CrossMark

REVIEW

The study of human Y chromosome variation through ancient DNA

Toomas Kivisild[1,2]

Abstract High throughput sequencing methods have completely transformed the study of human Y chromosome variation by offering a genome-scale view on genetic variation retrieved from ancient human remains in context of a growing number of high coverage whole Y chromosome sequence data from living populations from across the world. The ancient Y chromosome sequences are providing us the first exciting glimpses into the past variation of male-specific compartment of the genome and the opportunity to evaluate models based on previously made inferences from patterns of genetic variation in living populations. Analyses of the ancient Y chromosome sequences are challenging not only because of issues generally related to ancient DNA work, such as DNA damage-induced mutations and low content of endogenous DNA in most human remains, but also because of specific properties of the Y chromosome, such as its highly repetitive nature and high homology with the X chromosome. Shotgun sequencing of uniquely mapping regions of the Y chromosomes to sufficiently high coverage is still challenging and costly in poorly preserved samples. To increase the coverage of specific target SNPs capture-based methods have been developed and used in recent years to generate Y chromosome sequence data from hundreds of prehistoric skeletal remains. Besides the prospects of testing directly as how much genetic change in a given time period has accompanied changes in material culture the sequencing of ancient Y chromosomes allows us also to better understand the rate at which mutations accumulate and get fixed over time. This review considers genome-scale evidence on ancient Y chromosome diversity that has recently started to accumulate in geographic areas favourable to DNA preservation. More specifically the review focuses on examples of regional continuity and change of the Y chromosome haplogroups in North Eurasia and in the New World.

Background

Until recently, ancient DNA studies of human remain focused primarily on variation embedded in mitochondrial DNA (mtDNA). For decades, mtDNA had been a target of choice in population genetic studies because of its high mutation rate and high density of polymorphic markers (Wilson et al. 1985). Also, because for any unique sequence in the autosomal, X or Y chromosome locus there are many hundreds or even thousands of copies of mtDNA means that this maternally inherited locus was more likely to work in cases where only a very small number of molecules had survived (Krings et al. 1997). HTS technologies have significantly increased the rate at which sequence data can be generated and make accessible surviving chunks of DNA that are shorter than the size of a PCR (polymerase chain reaction) amplicon (Orlando et al. 2015). They have reduced the costs of sequencing and opened the prospects of assessing the variation of human populations, both modern and ancient, at the scale of entire genomes (Genomes Project et al. 2012; Gilbert et al. 2008; Green et al. 2010; Rasmussen et al. 2010). These technological advances (Orlando et al. 2015) have also made it possible now to study ancient Y chromosome (aY) variation in human populations at the scale of the entire accessible length of the male-specific and non-recombining regions of human Y chromosome.

✉ Toomas Kivisild
  tk331@cam.ac.uk

1 Department of Archaeology and Anthropology, University of Cambridge, Cambridge CB2 1QH, UK

2 Estonian Biocentre, 51010 Tartu, Estonia

Ancient DNA presents us the opportunity to directly examine which Y chromosome single nucleotide polymorphisms (SNPs) and haplotypes were present at different time periods in regions that support long-term survival of ancient DNA. It is perhaps not surprising that archaeological sites from high latitude areas (Hofreiter et al. 2015) have yielded the largest number of successfully sequenced samples in recent genome-scale studies that have reported on aY. These studies, as reviewed in further detail below, have provided us the first glimpses of the dynamics of aY haplogroup composition and frequency changes in transects of time and allow us to test hypotheses based on earlier phylogeographic inferences made from the Y chromosome data of presently living populations. In this review, 'haplogroups' and 'clades' are terms that are interchangeably used to refer to groups of closely-related Y chromosome sequences that share a common ancestor. While there is a general agreement in the definition of the basic haplogroups (A, B, C, etc.) multiple parallel nomenclatures are in use for sub-clades (e.g. see http://www.phylotree.org/Y/, http://isogg.org/tree/, https://www.yfull.com/tree/). For clarity, the names of sub-clades used here are suffixed by the defining SNP-marker name, as suggested previously (Karmin et al. 2015; van Oven et al. 2014). Similarly to the extent to which our earlier views on peopling of continental regions such as Europe based on inferences made from extant mtDNA variation have changed in the light of new ancient mtDNA evidence (Bollongino et al. 2013; Bramanti et al. 2009; Brandt et al. 2013; Haak et al. 2015; Posth et al. 2016; Thomas et al. 2013), it can be seen that models linking the spread of specific Y chromosome haplogroups with the spread of material culture may require substantial revision in the light of new aY evidence.

In this review, methodological aspects of ancient Y chromosome work will be first discussed with a focus on recent HTS research based on shotgun and capture approaches. Challenges common to all ancient DNA studies include those related to calling human polymorphic variants from short and damaged sequence reads that are derived from a mix of different organisms. Highly repetitive nature of the Y chromosome combined with its high sequence homology with the X chromosome (Skaletsky et al. 2003) further complicates variant calling from short ancient DNA sequence reads. These issues, together with the paternal inheritance of Y chromosome, its haploid nature and high linkage between physically distant SNPs, impact on choices of the bioinformatics methods of downstream data analysis. They define the restrictions and specifics how the aY sequence data should be processed and what are the limitations for the interpretations made from such data. Finally, the demographic histories European and North American populations will be briefly reviewed in the light

of recently emerged Y chromosome evidence from ancient DNA studies.

Methodological issues in dealing with aY

Two different approaches, shotgun and hybridization capture-based sequencing, have been used in recent ancient DNA studies to assess the variation of ancient Y chromosomes. In all genome-scale studies, shotgun sequencing is typically used firstly in the screening phase. In case of high percentage (~>10%) of reads mapping uniquely to the human reference genome combined with low clonality of the reads, further shotgun sequencing of the sample to a higher coverage can be generally considered to be an efficient and relatively cost-effective strategy. However, in cases where the content of human genome mapping reads is low (~1 to 10%), enrichment methods that target specifically human-specific DNA have been considered to be a practical solution for getting improved coverage of SNPs that are considered informative in downstream analyses (Haak et al. 2015; Lazaridis et al. 2016; Mathieson et al. 2015). Having sufficient coverage of the Y chromosome SNP data in ancient samples is important for mapping them correctly to the phylogenetic tree and for examining the relationship among multiple ancient samples in parallel. First, low coverage of data can have an undesirable effect on the inferences of the positioning of ancient samples in a tree in relation to high coverage modern Y chromosome sequence data, and, in particular, for the estimation of branch lengths of the phylogeny as exemplified in Fig. 1. Besides the coverage (a measure of how many sites of the reference genome are covered by data in a given sample), it is also the sequencing depth (the number of reads covering a site) at each given site that contributes to the branch length estimates as sites covered by only one or two reads tend to have a high rate of false positive sequencing errors. Second, when pooling from many ancient samples, data analysis methods that rely on allele frequency comparisons of individual SNPs require reasonable number of overlapping sites covered by data in multiple samples. Predictably, the more low coverage samples are included in the analyses the less overlapping SNPs remain available for downstream data analyses.

A number of Y chromosome sequences covered in this review have been generated with shotgun sequencing approach. These include the four oldest genomes of individuals dated to late Pleistocene as well as a number of remains from Europe and Americas dated to the Holocene period. A substantial proportion of the European and Middle Eastern Neolithic and Bronze Age remains have been sequenced, however, with hybridization-based capture technique. Capture-based methods focus on specific

Fig. 1 Effect of low coverage of the data and post-mortem damage on the inferences of branch lengths and on the phylogenetic mapping of mutations to the tree. A general example of phylogenetic relationships between one high coverage modern (M1) and two low coverage ancient (A1 and A2) samples is shown. The number of mutations mapping to each branch is shown on both trees. The change in numbers of mutations mismapped due to low coverage on each branch of the tree is explained with the *red arrows* that indicate the directionality of the mismapping. A1 is shown to carry more damage-induced mutations than A2. Modified from Poznik et al. (2016), Supplementary Figure 18
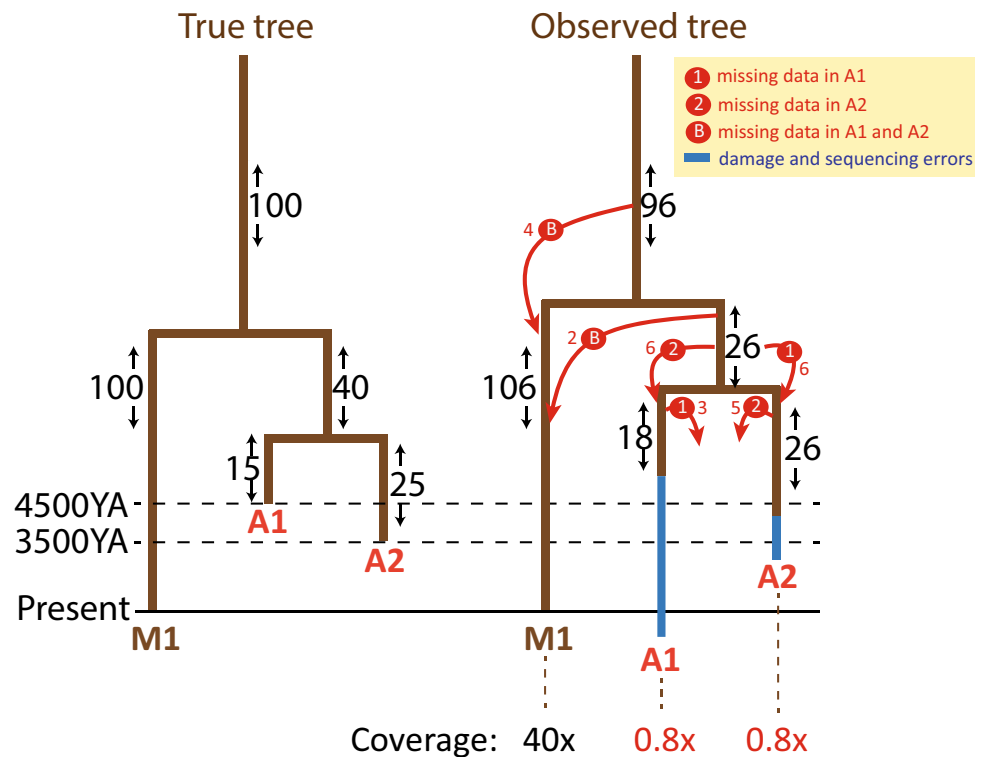
targets in the genome with the aim to raise their relative coverage in the resulting data. The human origins (HO) genome-wide SNP array (Patterson et al. 2012) has been used to genotype more than 600 K SNPs in more than two thousand living individuals from 203 populations (Lazaridis et al. 2014). A subset of 390 K SNPs from this array, including 2258 Y chromosome markers, were targeted later (Haak et al. 2015) in a hybridization-based capture design to assess variation in 69 Europeans living 8–3 thousand years ago. The extended panel of the HO hybridization design now targets 1240 K SNPs genome-wide including 32,681 from the Y chromosome (Mathieson et al. 2015). This approach has been used to assess at a high molecular resolution the phylogenetic affiliation of a large number of 110 ancient Y chromosomes from Europe and Near East (Lazaridis et al. 2016).

Several aspects of the capture-based data need to be considered when interpreting the phylogenetic reconstructions based on genotype rather than sequence data. First, the advantage of the capture approach, as mentioned above, is that it allows for the generation of data with higher coverage and higher overlaps among large number of individuals for a given set of SNPs (Pickrell and Reich 2014). This means that even in case of poorly preserved samples, robust haplogroup inferences, supported by multiple phylogenetically equivalent SNPs, can be made. The generic limitation of the approaches that focus on the enrichment-targeted SNPs is the ascertainment bias towards previously

discovered variants and that the capture approach does not enable the user to discover novel variants and clades that have either become extinct or that have been not presented in the ascertainment set used for the design of the panel of SNPs to be captured (Fig. 2). Although the 32,681 Y-SNPs of the HO-1240K design include a large number of Simons Genome Diversity Project (Mallick et al. 2016) and ISOGG (International Society of Genetic Genealogy, http://isogg.org/tree/) catalogued variants these would entail, by design, predominantly annotations of the clades that are common in present day populations.

