, Kutuev I, **Järve M**, Chaubey G, Ayub Q, Mohyuddin A, Mehdi SQ, Sengupta S, Rogaev EI, Khusnutdinova EK, Pshenichnov A, Balanovsky O, Balanovska E, Jeran N, Havas Augustin D, Baldovic M, Herrera RJ, Thangaraj K, Singh V, Singh L, Majumder P, Rudan P, Primorac D, Villems R, Kivisild T. (2010). *Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a*. Eur J Hum Genet 18, 479–484.

Järve M, Zhivotovsky LA, Rootsi S, Help H, Rogaev EI, Khusnutdinova EK, Kivisild T, Sanchez JJ. (2009). *Decreased Rate of Evolution in Y Chromosome STR Loci of Increased Size of the Repeat Unit*. PLoS ONE 4, e7276.

DISSERTATIONES BIOLOGICAE
UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand.** The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic micro-organisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous crassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.

41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.

61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.
63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligo-nucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p

82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.

83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.

84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.

85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.

86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.

87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.

88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.

89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004. 117 p.

90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004. 175 p.

91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004. 109 p.

92. **Kristiina Tambets.** Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylo-geographic approach. Tartu, 2004. 163 p.

93. **Arvi Jõers.** Regulation of p53-dependent transcription. Tartu, 2004. 103 p.

94. **Lilian Kadaja.** Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004. 103 p.

95. **Jaak Truu.** Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004. 128 p.

96. **Maire Peters.** Natural horizontal transfer of the *pheBA* operon. Tartu, 2004. 105 p.

97. **Ülo Maiväli.** Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004. 130 p.

98. **Merit Otsus.** Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004. 103 p.

99. **Mikk Heidemaa.** Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004. 167 p.

100. **Ilmar Tõnno.** The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004. 111 p.

101. **Lauri Saks.** Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004. 144 p.

102. **Siiri Rootsi.** Human Y-chromosomal variation in European populations. Tartu, 2004. 142 p.

103. **Eve Vedler.** Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.

104. **Andres Tover.** Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 126 p.

105. **Helen Udras.** Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005. 100 p.

106. **Ave Suija.** Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005. 162 p.

107. **Piret Lõhmus.** Forest lichens and their substrata in Estonia. Tartu, 2005. 162 p.

108. **Inga Lips.** Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005. 156 p.

109. **Kaasik, Krista.** Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005. 121 p.

110. **Juhan Javoiš.** The effects of experience on host acceptance in ovipositing moths. Tartu, 2005. 112 p.

111. **Tiina Sedman.** Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005. 103 p.

112. **Ruth Aguraiuja.** Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005. 112 p.

113. **Riho Teras.** Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 106 p.

114. **Mait Metspalu.** Through the course of prehistory in india: tracing the mtDNA trail. Tartu, 2005. 138 p.

115. **Elin Lõhmussaar.** The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006. 124 p.

116. **Priit Kupper.** Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006. 126 p.

117. **Heili Ilves.** Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006. 120 p.

118. **Silja Kuusk.** Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006. 126 p.

119. **Kersti Püssa.** Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006. 90 p.

120. **Lea Tummeleht.** Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006. 94 p.

121. **Toomas Esperk.** Larval instar as a key element of insect growth schedules. Tartu, 2006. 186 p.

122. **Harri Valdmann.** Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.

123. **Priit Jõers.** Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.

124. **Kersti Lilleväli.** Gata3 and Gata2 in inner ear development. Tartu, 2007. 123 p.

125. **Kai Rünk.** Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007. 143 p.

126. **Aveliina Helm.** Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007. 89 p.

127. **Leho Tedersoo.** Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007. 233 p.

128. **Marko Mägi.** The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007. 135 p.

129. **Valeria Lulla.** Replication strategies and applications of Semliki Forest virus. Tartu, 2007. 109 p.

130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007. 79 p.

131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007. 171 p.

132. **Tatjana Krama.** Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007. 112 p.

133. **Signe Saumaa.** The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida.* Tartu, 2007. 172 p.

134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007. 96 p.

135. **Priit Kilgas.** Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007. 129 p.

