

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/226875048>

# Anthropological Genetics: Inferring the History of Our Species Through the Analysis of DNA

**Article** in *Evolution Education and Outreach* · September 2010

DOI: 10.1007/s12052-010-0262-9

---

CITATION

1

READS

88

**2 authors**, including:



**Jason Hodgson**

Imperial College London

**34** PUBLICATIONS **659** CITATIONS

[SEE PROFILE](#)

# Anthropological Genetics: Inferring the History of Our Species Through the Analysis of DNA

Jason A. Hodgson · Todd R. Disotell

Published online: 19 August 2010  
© Springer Science+Business Media, LLC 2010

**Abstract** The genetic material, deoxyribonucleic acid (DNA), contains information about the evolutionary history of life. Both the relationships amongst organisms and the times of their divergence can be inferred from DNA sequences. Anthropological geneticists use DNA sequences to infer the evolutionary history of humans and their primate relatives. We review the basic methodology used to infer these relationships. We then review the anthropological genetic evidence for modern human origins. We conclude that modern humans evolved recently in Africa and then left to colonize the rest of the world within the last 50,000 years, largely replacing the other human groups that they encountered. Modern humans likely exchanged genes with Neanderthals prior to or early during their expansion out of Africa.

**Keywords** Human origins · Anthropological genetics · Phylogeny · Molecular clock · Ancient DNA

## Introduction

Among the most spectacular scientific discoveries of the twentieth century is Watson and Crick's elucidation of the structure of the molecule responsible for inheritance—deoxyribonucleic acid (DNA) (Watson and Crick 1953). This molecule is critically important for life, providing both the instructions that the organism needs to carry out

life's functions and the ability to copy itself to future generations. The property of inheritance also makes the DNA molecule a recorder of evolutionary history. Information about the relationships amongst species or populations within species and the time of their divergence from each other can be found in the DNA. It is the job of the evolutionary geneticist to interpret this information from DNA. A subset of anthropology—anthropological genetics—uses the evolutionary geneticist's tool kit to infer human evolutionary history from our and our closest relative's DNA.

## DNA: Form and Function

The DNA molecule is one of the most elegant structures found in nature. Its two major functions, transmission of information between generations and providing the information to build, maintain, and operate a living organism, follow directly from the molecule's structure. DNA is a long, linear, double-stranded molecule consisting of four nucleotide bases: adenine (A), cytosine (C), guanine (G), and thymine (T) joined by a sugar-phosphate backbone. The two strands of the molecule are joined together by weak chemical bonds between the nucleotide bases of the individual strands. Critically, A always binds to T and C always binds to G.

Watson and Crick (1953) concluded their famous paper revealing DNA's double helical structure by stating: "It has not escaped our notice that the specific [nucleotide base] pairing we have postulated immediately suggests a possible copying mechanism for the genetic material." What Watson and Crick realized was that the strict A–T and C–G base pairing meant that, if you split the double strand into two single strands, each individual strand serves as a template for rebuilding the original double-stranded molecule. Thus,

---

J. A. Hodgson · T. R. Disotell (✉)  
Department of Anthropology,  
Center for the Study of Human Origins, New York University,  
25 Waverly Place,  
New York, NY 10003, USA  
e-mail: todd.disotell@nyu.edu

by splitting and rebuilding you get two daughter strands both identical to the original. Watson and Crick's intuition was correct, as this is how DNA is copied in the process known as replication.

The second great question was how DNA stored the information necessary for the organism to carry out its life functions. DNA by itself cannot carry out any physical functions. To function, it provides cells the blueprints to build other molecules. These molecules, known as ribonucleic acids (RNA) and proteins, interact with each other to form the structures and chemical reactions necessary for life.

RNA is a very similar molecule to DNA except that the nucleotide thymine (T) is replaced by a nucleotide called uracil (U) and the sugar in the sugar–phosphate backbone contains an additional oxygen molecule. RNA is a very versatile molecule and is involved in a multitude of life functions, including protein synthesis, enzymatic activity, and gene regulation. The order of the nucleotide bases in the RNA molecule determines its structure and function.

Proteins are complex three-dimensional molecules comprised of amino acid subunits, of which there are 20. During protein synthesis, amino acids are joined together in a long chain. The chains then fold up to form a three-dimensional structure that is determined by the order of the amino acids. The final structure of the protein determines its function. Humans are estimated to produce 20,000–25,000 unique proteins (International Human Genome Sequencing Consortium 2004). Proteins are involved in all aspects of life.

The information required to build all of the RNA and protein molecules required for life is stored within the order of the nucleotide bases in the DNA molecule. Gene expression is the process of creating these molecules. First the DNA double-helix unwinds and, similar to DNA replication, a complementary RNA molecule is created from the DNA template, a process called transcription. Thus there is a one-to-one correspondence between the sequence of the DNA and RNA nucleotides. To produce proteins, some RNA molecules undergo an additional procedure known as translation. The RNA nucleotides are read in groups of three known as codons. There are 64 possible codons and each corresponds to one of the 20 amino acids or tells the cellular machinery to start or stop gene expression. The human genome consists of about three billion base pairs of DNA, of which less than 2% is expressed as protein or RNA genes (International Human Genome Sequencing Consortium 2004). This means the vast majority of the human genome has no currently known function.

Thus, the two necessary conditions for life—replication and information storage—follow directly from the relatively simple physical and chemical structure of the DNA. It is no

wonder that, after first working out DNA's structure, Francis Crick burst into The Eagle pub in Cambridge, England and screamed, "We've found the secret to life!" (Watson 1968).

### Evolution of DNA Sequences

At its most basic level, evolution is a change in allele frequency over time. An allele is a variant of a stretch of DNA, known as a gene, inherited from one parent. An allele's frequency in a population can change due to four forces: mutation, natural selection, random genetic drift, and gene flow. When DNA is copied from one parent strand to two daughter strands, errors sometimes occur. It is these errors in replication, or mutations, that are ultimately responsible for all of the variation that the other three forces of evolution act upon. These mutations also serve as guideposts that allow us to reconstruct evolutionary history (see below).

If a mutation occurs in the gametes, or sex cells, it can be passed on to future generations. The ultimate fate of any such mutation is either extinction, where it disappears completely from the gene pool, or fixation, where it replaces all other variants in the gene pool. Genetic drift, gene flow, and/or selection are the forces that bring about these alternative outcomes. Genetic drift and gene flow both change allele frequencies randomly. Genetic drift is random fluctuations in allele frequency between generations caused by mating in finite populations and can either increase or decrease the frequency. Allele frequencies drift up and down through the generations until the gene variant goes extinct or becomes fixed. Gene flow is a change in allele frequency when migrants are added to the population from some other population.

Natural selection, on the other hand, is deterministic. If a mutation occurs in a protein-coding or RNA gene, the change might effect the function of the gene product. The change might have no effect and could be detrimental, or, rarely, beneficial to the organism. If the change has no effect, then the mutation will evolve neutrally through genetic drift. If the change is detrimental or beneficial, it will evolve through natural selection. Beneficial mutations will increase in frequency and replace the less fit variants. Detrimental mutations will decrease in frequency and go extinct.