Because most personal genomes that have been annotated by ISOGG come from individuals of European or North American descent the capture enriched for ISOGG SNPs is best suited for the study of European Y chromosome diversity while being less efficient for the study of other regions. But also in Europe it should be noted that clades that have become infrequent or extinct over time due to extensive admixture or population replacement would have less chance to be recognised with the SNP-targeting capture approach. If the parental clade of an extinct branch has extant sister-clades that have been characterised in the annotated databases the aY sequence from that extinct clade will be characterised at the level of the parental clade (Fig. 2). While the most common Y chromosome haplogroups in Europe are represented in the HO-1240K SNP array by many phylogenetically equivalent SNPs they can be robustly recognised and called even in samples with
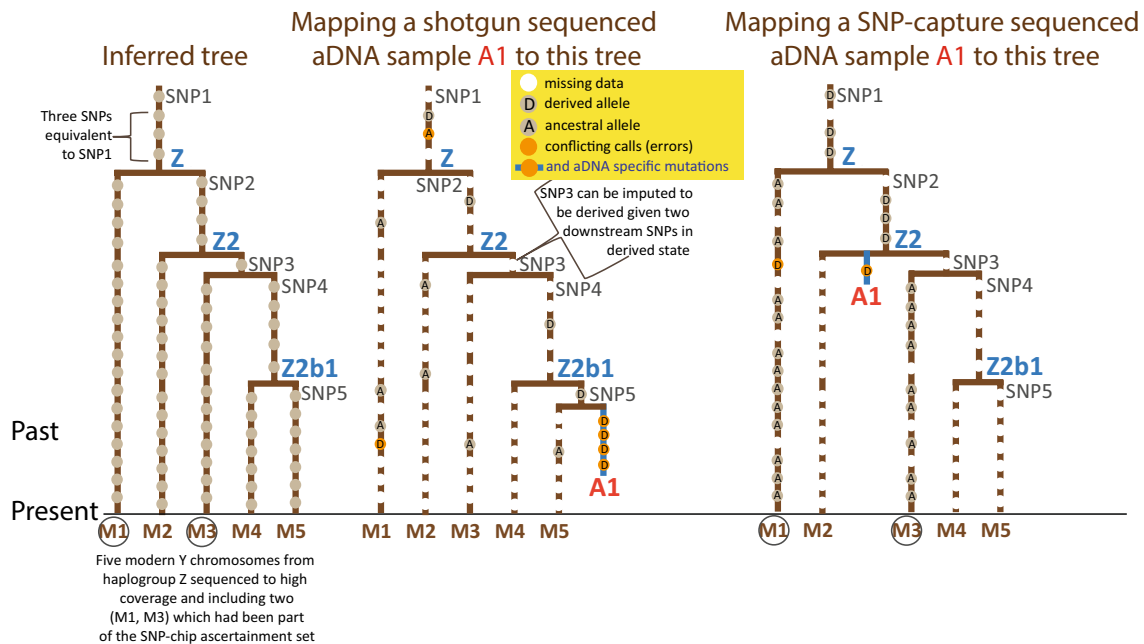
Fig. 2 Phylogenetic mapping of a low-coverage ancient DNA sequence to a tree drawn from a general example of a data set of high coverage sequences of modern samples. *Left panel* shows the tree inferred from modern samples. *Middle section* shows the mapping of a sample sequenced by shotgun approach. The *right panel* shows the mapping of the same sample to the tree following SNP-targeted capture approach where modern samples 1 and 3 had been used in the capture design

large number of missing data, the haplogroups of other continental regions are less completely characterised in SNP arrays and therefore shotgun approach may be preferable for the ancient DNA (aDNA) study of Y chromosome variation in those areas. As the accessible regions of the Y chromosome that are commonly used in genetic studies are haploid and do not undergo recombination imputation is more efficient in Y chromosome branches represented by multiple equivalent SNPs than it is in autosomal loci and therefore shotgun sequence data with fairly low coverage can be used for unbiased assessment of phylogenetic affinities of aY lineages. However, when using imputation, particularly in the more terminal branches of the tree, one should be cautioned that some of the ancient samples may in fact derive from clades that share only part of the SNPs defining the extant branch of the Y tree and not all of the SNPs.

Another complicating factor for determining Y chromosome haplogroups from ancient DNA data is that it can be challenging to distinguish true mutations from those induced by damage, particularly in case of C to T and G to A substitutions (Gilbert et al. 2003; Hofreiter et al. 2001). A commonly used strategy for dealing with post-mortem damage in ancient DNA studies is the removal of transitions which are the main targets of deamination-induced miscalls via uracil DNA glycosylase treatment or data filtering (Orlando et al. 2015). As transitions occur naturally

more frequently than transversions their total removal can lead, however, to dramatic losses of data and thereby loss of phylogenetic resolution, particularly in cases where the sub-clade-defining branches are short. Therefore, aY sequences with low coverage ($<0.1\times$) may not contain enough informative transversions to allow for robust haplogroup assignment at a resolution that would be useful for testing cases of genetic continuity or admixture. Capture approach may yield better coverage at targeted sites but it should be also cautioned that some SNP-targeting capture designs can be enriched for transitions due to the removal of strand ambiguous G<->C and A<->T transversions from the design of SNP-chips.

High coverage Y chromosome sequences, whether from modern or ancient samples, can be used to draw trees with informative branch lengths. With low coverage data generated from ancient human remains, it is typically not possible to make a clear distinction between damage-affected sites and true mutations (Fig. 1). This ambiguity makes it either highly problematic or impossible for the user to determine the lengths of the private branches of the ancient samples. Capture-based methods that are designed to target sequences surrounding a restricted number of known SNPs have a number of advantages, as reviewed earlier (Pickrell and Reich 2014), while at the same time being limited to detect only variants that are included in the capture design. This means that these methods do not allow for

the discovery of new variants nor to determine the length of the private branches of the Y chromosome tree. The capture designs targeting large regions of DNA that uniquely map to the Y chromosome, such as the *BigY* design of FamilyTree (https://www.familytreedna.com/learn/wp-content/uploads/2014/08/BIG_Y_WhitePager.pdf) which has 67,000 probes that enable the generation of 4.36 Mbp sequence data from "globally covered" regions of the Y chromosome, or the smaller scale capture design of 500 kb non-recombining regions of Lippold et al. (2014), have been successfully applied on modern samples. These capture approaches are able to detect new, previously uncharacterized variants within the captured regions but they are yet to be shown to work efficiently on ancient DNA. Hybridization enrichment approach targeting chromosome 21 has been successfully applied on the ~40 thousand year-old (KYA) human remains from Tianyuan Cave outside Beijing, China (Fu et al. 2013) suggesting that similar approaches can be applied in the future also on accessible (non-ampliconic) regions of the Y chromosome.

Although HTS methods enable us to generate sequence data from short ancient DNA molecules downstream analysis of extremely short Y chromosome mapping reads is limited by the complexity of the Y chromosome which includes large number of repeats and regions that map to other chromosomal regions. This means that female individuals may yield calls for Y chromosome SNPs surrounded by sequence highly homologous to the X chromosome. Inclusion of SNPs that map outside unique single-copy regions of the Y chromosome confounds phylogenetic mapping of ancient DNA samples and may also bias the analyses based on branch lengths and quantitative estimates of sequence divergence between samples. The approach most typically used in HTS studies of Y chromosome diversity is to filter out SNPs from high homology regions, typically converging to X chromosome degenerate regions of a total size <10 Mbp (Francalacci et al. 2015; Karmin et al. 2015; Poznik et al. 2013; Wei et al. 2013). Several attempts to infer Y chromosome mutation rate from these uniquely mapping regions using ancient DNA data have been made, as reviewed in the paper by Oleg Balanovsky in this volume.

Genetic history of Eurasian Y chromosomes

Consistent with generally higher genetic diversity in African populations, the highest number of deeply splitting branches of human Y chromosome tree can be found in African populations (Cruciani et al. 2011; Mendez et al. 2013; Poznik et al. 2013; Wei et al. 2013). The extant variation in other continents is mainly restricted to three branches of the M168 clade: D, C and F (Hallast et al.

2015; Karmin et al. 2015; Poznik et al. 2013, 2016; Scozzari et al. 2014; Wei et al. 2013). Genetic variation within each of these three clades coalesces to a single founding lineage within 40–60 KYA time depth, consistent with the Out of Africa (OOA) dispersal model (Stringer 2002). In addition to these three clades, the populations of Near East and Europe also show the presence of young sub-clades of haplogroup E which are likely to reflect recent episodic gene flow from Africa within the last 25 KYA and a range expansion of a sub-clade E2a1-V13 in Europe within the last few thousand years (Cruciani et al. 2007). The three basal Eurasian branches D, C and F split further into ~40 extant branches that are older than 30KYA (Fig. 3). Knowing the age of the ancient human remains whose genomes have been sequenced, we can map their Y chromosomes on the phylogenetic tree relating these extant branches and ask whether they are consistent with the inferences of ancestral variation made from modern data or whether they represent distantly related clades that have left no living descendants. In Fig. 3, aY sequences with known radiocarbon dates are mapped, on the basis of the derived alleles they share with modern sequences, on a tree summarizing variation in modern whole Y chromosome sequence data (Scozzari et al. 2014; Karmin et al. 2015; Hallast et al. 2015; Batini et al. 2015; Poznik et al. 2016). The nine oldest ancient Y chromosomes that have been sequenced so far at sufficiently high coverage to allow for phylogenetic mapping are from northern parts of Eurasia (Fu et al. 2014, 2015, 2016; Raghavan et al. 2014b; Seguin-Orlando et al. 2014). Notably, all nine of these aY sequences can be assigned to one of the three major founding lineages that have been inferred from the analyses of extant Y variation. This good match between the ancestral lineages that are observed and those expected from present-day variation to have existed in the time frame of 20–50 thousand years ago supports the view of a robust recovery of non-Africans from extremely low values of male effective population size during the OOA bottleneck (Lippold et al. 2014).

Two of the oldest human Y chromosomes sequenced so far, Ust'Ishim Man (Fu et al. 2014) and the Oase Man (Fu et al. 2015), are both placed near the root of haplogroup K, a sub-clade of F, which is globally the most frequent Y chromosome lineage alive today. K is an ancestral group that unites a number of regional haplogroups that are found widely spread today in Europe, East Asia, Oceania, and Americas. Notably, however, the Y chromosomes of these two ancient Eurasian colonists are not exactly equidistant to all living descendants of haplogroup K that are found today in the world. Both Ust'Ishim and Oase men share the derived allele of a marker M2308 (Poznik et al. 2016) which defines the basal root of two common haplogroups N and O (Fig. 3). These M2308-derived haplogroups have a wide spread today in Eurasia, extending from Finno-Ugric
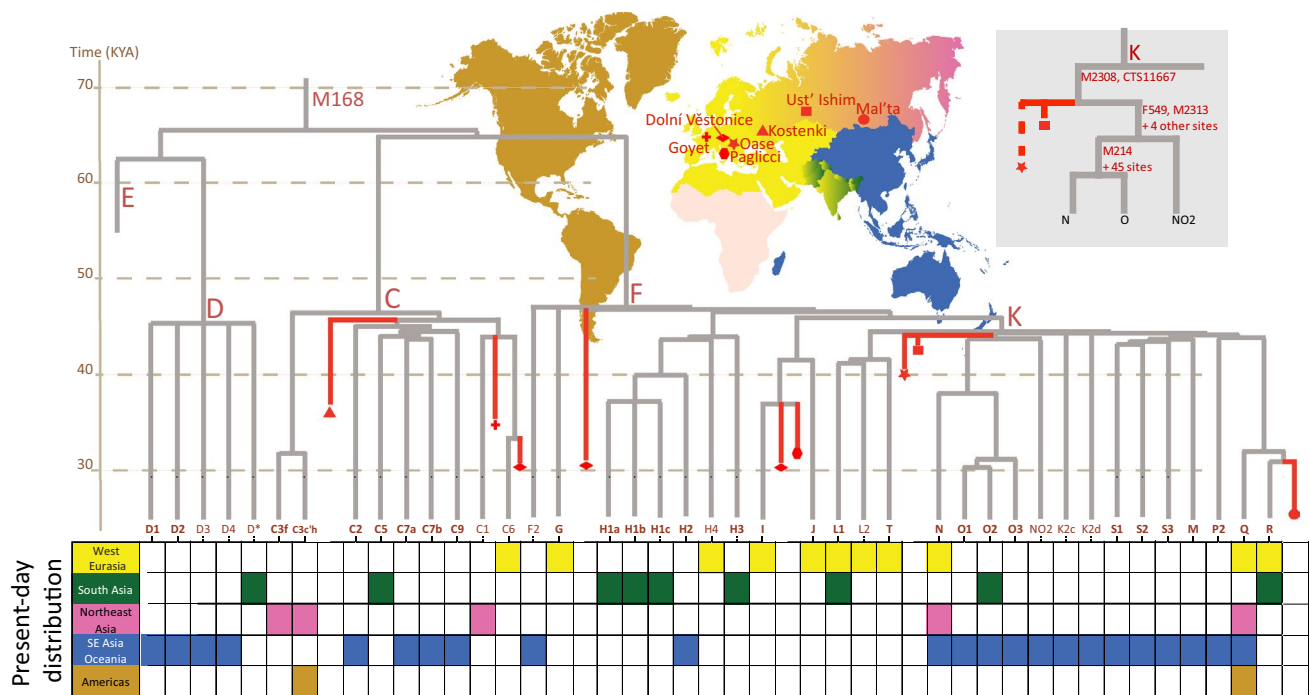
Fig. 3 Human Y chromosome diversity outside Africa 20–50 thousand years ago. The branching structure of 41 extant Y chromosome clades inferred to be older than 30,000 years is shown according to the tree based on high coverage Y chromosome sequences presented in Karmin et al. (2015), Figure S9. The phylogenetic mapping to this tree of nine late Pleistocene ancient Y chromosomes that predate the Last Glacial Maximum—from Ust'Ishim (Fu et al. 2014), Oase (Fu et al. 2015), Kostenki (Seguin-Orlando et al. 2014), Malt'a (Raghavan et al. 2014b), Dolní Věstonice, Goyet, and Paglicci (Fu et al. 2016) sites—is shown in red. The zoomed in version of the subtree relating Oase, Ust'Ishim with extant variation of the N, O and NO2 clades is shown with branch-defining marker names [modified from (Poznik et al. 2016)] next to the geogrpaphic map displaying the locations of the four ancient sites. The *colour-coded* distribution of Y chromosome haplogroups in present day populations ignores cases of recent admixture. Haplogroup names here and in other figures follow the nomenclature suggested in Karmin et al. (2015)

populations of Northeast Europe to Tibeto-Burman, Austroasiatic and Austronesian speaking groups of South and East Asia (Ilumae et al. 2016; Poznik et al. 2016). Although the Ust'Ishim and Oase men were separated from each other about 5 thousand years in time and about 5 thousand kilometres in space these two sampling points we have from the earliest period of peopling of Eurasia suggest some level of continuity in certain Y chromosome lineages that exist in present-day populations of Eurasia. In contrast, the analyses of autosomal genomes have shown that both Ust'Ishim and Oase men were equidistant to East and West Eurasian populations and therefore unlikely to have contributed substantially to later humans in geographic regions where they were living in (Fu et al. 2014, 2015). The phylogenetic affiliation of these early Eurasian men with an NO-related clade is notable because phylogeographic inferences made from present day Y chromosome variation have highlighted other haplogroups, I and R1-M173, as signatures of the genetic legacy of Palaeolithic humans in West Eurasia (Semino et al. 2000) while the origin of haplogroups N and O (Rootsi et al. 2007), as well as other subclades of haplogroup K (Karafet et al. 2015), was deemed

most likely to be, on the grounds of the highest genetic diversity, in Southeast Asia.