136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007. 95 p.

137. **Kärt Padari.** Protein transduction mechanisms of transportans. Tartu, 2008. 128 p.

138. **Siiri-Lii Sandre.** Selective forces on larval colouration in a moth. Tartu, 2008. 125 p.

139. **Ülle Jõgar.** Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008. 99 p.

140. **Lauri Laanisto.** Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008. 133 p.

141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008. 105 p.

142. **Birgot Paavel.** Bio-optical properties of turbid lakes. Tartu, 2008. 175 p.

143. **Kaire Torn.** Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.

144. **Vladimir Vimberg.** Peptide mediated macrolide resistance. Tartu, 2008, 190 p.

145. **Daima Örd.** Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.

146. **Lauri Saag.** Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.

147. **Ulvi Karu.** Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.

148. **Jaanus Remm.** Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.

149. **Epp Moks.** Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.

150. **Eve Eensalu.** Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.

151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.

152. **Marta Putrinš.** Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.

153. **Marina Semtšenko.** Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.

154. **Marge Starast.** Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.

155. **Age Tats.** Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.

156. **Radi Tegova.** The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.

157. **Tsipe Aavik.** Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2008, 112 p.

158. **Kaja Kiiver.** Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.

159. **Meelis Kadaja.** Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.

160. **Pille Hallast.** Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.

161. **Ain Vellak.** Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.

162. **Triinu Remmel.** Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.

163. **Jaana Salujõe.** Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.

164. **Ele Vahtmäe.** Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.

165. **Liisa Metsamaa.** Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.

166. **Pille Säälik.** The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.

167. **Lauri Peil.** Ribosome assembly factors in *Escherichia coli.* Tartu, 2009, 147 p.

168. **Lea Hallik.** Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.

169. **Mariliis Tark.** Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.

170. **Riinu Rannap.** Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.

171. **Maarja Adojaan.** Molecular variation of HIV-1 and the use of this knowledge in vaccine development. Tartu, 2009, 95 p.

172. **Signe Altmäe.** Genomics and transcriptomics of human induced ovarian folliculogenesis. Tartu, 2010, 179 p.

173. **Triin Suvi.** Mycorrhizal fungi of native and introduced trees in the Seychelles Islands. Tartu, 2010, 107 p.

174. **Velda Lauringson.** Role of suspension feeding in a brackish-water coastal sea. Tartu, 2010, 123 p.

175. **Eero Talts.** Photosynthetic cyclic electron transport – measurement and variably proton-coupled mechanism. Tartu, 2010, 121 p.

176. **Mari Nelis.** Genetic structure of the Estonian population and genetic distance from other populations of European descent. Tartu, 2010, 97 p.

177. **Kaarel Krjutškov.** Arrayed Primer Extension-2 as a multiplex PCR-based method for nucleic acid variation analysis: method and applications. Tartu, 2010, 129 p.

178. **Egle Köster.** Morphological and genetical variation within species complexes: *Anthyllis vulneraria* s. l. and *Alchemilla vulgaris* (coll.). Tartu, 2010, 101 p.

179. **Erki Õunap.** Systematic studies on the subfamily Sterrhinae (Lepidoptera: Geometridae). Tartu, 2010, 111 p.

180. **Merike Jõesaar.** Diversity of key catabolic genes at degradation of phenol and *p*-cresol in pseudomonads. Tartu, 2010, 125 p.

181. **Kristjan Herkül.** Effects of physical disturbance and habitat-modifying species on sediment properties and benthic communities in the northern Baltic Sea. Tartu, 2010, 123 p.

