Human Relationships Inferred from Genetic Variation

Oscar Lao, *Erasmus University Medical Center Rotterdam, The Netherlands*

Manfred Kayser, *Erasmus University Medical Center Rotterdam, The Netherlands*

Advanced article

Article Contents

Online posting date: 15th September 2009

We humans are a diverse species, both at the phenotypic and the genetic levels. The genetic diversity of current human populations has been shaped by various demographic and evolutionary as well as some cultural factors. Thus, studying how genetic variation is distributed through individuals around the world can provide insights into (i) when and where our human ancestors first appeared on the planet, (ii) the dynamics of admixture with other *Homo* species not existing anymore, (iii) migration waves that brought humans across the world and (iv) processes of adaptations towards environmental and other factors that shaped human genomes and phenotypic traits. Additionally, the quantification of genetic variation between human populations provides data evidence for elucidating whether humans can be classified according to genetically homogeneous groups or not, with implications for medical and forensic studies. Hence, studying human genetic diversity is important for better understanding of our past, presence and future.

Introduction

The genetic variation of extant human populations is a result of complex interactions between genetic, demographic, selective and even cultural factors in the recent history of mankind, on the basis of the genetic heritage we received from our ancestor species. In practice, this means that the amount of genetic variation is locus-dependent (Weir *et al*., 2005). On one hand, genetic signatures of events such as demographic ones that have shaped the entire genome should also be conserved at many (neutral) loci, but there are reasons why such signatures can be conserved more in some types of genetic markers than others. The human Y-chromosome and mitochondrial deoxyribonucleic acid (mtDNA) escape homologous recombination and thus conserve signatures much longer than autosomal DNA (which undergoes homologous recombination) usually does. Population histories, such as those of the Pacific Islanders, may have been sex-biased, which can only be detected by studying Y-chromosome and mtDNA (Kayser *et al*., 2006). Migration events may have been sex-dependent, for example, warrior-based elite dominance events such as the spread of Genghis Khan and his army (Zerjal *et al*., 2003), which is detectable by Y-chromosome but not by mtDNA markers. On the other, polymorphisms within functional motifs, such as genes or promoters, tend to show either smaller or larger genetic differences than neutral genomic regions, reflecting the presence of negative or positive selective pressures respectively (Barreiro *et al*., 2008), although they could also represent extreme examples of stochastic processes (Hofer *et al*., 2009). Hence, quantifying the relationships between human populations using a large number of neutral autosomal and/or informative Y/mtDNA loci can provide insights into the origins and migration history of humans (Cavalli-Sforza *et al*., 1994), whereas exploring human genetic diversity at particular loci that received selective pressure can provide a better understanding of the mechanisms how humans adapted to new environments during their history (Sabeti *et al*., 2006). Both fields of human genetic research have provided a highly valuable amount of information on the evolution of our species, with more to be expected in the future. **See also**: Genetic Variation: Polymorphisms and Mutations; Human Genetics and Languages

ELS subject area: Evolution and Diversity of Life

How to cite:
Lao, Oscar; and, Kayser, Manfred (September 2009) Human Relationships Inferred from Genetic Variation. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0021758

Genetic Relationships and Human Origins

One of the most fundamental questions in human history is about the place and time of modern human origins, whether modern humans first occurred as a single event or multiple events and whether there was admixture with other so-called archaic human species such as Neanderthals. There is a general consensus among palaeontologists

that the genus *Homo* originated in Africa in the Pliocene, but there still is controversial discussion among and between various scientific disciplines on how and where anatomically modern humans *Homo sapiens* originated. Different evolutionary models have been proposed for explaining the natural evolution of the human species (Excoffier, 2002). Among them, two extreme models have driven the scientific discussion (**Figure 1**). **See also**: *Homo erectus*; Human Evolution: Early Radiations; Human Evolution: Radiations in the Last 300 000 Years