### Inferring Evolutionary Relationships from DNA Sequences

When we observe a difference in two DNA sequences caused by a mutation, it is called a substitution. Substitutions record the evolutionary history of a gene. A

substitution shared between individuals indicates that these individuals share common ancestry with the individual within whom the original mutation occurred. It also indicates that all those sharing the substitution are likely to be more closely related to each other than they are to those without it at this particular gene. That is, they have shared an ancestor more recently. Thus, by tracking the pattern of shared substitutions, geneticists can reconstruct the phylogenetic relationships amongst a sample of organisms—their evolutionary tree. Below we describe one method of phylogenetic reconstruction known as cladistics. A comprehensive review of phylogenetic reconstruction is provided by Felsenstein (2004), a general review of evolutionary biology including phylogenetic reconstruction can be found in Ridley (2003), and a practical guide to phylogenetic reconstruction and commonly used software is found in Hall (2008).

Inferring evolutionary relationships through the pattern of shared substitutions is known as cladistics. However, not all shared substitutions are treated equally. In cladistics, substitutions are classified as uniquely derived (autapomorphies), shared derived (synapomorphies), and parallelly or convergently derived (homoplasies). Autapomorphies are substitutions that are unique to a single taxon (or unit of study) and are derived relative to the ancestral state of the character (i.e., has undergone some mutation). Autapomorphies are not useful for cladistic analysis because a unique character gives no information about shared ancestry. Synapomorphies are substitutions that are shared between two or more taxa and are derived relative to the ancestral state of the character. Synapomorphies are the only kind of substitution useful for cladistic analysis because they provide evidence for recent shared ancestry. Homoplasies, on the other hand, are the bane of cladistic analyses. Homoplasies are substitutions shared between two or more taxa but are independently derived. Homoplasies result from parallel or convergent evolution and result in traits that look similar or the same, not because of recent shared ancestry but because the character state evolved multiple times independently. Homoplasies make distantly related taxa appear closely related. The final class of characters is symplesiomorphies. These are shared ancestral characters and only provide evidence of ancient shared ancestry.

Cladistic analysis is difficult to operationalize because without a time machine we cannot be sure whether a character state is ancestral or derived. We also cannot be sure whether a character state has evolved once or multiple times. To deal with this, we rely on the principle of maximum parsimony. Maximum parsimony considers the evolutionary reconstruction that requires the least amount of evolutionary changes to be the most likely description of what actually happened. The evolutionary reconstruction with the fewest number of changes will have the fewest

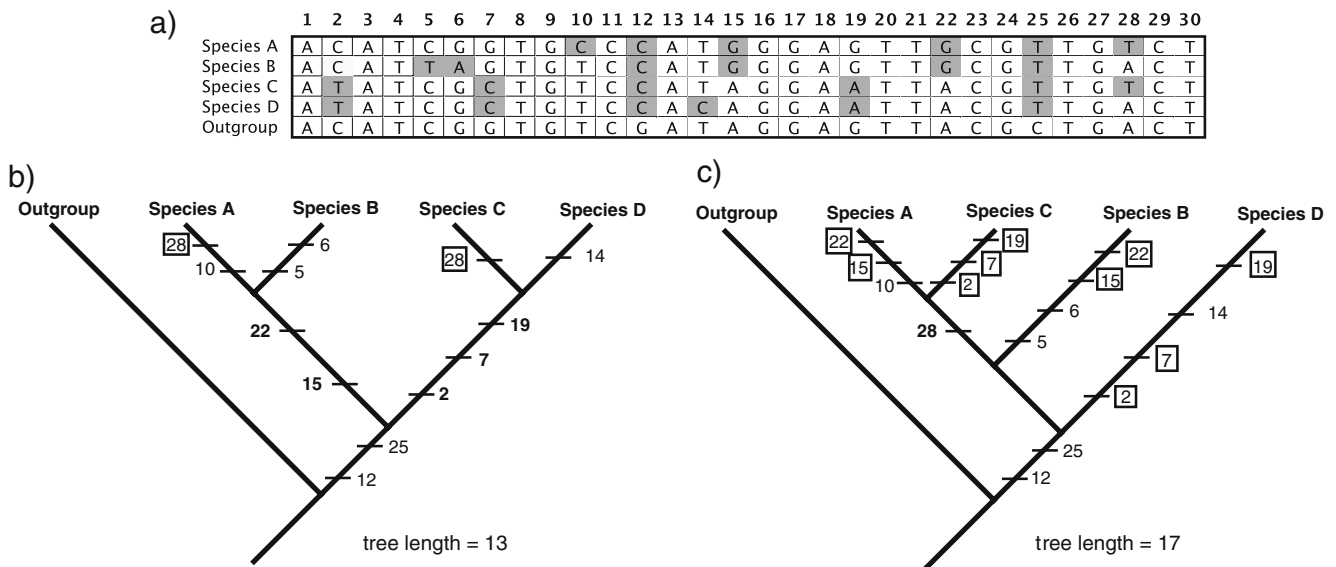
number of homoplasies and consequently the greatest number of inferred synapomorphies. Thus, to carry out a maximum parsimony analysis, we do not have to know the ancestral and derived conditions in advance.

Finding the evolutionary tree with the fewest number of changes involves evaluating each possible tree topology. The topology of a tree is merely the pattern of branching that describes the relationships between all the different lineages. Evaluating every possible topology can be a daunting task. The number of possible trees increases more than exponentially with the number of sequences used in the analysis. For example, there are three possible trees for four sequences, 15 possible trees for five sequences, and  $8.69 \times 10^{36}$  possible trees for 30 sequences! It is not possible to evaluate every possible tree in studies with many taxa, so computer algorithms have been developed to reduce the tree search space (see Felsenstein [2004] for a review).

Figure 1 demonstrates how phylogenies are reconstructed from DNA sequence data using maximum parsimony. First, DNA sequences are aligned to each other. Figure 1a shows a DNA alignment of five sequences. Alignment is amongst the most critical steps in any comparative DNA analysis (a current review of DNA alignment can be found here [Kumar and Filipinski 2007]). The purpose of alignment is to best determine sequence homology. Homologous nucleotide positions are those that are identical by descent. In other words, they were inherited from a common ancestor. It is critical that we only compare homologous nucleotide sequences. Comparing non-homologous nucleotides is apples to oranges and results in a meaningless and potentially misleading result. Alignment is a computationally difficult problem and many software applications implementing different strategies and methodologies are available.

Once aligned, maximum parsimony analysis can be performed. The analysis only considers parsimony informative columns of the DNA alignment. A parsimony informative column is one in which two or more taxa share a substitution (e.g., Fig. 1a, columns 2, 7, 15, 19, 22, and 28). The reason for this is that invariant columns require no evolutionary change and columns with an autapomorphy only ever require a single evolutionary change no matter the evolutionary reconstruction and are thus parsimony uninformative changes. Parsimony informative columns require at least one evolutionary change (in which case they are reconstructed as synapomorphies) but may require more depending upon the tree topology (in which case they are reconstructed as homoplasies).