The 37 KYA Kostenki-14 Man, found on the western bank of river Don near Voronezh in Southwest Russia, and whose genome has been sequenced to an average depth of 2.8×, descended from a population that was more closely related to modern Europeans than to East Eurasians (Seguin-Orlando et al. 2014). Yet, his Y chromosome belonged to haplogroup C which is extremely rare or absent in most European populations sampled today (Semino et al. 2000). So, although his autosomal genes show the highest affinity to European present-day populations his Y chromosome is different. Haplogroup C is common today in populations of Siberia, Southeast Asia and Oceania (Bergstrom et al. 2016; Karafet et al. 2001, 2002; Kayser 2010). This disparity may suggest, in line with similar findings of mitochondrial DNA haplogroup M lineages in pre-Holocene remains from Europe, that peopling of Europe involved several population replacements and turnovers (Posth et al. 2016). Furthermore, the occurrence of haplogroup C lineages different from the Kostenki type, those belonging to a rare C6-V20 (Scozzari et al. 2012) sub-clade (Fig. 4), in

Iberian hunter-gatherer sample from La Brana (Olalde et al. 2014) and in three Neolithic farmer samples from Anatolia and Central Europe (Mathieson et al. 2015) shows that a diverse set of haplogroup C lineages may have been common and widely spread throughout Eurasia before Middle Holocene. Overall, this newly emerging aDNA evidence from late Pleistocene shows that the Y chromosome pool of West Eurasia has undergone significant changes over time and that inferences made from present-day genetic variation about the time and place of origin of Y chromosome haplogroups can be imperfect.

Besides the haplogroup C lineages that are atypical to present-day populations living in the area the Y chromosome pool of the early Holocene hunter-gatherer and farmer populations of Europe and Middle East was characterised by a diverse set of haplogroups, such as G, H, I, J and R, which are restricted in their present-day distribution by and large to Europe, Middle East, North Africa, South and Central Asia. The geographic patterns of haplogroup G distribution in present-day populations have been suggested to reflect the spread of Anatolian farmers to Europe (Semino et al. 2000). Its sub-clades are most frequently

found today in the Caucasus and West Asia while being rare in Europe (Cinnioglu et al. 2004; Firasat et al. 2007; Haber et al. 2012; Nasidze et al. 2008; Rootsi et al. 2012; Yunusbayev et al. 2012). Ancient DNA evidence from Anatolia and Iran confirms that G, along with H, was the most common Y chromosome haplogroup of the early farmers in these areas (Fig. 5) as well as being characteristically frequent in European Early Neolithic populations who also show low autosomal genetic distances with Anatolian farmers (Broushaki et al. 2016; Hofmanova et al. 2016; Lazaridis et al. 2016; Mathieson et al. 2015). Coherent with the tight clustering of the autosomal genome of the Tyrolean Iceman, nicknamed Ötzi, together with present-day populations of Sardinia and Corsica (Keller et al. 2012), and the claims that Sardinians represent the genetic continuity from the early farmers of Europe (Sikora et al. 2014; Skoglund et al. 2012), Ötzi's Y chromosome lineage, G2a-L166, descends from the G2a-L91 lineage that was common among Anatolian Farmers 8 KYA (Lazaridis et al. 2016). Today, the G2a-L91 sub-clade has survived in Europe as a rare lineage with the highest incidence in Sardinia and Corsica (Francalacci et al. 2013; Keller et al. 2012).

Fig. 4 Major sub-clades of Y chromosome haplogroup C in ancient and present-day populations. The structure of the major sub-clades is drawn in proportion to their coalescent time (*the tip of each triangle*) estimated from high coverage genomes of present-day populations (Bergstrom et al. 2016; Karmin et al. 2015; Poznik et al. 2016; Scozzari et al. 2012). The phylogenetic mapping of ancient Y chromosomes (Gamba et al. 2014; Mathieson et al. 2015; Olalde et al. 2014; Seguin-Orlando et al. 2014) is shown with *red symbols*. Haplogroup names are shown in *brown font* and haplogroup-defining SNP-marker names in *grey font* next to relevant branches. The key areas of the present-day spread of the haplogroups are indicated with *colour* and *white text inside the triangles*. *PNG* Papua New Guinea
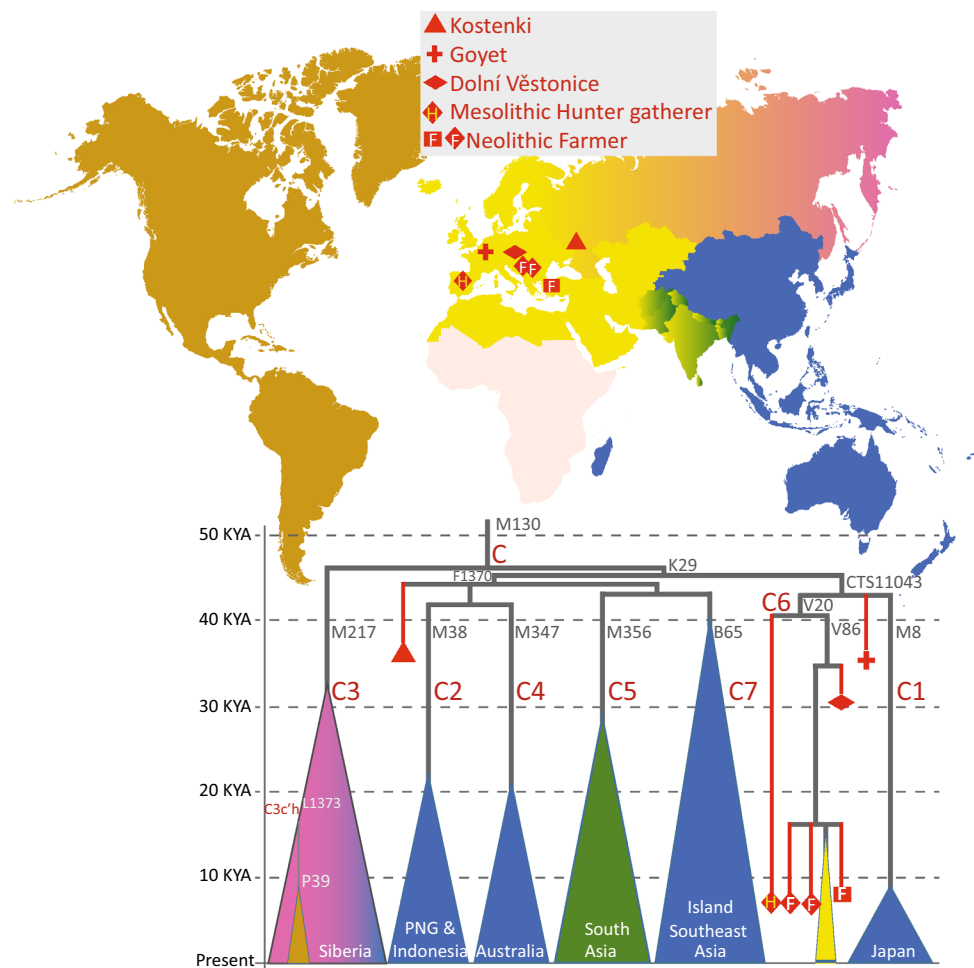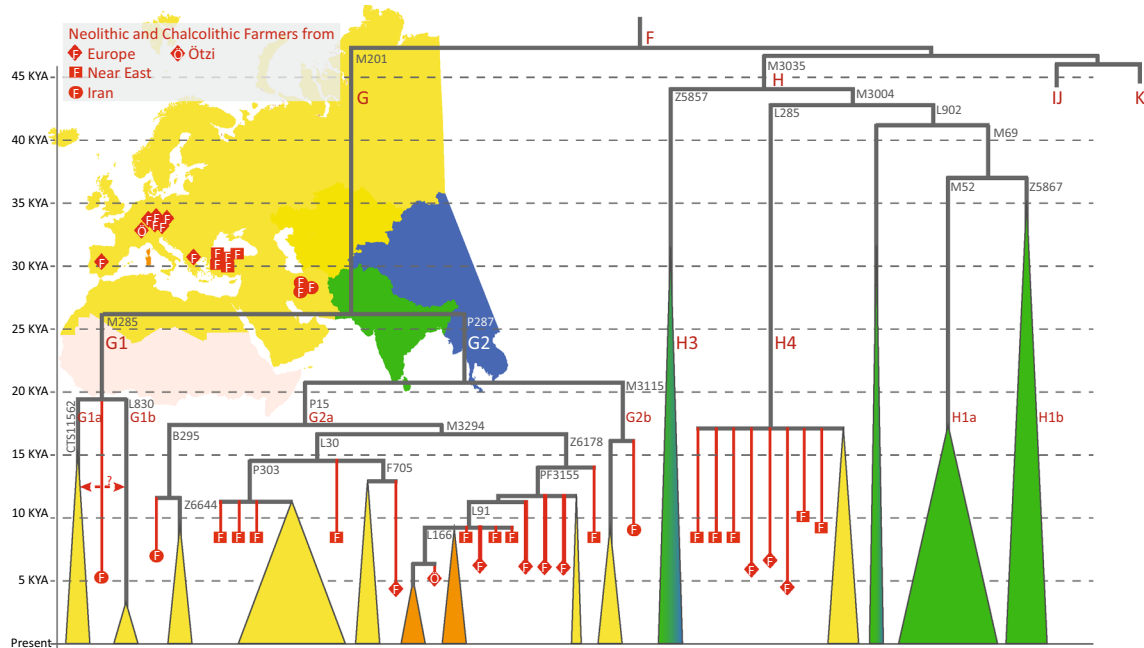
Fig. 5 Major sub-clades of Y chromosome haplogroups G and H in ancient and present-day populations. The structure of the major sub-clades is drawn in proportion to time according to estimates from high coverage genomes of present-day populations (Hallast et al. 2015; Karmin et al. 2015; Poznik et al. 2016) http://isogg.org/tree/, https://www.yfull.com/). The phylogenetic mapping of ancient Y chromosomes (Allentoft et al. 2015; Broushaki et al. 2016; Fu et al. 2016; Gunther et al. 2015; Haak et al. 2015; Hofmanova et al. 2016; Lazaridis et al. 2014, 2016; Mathieson et al. 2015) is shown with *red symbols*. Haplogroup-defining marker names are shown in *grey font* next to relevant branches

In contrast to haplogroup G, the geographic distribution of haplogroup H is presently almost entirely restricted to South Asia, while one of its sub-clades, H4-L285 (Fig. 5), can be detected as an extremely rare lineage in some European populations. H4-L285 has also been found in the aY sequences of the Anatolian and Levantine farmers as well as in Iberian Chalcolithic samples (Gunther et al. 2015; Lazaridis et al. 2016). Overall, the comparisons of early and middle Holocene versus present-day distributions of haplogroup G and H suggest that as characteristic markers of the early farmer populations of Middle East they were introduced to Europe by the expanding Anatolian farming populations. Their frequency has remained high in some geographically isolated areas such as the Caucasus, Sardinia, Corsica, whereas their frequency in main parts of Europe dropped later due to inflow of other Y chromosome lineages.