The MultiREgional model or MRE (Wolpoff *et al*., 2000) suggests that anatomically modern humans evolved by anagenesis or gradual evolution in the different continents from the *Homo erectus* specimens that spread between 1.7 and 1.9 million years ago from Africa (see **Figure 1b**). According to this model, individuals from different continents would have shared a continuous spatial and temporal gene flow that prevented the phenomena of speciation, but allowed genetic adaptation with regard to local environmental factors. It should be noted that the MRE model is different from the *Candelabra* model of human origins, which suggests (see **Figure 1a**) the complete lack of genetic flux between the continental groups of different origins (Templeton, 2007), nowadays believed as quite unlikely a scenario. Many scientists today, including most geneticists, favour the Recent African Origin (RAO) also called Out of Africa (OAA) or Replacement model (Excoffier, 2002), which proposes (see **Figure 1c**) that anatomically modern humans would have evolved by cladogenesis from a small hominid population in Africa, and would have spread over the world approximately 100 000 years ago replacing 'archaic' humans at the different continental regions outside Africa, for example, Neanderthals in Europe. Can the genetic information present in current populations help to distinguish among them? If the MRE was correct, estimates of the time of the most recent common ancestor (TMRA) of modern humans from genetic diversity present in current human populations should reflect the first spread of *H. erectus* out of Africa. Furthermore, the place of the most recent common ancestor (PMRCA; Takahata *et al*., 2001) according to the MRE model could be in regions other than Africa (Takahata *et al*., 2001). In contrast, according to the RAO/OAA model, the genetic diversity present in populations other than Africa should be a subset of that present in the African continent, and the TMRA estimated from genetic data should reflect the first spread of *H. sapiens* out of Africa, and the PMRCA should be placed in Africa (Stoneking, 2008). Moreover, the overall human genetic diversity should be relatively small due to the recent origin. **See also**: Coalescent Theory

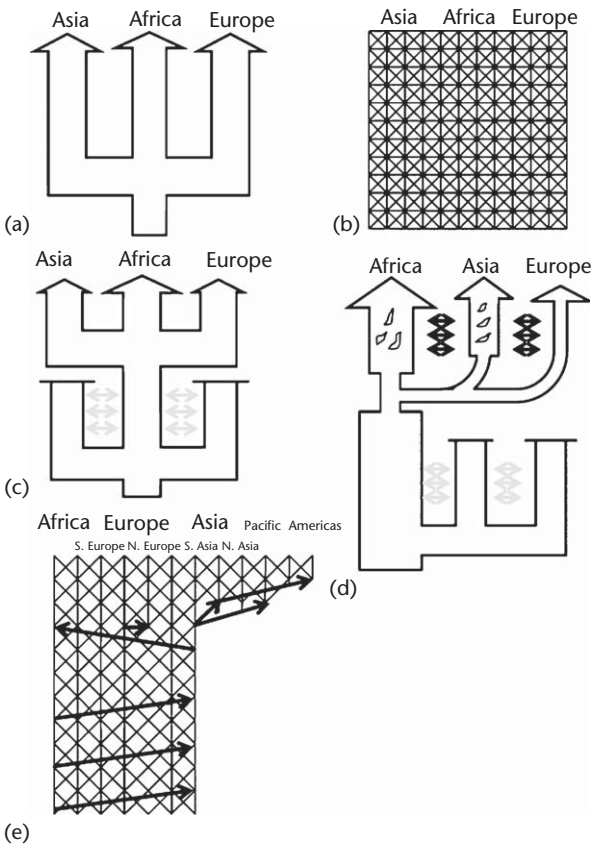

Figure 1 Popular models proposed for the origin of anatomically modern humans. (a) Candelabra model. (b) Multiregional trellis model. (c) Out of Africa model. (d) Out of Africa model with population substratification. (e) Out of Africa again and again model. Adapted from Templeton (2007) and Excoffier (2002).

Current knowledge of the human variation suggests that the true model is somewhat more complex than these simple models depict. Despite *H. sapiens* being a ubiquitous species with more than 6000 million representatives, its genetic variability is quite modest (percentage of variation between two sequences approximately 0.12) when compared with that found in other primate species such as the Orangutan (percentage of variation approximately 0.36) (Fischer *et al*., 2006). In addition, as a general rule, it has been found that African subSaharan populations tend to be the repository of genetic diversity for non-African populations both in number of polymorphic variants, ancestral alleles, allelic frequencies and the number of combinations of linked genetic variants usually referred to as haplotypes (Campbell and Tishkoff, 2008). These results have been interpreted as evidence for supporting the RAO/OOA model against the MRE model. Nevertheless, some authors have pointed out that these results could also be interpreted in favour of the MRE model, implying a historically larger effective population size in the African continent (Templeton, 2007). Estimations of TMRA and PMRCA based on single-locus analyses have revealed a complex picture of human genetic evolution (Excoffier, 2002; Harding and McVean, 2004). Although, the TMCA and PMRCA estimations of some loci such as mtDNA and Y-chromosome are compatible with the RAO/OOA model (Stoneking, 2008), some autosomal loci provide TMRCAs ranging from 41 KY to more than 1780 KY, and PMRCAs outside the African continent (Excoffier, 2002). This has been interpreted as evidence for an alternative hybrid model (see **Figure 1e**) between the pure RAO/OOA