Next tree topologies are systematically chosen. The parsimony informative changes are mapped onto the topologies one by one and the minimum number of required changes is counted. Figures 1b and 1c show two different tree arrangements with parsimony informative changes mapped



**Fig. 1** This figure illustrates inferring evolutionary history with the principle of maximum parsimony: **a)** DNA alignment (substitutions *highlighted*), **b)** the most parsimonious tree topology (inferred synapomorphies are in **bold**; homoplasies are *boxed*), **c)** less parsimonious topology

onto the trees. Topology 1b has informative characters 2, 7, 15, 19, and 22 mapped as a single evolutionary change (synapomorphies) and character 28 mapped as two evolutionary changes (homoplasy). Topology 1c has character 28 mapped as a single change (synapomorphy) and characters 2, 7, 15, 19, and 22 mapped as two changes apiece (homoplasies). The overall tree length for topology 1b is 13 evolutionary changes versus 17 evolutionary changes for 1c. Tree length is calculated as the number of parsimony informative changes (which changes depending upon the tree topology) added to the number of parsimony uninformative changes (which is constant for all topologies). All possible trees would be compared in this manner. Topology 1b is the most parsimonious evolutionary reconstruction for the DNA alignment shown in Fig. 1a.

We have chosen to include an outgroup in our DNA alignment. An outgroup is a taxon chosen because it is thought to be equally distantly related to each study taxa. For example, a chimpanzee can be used as an outgroup if we want to compare humans to each other since all living humans are obviously more closely related to each other than any is to a chimpanzee. Outgroups are useful for inferring whether substitutions are ancestral or derived. To do so, we assume that the outgroup is likely to carry the ancestral condition for any sites that vary within the study group. This assumption is not always warranted and the appropriate choice of an outgroup can be problematic.

There are several other approaches to phylogenetic inference (see Felsenstein [2004] for a comprehensive review). Phenetic or distance-based methods count the number of pairwise differences between sequences. The sequences with the fewest differences to each other are

then considered to be the closest relatives. Statistical methods use a model of DNA sequence evolution and try to find the tree amongst all the possible trees with the maximum likelihood (Felsenstein 1981) or the greatest posterior probability (Yang and Rannala 1997) given the evolutionary model. Distance-based methods enjoy the benefit of being computationally simple and therefore fast, however they are known to be less reliable because they consider all similarities as equal (e.g., synapomorphies, symplesiomorphies, and homoplasies). Statistical and cladistic phylogenetic inference methods are preferable to distance-based approaches but suffer from being computationally expensive and can take hours, days, and even weeks.

### Telling Time with DNA

Not only can we infer the branching order of divergence between taxa but we can estimate the times of their divergence from DNA sequences. This information allows us to better understand the fossil record and to make inferences about the mode and tempo of evolutionary change.

The idea of a “molecular clock” for genetic changes predates our ability to sequence DNA. The idea was first proposed for protein amino acid sequences (Zuckerandl and Pauling 1965). Zuckerandl and Pauling observed that the number of amino acid differences between species increased with decreasing relatedness. They reasoned that if the amino acid substitutions occurred regularly during the course of evolution then the number of substitutions

between two species is related to the amount of time that has passed since their divergence. They further noted that the rate of amino acid substitution can be calibrated to estimate the times of unknown divergences from divergence times better understood from a robust fossil record.

When this idea was first proposed, it was unknown whether molecular substitutions occurred with any clock-like regularity or if they tracked the irregular path of morphological evolution (Zuckerkindl 1987). Motoo Kimura’s theory of neutral evolution provided the theoretical framework for expecting substitutions to occur with some regularity (Kimura 1968). Kimura noted that mutations with no selective advantage or disadvantage evolve solely through genetic drift. The expected number of neutral mutations produced per generation is equal to the rate at which mutations are produced, multiplied by the population size. Thus, a population of  $N$  individuals with a mutation rate of  $\mu$  will produce  $N\mu$  mutations per generation. The probability that any neutral mutation eventually replaces all other types in the population and becomes fixed is equal to its frequency in the population. Thus, for a population of  $N$  individuals, the probability that a new mutation becomes fixed is  $1/N$ . The rate at which substitutions occur in a population is found by multiplying the number of mutations produced per generation with the probability that these mutations are fixed. Thus, the substitution rate  $K$  is equal to  $N\mu$  multiplied by  $1/N$ . Kimura noted the amazing fact that in this calculation the population size  $N$  cancels itself out (Kimura 1968). Critically, we are left with the substitution rate being equal to the mutation rate ( $K = \mu$ ) regardless of population dynamics. This makes intuitive sense; large populations will produce more mutations but fix fewer of them while small populations will produce few mutations but fix more of them. So long as mutations are produced at a steady rate, then the neutral substitutions that we observe will be produced at a steady rate as well.

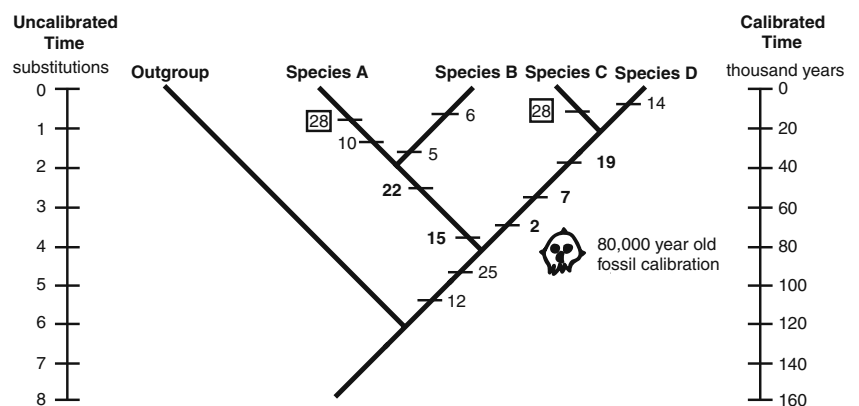
We now know that the mutation rate has not remained constant through time. Some species have evolved at different rates. For example, apes (including humans) have evolved at a slower rate than their Old World (Goodman 1961) and New World monkey (Hodgson et al. 2009) cousins. The “molecular clock” is not strictly clock-like, however it is regular enough that substitutions do relate to time. There are now sophisticated methods available using Bayesian (Drummond et al. 2006; Thorne and Kishino 2002; Thorne et al. 1998) or maximum likelihood (Sanderson 1997) statistics to deal with changes in the rate of evolution amongst lineages when estimating divergence times.

These methods all estimate divergence times very similar to Zuckerkindl and Pauling’s original conception of the molecular clock. First, a phylogeny is inferred for the data set as described in the previous section. Then the lengths of each branch in the tree are estimated according to the number of substitutions that have occurred along each branch. Finally, the substitution rate is calibrated with some information from the fossil record. Figure 2 illustrates this process. In molecular clock studies, choice of fossil calibration is critical but will not be discussed further here (see Steiper and Young [2008] for a review).

### Anthropological Genetics: Zeroing in on Our Evolutionary Past

These are exciting times for anthropological genetics. The amount of DNA sequence data currently available is staggering and new data are being created at a mind-boggling pace. Next-generation sequencing technologies allow a single researcher to produce in an afternoon what used to take multiple laboratories and armies of technicians years to produce (Mardis 2008). Below we review the anthropological genetic evidence for modern human origins.