Haplogroup I frequency distribution is largely restricted to Europe where two major sub-clades I1 and I2 describe most of the extant variation (Fig. 6). It has been proposed that this restricted geographic area of its spread reflects local continuity since the Palaeolithic hunter-gatherers (Rootsi et al. 2004) standing in contrast to the geographic area of the highest frequency and diversity of its sister-clade J which alongside with the cline of haplogroup G frequency has been interpreted as a reflection of demic diffusion of farmers from Middle East to Europe (Semino et al. 2000). The ancient DNA evidence that has started to emerge recently confirms that haplogroup I was indeed common in Palaeolithic hunter-gatherers of Europe (Fu et al. 2016; Lazaridis et al. 2014); but it also points to a more complex picture with both hunter-gathers and farmers carrying haplogroup I as well as J lineages. The minor sub-clade of haplogroup I, I3-L596 (Fig. 6), has been found spread across a wide geographic area in Early and Middle Holocene samples, being found in Anatolian Farmers (Lazaridis et al. 2016) as well as in Scandinavian hunter-gatherers from Motala (Mathieson et al. 2015). I1-M253 lineages which are common in Scandinavia today, at 25–35% frequency (Rootsi et al. 2004), coalesce to a recent common founder at 5KYA. Similarly to its present day peak frequency area aY sequences falling to this clade are by and large restricted to three Nordic Late Neolithic and Bronze Age samples (Allentoft et al. 2015). Interestingly, one Trans-Danubian Early Neolithic (7.6–6.9KYA), LBK, sample from Hungary has also been recovered with I1-M253 affiliation (Szecsenyi-Nagy et al. 2015), suggesting that I1-M253 lineages may have been brought to Scandinavia by Neolithic farmers rather than representing local continuity of a pre-Holocene pool of Y chromosome lineages. I2-PF3835 lineages were, second only to haplogroup G, among the most common Y chromosome groups in early
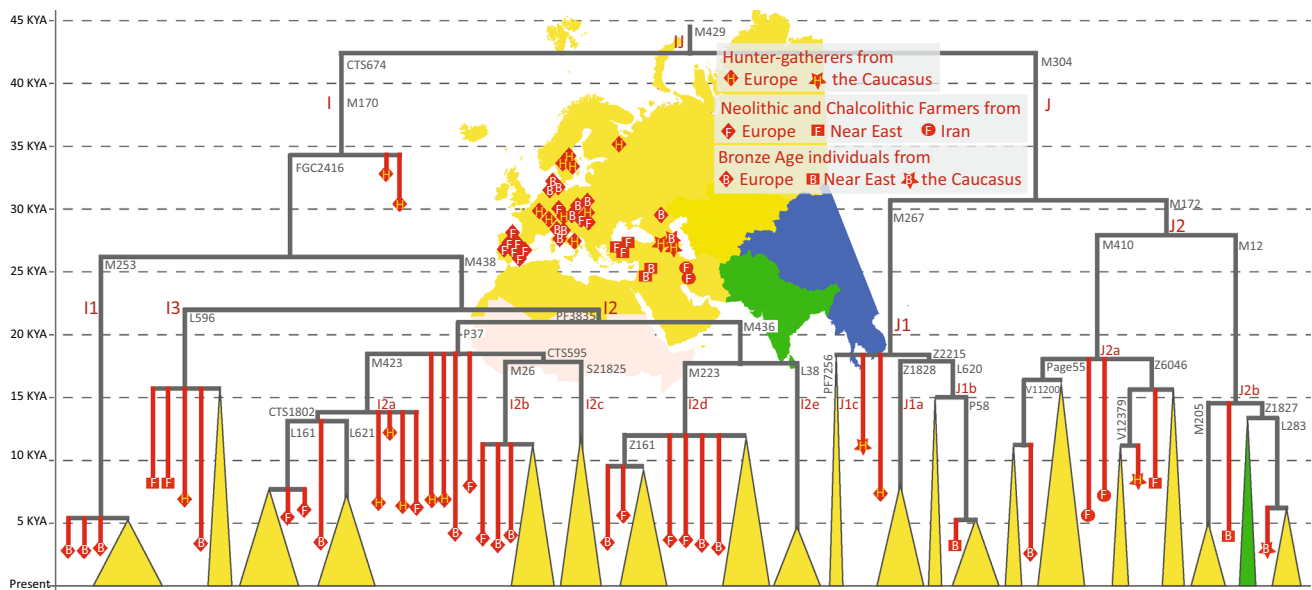
Fig. 6 Major sub-clades of Y chromosome haplogroups I and J in ancient and present-day populations. The structure of the major extant sub-clades of haplogroups I and J is shown by *triangles* the tips of which are drawn in proportion to time according to coalescent time estimates from high coverage genomes of present-day populations (Hallast et al. 2015; Karmin et al. 2015; Poznik et al. 2016) http://isogg.org/tree/, https://www.yfull.com/). The *colour of each* *triangle* reflects the main geographic areas of its spread according to the map shown in the *middle of the plot*. The phylogenetic affiliations of ancient Y chromosomes (Allentoft et al. 2015; Gamba et al. 2014; Gunther et al. 2015; Haak et al. 2015; Hofmanova et al. 2016; Jones et al. 2015; Lazaridis et al. 2016; Mathieson et al. 2015; Skoglund et al. 2012) are shown with *red symbols*. Haplogroup-defining marker names are shown in *grey font* next to relevant branches

farmers of Central Europe (Gamba et al. 2014; Mathieson et al. 2015; Szecsenyi-Nagy et al. 2015). Lineages distantly related to the extant sub-clades of haplogroup I2a-M423, which are common in Western (I2a-L161) and Eastern (I2a-L621) Europe today have been found in the aY sequences of hunter-gatherers from Switzerland, Hungary and Scandinavia as well as in Neolithic and Bronze Age samples from Hungary, Germany and Iberia suggesting some level of regional continuity (Haak et al. 2015; Jones et al. 2015; Mathieson et al. 2015).

Haplogroup J, which, on the basis of its present-day clinal frequency distribution has been associated with the early spread of farming to Europe (Rosser et al. 2000; Semino et al. 2000), has not been detected so far in European Neolithic context, instead, this haplogroup is found in hunter-gatherers from geographically distant areas, from the Caucasus and Karelia (Fig. 6), as well as in two early farmers from Iran and one from Anatolia (Jones et al. 2015; Lazaridis et al. 2016; Mathieson et al. 2015). In Central and Western Europe, J lineages start to emerge in the Bronze Age, likely being part of the demographic processes and population movements initiated from the North Caucasus area during that period. It is possible that these recent processes also introduced to Europe sub-clades of haplogroup E (Cinnioglu et al. 2004; Cruciani et al. 2007; Trombetta et al. 2015) which according to recent ancient DNA evidence was a characteristic haplogroup of the Natufians,

pre-pottery farmers of Levant and Ethiopians (Gallego Llorente et al. 2015; Lazaridis et al. 2016).

Y chromosome haplogroup R1b-M343 is currently the most common haplogroup among present day populations of Western Europe. It has a peak frequency of ~90% in the Basques who have been considered to represent a relict descendant group, by its genes and languages, of a pre-Neolithic population of Europe (Cavalli-Sforza et al. 1994; Richards et al. 1996). The decreasing northwest to southeast frequency gradient of haplogroup R1b-M343 was interpreted first to have resulted from admixture of Meso-lithic hunter-gatherers with Neolithic farmers carrying haplogroups J and G (Rosser et al. 2000; Semino et al. 2000). We now know from whole Y chromosome sequencing studies of modern samples that the coalescent time of the most common European sub-clade of R1b-M269 is shallow, 5–7 thousand years (Batini et al. 2015; Hallast et al. 2015; Karmin et al. 2015; Poznik et al. 2016). From the aDNA studies we have learned that the oldest R1b-M343 lineages, including 14 KYA Villabruna Man from Italy (Fu et al. 2016) and three European hunter-gatherers and three early farmer samples (Fig. 7), did not belong to the R1b-M269 sub-clade. According to the ancient DNA evidence, the R1b-M269 lineages did not, in fact, become common in Europe before the Late Neolithic/Bronze Age (Allentoft et al. 2015; Haak et al. 2015; Mathieson et al. 2015). We also know that modern-day Basques have the highest
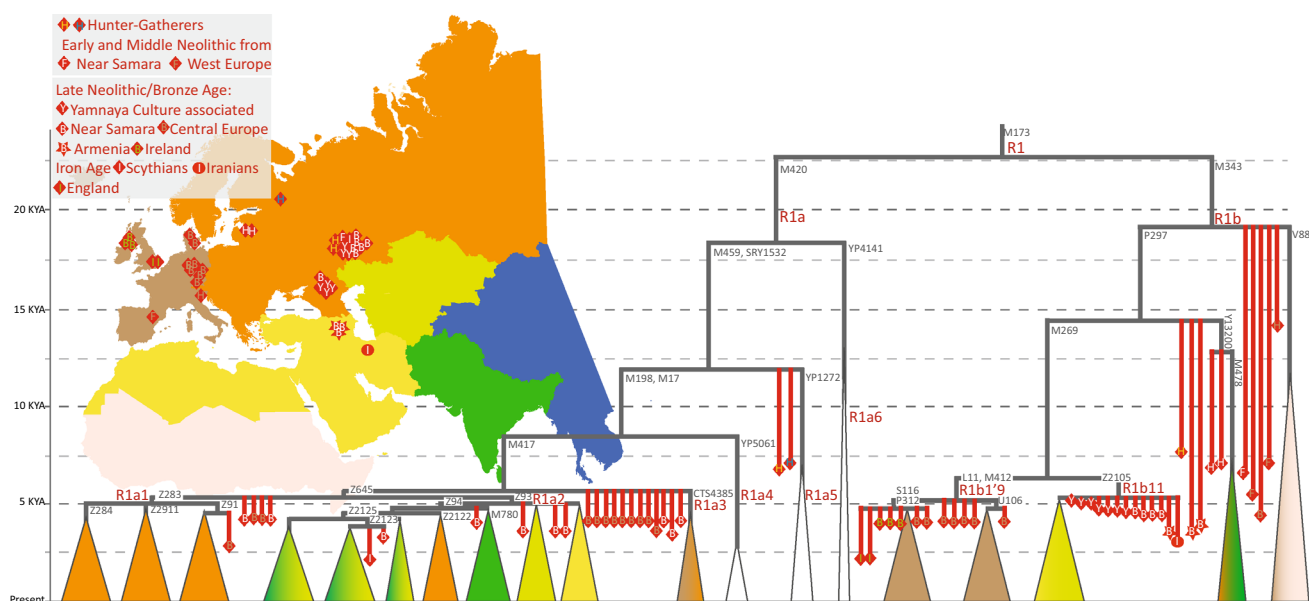
Fig. 7 Major sub-clades of Y chromosome haplogroups R1a and R1b in ancient and present-day populations. The structure of the major extant sub-clades of haplogroups R1a and R1b is shown by *triangles* the tips of which are drawn in proportion to time according to coalescent time estimates from high coverage genomes of present-day populations (Batini et al. 2015; Cruciani et al. 2010; Hallast et al. 2015; Karmin et al. 2015; Poznik et al. 2016), http://isogg.org/tree/, https://www.yfull.com/). The *colour of each triangle* reflects the main geographic areas of its spread according to the map shown in the *middle of the plot*. The phylogenetic affiliations of ancient Y chromosomes chromosomes (Allentoft et al. 2015; Broushaki et al. 2016; Cassidy et al. 2016; Haak et al. 2015; Jones et al. 2017; Lazaridis et al. 2016; Mathieson et al. 2015; Schiffels et al. 2016) are shown with *red symbols*. Haplogroup-defining marker names are shown in *grey font* next to relevant branches. Four Yamnaya, one Samara Bronze Age and two Nordic Bronze Age R1b-M269 samples could not be placed on the tree as they lacked coverage at M412, L11, and Z2105 SNPs

affinity in their autosomal genes to the early farmers of Atapuerca from Spain (Gunther et al. 2015) while their Y chromosomes arguably reflect a more recent male-specific admixture from Eastern Europe, from the areas of the distribution of Yamnaya Culture (Goldberg et al. 2017; Haak et al. 2015).

In contrast to preceding Early and Middle Neolithic sections of time, a large proportion of the Y chromosomes recovered from Bronze Age remains of Central Europe, Northern Caucasus and the Steppe belt of Russia belong to a couple of sub-clades of haplogroups R1a-M420 and R1b-M343 (Fig. 7). Late Neolithic, Early Bronze Age and Iron Age samples from Central and Western Europe have typically the R1b-L11, R1a1-Z283 and R1a-M417 (xZ645) affiliation while the samples from the Yamnaya and Samara neighbourhood are different and belong to sub-clades R1b11-Z2105 and R1a2-Z93 (Allentoft et al. 2015; Cassidy et al. 2016; Haak et al. 2015; Mathieson et al. 2015; Schiffels et al. 2016). The R1b11-Z2015 lineage is today common in the Caucasus and Volga-Uralic region while being virtually absent in Central and Western Europe (Broushaki et al. 2016). Interestingly, the earliest offshoot of extant haplogroup R1b-M343 variation, the V88 sub-clade, which is currently most common in Fulani speaking populations in Africa (Cruciani et al. 2010) has distant

relatives in Early Neolithic samples from across wide geographic area from Iberia, Germany to Samara (Fig. 7). In a similar way, early offshoots of the R1b and R1a phylogenies, including R1b lineages derived at P297 and ancestral at M269, and R1a lineages which are derived at M459 while ancestral at M198 and M417 markers have been found in mid-Holocene hunter-gatherer samples in a wide area in Eastern Europe, from Karelia, Latvia and Samara region (Haak et al. 2015; Jones et al. 2017; Mathieson et al. 2015). Extremely rare extant sub-clades of R1a, such as R1a4-YP5061, R1a5-YP1272, and R1a6-YP4141 (Fig. 7), may bear witness to a long-term continuity of such old genetic lineages while the majority of present-day R1a and R1b lineages in West Eurasia derives from just a handful of Late Neolithic/Early Bronze Age male founders.