and the MRE models (Templeton, 2007), with migratory waves into Africa as well as out of Africa, and recurrent gene flow. Others have suggested that this complex scenario suggests birth–death population processes (see **Figure 1d**) as well as a considerable amount of ancestral population substructure (Harding and McVean, 2004). Notwithstanding, a recent study (Fagundes *et al*., 2007) statistically evaluated different models of human evolution, including different variants of the RAO/OOA and MRE models, by means of analysing the genetic diversity in 50 autosomal loci of approximayely 500 bp each in current individuals from African, Asian and American geographic origin, concluding that the model most supported by these genetic data is the RAO/OOA model with exponential population growth. The time of speciation of modern humans and spread out of Africa are in agreement with estimations obtained from the oldest archaeological remains of anatomically modern humans found so far, which are placed in Africa (White *et al*., 2003). Furthermore, according to this model, no admixture with ancient populations would have taken place. However, it has been pointed out (Garrigan and Hammer, 2008) that the DNA fragments analysed so far may not be long enough to detect the signals of genetic introgression. Indeed, in particular cases the pattern of genetic diversity has been interpreted as evidence of such introgression, mainly in particular genomic regions that could have been under selective pressures. This has been explained because in the case of positive events, these genomic regions would have had fewer chances to get lost by means of genetic drift (Garrigan and Hammer, 2006; Hawks *et al*., 2008). Nevertheless, so far the largely incomplete analysis of the genome of one archaic human species, the Neanderthal (Hodgson and Disotell, 2008), has suggested that there was a genetic discontinuity between archaic and anatomically modern humans, at least in respect of Neanderthals. However, it has been pointed out that these conclusions so far are based on only one particular locus, namely mtDNA, and may be troubled by problems of contaminations with modern human DNA (Hodgson and Disotell, 2008). The draft sequence of the entire Neanderthal genome, to become available very soon (Dalton, 2009) may shed more light on this question. **See also**: Ancient DNA: Phylogenetic Applications; Genetic Diversity in Africa; Human and Chimpanzee Nucleotide Diversity; Mitochondria: Origin; Y Chromosome

Genetic Relationships and Demographic History: Europe as an Example

Reaching this point, the reader should have noticed that interpreting genetic variability in an evolutionary respect can be a quite hazardous issue and that special care should be taken when drawing conclusions. The latter point becomes even more crucial when focusing on the recent (<40 KYA) human history. It is not unusual to find in the scientific literature genetic variants with denomination of origin, particularly if it is associated to a disease, such as 'Viking', 'Celtic' or 'Phoenician'. Although these populations could indeed be the origin and distributor of the respective genetic variants, it should be noted that the recent history of human populations has been ruled by a large number of migration events at different periods of time. Consequently, it is likely to find historical population movements that match (just by chance) the geographic pattern of a genetic variant (Sokal *et al*., 1996). Moreover, it can be expected that only migration events involving a large number of people would lead to a quantifiable fingerprint in the genome. However, in particular cases such signatures could be recovered with reasonable certainty especially for relatively recent events (Sokal *et al*., 1996). To highlight some of the problems in reliably inferring human migration history from genetic diversity data, we will consider the recent migration history of Europe as an example. Noteworthy, similar discourses could be written for all major (as well as many minor) geographic regions worldwide based on genetic knowledge that has accumulated over the last two decades or so. **See also**: Genetics and the Origins of the Chinese; Genetics and the Origins of the Polynesians; Migration; Origins of the Austro-Asiatic Populations; The Peopling of the Americas as Revealed by Molecular Genetic Studies

As previously seen with the origins of humankind, there are different models explaining the origin of the European population (**Figure 2**). All of them assume that interbreeding with Neanderthals is not relevant for the interpretation of contemporary European genetic diversity (Barbujani and Goldstein, 2004).

The Palaeolithic model proposes that the genetic variation present in current Europeans comes from that present in the first anatomically modern human settlers that came out of the African continent approximately

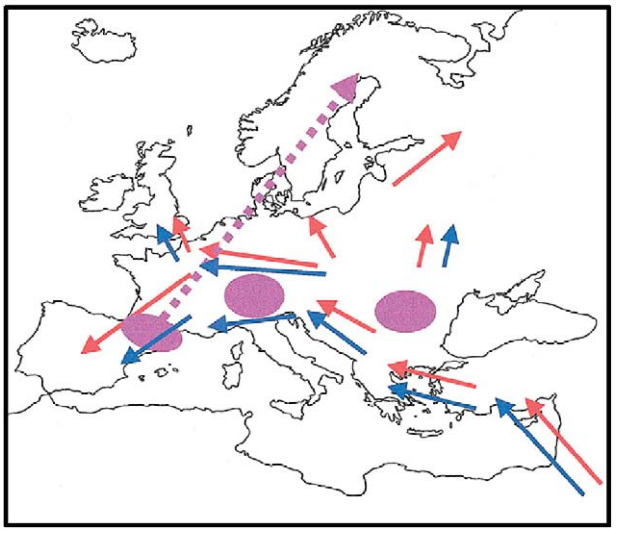

Figure 2 Scheme of the main demographic processes documented in the archaeological record of Europe. Reprinted from Simoni *et al.*, 2000. Copyright 2000, with permission from Elsevier.