**Fig. 2** This figure illustrates the “molecular clock” using the most parsimonious evolutionary reconstruction shown in Fig. 1a. Branch lengths are estimated according to the number of inferred substitutions along each branch. Substitutions are then converted to calendar time by referring to a fossil calibration



## Modern Human Origins

The fossil record from the last two million years shows many human forms across the Old World. During this time, *Homo erectus* and *Homo heidelbergensis* are found in Africa, *H. heidelbergensis* and *Homo neanderthalensis* (the Neanderthals) are found in Europe, and *Homo erectus* and *Homo floresiensis* are found in East Asia. Then around 200,000 years ago, fossils that are essentially indistinguishable from living people are found in Africa (McDougall et al. 2005). These essentially modern forms can then be found in Europe by 31,000 years ago (Wild et al. 2005) and in East Asia by 35,000 years ago (Trinkaus 2005). Once the modern forms appear in the record, the earlier forms rapidly disappear.

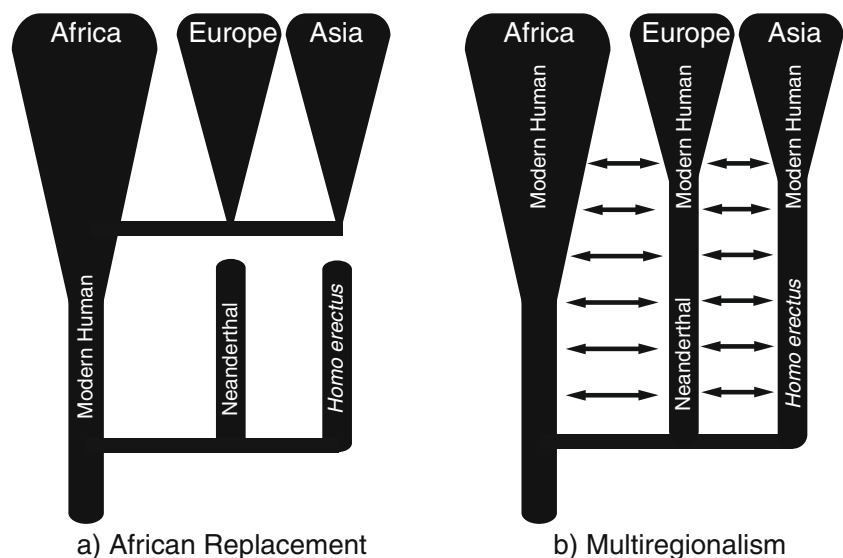
There is much debate amongst paleoanthropologists about the relationship between the earlier fossil forms and living people. There are two predominant hypotheses describing this relationship. The first hypothesis, known as multiregionalism, suggests that all of the fossil forms over the last two million years can best be described as a single polytypic species united by gene flow throughout the world (Weidenreich 1946; Wolpoff et al. 1984). In this view, anatomically modern *Homo sapiens* evolved from the various more primitive populations around the world with Europeans having significant *H. neanderthalensis* ancestry and East Asians having significant *H. erectus* ancestry (Hawks et al. 2000; Wolpoff et al. 2001). The opposing view, known as recent African replacement, suggests that anatomically modern humans evolved recently in Africa and then spread around the world replacing *H. neanderthalensis* and *H. erectus* with little or no interbreeding (Howells 1976; Stringer and Andrews 1988). Figure 3 illustrates these hypotheses.

Anthropological genetics have made great contributions towards resolving this debate. Both models of modern human origins make clear predictions about the genetic diversity we should see amongst living people and between living people and ancient fossils. Worldwide genetic surveys of living people have now tested these hypotheses. Exciting advances in our ability to analyze ancient DNA preserved in some fossils has also allowed us to assess the relationship between some fossils and living people.

The first major test of the hypotheses was a pioneering study by Cann et al. (1987) surveying worldwide diversity in mitochondrial DNA (mtDNA). Mitochondrial DNA was chosen for study for two important reasons: First, mitochondria are plentiful in cells and are therefore easy to study. More importantly, the mitochondrial genome is inherited as a unit strictly through the maternal line. This means that the mitochondrial genome is not subject to recombination (the process by which maternal and paternal chromosomes swap regions of DNA) as is the nuclear genome. This greatly simplifies the analysis of mitochondrial DNA.

Cann et al. (1987) compared mtDNA sequence diversity from 147 people drawn from a worldwide sample including Africans, Asians, Europeans, Australians, and New Guineans. The phylogeny they reconstructed had the deepest divergences within Africa, while some African and all non-African populations were closely related and formed a single branch of the tree. This means that most of the mitochondrial diversity is found within African populations while the diversity in non-African populations is restricted. They estimated that all living humans share a common maternal ancestor that lived in Africa about 200,000 years ago, while all non-Africans share a maternal ancestor that lived more recently (Cann et al. 1987; Fig. 4a). Though

**Fig. 3** Alternative models of modern human origins: **a)** modern humans evolve in Africa and replace archaic forms around the world, **b)** archaic populations around the world maintain gene flow and evolve towards the modern form multiregionally



these findings were initially criticized for methodological reasons, they have been confirmed by new methods of analysis along with complete mitochondrial genome sequences from a worldwide sample (Ingman et al. 2000).

The observed pattern of mitochondrial diversity is consistent with the predictions of the recent African replacement model of modern human origins. These findings, on their own, were not enough to rule out multiregional evolution, however. The mitochondria are only a single genetic lineage and it was not clear if the mitochondrial pattern was typical of the genome at large or if the pattern was anomalous perhaps due to its maternal pattern of inheritance. Scientists turned their attention next to the Y-chromosome, the paternal counterpart of the mitochondria. Like the mitochondria, the Y-chromosome is (largely) inherited as a unit and not subject to the recombination that shuffles genomic variants in the rest of the nuclear genome. However, unlike the mitochondria, Y-chromosomes are passed strictly from fathers to sons. While mitochondria tell the female story, Y-chromosomes tell the male story.

The Y-chromosome pattern is in many ways a mirror image of the mitochondrial story (Underhill and Kivisild 2007). Worldwide population sampling of Y-chromosomes reveals a phylogeny with the deepest divergences within African populations and all non-Africans closely related and restricted to a single branch of the tree (Hammer et al. 1998; Underhill et al. 2000). All living males are estimated to have shared a common ancestor around 100,000 years ago, while all non-African males are estimated to have shared an ancestor around 40,000 years ago (Tang et al. 2002; Underhill et al. 2000; Fig. 4b). The difference in age estimates between mitochondrial and Y-chromosome lineages is expected given the wide variance in times of

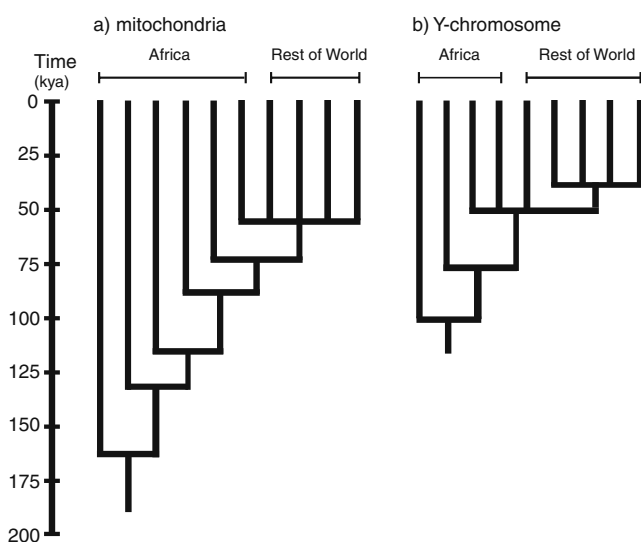
most recent common ancestry between independent genetic lineages and given the different inheritance patterns of the two loci. The important point is that both markers are young and show that non-African diversity is a subset of African diversity.