Ancient Y chromosomes of the Native Americans

The present-day pool of Native American Y chromosomes is a mixture of haplogroups that derive from pre-Columbian dispersals from Siberia and more recent gene flow from Europe and Africa (Grugni et al. 2015; Kimura et al. 2016; Roewer et al. 2013; Zegura et al. 2004). The diversity derived from the first dispersals is restricted to just two

founding lineages within haplogroup Q and one or two in haplogroup C3-M217 (Fig. 3). The lineages specific to Native Americans within these two ancient haplogroups Q and C that have also branches that are commonly found in different parts of Eurasia have been suggested to have reached America by multiple independent dispersal events from Siberia (Lell et al. 2002). Lell et al. considered haplogroup Q, which is the most common clade in both North and South American Native populations, to derive from a migration from 'Middle' Siberia where it is highly frequent today. C3-M217, which in Americas is restricted, as a minor haplogroup, to a small number of populations, is the most common Y chromosome haplogroup in Northeast Siberians. Lell et al. (Lell et al. 2002) argued that the different spread patterns of these two haplogroups in both America and in Siberia are the outcome of dual origins of Native Americans implying two early dispersal events with two distinct source populations. Similar levels of STR diversities, however, that have been observed in Native American C3-P39 and Q1a-M3 lineages have been interpreted in favour of both of these haplogroups being part of the initial pool of the first expansion of Native Americans some 10–17 thousand years ago and thus can be viewed as being in line with what is called the single wave model (Zegura et al. 2004). The finding of a rare group of lineages of C3-M217 that share the ancestral allele for the P39 marker in Ecuador has recently reignited the debate about dual origins of Native American haplogroup C and Q lineages (Roewer et al. 2013). The C3-M217 Y chromosomes without the P39 marker were found in association with STR diversity coalescing to 6000 years and Roewer et al. (2013) suggested that their presence could be explained by a secondary wave from East Asia, and possibly, more specifically from Japan, considering that certain similarities in material culture exist between the areas where the C3-M217 lineages are found. However, analyses of autosomal SNPs have not supported the model of additional gene flow from East Asia to Ecuador suggesting that instead the presence of the rare C3-M217 lineages without the P39 marker represents a case of a rare founding lineage that has been lost elsewhere by drift (Mezzavilla et al. 2015).

Ancient DNA studies have cast new light on the debate of Native American origins and shown that the loss of rare lineages in post-contact Native Americans is not unusual and possibly part of the extensive lineage extinction process that has been observed in mitochondrial lineages (Llamas et al. 2016). The analyses of the genome of 24 KYA human remains, recovered from the Mal'ta site near Lake Baikal and shotgun-sequenced to an average depth of $1\times$ (Raghavan et al. 2014b), have shown that Native Americans do have 'dual ancestry' but not in the sense of the dual ancestry model of Lell et al. (2002). The autosomal genome of the Mal'ta Boy has close affinity to modern European and Native American populations while being more distant to East Asians which suggests, considering that overall Native Americans are more closely related to East Asians than to Europeans, that the Native Americans derive approximately one-third of their genetic ancestry from a population to which the Mal'ta Boy was related to while two-thirds of their ancestry derives from a different source, closely related to modern East Asian populations. The Y chromosome of the Mal'ta Boy (Fig. 3) is more closely affiliated to West Eurasian R lineages than to East Asian D, C or O lineages. It represents an extinct lineage that derives from the base of haplogroup R closely after the split of the ancestors of haplogroups Q and R. Because most Native American Y chromosomes belong to haplogroup Q, they are more closely related to European Y chromosomes while their maternal lineages (with the exception of rare haplogroup X2a) are all nested within East Asian variation. It should be noted, though, that since both mtDNA and Y are effectively single loci no firm conclusions about the sex-specific nature of the admixture process that lies at the foundation of Native American ancestry can be made from these observations.

Haplogroup Q has two ancient sub-clades, Q1a-M3 and Q1b-M971, which were likely born somewhere in Siberia before the first dispersal into Americas, and which together capture the overwhelming majority of extant Native American Y chromosomes today (Jota et al. 2016; Zegura et al. 2004). Besides these two major clades a number of rare sub-clades of Q that are geographically restricted to Europe, Central and South Asia, or Siberia have been identified (Jota et al. 2016; Karmin et al. 2015; Mallick et al. 2016; Poznik et al. 2016). Targeted PCR-based sequencing of a region surrounding the Q-M242 and Q1a-M3 markers confirmed the affiliation of the 10.3 KYA On Your Knees Cave Man (OYKCM) to haplogroup Q3-L275 (Kemp et al. 2007). Shotgun sequences of two ancient genomes from the Americas, the Anzick Boy (Rasmussen et al. 2014) associated with Clovis Culture, and the Kennewick Man (Rasmussen et al. 2015), have revealed at a finer resolution Y chromosomes that are representative of the present day diversity of haplogroup Q in Americas and together with the OYKCM evidence provide, thus, a direct support to the case that the two extant clades Q1a-M3 and Q1b-M971 have been in the Americas for at least 10,300 years. The Kennewick Man has a Y chromosome that belongs to the most common sub-clade Q1a-M3 while the Anzick's Y chromosome belongs to the minor Q1b-M971 lineage (Fig. 8).

The third ancient Y chromosome sequence from the Americas, or in fact, technically, from Greenland comes from the Saqqaq site and is dated to 4 KYA. The Saqqaq Man's mtDNA (Gilbert et al. 2008) and his whole genome, shotgun sequenced to an average depth of $20\times$

Fig. 8 The main branches of Y chromosome haplogroup Q in ancient and present-day populations. The structure of the major sub-clades is drawn in proportion to time according to estimates from high coverage genomes of present-day populations (Karmin et al. 2015; Poznik et al. 2016), http://isogg. org/tree/ISOGG_HapgrpQ.html, https://www.yfull.com/tree/Q/). The phylogenetic mapping of ancient Y chromosomes (Kemp et al. 2007; Rasmussen et al. 2010, 2014, 2015) is shown with *red symbols*. Haplogroup-defining marker names are shown in *grey font* next to relevant branches
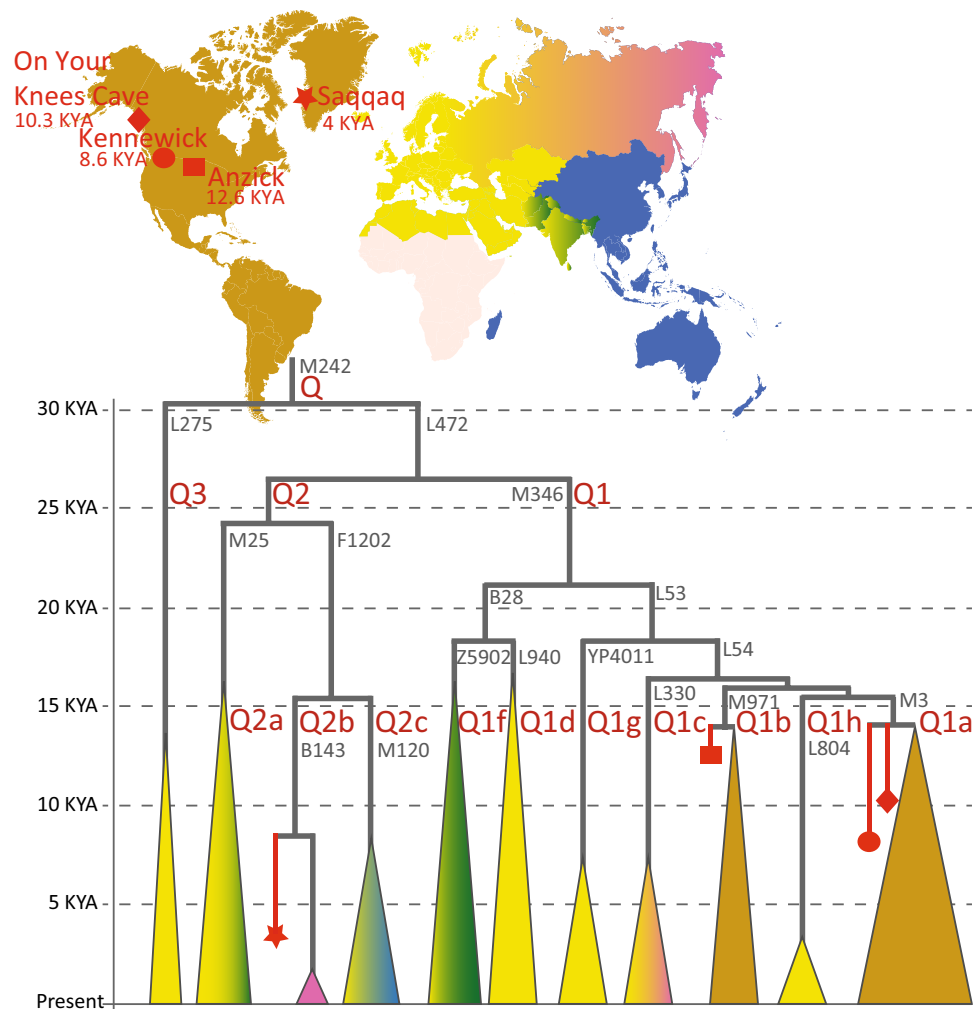
(Rasmussen et al. 2010) provided the first direct evidence of a separate Palaeo-Eskimo dispersal event into the Arctic North Americas. The Saqqaq Man's Y chromosome belongs to Q2b-B143, which is a sub-clade of haplogroup Q that is only distantly related, at time depth >25KYA, to the Q1a-M3 and Q1b-M971 lineages (Fig. 8). The Saqqaq's Q2b-B143 lineage was not found to be present in a survey of 1863 haplogroup Q lineages from South America making it unlikely to have been among the initial founding pool of Beringian Y chromosomes (Jota et al. 2016). From further ancient DNA analyses, we know that the Palaeo-Eskimos had extremely low genetic diversity, with only a single characteristic mtDNA lineage of haplogroup D2a1 being found in a wide range of sites from Northeast Canada and Greenland dated between 5000 to 700 years before present (Raghavan et al. 2014a). Analyses of autosomal genes and mtDNA of more recent remains associated with the Thule Culture suggest that the Palaeo-Eskimo population was completely replaced by the Neo-Eskimos within less than thousand years ago. Interestingly, though the Saqqaq Man's Y chromosome

lineage may have some continuity in the present day descendants of the Neo-Eskimo dispersal. Greenland Inuits have been shown to carry, at frequencies up to 54% in East Sermersooq, a lineage which is characterised by the NWT01 mutation (Olofsson et al. 2015), a SNP which separates the Y chromosomes of Inuits, who have it, from Athabascans, who do not (Dulik et al. 2012), and which is equivalent to the F1202 SNP that defines the clade that unites the Q2b-B143 and Q2c-M120 sub-clades (Fig. 8). The Y chromosome of the Saqqaq Man has been shown to share a number of SNPs equivalent to B143 with a group of Koryaks from Northeast Siberia (Karmin et al. 2015). It is yet to be revealed, however, whether the Neo-Eskimo Y chromosomes are derived at the SNPs defining the Q2b-B143 branch or whether they represent a yet another sub-clade of Q2-F1202. Further analyses of aY variation across the Americas would help to broaden our understanding of the past dynamics of the male effective population size and show to what extent the lineages present in the past have survived up to the present.

In sum, a number of aDNA studies have already started to reveal the potential of human Y chromosome to inform us about the demographic past complementing the study of autosomal genome and the X chromosome. When compared against the results derived from the analyses of autosomes the unique inheritance patterns of Y, as well as the X chromosome and mtDNA will provide us the opportunity to explore sex-specific dispersal and admixture processes in the future when further sampling will provide larger sample sizes and better coverage of the same geographic areas in time. Further breakthroughs of ancient DNA success in regions like Africa, Southeast Asia and Oceania will be most desirable to tackle broader range of questions about the continuity and nature of sex-specific dispersals and admixture in human evolutionary history. Recent cases of successful retrieval of ancient DNA from Ethiopia (Gallego Llorente et al. 2015), Vanuatu and Tonga (Skoglund et al. 2016) provide some optimism for the retrieval of aY sequence data from warmer climate regions in the future.

Compliance with ethical standards

Conflict of interest I declare no conflict of interests.

References

Allentoft ME, Sikora M, Sjogren KG, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlstrom T, Vinner L, Malaspinas AS, Margaryan A, Higham T, Chivall D, Lynnerup N, Harvig L, Baron J, Della Casa P, Dabrowski P, Duffy PR, Ebel AV, Epimakhov A, Frei K, Furmanek M, Gralak T, Gromov A, Gronkiewicz S, Grupe G, Hajdu T, Jarysz R, Khartanovich V, Khokhlov A, Kiss V, Kolar J, Kriiska A, Lasak I, Longhi C, McGlynn G, Merkevicius A, Merkyte I, Metspalu M, Mkrtchyan R, Moiseyev V, Paja L, Palfi G, Pokutta D, Pospieszny L, Price TD, Saag L, Sablin M, Shishlina N, Smrcka V, Soenov VI, Szeverenyi V, Toth G, Trifanova SV, Varul L, Vicze M, Yepiskoposyan L, Zhitenev V, Orlando L, Sicheritz-Ponten T, Brunak S, Nielsen R, Kristiansen K, Willerslev E (2015) Population genomics of Bronze Age Eurasia. Nature 522:167–172. doi:10.1038/nature14507

Batini C, Hallast P, Zadik D, Delser PM, Benazzo A, Ghirotto S, Arroyo-Pardo E, Cavalleri GL, de Knijff P, Dupuy BM, Eriksen HA, King TE, Lopez de Munain A, Lopez-Parra AM, Loutradis A, Milasin J, Novelletto A, Pamjav H, Sajantila A, Tolun A, Winney B, Jobling MA (2015) Large-scale recent expansion of European patrilineages shown by population resequencing. Nat Commun 6:7152. doi:10.1038/ncomms8152

Bergstrom A, Nagle N, Chen Y, McCarthy S, Pollard MO, Ayub Q, Wilcox S, Wilcox L, van Oorschot RA, McAllister P, Williams L, Xue Y, Mitchell RJ, Tyler-Smith C (2016) Deep roots for aboriginal Australian Y chromosomes. Curr Biol 26:809–813. doi:10.1016/j.cub.2016.01.028