40 KYA (Barbujani and Goldstein, 2004). According to this model, the cultural/technological advances such as animal domestication and farming developed in the Fertile Crescent during the Neolithic approximately 10 KYA would have been culturally spread through the populations of the initial hunter–gatherers without additional population influx (Barbujani and Goldstein, 2004). In contrast, the Neolithic model, also known as the demic diffusion hypothesis (Barbujani and Goldstein, 2004), suggests that current European populations are direct descendents from the farming populations of the Fertile Crescent, which would have colonized the European continent starting from the Levant and replaced the first *H. sapiens* settlers without admixing with them. The third model, called the Postglacial Expansion or Mesolithic model (Barbujani and Goldstein, 2004), proposes an additional large human population migration during the last maximum glacial approximately 18 KYA to different glacial refuges in the south of Europe and a subsequent population re-expansion covering large territories of central and northern Europe after the ice melted. Perhaps, a more realistic model would be a compromise between the three models, considering that Neolithic farmers did spread, taking with them culture and technology, and that genetic admixture with initial European hunter–gatherer populations, and/or postglacial refugees who repopulated Europe after the last glacial maximum, did happen (Barbujani and Goldstein, 2004).

The amount of divergence in the genetic variation of the European population expected according to the three models should be somewhat different (Barbujani and Goldstein, 2004), which allows the use of genetic data as a tool to test these hypotheses. Initial studies based on a statistical technique known as principal component (PC) analysis and synthetic maps (Cavalli-Sforza *et al*., 1994) applied to the variation observed at classical autosomal genetic markers (such as blood groups and other plasma proteins as well as variation at the human histocompatibility loci), in the European continent showed clinal patterns and gradients from SE to NW in the first PC, which was interpreted as supporting the Neolithic model (Cavalli-Sforza *et al*., 1994). Despite these analyses attracting some criticism because of both the massive use of data interpolation in the maps (Sokal *et al*., 1999) and the interpretation of PC (Novembre and Stephens, 2008), further analyses based on similar types of markers but different statistical approaches (Sokal *et al*., 1991) also supported the presence of soft gradients in the European genetic variation in the same direction as the migratory route predicted by the Neolithic model. Recent studies performed with approximately 500 000 genome-wide autosomal single nucleotide polymorphisms (SNPs) in thousands of individuals sampled across Europe have shown a similar geographic pattern as previously observed with classical markers (Lao *et al*., 2008; Novembre *et al*., 2008). Furthermore, higher amounts of genetic diversity were observed in Southern European populations compared with Northern populations and the distribution of genetic diversity followed a clinal pattern roughly from the south to the north (Lao *et al*., 2008). This most recent genetic evidence is not in disagreement with any of the three models described earlier, as they all assume major migration waves into Europe from the south (perhaps least so for the Mesolithic model as glacial refugee areas were in Iberia, the Alps and the Balkans, but not further south). Conclusions obtained from analysing the genetic diversity of mtDNA and Y-chromosome in the European population differ depending on the type of statistical approach applied. Spatial distributions of particular mtDNA haplogroups have been interpreted as a remnant of the Mesolithic expansion (Torroni *et al*., 2001) and Palaeolithic ancestry (Richards *et al*., 2002). Others have concluded, based on the overall spatial distribution of mtDNA haplogroups, that the Neolithic hypothesis is the best supported (Simoni *et al*., 2000). In a similar way, Y-chromosome data interpretation has been controversial. Estimations of the Neolithic component in European populations have ranged from 22% as estimated by some authors (Semino *et al*., 2000) to up to 80% by others (Chikhi *et al*., 2002) even based on the same Y-chromosome data. Interpretation of the spatial patterns and biological significance of different Y-chromosome haplogroups appears more complex than those for mtDNA and autosomal markers. This may not be unexpected, as the human Y-chromosome is highly sensitive to demographic effects given its small effective population size but also because of cultural effects, such as patrilocal residence pattern, that most likely played a role also in the European history (Seielstad *et al*., 1998). Thus, Y-chromosomal data may indicate more than the major migration waves detectable by means of autosomal DNA markers, as it has been suggested previously (Rosser *et al*., 2000; Semino *et al*., 2000).
See also: Single Nucleotide Polymorphism (SNP)