Both the mitochondrial and Y-chromosome data suggest that modern humans have been living in Africa longer than any other place in the world and that expansion into Europe and East Asia was a more recent event. Despite sampling thousands of individuals from European and Asian populations, no ancient lineages that could have been inherited from either *H. neanderthalensis* or *H. erectus* have been found. Also, the estimated times of shared ancestry seem to correspond closely to the fossil and archaeological record, with the oldest modern human fossils found in Africa and not appearing in Europe and Asia until more recently. Traces in the archaeological record attributed to behavioral modernity show the same pattern, appearing in Africa first and only in Eurasia around the time that modern human fossils appear (Mellars 2006). Most authors interpret these findings as a strong support for the recent African replacement hypothesis (Cann et al. 1987; Hammer et al. 1998; Ingman et al. 2000; Underhill and Kivisild 2007; Underhill et al. 2000), with the inferred common ancestry of all humans in Africa between 100,000 and 200,000 years ago in Africa corresponding to the first appearance of modern humans during this time and the inferred common ancestry of Eurasians between 40,000 and 60,000 years ago coinciding with the movement of modern humans out of Africa to replace the populations that had moved out of Africa earlier.

However, not everyone agrees with this interpretation of the genetic data. Some authors have noted that two genetic loci are insufficient to generalize about overall genomic patterns. These authors note that the patterns seen with the mitochondria and Y-chromosome are possible with multi-regional evolution even if unlikely and are therefore insufficient to rule out multiregionalism (Templeton 2005, 2007; Wall 2000). Consequently, several autosomal and X-chromosome loci have now been looked at.

Studies of overall genomic diversity clearly indicate that Africans are more genetically diverse than non-African populations (e.g., Li et al. 2008; Prugnolle et al. 2005; Tishkoff et al. 2009; Witherspoon et al. 2006). Also, genetic (Li et al. 2008; Prugnolle et al. 2005) and phenotypic (Manica et al. 2007) variation decreases the farther the population gets from East Africa. These patterns are clearly in accord with the predictions of the recent African replacement hypothesis.

However, individual nuclear loci show a variable pattern. For example, CD4, a gene located on human chromosome 12, shows a pattern very similar to that seen with mitochondria and Y-chromosomes (Tishkoff et al. 1996).



**Fig. 4** Global distribution and divergence times of human (a) mitochondrial and (b) Y-chromosome gene lineages



On the other hand, RRM2P4, a gene on the human X-chromosome, shows a very different pattern. This gene has the deepest divergences in Asia rather than Africa, and the most recent common ancestor was estimated to have lived approximately two million years ago. This pattern has been interpreted as evidence for admixture from *H. erectus* (Garrigan et al. 2005). Because of the wide variance in expected genealogical histories under any evolutionary scenario, it has been estimated that at least 50 nuclear loci are needed to evaluate the hypotheses of modern human origins (Wall 2000). Recently, Fagundes et al. (2007) sampled 50 nuclear loci from a worldwide sample of humans. Then, using sophisticated computer modeling, they compared their observed data to data simulated under a variety of recent African replacement and multiregional scenarios. They found that the data best fit the recent African replacement model. Interestingly, they found that a small but significant percentage of nuclear loci should show patterns similar to the anomalous pattern seen with RRM2P4.

On balance, the distribution of human genetic variation clearly follows the predictions of the recent African replacement model of human origins. We clearly trace most of our ancestry through a population that lived in Africa between 100,000 and 200,000 years ago and the archaic populations outside of Africa contributed little if any to present day diversity. However, it is very difficult to rule out small contributions from Neanderthals or *H. erectus* without looking at their genes directly.

### Ancient DNA: the Fate of the Neanderthals

New developments in our ability to recover DNA from fossils now make it possible to actually look at DNA extracted from fossils. The study of ancient DNA is rapidly refining our view of modern human origins.

Another amazing property of DNA is its durability. Most biomolecules quickly degrade upon death. However, under certain conditions, DNA can persist for long afterwards. Svante Pääbo first demonstrated this fact by recovering DNA from the tissue of a 2,400-year-old Egyptian mummy (Pääbo 1985) and confirmed DNA sequences that are older than 110,000 years have now been recovered (Lindqvist et al. 2010). Amazingly, a draft sequence of the entire *H. neanderthalensis* genome is now complete (Green et al. 2010). Recently, DNA was also recovered from an isolated finger bone found in 30,000- to 48,000-year-old sediment in Denisova Cave, Southern Siberia, and the DNA sequence suggests that the finger bone comes from a yet unknown human group (Krause et al. 2010).

The first *H. neanderthalensis* DNA sequence was recovered from the original specimen discovered in 1856

(Krings et al. 1997). Krings and colleagues sequenced a 379-base-pair section of the mitochondrial genome and found that the *H. neanderthalensis* sequence is very different from living human sequences. This finding has now been corroborated with mitochondrial sequences from 14 additional Neanderthal specimens (reviewed in Hodgson and Disotell [2008]). This now includes complete mitochondrial genome sequences from nine Neanderthal individuals (Briggs et al. 2009; Green et al. 2008, 2010). It is now clear that all sampled Neanderthals form a cluster outside the cluster formed by living humans. Based on the complete mitochondrial genome sequences, living humans and Neanderthals are estimated to have last shared a maternal ancestor 660,000 years ago (Green et al. 2008).

These findings have been interpreted to indicate that little or no interbreeding between Neanderthal females and modern human men occurred during the colonization of Europe by modern humans (Belle et al. 2009; Currat and Excoffier 2004; Serre et al. 2004). However, others have pointed out that with only data from the mitochondria we still have the single locus problem. The mitochondrial pattern does not rule out Neanderthal male to modern human female gene flow. Also, if modern human mitochondria have a selective advantage over Neanderthal mitochondria, we should not expect to see Neanderthal mitochondria in European populations even if interbreeding between Neanderthals and modern humans was extensive (Hawks 2006).

For this reason and others, sequence data from the Neanderthal nuclear genome has been sought. New high-throughput DNA sequencing methods developed to increase our genome sequencing capacity (Mardis 2008) are coincidentally particularly well suited to ancient DNA (Poinar et al. 2006). These methods allow researchers to recover millions of base pairs of ancient DNA sequence; however, the sequence recovered is not targeted but randomly chosen from throughout the genome. These methods have now been extended to allow targeted sequencing in addition to random shotgun sequencing (Briggs et al. 2009; Burbano et al. 2010).

The first two studies of untargeted Neanderthal genomic DNA recovered 65,000 (Noonan et al. 2006) and 1,000,000 base pairs of DNA sequence, respectively (Green et al. 2006). Unfortunately, the larger study is known to be largely contaminated with modern human DNA probably introduced by laboratory workers (Wall and Kim 2007) and the smaller study may have significant contamination as well. Neither data set showed evidence of a close relationship between Neanderthals and Europeans or any other living human population (Noonan et al. 2006; Wall and Kim 2007).