Bollongino R, Nehlich O, Richards MP, Orschiedt J, Thomas MG, Sell C, Fajkosova Z, Powell A, Burger J (2013) 2000 years of parallel societies in Stone Age Central Europe. Science 342:479–481. doi:10.1126/science.1245049

Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P, Tambets K, Antanaitis-Jacobs I, Haidle MN, Jankauskas R, Kind CJ, Lueth F, Terberger T, Hiller J, Matsumura S, Forster P, Burger J (2009) Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. Science 326:137–140. doi:10.1126/science.1176869

Brandt G, Haak W, Adler CJ, Roth C, Szecsenyi-Nagy A, Karimnia S, Moller-Rieker S, Meller H, Ganslmeier R, Friederich S, Dresely V, Nicklisch N, Pickrell JK, Sirocko F, Reich D, Cooper A, Alt KW, Genographic C (2013) Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. Science 342:257–261. doi:10.1126/science.1241844

Broushaki F, Thomas MG, Link V, Lopez S, van Dorp L, Kirsanow K, Hofmanova Z, Diekmann Y, Cassidy LM, Diez-del-Molino D, Kousathanas A, Sell C, Robson HK, Martiniano R, Blocher J, Scheu A, Kreutzer S, Bollongino R, Bobo D, Davoudi H, Munoz O, Currat M, Abdi K, Biglari F, Craig OE, Bradley DG, Shennan S, Veeramah KR, Mashkour M, Wegmann D, Hellenthal G, Burger J (2016) Early Neolithic genomes from the eastern fertile crescent. Science 353:499–503. doi:10.1126/science.aaf7943

Cassidy LM, Martiniano R, Murphy EM, Teasdale MD, Mallory J, Hartwell B, Bradley DG (2016) Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. Proc Natl Acad Sci USA 113:368–373. doi:10.1073/pnas.1518445113

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton University Press, Princeton

Cinnioglu C, King R, Kivisild T, Kalfoglu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K, Oefner PJ, Shen P, Semino O, Cavalli-Sforza LL, Underhill PA (2004) Excavating Y-chromosome haplotype strata in Anatolia. Hum Genet 114:127–148. doi:10.1007/s00439-003-1031-4

Cruciani F, La Fratta R, Trombetta B, Santolamazza P, Sellitto D, Colomb EB, Dugoujon JM, Crivellaro F, Benincasa T, Pascone R, Moral P, Watson E, Melegh B, Barbujani G, Fuselli S, Vona G, Zagradisnik B, Assum G, Brdicka R, Kozlov AI, Efremov GD, Coppa A, Novelletto A, Scozzari R (2007) Tracing past human male movements in northern/eastern Africa and western Eurasia: new clues from Y-chromosomal haplogroups E-M78 and J-M12. Mol Biol Evol 24:1300–1311. doi:10.1093/molbev/msm049

Cruciani F, Trombetta B, Sellitto D, Massaia A, Destro-Bisol G, Watson E, Beraud Colomb E, Dugoujon JM, Moral P, 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. doi:10.1038/ejhg.2009.231

Cruciani F, Trombetta B, Massaia A, Destro-Bisol G, Sellitto D, 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. doi:10.1016/j.ajhg.2011.05.002

Dulik MC, Owings AC, Gaieski JB, Vilar MG, Andre A, Lennie C, Mackenzie MA, Kritsch I, Snowshoe S, Wright R, Martin J,

Gibson N, Andrews TD, Schurr TG, Genographic C (2012) Y-chromosome analysis reveals genetic divergence and new founding native lineages in Athapaskan- and Eskimoan-speaking populations. Proc Natl Acad Sci USA 109:8471–8476. doi:10.1073/pnas.1118760109

Firasat S, Khaliq S, Mohyuddin A, Papaioannou M, Tyler-Smith C, Underhill PA, Ayub Q (2007) Y-chromosomal evidence for a limited Greek contribution to the Pathan population of Pakistan. Eur J Hum Genet 15:121–126. doi:10.1038/sj.ejhg.5201726

Francalacci P, Morelli L, Angius A, Berutti R, Reinier F, Atzeni R, Pilu R, Busonero F, Maschio A, Zara I, Sanna D, Useli A, Urru MF, Marcelli M, Cusano R, Oppo M, Zoledziewska M, Pitzalis M, Deidda F, Porcu E, Poddie F, Kang HM, Lyons R, Tarrier B, Gresham JB, Li B, Tofanelli S, Alonso S, Dei M, Lai S, Mulas A, Whalen MB, Uzzau S, Jones C, Schlessinger D, Abecasis GR, Sanna S, Sidore C, Cucca F (2013) Low-pass DNA sequencing of 1200 Sardinians reconstructs European Y-chromosome phylogeny. Science 341:565–569. doi:10.1126/science.1237947

Francalacci P, Sanna D, Useli A, Berutti R, Barbato M, Whalen MB, Angius A, Sidore C, Alonso S, Tofanelli S, Cucca F (2015) Detection of phylogenetically informative polymorphisms in the entire euchromatic portion of human Y chromosome from a Sardinian sample. BMC Res Notes 8:174. doi:10.1186/s13104-015-1130-z

Fu Q, Meyer M, Gao X, Stenzel U, Burbano HA, Kelso J, Paabo S (2013) DNA analysis of an early modern human from Tianyuan Cave, China. Proc Natl Acad Sci USA 110:2223–2227. doi:10.1073/pnas.1221359110

Fu Q, Li H, Moorjani P, Jay F, Slepchenko SM, Bondarev AA, Johnson PL, Aximu-Petri A, Prufer K, de Filippo C, Meyer M, Zwyns N, Salazar-Garcia DC, Kuzmin YV, Keates SG, Kosintsev PA, Razhev DI, Richards MP, Peristov NV, Lachmann M, Douka K, Higham TF, Slatkin M, Hublin JJ, Reich D, Kelso J, Viola TB, Paabo S (2014) Genome sequence of a 45,000-year-old modern human from western Siberia. Nature 514:445–449. doi:10.1038/nature13810

Fu Q, Hajdinjak M, Moldovan OT, Constantin S, Mallick S, Skoglund P, Patterson N, Rohland N, Lazaridis I, Nickel B, Viola B, Prufer K, Meyer M, Kelso J, Reich D, Paabo S (2015) An early modern human from Romania with a recent Neanderthal ancestor. Nature 524:216–219. doi:10.1038/nature14558

Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, Furtwangler A, Haak W, Meyer M, Mittnik A, Nickel B, Peltzer A, Rohland N, Slon V, Talamo S, Lazaridis I, Lipson M, Mathieson I, Schiffels S, Skoglund P, Derevianko AP, Drozdov N, Slavinsky V, Tsybankov A, Cremonesi RG, Mallegni F, Gely B, Vacca E, Morales MR, Straus LG, Neugebauer-Maresch C, Teschler-Nicola M, Constantin S, Moldovan OT, Benazzi S, Peresani M, Coppola D, Lari M, Ricci S, Ronchitelli A, Valentin F, Thevenet C, Wehrberger K, Grigorescu D, Rougier H, Crevecoeur I, Flas D, Semal P, Mannino MA, Cupillard C, Bocherens H, Conard NJ, Harvati K, Moiseyev V, Drucker DG, Svoboda J, Richards MP, Caramelli D, Pinhasi R, Kelso J, Patterson N, Krause J, Paabo S, Reich D (2016) The genetic history of Ice Age Europe. Nature 534:200–205. doi:10.1038/nature17993

Gallego Llorente M, Jones ER, Eriksson A, Siska V, Arthur KW, Arthur JW, Curtis MC, Stock JT, Coltorti M, Pieruccini P, Stretton S, Brock F, Higham T, Park Y, Hofreiter M, Bradley DG, Bhak J, Pinhasi R, Manica A (2015) Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent. Science 350:820–822. doi:10.1126/science.aad2879

Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboroczki L, Kovari I, Pap I, Anders A, Whittle A, Dani J, Raczky P, Higham TF, Hofreiter M, Bradley DG, Pinhasi R (2014) Genome flux and stasis in a five millennium transect of European prehistory. Nat Commun 5:5257. doi:10.1038/ncomms6257

Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491:56–65. doi:10.1038/nature11632

Gilbert MT, Willerslev E, Hansen AJ, Barnes I, Rudbeck L, Lynnerup N, Cooper A (2003) Distribution patterns of postmortem damage in human mitochondrial DNA. Am J Hum Genet 72:32–47. doi:10.1086/345378

Gilbert MT, Kivisild T, Gronnow B, Andersen PK, Metspalu E, Reidla M, Tamm E, Axelsson E, Gotherstrom A, Campos PF, Rasmussen M, Metspalu M, Higham TF, Schwenninger JL, Nathan R, De Hoog CJ, Koch A, Moller LN, Andreasen C, Meldgaard M, Villems R, Bendixen C, Willerslev E (2008) Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. Science 320:1787–1789. doi:10.1126/science.1159750

Goldberg A, Günther T, Rosenberg NA, Jakobsson M (2017) Ancient X chromosomes reveal contrasting sex bias in Neolithic and Bronze Age migrations. Proc Natl Acad Sci USA. doi:10.1073/pnas.1616392114

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, Hansen NF, Durand EY, Malaspinas AS, Jensen JD, Marques-Bonet T, Alkan C, Prufer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Hober B, Hoffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Z, Gusic I, Doronichev VB, Golovanova LV, Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PL, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D, Paabo S (2010) A draft sequence of the Neandertal genome. Science 328:710–722. doi:10.1126/science.1188021

Grugni V, Battaglia V, Perego UA, Raveane A, Lancioni H, Olivieri A, Ferretti L, Woodward SR, Pascale JM, Cooke R, Myres N, Motta J, Torroni A, Achilli A, Semino O (2015) Exploring the Y chromosomal ancestry of modern panamanians. PLoS One 10:e0144223. doi:10.1371/journal.pone.0144223

Gunther T, Valdiosera C, Malmstrom H, Urena I, Rodriguez-Varela R, Sverrisdottir OO, Daskalaki EA, Skoglund P, Naidoo T, Svensson EM, Bermudez de Castro JM, Carbonell E, Dunn M, Stora J, Iriarte E, Arsuaga JL, Carretero JM, Gotherstrom A, Jakobsson M (2015) Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. Proc Natl Acad Sci USA 112:11917–11922. doi:10.1073/pnas.1509851112

Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney E, Stewardson K, Fu Q, Mittnik A, Banffy E, Economou C, Francken M, Friederich S, Pena RG, Hallgren F, Khartanovich V, Khokhlov A, Kunst M, Kuznetsov P, Meller H, Mochalov O, Moiseyev V, Nicklisch N, Pichler SL, Risch R, Rojo Guerra MA, Roth C, Szecsenyi-Nagy A, Wahl J, Meyer M, Krause J, Brown D, Anthony D, Cooper A, Alt KW, Reich D (2015) Massive migration from the steppe was a source for Indo-European languages in Europe. Nature 522:207–211. doi:10.1038/nature14317

Haber M, Platt DE, Ashrafian Bonab M, Youhanna SC, Soria-Hernanz DF, Martinez-Cruz B, Douaihy B, Ghassibe-Sabbagh M, Rafatpanah H, Ghanbari M, Whale J, Balanovsky O, Wells RS, Comas D, Tyler-Smith C, Zalloua PA, Genographic C (2012) Afghanistan's ethnic groups share a Y-chromosomal heritage structured by historical events. PLoS One 7:e34288. doi:10.1371/journal.pone.0034288

Hallast P, Batini C, Zadik D, Maisano Delser P, Wetton JH, Arroyo-Pardo E, Cavalleri GL, de Knijff P, Destro Bisol G, Dupuy BM, Eriksen HA, Jorde LB, King TE, Larmuseau MH, Lopez de Munain A, Lopez-Parra AM, Loutradis A, Milasin J, Novelletto

A, Pamjav H, Sajantila A, Schempp W, Sears M, Tolun A, Tyler-Smith C, Van Geystelen A, Watkins S, Winney B, Jobling MA (2015) The Y-chromosome tree bursts into leaf: 13,000 high-confidence SNPs covering the majority of known clades. Mol Biol Evol 32:661–673. doi:10.1093/molbev/msu327

Hofmanova Z, Kreutzer S, Hellenthal G, Sell C, Diekmann Y, Diez-Del-Molino D, van Dorp L, Lopez S, Kousathanas A, Link V, Kirsanow K, Cassidy LM, Martiniano R, Strobel M, Scheu A, Kotsakis K, Halstead P, Triantaphyllou S, Kyparissi-Apostolika N, Urem-Kotsou D, Ziota C, Adaktylou F, Gopalan S, Bobo DM, Winkelbach L, Blocher J, Unterlander M, Leuenberger C, Cilingiroglu C, Horejs B, Gerritsen F, Shennan SJ, Bradley DG, Currat M, Veeramah KR, Wegmann D, Thomas MG, Papageorgopoulou C, Burger J (2016) Early farmers from across Europe directly descended from Neolithic Aegeans. Proc Natl Acad Sci USA 113:6886–6891. doi:10.1073/pnas.1523951113

Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Paabo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res 29:4793–4799

Hofreiter M, Paijmans JL, Goodchild H, Speller CF, Barlow A, Fortes GG, Thomas JA, Ludwig A, Collins MJ (2015) The future of ancient DNA: technical advances and conceptual shifts. BioEssays 37:284–293. doi:10.1002/bies.201400160