Genetic Structure of Human Populations and Implications

The existence of groups of genetically homogeneous individuals has been traditionally assumed by classical anthropologists up to the twentieth century, despite the fact it was never properly tested by scientific means (Barbujani *et al*., 1997). In part, this has led to the assumption of the existence of distinct human groups, formerly called human races, including the terms' unfortunate misuse in the recent history (Kittles and Weiss, 2003). Classically, human externally visible traits, for example, pigmentation traits and facial characteristics were used for 'racial' inferences, ignoring that genetic variation at these traits only represents small parts of human genomic variation. Obviously, the most appropriate way of testing whether there indeed are biologically distinct human groups is by directly analysing human genomic variation, which will be possible in the near future initiated by the 1000 Genomes Project currently underway (http://www.1000genomes.org/page.php). So far, studies performed at various loci have shown that the proportion of genetic variation obtained when individuals were

clustered according to their geographic continent of origin is quite small (ranging from only 5% up to 15%) compared to that seen when all humans were considered as a single group (approximately 80%) (Romualdi *et al*., 2002). For comparison: a biological criterion (despite subjective) to define the presence of subspecies is finding estimations of genetic differentiation greater than approximately 25% (Kittles and Weiss, 2003). The largest analysis of DNA polymorphisms in respect to genomic coverage available so far was performed in four populations of different geographic origin (Yoruba from Africa, Chinese, Japanese and individuals of North/Northwest European ancestry living in Utah (Lao *et al*., 2008)) in the International HapMap Project comprising more than 3 million SNPs (International HapMap Project, 2005). This study has shown a mean degree of differentiation between three continental populations (Japanese and Chinese were combined) of approximately 15% (Weir *et al*., 2005). **See also**: HapMap Project; Population Differentiation: Measures

A slightly different question is whether the relatively small amount of genetic differentiation usually observed would be sufficient to correctly classify human individuals according to genetically homogeneous groups. The answer strongly depends on the underlying assumptions of the clustering algorithms applied, as well as on the genetic loci considered. HapMap populations can be genetically clustered according to the continent of origin but an additional population substructure can be observed within continents depending on the algorithm (Paschou *et al*., 2007). In a similar way, results of genetic clustering of individuals of the Human Genome Diversity Panel (HGPD-CEPH) representing a set of 53 populations is highly dependent on the method and the markers used (e.g. compare the results obtained by Rosenberg *et al*. (2002) with those obtained by (Corander *et al*., 2004)). Data from more than 650 000 SNPs in these individuals also attained different degrees of population substructure, depending on the statistical tool that is applied (Jakobsson *et al*., 2008; Li *et al*., 2008). However, worldwide sampling for genetic studies such as in HGDP is highly limited and it remains to be seen whether population differentiation as seen in currently available data holds when more densely collected samples will be analysed. At a regional geographic level, for example, in Europe, and with more densely collected samples, the presence of clinal patterns becomes more evident (Lao *et al*., 2008; Novembre *et al*., 2008; Dalton, 2009; **Figure 3**) and genetic discontinuities are restricted to particular populations genetically known as outliers (i.e. Finns). **See also**: Human Genome Diversity Project (HGDP)

For some investigators all these results prove that humans should be considered as one metapopulation and that only interindividual variation is important (Ng *et al*., 2008). Others have argued that even if it is small, the interpopulation variation is large enough to ascertain the geographic origin of an individual and could be of medical importance (Edwards, 2003). Analyses of common genetic variants associated to complex diseases that have been robustly replicated tend to show lower amounts of continental diversity compared with other genetic variants (Lohmueller *et al*., 2006) supporting the hypothesis of a lack of population substructure in relevant medical variants. In contrast, common variants associated to eye (iris) pigmentation in European populations showed particularly strong differences between regional groups (Kayser *et al*., 2008). The main difference between these examples is the phenotype considered. Pigmentation is known to be one of the most likely traits to be shaped by selective pressures, due to environmental adaptation and/or mate choice preferences, and therefore shows a strong continental distribution (Parra, 2007) whereas genetic variants associated to common diseases that appear after the reproductive age (such as coronary heart diseases) could be shaped by neutral evolution (Reich and Lander, 2001). Again, this demonstrates that the importance of population substructure is highly dependent on the locus considered and its particular evolutionary history, and the practical relevance of genetic substructure depends on whether such particular genetic loci are indeed applied. Nevertheless, overall there is no reason to believe that all the selective pressures are going to follow the same geographic patterns, and so the necessity to define fixed phenotypic groups of individuals. Genetic differences between human populations are also important in the forensic context and need to be controlled in estimating matching probabilities of DNA profiles from crime scenes with those of suspects. Additionally, ancestry-sensitive genetic markers may also become applied directly in forensics to predict the geographic region of genetic origin of an unknown person, which in principle should provide extra information to better find unknown persons (Lao *et al*., 2006). **See also**: Evolution of skin pigmentation differences in humans; Mutations in Human Genetic Disease