A team led by Svante Pääbo has now sequenced the genomes of three Neanderthal individuals to 1.3-fold genomic coverage (Green et al. 2010). This means that

each nucleotide in the Neanderthal genome has been read 1.3 times on average. Because the sequencing was performed randomly, some nucleotides have been read several times and some remain unread. In comparison, recent human genomes have been sequenced to 7- to 28-fold coverage (Kim et al. 2009; Pushkarev et al. 2009; Schuster et al. 2010; Wheeler et al. 2008). In these studies, each nucleotide is read multiple times, resulting in fuller genomic coverage, fewer errors, and a better assessment of heterozygous positions. The Neanderthal genome has not yet been sequenced to the level desired for modern high-quality DNA sources and many gaps and errors remain. However, new stringent methodologies have been employed that have reduced the modern human contamination that plagued previous studies to very low levels, and more than four billion nucleotides of Neanderthal DNA have now been sequenced and analyzed (Green et al. 2010).

The findings of the Neanderthal genome project were not anticipated by any previous analysis of fossil or genetic data and do not meet the predictions of either the multiregional or the recent African replacement model of human origins. To determine the relationship between Neanderthals and living humans, five modern human genomes were sequenced as a comparative sample. The researchers chose two African samples (San and Yoruba), a European (French), an East Asian (Chinese), and an Oceanian (Papuan) to represent worldwide genetic diversity. First, it was found that, though Neanderthals and modern humans are very closely related, living humans are on average more closely related to each other than any is to Neanderthals. However, the divergence between Neanderthals and modern humans was recent enough that much of the genetic variation that exists now also existed in the population that split to give rise to both populations. This means that Neanderthals share derived changes with some living humans to the exclusion of others. In other words, all of us have some gene variants that are more closely related to those possessed by some Neanderthals than to those of other living people. It is estimated that the population divergence between Neanderthals and modern humans occurred between 270 and 440 thousand years ago (Green et al. 2010). To put this in perspective, the pygmy chimpanzee (*Pan paniscus*) and common chimpanzee (*Pan troglodytes*) are estimated to have diverged 930 thousand years ago, and three common chimpanzee subspecies (*Pan troglodytes troglodytes*, *Pan troglodytes verus*, and *Pan troglodytes schweinfurthii*) are estimated to have diverged 460 thousand years ago (Hey 2010). Thus, Neanderthals and modern humans likely shared an ancestral population more recently than did the common chimpanzee subspecies.

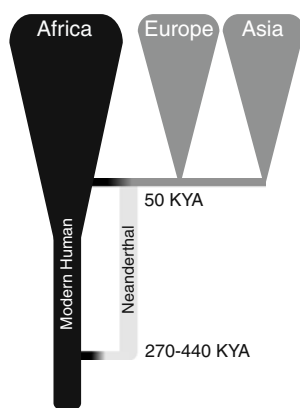
Because of the closeness of Neanderthals and living humans, the question of recent interbreeding after the initial split is a difficult one. Neanderthals and living people share derived polymorphisms regardless of recent interbreeding,

yet these are the markers we look for as evidence of interbreeding. However, through a series of rigorous analyses, Green et al. (2010) find a surprising genetic relationship between Neanderthals and the comparative genomes. They show that the Neanderthals are more closely related to the three Eurasian genomes (French, Chinese, and Papuan) than they are to the two African genomes (San and Yoruba). Though all humans share derived polymorphisms with Neanderthals, the Eurasian genomes have an excess of shared derived polymorphisms when compared to Africans. They interpret these results as recent admixture between Neanderthals and non-African modern humans and estimate that Neanderthals contributed 1–4% of the diversity seen in non-Africans. The direction of gene flow appears to be exclusively from Neanderthals to modern humans, as there is no evidence of modern human genes in the Neanderthal genome.

This pattern conflicts with both the multiregional and recent African replacement models of modern human origins. The multiregional model predicts that Neanderthals are more closely related to Europeans than to any other population. This was not found to be the case, as the Papuan and Chinese individuals have as much Neanderthal ancestry as the French individual. On the other hand, the recent African replacement model predicts that Neanderthals are equally distantly related to all living humans. This also was not found to be the case, as the three Eurasian genomes are more closely related to the Neanderthals than are the Africans. Green et al. (2010) suggest that admixture occurred between Neanderthals and modern humans in the Middle East early during the expansion of modern humans out of Africa between 50,000 and 100,000 years ago. Further sampling of African populations is required to verify the absence of Neanderthal admixture in these populations (Hodgson et al. 2010).

The findings of the Neanderthal genome project require us to reevaluate our ideas about modern human origins. Though the majority of our genome follows the predictions of the recent African replacement model, a small but significant percentage of the genomes of Eurasians appear to descend from Neanderthals. Thus, we all primarily descend from a recent African population. However, early during the expansion out of Africa, a small amount of interbreeding occurred between these early modern humans and Neanderthals (Fig. 5). The fact that this pattern of admixture was not anticipated by any surveys of extant genetic diversity suggests the possibility of small amounts of admixture from other archaic populations such as *H. erectus* in East Asia. Ancient DNA from these other archaic populations may be required to conclusively answer this question. The finding that the European sample is equally related to Neanderthal as the Chinese and Papuan sample raises interesting questions about what

**Fig. 5** This shows the pattern of Neanderthal admixture in modern humans found by the Neanderthal genome project. Neanderthals and modern humans admixed early during the latter's exodus from Africa. Eurasians are equally related to Neanderthals, with 1–4% of their genomes descending from them



went on in Pleistocene Europe. Why did not Neanderthals and modern humans continue to interbreed in Europe despite overlapping in range for thousands of years? It is possible that Neanderthals and modern humans rarely came into contact with each other and avoided each other because their population sizes were small. It is also possible that Neanderthals and modern humans did continue to interbreed in Europe but that this admixture was erased by subsequent population replacements in Europe such as that associated with the spread of agriculture (Ammerman and Cavalli-Sforza 1984; Barbujani and Bertorelle 2001). Further study of living and ancient European genomes will be required to answer these questions.

Because DNA preserves best in cold, dry environments, ancient DNA studies of Pleistocene humans had been limited to Neanderthals and anatomically modern humans from Western Eurasia. DNA has not yet been successfully recovered from other human forms that lived in more temperate climates such as *H. erectus* in East Asia or *H. floresiensis* from Indonesia. However, mitochondrial DNA sequence was recently recovered from an isolated human finger bone found in Denisova Cave in the Altai Mountains of Southern Siberia. The finger bone is clearly human, however it is not possible to tell from which human group it comes without further paleontological evidence. Krause et al. (2010) used a similar DNA sequencing method to that used to sequence the Neanderthal genome but designed to specifically target the mitochondrial genome (Briggs et al. 2009). The mitochondrial genome sequence they recovered was unlike any that had been seen before. They compared the Denisova finger bone sequence to 54 modern humans, one anatomically modern Pleistocene human and six Neanderthals and found that the Denisova sequence is equally distantly related to both Neanderthals and modern humans. This means that living humans shared a common ancestor with Neanderthals more recently than they did with the Denisova human. They estimate that Neanderthals and modern humans last shared a mitochondrial ancestor 466,000 years ago and the Denisova human last shared a common ancestor with Neanderthals and modern humans

1.04 million years ago. It remains unclear what group the Denisova human comes from, whether *H. erectus*, Neanderthals, or some as yet unknown group. Genomic sequence from the finger bone and more informative fossil findings from the region should clarify the place of Denisova in the future. For the time being, Denisova is the first and only ancient human known exclusively from its DNA sequence.