Ilumae AM, Reidla M, Chukhryaeva M, Jarve M, Post H, Karmin M, Saag L, Agdzhoyan A, Kushniarevich A, Litvinov S, Ekomasova N, Tambets K, Metspalu E, Khusainova R, Yunusbayev B, Khusnutdinova EK, Osipova LP, Fedorova S, Utevska O, Koshel S, Balanovska E, Behar DM, Balanovsky O, Kivisild T, Underhill PA, Villems R, Rootsi S (2016) Human Y chromosome haplogroup N: a non-trivial time-resolved phylogeography that cuts across language families. Am J Hum Genet 99:163–173. doi:10.1016/j.ajhg.2016.05.025

Jones ER, Gonzalez-Fortes G, Connell S, Siska V, Eriksson A, Martiniano R, McLaughlin RL, Gallego Llorente M, Cassidy LM, Gamba C, Meshveliani T, Bar-Yosef O, Muller W, Belfer-Cohen A, Matskevich Z, Jakeli N, Higham TF, Currat M, Lordkipanidze D, Hofreiter M, Manica A, Pinhasi R, Bradley DG (2015) Upper Palaeolithic genomes reveal deep roots of modern Eurasians. Nat Commun 6:8912. doi:10.1038/ncomms9912

Jones ER, Zarina G, Moiseyev V, Lightfoot E, Nigst PR, Manica A, Pinhasi R, Bradley DG (2017) The neolithic transition in the baltic was not driven by admixture with early European farmers. Curr Biol. doi:10.1016/j.cub.2016.12.060

Jota MS, Lacerda DR, Sandoval JR, Vieira PP, Ohasi D, Santos-Junior JE, Acosta O, Cuellar C, Revollo S, Paz YMC, Fujita R, Vallejo GA, Schurr TG, Tarazona-Santos EM, Pena S, Ayub Q, Tyler-Smith C, Santos FR, Genographic C (2016) New native South American Y chromosome lineages. J Hum Genet 61:593–603. doi:10.1038/jhg.2016.26

Karafet T, Xu L, Du R, Wang W, Feng S, Wells RS, Redd AJ, Zegura SL, Hammer MF (2001) Paternal population history of East Asia: sources, patterns, and microevolutionary processes. Am J Hum Genet 69:615–628. doi:10.1086/323299

Karafet TM, Osipova LP, Gubina MA, Posukh OL, Zegura SL, Hammer MF (2002) High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. Hum Biol 74:761–789

Karafet TM, Mendez FL, Sudoyo H, Lansing JS, Hammer MF (2015) Improved phylogenetic resolution and rapid diversification of Y-chromosome haplogroup K-M526 in Southeast Asia. Eur J Hum Genet 23:369–373. doi:10.1038/ejhg.2014.106

Karmin M, Saag L, Vicente M, Wilson Sayres MA, Jarve M, Talas UG, Rootsi S, Ilumae AM, Magi R, Mitt M, Pagani L, Puurand T, Faltyskova Z, Clemente F, Cardona A, Metspalu E, Sahakyan H, Yunusbayev B, Hudjashov G, DeGiorgio M, Loogvali EL,

Eichstaedt C, Eelmets M, Chaubey G, Tambets K, Litvinov S, Mormina M, Xue Y, Ayub Q, Zoraqi G, Korneliussen TS, Akhatova F, Lachance J, Tishkoff S, Momynaliev K, Ricaut FX, Kusuma P, Razafindrazaka H, Pierron D, Cox MP, Sultana GN, Willerslev R, Muller C, Westaway M, Lambert D, Skaro V, Kovacevic L, Turdikulova S, Dalimova D, Khusainova R, Trofimova N, Akhmetova V, Khidiyatova I, Lichman DV, Isakova J, Pocheshkhova E, Sabitov Z, Barashkov NA, Nymadawa P, Mihailov E, Seng JW, Evseeva I, Migliano AB, Abdullah S, Andriadze G, Primorac D, Atramentova L, Utevska O, Yepiskoposyan L, Marjanovic D, Kushniarevich A, Behar DM, Gilissen C, Vissers L, Veltman JA, Balanovska E, Derenko M, Malyarchuk B, Metspalu A, Fedorova S, Eriksson A, Manica A, Mendez FL, Karafet TM, Veeramah KR, Bradman N, Hammer MF, Osipova LP, Balanovsky O, Khusnutdinova EK, Johnsen K, Remm M, Thomas MG, Tyler-Smith C, Underhill PA, Willerslev E, Nielsen R, Metspalu M, Villems R, Kivisild T (2015) A recent bottleneck of Y chromosome diversity coincides with a global change in culture. Genome Res 25:459–466. doi:10.1101/gr.186684.114

Kayser M (2010) The human genetic history of Oceania: near and remote views of dispersal. Curr Biol 20:R194–R201. doi:10.1016/j.cub.2009.12.004

Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, Stade B, Franke A, Mayer J, Spangler J, McLaughlin S, Shah M, Lee C, Harkins TT, Sartori A, Moreno-Estrada A, Henn B, Sikora M, Semino O, Chiaroni J, Rootsi S, Myres NM, Cabrera VM, Underhill PA, Bustamante CD, Vigl EE, Samadelli M, Cipollini G, Haas J, Katus H, O'Connor BD, Carlson MR, Meder B, Blin N, Meese E, Pusch CM, Zink A (2012) New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nat Commun 3:698. doi:10.1038/ncomms1701

Kemp BM, Malhi RS, McDonough J, Bolnick DA, Eshleman JA, Rickards R, Martinez-Labarga C, Johnson JR, Lorenz JG, Dixon EJ, Fifield TE, Heaton TH, Worl R, Smith DG (2007) Genetic analysis of early holocene skeletal remains from Alaska and its implications for the settlement of the Americas. Am J Phys Anthropol 132:605–621. doi:10.1002/ajpa.20543

Kimura L, Nunes K, Macedo-Souza LI, Rocha J, Meyer D, Mingroni-Netto RC (2016) Inferring paternal history of rural African-derived Brazilian populations from Y chromosomes. Am J Hum Biol. doi:10.1002/ajhb.22930

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Paabo S (1997) Neandertal DNA sequences and the origin of modern humans. Cell 90:19–30

Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, Berger B, Economou C, Bollongino R, Fu Q, Bos KI, Nordenfelt S, Li H, de Filippo C, Prufer K, Sawyer S, Posth C, Haak W, Hallgren F, Fornander E, Rohland N, Delsate D, Francken M, Guinet JM, Wahl J, Ayodo G, Babiker HA, Bailliet G, Balanovska E, Balanovsky O, Barrantes R, Bedoya G, Ben-Ami H, Bene J, Berrada F, Bravi CM, Brisighelli F, Busby GB, Cali F, Churnosov M, Cole DE, Corach D, Damba L, van Driem G, Dryomov S, Dugoujon JM, Fedorova SA, Gallego Romero I, Gubina M, Hammer M, Henn BM, Hervig T, Hodoglugil U, Jha AR, Karachanak-Yankova S, Khusainova R, Khusnutdinova E, Kittles R, Kivisild T, Klitz W, Kucinskas V, Kushniarevich A, Laredj L, Litvinov S, Loukidis T, Mahley RW, Melegh B, Metspalu E, Molina J, Mountain J, Nakkalajarvi K, Nesheva D, Nyambo T, Osipova L, Parik J, Platonov F, Posukh O, Romano V, Rothhammer F, Rudan I, Ruizbakiev R, Sahakyan H, Sajantila A, Salas A, Starikovskaya EB, Tarekegn A, Toncheva D, Turdikulova S, Uktveryte I, Utevska O, Vasquez R, Villena M, Voevoda M, Winkler CA, Yepiskoposyan L, Zalloua P

et al (2014) Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513:409–413. doi:10.1038/nature13673

Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, Fernandes D, Novak M, Gamarra B, Sirak K, Connell S, Stewardson K, Harney E, Fu Q, Gonzalez-Fortes G, Jones ER, Roodenberg SA, Lengyel G, Bocquentin F, Gasparian B, Monge JM, Gregg M, Eshed V, Mizrahi AS, Meiklejohn C, Gerritsen F, Bejenaru L, Bluher M, Campbell A, Cavalleri G, Comas D, Froguel P, Gilbert E, Kerr SM, Kovacs P, Krause J, McGettigan D, Merrigan M, Merriwether DA, O'Reilly S, Richards MB, Semino O, Shamoon-Pour M, Stefanescu G, Stumvoll M, Tonjes A, Torroni A, Wilson JF, Yengo L, Hovhannisyan NA, Patterson N, Pinhasi R, Reich D (2016) Genomic insights into the origin of farming in the ancient Near East. Nature 536:419–424. doi:10.1038/nature19310

Lell JT, Sukernik RI, Starikovskaya YB, Su B, Jin L, Schurr TG, Underhill PA, Wallace DC (2002) The dual origin and Siberian affinities of Native American Y chromosomes. Am J Hum Genet 70:192–206. doi:10.1086/338457

Lippold S, Xu H, Ko A, Li M, Renaud G, Butthof A, Schroder R, Stoneking M (2014) Human paternal and maternal demographic histories: insights from high-resolution Y chromosome and mtDNA sequences. Investig Genet 5:13. doi:10.1186/2041-2223-5-13

Llamas B, Fehren-Schmitz L, Valverde G, Soubrier J, Mallick S, Rohland N, Nordenfelt S, Valdiosera C, Richards SM, Rohrlach A, Romero MI, Espinoza IF, Cagigao ET, Jimenez LW, Makowski K, Reyna IS, Lory JM, Torrez JA, Rivera MA, Burger RL, Ceruti MC, Reinhard J, Wells RS, Politis G, Santoro CM, Standen VG, Smith C, Reich D, Ho SY, Cooper A, Haak W (2016) Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Sci Adv 2:e1501385. doi:10.1126/sciadv.1501385

Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, Zhao M, Chennagiri N, Nordenfelt S, Tandon A, Skoglund P, Lazaridis I, Sankararaman S, Fu Q, Rohland N, Renaud G, Erlich Y, Willems T, Gallo C, Spence JP, Song YS, Poletti G, Balloux F, van Driem G, de Knijff P, Romero IG, Jha AR, Behar DM, Bravi CM, Capelli C, Hervig T, Moreno-Estrada A, Posukh OL, Balanovska E, Balanovsky O, Karachanak-Yankova S, Sahakyan H, Toncheva D, Yepiskoposyan L, Tyler-Smith C, Xue Y, Abdullah MS, Ruiz-Linares A, Beall CM, Di Rienzo A, Jeong C, Starikovskaya EB, Metspalu E, Parik J, Villems R, Henn BM, Hodoglugil U, Mahley R, Sajantila A, Stamatoyannopoulos G, Wee JT, Khusainova R, Khusnutdinova E, Litvinov S, Ayodo G, Comas D, Hammer MF, Kivisild T, Klitz W, Winkler CA, Labuda D, Bamshad M, Jorde LB, Tishkoff SA, Watkins WS, Metspalu M, Dryomov S, Sukernik R, Singh L, Thangaraj K, Paabo S, Kelso J, Patterson N, Reich D (2016) The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature 538:201–206. doi:10.1038/nature18964

Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, Sirak K, Gamba C, Jones ER, Llamas B, Dryomov S, Pickrell J, Arsuaga JL, de Castro JM, Carbonell E, Gerritsen F, Khokhlov A, Kuznetsov P, Lozano M, Meller H, Mochalov O, Moiseyev V, Guerra MA, Roodenberg J, Verges JM, Krause J, Cooper A, Alt KW, Brown D, Anthony D, Lalueza-Fox C, Haak W, Pinhasi R, Reich D (2015) Genome-wide patterns of selection in 230 ancient Eurasians. Nature 528:499–503. doi:10.1038/nature16152

Mendez FL, Krahn T, Schrack B, Krahn AM, Veeramah KR, Woerner AE, Fomine FL, Bradman N, Thomas MG, Karafet TM, Hammer MF (2013) An African American paternal lineage adds an extremely ancient root to the human Y chromosome

phylogenetic tree. Am J Hum Genet 92:454–459. doi:10.1016/j.ajhg.2013.02.002

Mezzavilla M, Geppert M, Tyler-Smith C, Roewer L, Xue Y (2015) Insights into the origin of rare haplogroup C3* Y chromosomes in South America from high-density autosomal SNP genotyping. Forensic Sci Int Genet 15:115–120. doi:10.1016/j.fsigen.2014.11.005

Nasidze I, Quinque D, Rahmani M, Alemohamad SA, Stoneking M (2008) Close genetic relationship between Semitic-speaking and Indo-European-speaking groups in Iran. Ann Hum Genet 72:241–252. doi:10.1111/j.1469-1809.2007.00413.x

Olalde I, Allentoft ME, Sanchez-Quinto F, Santpere G, Chiang CW, DeGiorgio M, Prado-Martinez J, Rodriguez JA, Rasmussen S, Quilez J, Ramirez O, Marigorta UM, Fernandez-Callejo M, Prada ME, Encinas JM, Nielsen R, Netea MG, Novembre J, Sturm RA, Sabeti P, Marques-Bonet T, Navarro A, Willerslev E, Lalueza-Fox C (2014) Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature 507:225–228. doi:10.1038/nature12960

Olofsson JK, Pereira V, Borsting C, Morling N (2015) Peopling of the North Circumpolar Region–insights from Y chromosome STR and SNP typing of Greenlanders. PLoS ONE 10:e0116573. doi:10.1371/journal.pone.0116573

Orlando L, Gilbert MT, Willerslev E (2015) Reconstructing ancient genomes and epigenomes. Nat Rev Genet 16:395–408. doi:10.1038/nrg3935