In summary, to us it appears as unfortunate to ignore existing genetic differences between individuals from different geographic regions as it is to highlight them in extrapolating assumptions about the entire genome. Genetic differences between human individuals have to be seen in an overall quantitative respect, and, if done so, it becomes evident that they are on average very small. On one hand, there appears to be very limited evidence to support a human grouping according to the geographic place of origin on the overall genetic level in agreement with a single common origin of all modern humans and a recent spread around the world as assumed by the OOA model. On the other, particular regions in our genome have been shaped by effects such as local positive selection starting after humans had occupied different worldwide regions, or migration history and genetic drift such as the human Y-chromosome, which cause genetic differences at these particular loci between extant human populations of different regions. When the use of these genomic regions is implemented in medical and forensic applications, their amount of population substructure needs to be considered carefully. Thus, as with many aspects of science (and of life in general) there is no simple yes or no when answering the question of human genetic differentiation, as things are more complex.

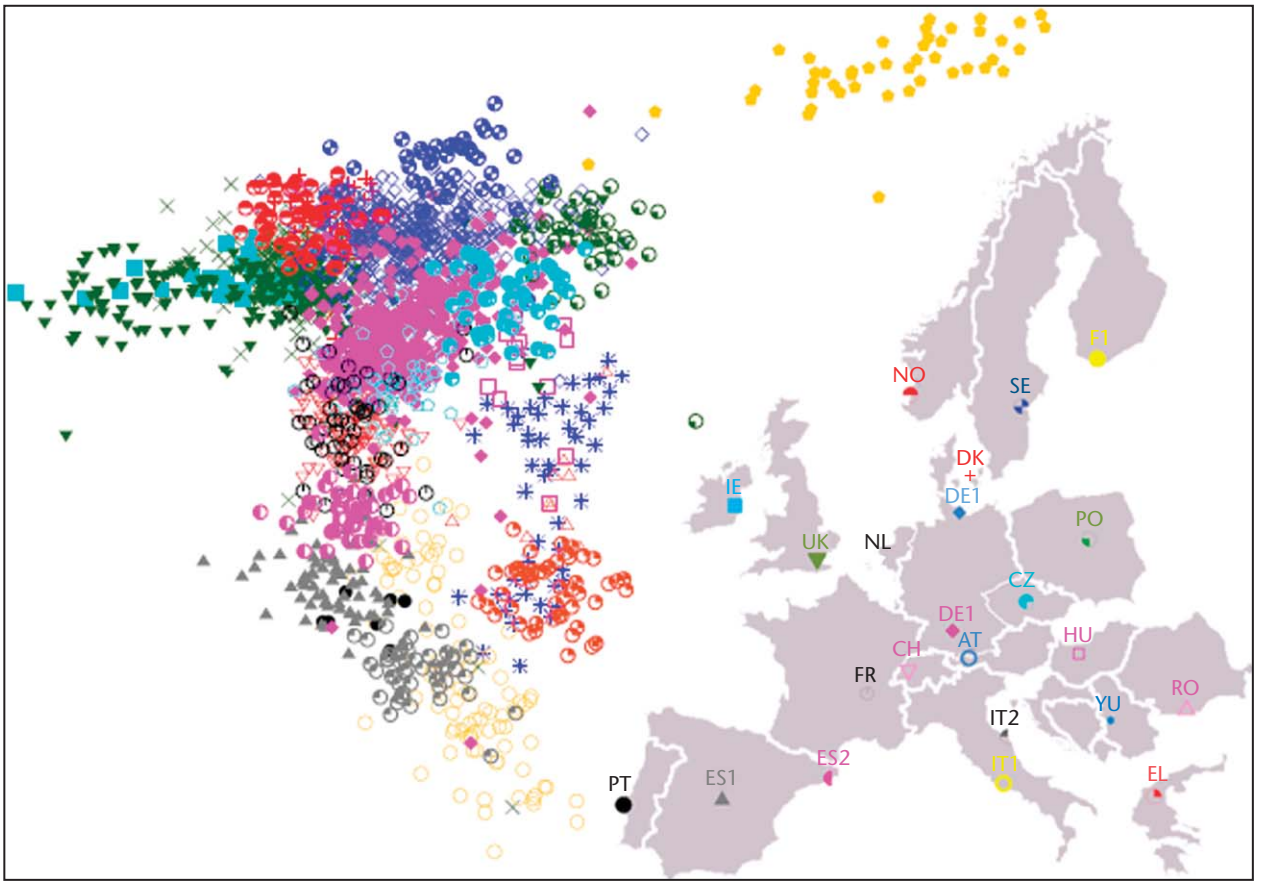

Figure 3 SNP-based principal component analysis (PCA) of 23 European subpopulations using 309 790 SNPs from The GeneChip® Human Mapping 500 K Array Set (Affymetrix) that passed quality control in 2457 European individuals. Each individual is a dot which is placed in the two genetic dimensions defined by the PCA. Individuals genetically close related will be placed closely. Geographic origin of the sampling is also provided. Adapted from Lao *et al.* (2008) with agreement from Current Biology/Cell Press.