In addition to answering questions about the relationship between Neanderthals and ourselves, Neanderthal DNA is telling us more about what the Neanderthals were like. The Neanderthal genome project has revealed the sequence of several Neanderthal protein-coding genes that have evolved adaptively recently in modern humans (Burbano et al. 2010; Green et al. 2010). The Neanderthal sequence will help to determine the timing of these selection events and may play a role in differentiating Neanderthals and modern humans biologically (Green et al. 2010).

For example, FOXP2 is a gene known to have experienced a recent selective sweep in modern humans, with all living humans sharing two unique amino acid substitutions (Enard et al. 2002). The FOXP2 gene product is required for normal speech development in humans, suggesting that it may have been important in the evolution of language ability (Lai et al. 2001). Neanderthals were found to have the derived form of the FOXP2 gene found in modern humans (Krause et al. 2007). Though it is premature to conclude from this that Neanderthals shared our language ability, the findings are suggestive.

Some Neanderthals also may have had red hair and light skin. Neanderthals have a mutation in the pigmentation gene MC1R that results in lower activity of the protein (Lalueza-Fox et al. 2007). People with red hair and very fair skin also have a low activity variant of MC1R (Rees 2003). Interestingly, Neanderthals and living humans evolved the low activity variant independently. Light skin is favored in environments with low UV light incidence (Jablonski and Chaplin 2000) and it is likely that the convergent evolution of skin color in Neanderthals and Europeans was in response to the European climate.

Finally, the Neanderthal genome has revealed 212 genomic regions that appear to have undergone adaptive evolution in modern humans since the split with Neanderthals (Green et al. 2010). These genes may play a role in the evolution of the things that make modern humans different from the archaic human populations. Interestingly, amongst these genes are those that are known to play a role in bone development. Perhaps these genes are responsible for the unique skeletal features that distinguish modern humans from Neanderthals and our other relatives.

The Neanderthal genome project promises to reveal many more secrets about Neanderthal biology that were not evident from study of the fossils and archaeology.

Also, the findings of the Neanderthal genome project demonstrate the necessity of ancient genomes for us to fully understand our origins. It is hoped that additional ancient genomes, such as that of Denisova finger or of ancient anatomically modern humans, will continue to refine our understanding of modern human origins.

## Conclusions and Future Directions

Anthropological genetics have made great contributions to our understanding of human origins and human evolution. DNA sequences allow us to infer the evolutionary relationships amongst organisms and the times of divergence between them. The distribution of human genetic variation suggests that modern humans evolved recently in Africa and then quickly spread around the globe, largely replacing the human groups that they encountered during their dispersal. The Neanderthal genome project has clarified our understanding of modern human origins and suggests that low levels of interbreeding occurred between Neanderthals and modern humans early during the expansion of modern humans out of Africa. Genomic sequences from additional ancient humans, such as the Denisova finger bone, promise to further refine our understanding of human evolution.

The future goals of anthropological genetics will include further refining our understanding of human migration patterns and demography on a local and global scale. Also anthropological geneticists will be increasingly concerned with identifying and describing the genes that have been important in the evolution of the modern human phenotype and in the phenotypic variation within and between human groups.

**Acknowledgements** We thank Christina Bergey for help with figures and Luca Pozzi for helpful comments on a draft of this manuscript.

## References

- Ammerman AJ, Cavalli-Sforza LL. The neolithic transition and the genetics of populations in Europe. Princeton, NJ: Princeton University Press; 1984.
- Barbujani G, Bertorelle G. Genetics and the population history of Europe. *Proc Natl Acad Sci USA*. 2001;98(1):22–5.
- Belle EM, Benazzo A, Ghirrotto S, Colonna V, Barbujani G. Comparing models on the genealogical relationships among Neandertal, Cro-Magnoid and modern Europeans by serial coalescent simulations. *Heredity*. 2009;102(3):218–25.
- Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, et al. Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science*. 2009;325(5938):318–21.
- Burbano HA, Hodges E, Green RE, Briggs AW, Krause J, Meyer M, et al. Targeted investigation of the Neandertal genome by array-based sequence capture. *Science*. 2010;328(5979):723–5.
- Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. *Nature*. 1987;325(6099):31–6.
- Currat M, Excoffier L. Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol*. 2004;2(12):e421.
- Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. *PLoS Biol*. 2006;4(5):e88.
- Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, et al. Molecular evolution of Foxp2, a gene involved in speech and language. *Nature*. 2002;418(6900):869–72.
- Fagundes NJ, Ray N, Beaumont M, Neuenschwander S, Salzano FM, Bonatto SL, et al. Statistical evaluation of alternative models of human evolution. *Proc Natl Acad Sci USA*. 2007;104(45):17614–9.
- Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol*. 1981;17(6):368–76.
- Felsenstein J. *Inferring phylogenies*. Sunderland, MA: Sinauer Associates; 2004.
- Garrigan D, Mobasher Z, Severson T, Wilder JA, Hammer MF. Evidence for archaic Asian ancestry on the human X chromosome. *Mol Biol Evol*. 2005;22(2):189–92.
- Goodman M. The role of immunochemical differences in the phyletic development of human behavior. *Hum Biol*. 1961;33:131–62.
- Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF, et al. Analysis of one million base pairs of Neandertal DNA. *Nature*. 2006;444(7117):330–6.
- Green RE, Malaspina A, Krause J, Briggs AW, Johnson PLF, Uhler C, et al. A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell*. 2008;134:416–26.
- Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, et al. A draft sequence of the Neandertal genome. *Science*. 2010;328(5979):710–22.
- Hall BG. *Phylogenetic trees made easy: a how-to manual*. 3rd ed. Sunderland, MA: Sinauer Associates; 2008.
- Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, et al. Out of Africa and back again: nested clastic analysis of human Y chromosome variation. *Mol Biol Evol*. 1998;15(4):427–41.
- Hawks J. Selection on mitochondrial DNA and the Neandertal problem. In: Harvati K, Harrison T, editors. *Dordrecht, The Netherlands: Springer*; 2006. p. 221–238.
- Hawks J, Oh S, Hunley K, Dobson S, Cabana G, Dayalu P, et al. An Australasian test of the recent African origin theory using the Wlh-50 calvarium. *J Hum Evol*. 2000;39(1):1–22.
- Hey J. The divergence of chimpanzee species and subspecies as revealed in multipopulation isolation-with-migration analyses. *Mol Biol Evol*. 2010;27(4):921–33.
- Hodgson JA, Disotell TR. No evidence of a Neandertal contribution to modern human diversity. *Genome Biol*. 2008;9(2):206.
- Hodgson JA, Sterner KN, Matthews LJ, Burrell AS, Jani RA, Raam RL, et al. Successive radiations, not stasis, in the South American primate fauna. *Proc Natl Acad Sci USA*. 2009;106(14):5534–9.
- Hodgson JA, Bergey CM, Disotell TR. Neandertal genome: the ins and outs of African genetic diversity. *Curr Biol*. 2010;20(12):R517–9.
- Howells WW. Explaining modern man: evolutionists versus migrationists. *J Hum Evol*. 1976;5:477–95.
- Ingman M, Kaessmann H, Pääbo S, Gyllenstein U. Mitochondrial genome variation and the origin of modern humans. *Nature*. 2000;408(6813):708–13.
- International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 2004;431(7011):931–45.
- Jablonski NG, Chaplin G. The evolution of human skin coloration. *J Hum Evol*. 2000;39(1):57–106.
- Kim JI, Ju YS, Park H, Kim S, Lee S, Yi JH, et al. A highly annotated whole-genome sequence of a Korean individual. *Nature*. 2009;460(7258):1011–5.