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D (2012) Ancient admixture in human history. Genetics 192:1065–1093. doi:10.1534/genetics.112.145037

Pickrell JK, Reich D (2014) Toward a new history and geography of human genes informed by ancient DNA. Trends Genet 30:377–389. doi:10.1016/j.tig.2014.07.007

Posth C, Renaud G, Mittnik A, Drucker DG, Rougier H, Cupillard C, Valentin F, Thevenet C, Furtwangler A, Wissing C, Francken M, Malina M, Bolus M, Lari M, Gigli E, Capecchi G, Crevecoeur I, Beauval C, Flas D, Germonpre M, van der Plicht J, Cottiaux R, Gely B, Ronchitelli A, Wehrberger K, Grigorescu D, Svoboda J, Semal P, Caramelli D, Bocherens H, Harvati K, Conard NJ, Haak W, Powell A, Krause J (2016) Pleistocene mitochondrial genomes suggest a single major dispersal of non-Africans and a late glacial population turnover in Europe. Curr Biol 26:827–833. doi:10.1016/j.cub.2016.01.037

Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, Ayub Q, McCarthy SA, Narechania A, Kashin S, Chen Y, Banerjee R, Rodriguez-Flores JL, Cerezo M, Shao H, Gymrek M, Malhotra A, Louzada S, Desalle R, Ritchie GR, Cerveira E, Fitzgerald TW, Garrison E, Marcketta A, Mittelman D, Romanovitch M, Zhang C, Zheng-Bradley X, Abecasis GR, McCarroll SA, Flicek P, Underhill PA, Coin L, Zerbino DR, Yang F, Lee C, Clarke L, Auton A, Erlich Y, Handsaker RE, Genomes Project C, Bustamante CD, Tyler-Smith C (2016) Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. Nat Genet 48:593–9. doi:10.1038/ng.3559

Poznik GD, Henn BM, Yee MC, Sliwerska E, Euskirchen GM, Lin AA, Snyder M, Quintana-Murci L, Kidd JM, Underhill PA, Bustamante CD (2013) Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. Science 341:562–565. doi:10.1126/science.1237619

Raghavan M, DeGiorgio M, Albrechtsen A, Moltke I, Skoglund P, Korneliussen TS, Gronnow B, Appelt M, Gullov HC, Friesen TM, Fitzhugh W, Malmstrom H, Rasmussen S, Olsen J, Melchior L, Fuller BT, Fahrni SM, Stafford T Jr, Grimes V, Renouf MA, Cybulski J, Lynnerup N, Lahr MM, Britton K, Knecht R, Arneborg J, Metspalu M, Cornejo OE, Malaspinas AS, Wang Y, Rasmussen M, Raghavan V, Hansen TV, Khusnutdinova E,

Pierre T, Dneprovsky K, Andreasen C, Lange H, Hayes MG, Coltrain J, Spitsyn VA, Gotherstrom A, Orlando L, Kivisild T, Villems R, Crawford MH, Nielsen FC, Dissing J, Heinemeier J, Meldgaard M, Bustamante C, O'Rourke DH, Jakobsson M, Gilbert MT, Nielsen R, Willerslev E (2014a) The genetic prehistory of the New World Arctic. Science 345:1255832. doi:10.1126/science.1255832

Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford TW Jr, Orlando L, Metspalu E, Karmin M, Tambets K, Rootsi S, Magi R, Campos PF, Balanovska E, Balanovsky O, Khusnutdinova E, Litvinov S, Osipova LP, Fedorova SA, Voevoda MI, DeGiorgio M, Sicheritz-Ponten T, Brunak S, Demeshchenko S, Kivisild T, Villems R, Nielsen R, Jakobsson M, Willerslev E (2014b) Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature 505:87–91. doi:10.1038/nature12736

Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R, Bertalan M, Nielsen K, Gilbert MT, Wang Y, Raghavan M, Campos PF, Kamp HM, Wilson AS, Gledhill A, Tridico S, Bunce M, Lorenzen ED, Binladen J, Guo X, Zhao J, Zhang X, Zhang H, Li Z, Chen M, Orlando L, Kristiansen K, Bak M, Tommerup N, Bendixen C, Pierre TL, Gronnow B, Meldgaard M, Andreasen C, Fedorova SA, Osipova LP, Higham TF, Ramsey CB, Hansen TV, Nielsen FC, Crawford MH, Brunak S, Sicheritz-Ponten T, Villems R, Nielsen R, Krogh A, Wang J, Willerslev E (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463:757–762. doi:10.1038/nature08835

Rasmussen M, Anzick SL, Waters MR, Skoglund P, DeGiorgio M, Stafford TW Jr, Rasmussen S, Moltke I, Albrechtsen A, Doyle SM, Poznik GD, Gudmundsdottir V, Yadav R, Malaspinas AS, White SSt, Allentoft ME, Cornejo OE, Tambets K, Eriksson A, Heintzman PD, Karmin M, Korneliussen TS, Meltzer DJ, Pierre TL, Stenderup J, Saag L, Warmuth VM, Lopes MC, Malhi RS, Brunak S, Sicheritz-Ponten T, Barnes I, Collins M, Orlando L, Balloux F, Manica A, Gupta R, Metspalu M, Bustamante CD, Jakobsson M, Nielsen R, Willerslev E (2014) The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature 506:225–9. doi:10.1038/nature13025

Rasmussen M, Sikora M, Albrechtsen A, Korneliussen TS, Moreno-Mayar JV, Poznik GD, Zollikofer CP, Ponce de Leon MS, Allentoft ME, Moltke I, Jonsson H, Valdiosera C, Malhi RS, Orlando L, Bustamante CD, Stafford TW Jr, Meltzer DJ, Nielsen R, Willerslev E (2015) The ancestry and affiliations of Kennewick Man. Nature 523:455–458. doi:10.1038/nature14625

Richards M, Corte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, Hedges R, Bandelt HJ, Sykes B (1996) Paleolithic and neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203

Roewer L, Nothnagel M, Gusmao L, Gomes V, Gonzalez M, Corach D, Sala A, Alechine E, Palha T, Santos N, Ribeiro-Dos-Santos A, Geppert M, Willuweit S, Nagy M, Zweynert S, Baeta M, Nunez C, Martinez-Jarreta B, Gonzalez-Andrade F, Fagundes de Carvalho E, da Silva DA, Builes JJ, Turbon D, Lopez Parra AM, Arroyo-Pardo E, Toscanini U, Borjas L, Barletta C, Ewart E, Santos S, Krawczak M (2013) Continent-wide decoupling of Y-chromosomal genetic variation from language and geography in native South Americans. PLoS Genet 9:e1003460. doi:10.1371/journal.pgen.1003460

Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barac L, Pericic M, Balanovsky O, Pshenichnov A, Dion D, Grobei M, Zhivotovsky LA, Battaglia V, Achilli A, Al-Zahery N, Parik J, King R, Cinnioglu C, Khusnutdinova E, Rudan P, Balanovska E, Scheffrahn W, Simonescu M, Brehm A, Goncalves R, Rosa A, Moisan JP, Chaventre A, Ferak V, Furedi S, Oefner PJ, Shen P, Beckman L, Mikerezi I, Terzic R, Primorac

D, Cambon-Thomsen A, Krumina A, Torroni A, Underhill PA, Santachiara-Benerecetti AS, Villems R, Semino O (2004) Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in europe. Am J Hum Genet 75:128–137. doi:10.1086/422196

Rootsi S, Zhivotovsky LA, Baldovic M, Kayser M, Kutuev IA, Khusainova R, Bermisheva MA, Gubina M, Fedorova SA, Ilumae AM, Khusnutdinova EK, Voevoda MI, Osipova LP, Stoneking M, Lin AA, Ferak V, Parik J, Kivisild T, Underhill PA, Villems R (2007) A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. Eur J Hum Genet 15:204–211. doi:10.1038/sj.ejhg.5201748

Rootsi S, Myres NM, Lin AA, Jarve 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, Hooshiar Kashani B, Houshmand M, Sanati MH, Toncheva D, Lisa A, Semino O, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA (2012) Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. Eur J Hum Genet 20:1275–1282. doi:10.1038/ejhg.2012.86

Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet 67:1526–1543. doi:10.1086/316890

Schiffels S, Haak W, Paajanen P, Llamas B, Popescu E, Loe L, Clarke R, Lyons A, Mortimer R, Sayer D, Tyler-Smith C, Cooper A, Durbin R (2016) Iron Age and Anglo-Saxon genomes from East England reveal British migration history. Nat Commun 7:10408. doi:10.1038/ncomms10408

Scozzari R, Massaia A, D'Atanasio E, Myres NM, Perego UA, Trombetta B, Cruciani F (2012) Molecular dissection of the basal clades in the human Y chromosome phylogenetic tree. PLoS One 7:e49170. doi:10.1371/journal.pone.0049170

Scozzari R, Massaia A, Trombetta B, Bellusci G, Myres NM, Novelletto A, Cruciani F (2014) An unbiased resource of novel SNP markers provides a new chronology for the human Y chromosome and reveals a deep phylogenetic structure in Africa. Genome Res 24:535–544. doi:10.1101/gr.160788.113

Seguin-Orlando A, Korneliussen TS, Sikora M, Malaspinas AS, Manica A, Moltke I, Albrechtsen A, Ko A, Margaryan A, Moiseyev V, Goebel T, Westaway M, Lambert D, Khartanovich V, Wall JD, Nigst PR, Foley RA, Lahr MM, Nielsen R, Orlando L, Willerslev E (2014) Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. Science 346:1113–1118. doi:10.1126/science.aaa0114

Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA (2000) The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science 290:1155–1159

Sikora M, Carpenter ML, Moreno-Estrada A, Henn BM, Underhill PA, Sanchez-Quinto F, Zara I, Pitzalis M, Sidore C, Busonero F, Maschio A, Angius A, Jones C, Mendoza-Revilla J, Nekhrizov G, Dimitrova D, Theodossiev N, Harkins TT, Keller A, Maixner F, Zink A, Abecasis G, Sanna S, Cucca F, Bustamante CD (2014) Population genomic analysis of ancient and modern genomes yields new insights into the genetic ancestry of the Tyrolean Iceman and the genetic structure of Europe. PLoS Genet 10:e1004353. doi:10.1371/journal.pgen.1004353

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, Page DC (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837. doi:10.1038/nature01722

Skoglund P, Malmstrom H, Raghavan M, Stora J, Hall P, Willerslev E, Gilbert MT, Gotherstrom A, Jakobsson M (2012) Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science 336:466–469. doi:10.1126/science.1216304

Skoglund P, Posth C, Sirak K, Spriggs M, Valentin F, Bedford S, Clark GR, Reepmeyer C, Petchey F, Fernandes D, Fu Q, Harney E, Lipson M, Mallick S, Novak M, Rohland N, Stewardson K, Abdullah S, Cox MP, Friedlaender FR, Friedlaender JS, Kivisild T, Koki G, Kusuma P, Merriwether DA, Ricaut FX, Wee JT, Patterson N, Krause J, Pinhasi R, Reich D (2016) Genomic insights into the peopling of the Southwest Pacific. Nature 538:510–513. doi:10.1038/nature19844

Stringer C (2002) Modern human origins: progress and prospects. Philos Trans R Soc Lond B Biol Sci 357:563–579. doi:10.1098/rstb.2001.1057

Szecsenyi-Nagy A, Brandt G, Haak W, Keerl V, Jakucs J, Moller-Rieker S, Kohler K, Mende BG, Oross K, Marton T, Osztas A, Kiss V, Fecher M, Palfi G, Molnar E, Sebok K, Czene A, Paluch T, Slaus M, Novak M, Pecina-Slaus N, Osz B, Voicsek V, Somogyi K, Toth G, Kromer B, Banffy E, Alt KW (2015) Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. Proc Biol Sci. doi:10.1098/rspb.2015.0339

Thomas MG, Kivisild T, Chikhi L, Burger J (2013) Europe and western Asia: genetics and population history. In: Ness I (ed) The encyclopedia of global human migration. Blackwell Publishing Ltd, pp 1–11

Trombetta B, D'Atanasio E, Massaia A, Ippoliti M, Coppa A, Candilio F, Coia V, Russo G, Dugoujon JM, Moral P, Akar N, Sellitto D, Valesini G, Novelletto A, Scozzari R, Cruciani F (2015) Phylogeographic refinement and large scale genotyping of human Y chromosome haplogroup E provide new insights into the dispersal of early pastoralists in the African continent. Genome Biol Evol 7:1940–1950. doi:10.1093/gbe/evv118

van Oven M, Van Geystelen A, Kayser M, Decorte R, Larmuseau MH (2014) Seeing the wood for the trees: a minimal reference phylogeny for the human Y chromosome. Hum Mutat 35:187–191. doi:10.1002/humu.22468

Wei W, Ayub Q, Chen Y, McCarthy S, Hou Y, Carbone I, Xue Y, Tyler-Smith C (2013) A calibrated human Y-chromosomal phylogeny based on resequencing. Genome Res 23:388–395. doi:10.1101/gr.143198.112

Wilson A, Cann R, Carr S, George M, Gyllensten U, Helm-Bychowski K, Higuchi RG, Palumbi S, Prager E, Sage R, Stoneking M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. Biol J Linn Soc 26:375–400

Yunusbayev B, Metspalu M, Jarve 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. doi:10.1093/molbev/msr221

Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF (2004) High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. Mol Biol Evol 21:164–175. doi:10.1093/molbev/msh009