Acknowledgements

Original work by the authors on genetic relationships and migration history of human populations including their forensic relevance is supported by the Erasmus University Medical Center Rotterdam, the Netherlands Forensic Institute (NFI), the Netherlands Genomics Initiative (NGI)/ Netherlands Organization for Scientific Research (NWO) within the framework of the Forensic Genomics Consortium Netherlands (FGCN) and the Deutsche Forschungsgemeinschaft (DFG) within the Sonderforschungsbereich (SFB) 680 Molecular Basis of Evolutionary Innovations at the University of Cologne.

References

Barbujani G and Goldstein DB (2004) Africans and Asians abroad: genetic diversity in Europe. *Annual Review of Genomics and Human Genetics* **5**: 119–150.

Barbujani G, Magagni A, Minch E and Cavalli-Sforza LL (1997) An apportionment of human DNA diversity. *Proceedings of the National Academy of Sciences of the USA* **94**: 4516–4519.

Barreiro LB, Laval G, Quach H, Patin E and Quintana-Murci L (2008) Natural selection has driven population differentiation in modern humans. *Nature Genetics* **40**: 340–345.

Campbell MC and Tishkoff SA (2008) African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annual Review of Genomics and Human Genetics* **9**: 403–433.

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

Chikhi L, Nichols RA, Barbujani G and Beaumont MAV (2002) Y genetic data support the Neolithic demic diffusion model. *Proceedings of the National Academy of Sciences of the USA* **99**: 11008–11013.

Corander J, Waldmann P, Marttinen P and Sillanpaa MJ (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* **20**: 2363–2369.

Dalton R (2009) Neanderthal genome to be unveiled. *Nature* **457**: 645.

Edwards AWV (2003) Human genetic diversity: Lewontin's fallacy. *BioEssays* **25**: 798–801.

Excoffier LV (2002) Human demographic history: refining the recent African origin model. *Current Opinion in Genetics & Development* **12**: 675–682.

Fagundes NJ, Ray N, Beaumont M *et al*. (2007) Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences of the USA* **104**: 17614–17619.

Fischer A, Pollack J, Thalmann O, Nickel B and Paabo S (2006) Demographic history and genetic differentiation in apes. *Current Biology* **16**: 1133–1138.

Garrigan D and Hammer MF (2008) Ancient lineages in the genome: a response to Fagundes *et al*. *Proceedings of the National Academy of Sciences of the USA* **105**: E3; author reply E4.

Garrigan D and Hammer MF (2006) Reconstructing human origins in the genomic era. *Nature Reviews. Genetics* **7**: 669–680.

Harding RM and McVean G (2004) A structured ancestral population for the evolution of modern humans. *Current Opinion in Genetics & Development* **14**: 667–674.

Hawks J, Cochran G, Harpending HC and Lahn BT (2008) A genetic legacy from archaic *Homo*. *Trends in Genetics* **24**: 19–23.

Hodgson JA and Disotell TR (2008) No evidence of a Neanderthal contribution to modern human diversity. *Genome Biology* **9**: 206.

Hofer T, Ray N, Wegmann D and Excoffier L (2009) Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Annals of Human Genetics* **73**: 95–108.

International HapMap Project (2005) A haplotype map of the human genome. *Nature* **437**: 1299–1320.

Jakobsson M, Scholz SW, Scheet P *et al*. (2008) Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* **451**: 998–1003.

Kayser M, Brauer S, Cordaux R *et al*. (2006) Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. *Molecular Biology and Evolution* **23**: 2234–2244.

Kayser M, Lao O, Saar K *et al*. (2008) Genome-wide analysis indicates more Asian than Melanesian ancestry of Polynesians. *American Journal of Human Genetics* **82**: 194–198.

Kittles RA and Weiss KMV (2003) Race, ancestry, and genes: implications for defining disease risk. *Annual Review of Genomics and Human Genetics* **4**: 33–67.

Lao O, van Duijn K, Kersbergen P, de Knijff P and Kayser M (2006) Proportioning whole-genome single-nucleotide-polymorphism diversity for the identification of geographic population structure and genetic ancestry. *American Journal of Human Genetics* **78**: 680–690.

Lao O, Lu TT, Nothnagel M *et al*. (2008) Correlation between genetic and geographic structure in Europe. *Current Biology* **18**: 1241–1248.