- Kimura M. Evolutionary rate at the molecular level. *Nature*. 1968;217(5129):624–6.
- Krause J, Lalueza-Fox C, Orlando L, Enard W, Green RE, Burbano HA, et al. The derived *Foxp2* variant of modern humans was shared with Neandertals. *Curr Biol*. 2007;17(21):1–5.
- Krause J, Fu Q, Good JM, Viola B, Shunkov MV, Derevianko AP, et al. The complete mitochondrial DNA genome of an unknown hominin from Southern Siberia. *Nature*. 2010;464(7290):894–7.
- Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S. Neandertal DNA sequences and the origin of modern humans. *Cell*. 1997;90(1):19–30.
- Kumar S, Filipinski A. Multiple sequence alignment: in pursuit of homologous DNA positions. *Genome Res*. 2007;17(2):127–35.
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. 2001;413(6855):519–23.
- Lalueza-Fox C, Rompler H, Caramelli D, Staubert C, Catalano G, Hughes D, et al. A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. *Science*. 2007;318(5855):1453–6.
- Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*. 2008;319(5866):1100–4.
- Lindqvist C, Schuster SC, Sun Y, Talbot SL, Qi J, Ratan A, et al. Complete mitochondrial genome of a pleistocene jawbone unveils the origin of polar bear. *Proc Natl Acad Sci USA*. 2010;107(11):5053–7.
- Manica A, Amos W, Balloux F, Hanihara T. The effect of ancient population bottlenecks on human phenotypic variation. *Nature*. 2007;448(7151):346–8.
- Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet*. 2008;9:387–402.
- McDougall I, Brown FH, Fleagle JG. Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature*. 2005;433(7027):733–6.
- Mellars P. Going east: new genetic and archaeological perspectives on the modern human colonization of Eurasia. *Science*. 2006;313(5788):796–800.
- Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, Alessi J, et al. Sequencing and analysis of Neanderthal genomic DNA. *Science*. 2006;314(5802):1113–8.
- Pääbo S. Molecular cloning of ancient Egyptian mummy DNA. *Nature*. 1985;314(6012):644–5.
- Poinar HN, Schwarz C, Qi J, Shapiro B, Macphee RD, Buigues B, et al. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science*. 2006;311(5759):392–4.
- Prugnolle F, Manica A, Balloux F. Geography predicts neutral genetic diversity of human populations. *Curr Biol*. 2005;15(5):R159–60.
- Pushkarev D, Neff NF, Quake SR. Single-molecule sequencing of an individual human genome. *Nat Biotechnol*. 2009;27(9):847–52.
- Rees JL. Genetics of hair and skin color. *Annu Rev Genet*. 2003;37:67–90.
- Ridley M. *Evolution*. 3rd ed. Cambridge, MA: Wiley-Blackwell; 2003.
- Sanderson M. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol Biol Evol*. 1997;14(12):1218–31.
- Schuster SC, Miller W, Ratan A, Tomsho LP, Giardine B, Kasson LR, et al. Complete Khoisan and Bantu genomes from Southern Africa. *Nature*. 2010;463(7283):943–7.
- Serre D, Langaney A, Chech M, Teschler-Nicola M, Paunovic M, Mennecier P, et al. No evidence of Neandertal mtDNA contribution to early modern humans. *PLoS Biol*. 2004;2(3):E57.
- Steiper ME, Young NM. Timing primate evolution: lessons from the discordance between molecular and paleontological estimates. *Evol Anthropol*. 2008;17(4):179–88.
- Stringer CB, Andrews P. Genetic and fossil evidence for the origin of modern humans. *Science*. 1988;239(4845):1263–8.
- Tang H, Siegmund DO, Shen P, Oefner PJ, Feldman MW. Frequentist estimation of coalescence times from nucleotide sequence data using a tree-based partition. *Genetics*. 2002;161(1):447–59.
- Templeton AR. Haplotype trees and modern human origins. *Am J Phys Anthropol Suppl*. 2005;41:33–59.
- Templeton AR. Genetics and recent human evolution. *Evolution*. 2007;61(7):1507–19.
- Thorne JL, Kishino H. Divergence time and evolutionary rate estimation with multilocus data. *Syst Biol*. 2002;51(5):689–702.
- Thorne JL, Kishino H, Painter IS. Estimating the rate of evolution of the rate of molecular evolution. *Mol Biol Evol*. 1998;15(12):1647–57.
- Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, Cheung K, et al. Global patterns of linkage disequilibrium at the *Cd4* locus and modern human origins. *Science*. 1996;271(5254):1380–7.
- Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, et al. The genetic structure and history of Africans and African Americans. *Science*. 2009;324(5930):1035–44.
- Trinkaus E. Early modern humans. *Annu Rev Anthropol*. 2005;34(1):207–30.
- Underhill PA, Kivisild T. Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet*. 2007;41(1):539–64.
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, et al. Y chromosome sequence variation and the history of human populations. *Nat Genet*. 2000;26(3):358–61.
- Wall JD. Detecting ancient admixture in humans using sequence polymorphism data. *Genetics*. 2000;154(3):1271–9.
- Wall JD, Kim SK. Inconsistencies in Neanderthal genomic DNA sequences. *PLoS Genet*. 2007;3(10):e175.
- Watson JD. The double helix: a personal account of the discovery of the structure of DNA. New York, NY: Atheneum; 1968.
- Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*. 1953;171(4356):737–8.
- Weidenreich F. *Apes, giants and man*. Chicago, IL: University of Chicago Press; 1946.
- Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature*. 2008;452(7189):872–6.
- Wild EM, Teschler-Nicola M, Kutschera W, Steier P, Trinkaus E, Wanek W. Direct dating of early upper palaeolithic human remains from Mladec. *Nature*. 2005;435(7040):332–5.
- Witherspoon DJ, Marchani EE, Watkins WS, Ostler CT, Wooding SP, Anders BA, et al. Human population genetic structure and diversity inferred from polymorphic L1 (Line-1) and Alu insertions. *Hum Hered*. 2006;62(1):30–46.
- Wolpoff MH, Zhi WX, Thorne AG. Modern *Homo Sapiens* origins: a general theory of hominid evolution involving the fossil evidence from East Asia. Smith FH and Spencer F, editor. Alan R. Liss, Inc.: New York, NY. 1984. 411–83.
- Wolpoff MH, Hawks J, Frayer DW, Hunley K. Modern human ancestry at the peripheries: a test of the replacement theory. *Science*. 2001;291(5502):293–7.
- Yang Z, Rannala B. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. *Mol Biol Evol*. 1997;14(7):717–24.
- Zuckerkandl E. On the molecular evolutionary clock. *J Mol Evol*. 1987;26(1–2):34–46.
- Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. Bryson V, Vogel HJ, editors. New York: Academic; 1965. p. 97–165.