182. **Arto Pulk.** Studies on bacterial ribosomes by chemical modification approaches. Tartu, 2010, 161 p.
183. **Maria Põllupüü.** Ecological relations of cladocerans in a brackish-water ecosystem. Tartu, 2010, 126 p.
184. **Toomas Silla.** Study of the segregation mechanism of the Bovine Papillomavirus Type 1. Tartu, 2010, 188 p.
185. **Gyaneshwer Chaubey.** The demographic history of India: A perspective based on genetic evidence. Tartu, 2010, 184 p.
186. **Katrin Kepp.** Genes involved in cardiovascular traits: detection of genetic variation in Estonian and Czech populations. Tartu, 2010, 164 p.
187. **Virve Sõber.** The role of biotic interactions in plant reproductive performance. Tartu, 2010, 92 p.
188. **Kersti Kangro.** The response of phytoplankton community to the changes in nutrient loading. Tartu, 2010, 144 p.
189. **Joachim M. Gerhold.** Replication and Recombination of mitochondrial DNA in Yeast. Tartu, 2010, 120 p.
190. **Helen Tammert.** Ecological role of physiological and phylogenetic diversity in aquatic bacterial communities. Tartu, 2010, 140 p.
191. **Elle Rajandu.** Factors determining plant and lichen species diversity and composition in Estonian *Calamagrostis* and *Hepatica* site type forests. Tartu, 2010, 123 p.
192. **Paula Ann Kivistik.** ColR-ColS signalling system and transposition of Tn*4652* in the adaptation of *Pseudomonas putida.* Tartu, 2010, 118 p.
193. **Siim Sõber.** Blood pressure genetics: from candidate genes to genome-wide association studies. Tartu, 2011, 120 p.
194. **Kalle Kipper.** Studies on the role of helix 69 of 23S rRNA in the factor-dependent stages of translation initiation, elongation, and termination. Tartu, 2011, 178 p.
195. **Triinu Siibak.** Effect of antibiotics on ribosome assembly is indirect. Tartu, 2011, 134 p.
196. **Tambet Tõnissoo.** Identification and molecular analysis of the role of guanine nucleotide exchange factor RIC-8 in mouse development and neural function. Tartu, 2011, 110 p.
197. **Helin Räägel.** Multiple faces of cell-penetrating peptides – their intra-cellular trafficking, stability and endosomal escape during protein trans-duction. Tartu, 2011, 161 p.
198. **Andres Jaanus.** Phytoplankton in Estonian coastal waters – variability, trends and response to environmental pressures. Tartu, 2011, 157 p.
199. **Tiit Nikopensius.** Genetic predisposition to nonsyndromic orofacial clefts. Tartu, 2011, 152 p.
200. **Signe Värv.** Studies on the mechanisms of RNA polymerase II-dependent transcription elongation. Tartu, 2011, 108 p.
201. **Kristjan Välk.** Gene expression profiling and genome-wide association studies of non-small cell lung cancer. Tartu, 2011, 98 p.

202. **Arno Põllumäe.** Spatio-temporal patterns of native and invasive zoo-plankton species under changing climate and eutrophication conditions. Tartu, 2011, 153 p.

203. **Egle Tammeleht.** Brown bear (*Ursus arctos*) population structure, demographic processes and variations in diet in northern Eurasia. Tartu, 2011, 143 p.

205. **Teele Jairus.** Species composition and host preference among ectomycorrhizal fungi in Australian and African ecosystems. Tartu, 2011, 106 p.

206. **Kessy Abarenkov.** PlutoF – cloud database and computing services supporting biological research. Tartu, 2011, 125 p.

207. **Marina Grigorova.** Fine-scale genetic variation of follicle-stimulating hormone beta-subunit coding gene (*FSHB*) and its association with reproductive health. Tartu, 2011, 184 p.

208. **Anu Tiitsaar.** The effects of predation risk and habitat history on butterfly communities. Tartu, 2011, 97 p.

209. **Elin Sild.** Oxidative defences in immunoecological context: validation and application of assays for nitric oxide production and oxidative burst in a wild passerine. Tartu, 2011, 105 p.

210. **Irja Saar**. The taxonomy and phylogeny of the genera *Cystoderma* and *Cystodermella* (Agaricales, Fungi). Tartu, 2012, 167 p.

211. **Pauli Saag.** Natural variation in plumage bacterial assemblages in two wild breeding passerines. Tartu, 2012, 113 p.

212. **Aleksei Lulla.** Alphaviral nonstructural protease and its polyprotein substrate: arrangements for the perfect marriage. Tartu, 2012, 143 p.