Li JZ, Absher DM, Tang H *et al*. (2008) Worldwide human relationships inferred from genome-wide patterns of variation. *Science* **319**: 1100–1104.

Lohmueller K, Mauney MM, Reich D and Braverman JM (2006) Variants associated with common disease are not unusually differentiated in frequency across populations. *American Journal of Human Genetics* **78**: 130–136.

Ng PC, Zhao Q, Levy S, Strausberg RL and Venter JC (2008) Individual genomes instead of race for personalized medicine. *Clinical Pharmacology and Therapeutics* **84**: 306–309.

Novembre J, Johnson T, Bryc K *et al*. (2008) Genes mirror geography within Europe. *Nature* **456**: 98–101.

Novembre J and Stephens M (2008) Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics* **40**: 646–649.

Parra EJ (2007) Human pigmentation variation: evolution, genetic basis, and implications for public health. *American Journal of Physical Anthropology. Supplement* **45**: 85–105.

Paschou P, Ziv E, Burchard EG *et al*. (2007) PCA-correlated SNPs for structure identification in worldwide human populations. *PLoS Genetics* **3**: 1672–1686.

Reich DE and Lander ESV (2001) On the allelic spectrum of human disease. *Trends in Genetics* **17**: 502–510.

Richards M, Macaulay V, Torroni A and Bandelt HJ (2002) In search of geographical patterns in European mitochond rial DNA. *American Journal of Human Genetics* **71**: 1168–1174.

Romualdi C, Balding D, Nasidze IS *et al*. (2002) Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms. *Genome Research* **12**: 602–612.

Rosenberg NA, Pritchard JK, Weber JL *et al*. (2002) Genetic structure of human populations. *Science* **298**: 2381–2385.

Rosser ZH, Zerjal T, Hurles ME *et al*. (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *American Journal of Human Genetics* **67**: 1526–1543.

Sabeti PC, Schaffner SF, Fry B *et al*. (2006) Positive natural selection in the human lineage. *Science* **312**: 1614–1620.

Seielstad MT, Minch E and Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. *Nature Genetics* **20**: 278–280.

Semino O, Passarino G, Oefner PJ *et al*. (2000) The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* **290**: 1155–1159.

Simoni L, Calafell F, Pettener D, Bertranpetit J and Barbujani GV (2000) Geographic patterns of mtDNA diversity in Europe. *American Journal of Human Genetics* **66**: 262–278.

Sokal RR, Oden NL and Thomson BAV (1999) A problem with synthetic maps. *Human Biology* **71**: 1–13; discussion 15–25.

Sokal RR, Oden NL, Walker J, Di Giovanni D and Thomson BA (1996) Historical population movements in Europe influence genetic relationships in modern samples. *Human Biology* **68**: 873–898.

Sokal RR, Oden NL and Wilson C (1991) Genetic evidence for the spread of agriculture in Europe by demic diffusion. *Nature* **351**: 143–145.

Stoneking M (2008) Human origins. The molecular perspective. *EMBO Reports* **9**(suppl. 1): S46–S50.

Takahata N, Lee SH and Satta Y (2001) Testing multiregionality of modern human origins. *Molecular Biology and Evolution* **18**: 172–183.

Templeton AR (2007) Genetics and recent human evolution. *Evolution* **61**: 1507–1519.

Torroni A, Bandelt HJ, Macaulay V *et al*. (2001) A signal, from human mtDNA, of postglacial recolonization in Europe. *American Journal of Human Genetics* **69**: 844–852.

Weir BS, Cardon LR, Anderson AD, Nielsen DM and Hill WG (2005) Measures of human population structure show heterogeneity among genomic regions. *Genome Research* **15**: 1468–1476.

White TD, Asfaw B, DeGusta D *et al*. (2003) Pleistocene Homo sapiens from Middle Awash, Ethiopia. *Nature* **423**: 742–747.

Wolpoff MH, Hawks J and Caspari R (2000) Multiregional, not multiple origins. *American Journal of Physical Anthropology* **112**: 129–136.

Zerjal T, Xue Y, Bertorelle G *et al*. (2003) The genetic legacy of the Mongols. *American Journal of Human Genetics* **72**: 717–721.

Further Reading

Crawford MH (ed.) (2006) *Anthropological Genetics: Theory, Methods and Applications*. New York: Cambridge University Press.

Hartl DL and Clark AG (1997) *Principles of population genetics*, 3rd edn. Sunderland, MA: Sinauer Associates Inc.

Jobling MA, Hurles ME and Tyler-Smith C (2003) *Human Evolutionary Genetics. Origins, Peoples & Disease*. New York: Garland Science.

Stone L, Lurquin PF, Cavalli-Sforza LL (2006) *Genes, Culture, and Human Evolution: A Synthesis*. Oxford, UK: Blackwell Publishing Ltd.