

Human Evolution

1. Origin of humans

Humans belong to the Great Apes (orangutans, gorillas, chimpanzees, bonobos, and humans). Molecular evidence indicates that the hominin lineage (the subfamily that includes the genus *Homo*) split from our closest relatives (the chimpanzees and bonobos) about 5–6 million years ago.

The oldest “good” fossils of hominins (species that are more closely related to us than to chimps) are about 4.4 million years old. They are found in east Africa (Kenya and Ethiopia) and belong to two species, *Australopithecus anamensis* and *A. afarensis* (“Lucy”). Around 2.5 million years ago the *Homo* genus first appeared with *Homo habilis*. *Homo erectus* is the first species that was also found outside of Africa (in Asia and Europe, with colonization of these regions occurring about 1.8 million years ago).

Adaptive evolution in the hominin lineage is driven by three main trends:

i) Bipedality: walking upright may have evolved after our ancestors moved out of the forest into more savannah-type habitats. It already evolved in early *Australopithecus*, which had a brain size comparable to chimps. A large number of morphological changes in modern humans are direct or indirect consequences of bipedality.

ii) Brain enlargement: brain size increased in *Homo* from 400 cm³ (as in modern chimps) to 1400 cm³ (in modern humans). Enlarged brain size is connected to a number of morphological changes, such as brain case and pelvic size, which evolved despite considerable costs due to birth problems.

iii) Social behavior: humans are intensely social and our cultural behavior is a major trait in which we differ from the other apes. There has probably been strong co-evolution of brain size and behavior. The main innovation underlying modern human culture is language. A larger role of teaching and learning in human social groups leads to the onset of cultural evolution that dominates all evolutionary change in *Homo sapiens*.

2. Origin of modern humans

500,000 years ago, *Homo erectus* was present in Africa, Europe, and Asia. In Europe, it was the ancestor of the Neanderthals (*H. neanderthalensis*), who lived from 200,000–40,000 years ago. However, after this period, only anatomically modern humans (*Homo sapiens*) are found. Where did the modern humans come from? Over the past decades, this has been a topic of debate with three main hypotheses:

1) “Candelabra model” – modern humans evolved from *H. erectus*, but did so independently in Europe, Asia, and Africa without gene flow.

2) “Multiregional model” – modern humans evolved from *H. erectus*, but there was sufficient gene flow (migration) among Africa, Asia, and Europe to maintain continuity as a single species.

3) “Out-of-Africa” (or “replacement”) model – modern humans evolved in Africa and then spread out from Africa about 100,000–60,000 years ago, migrating into Asia and Europe (first finds in Europe 40,000 years), replacing any *H. erectus* and *H. neanderthalensis* without interbreeding.

The models have differing implications regarding the origin and antiquity of the present human races. At the one extreme, under the candelabra model, racial variation derives from geographic variation among *H. erectus* populations that could have begun 1.5–2 million years ago. At other extreme, under the out-of-Africa model, present races are the result of recent geographic differentiation that occurred after modern humans migrated out of Africa (i.e., races are less than 100,000 years old).

There is little support for the Candelabra model. First, it is implausible that same species, *H. sapiens*, would evolve at about the same time in many different places with no gene flow among populations. Second, the molecular evidence is inconsistent with this model (see below).

It has been difficult to distinguish the “Multiregional” and “Out-of-Africa” models. The fossil evidence is heatedly debated, but somewhat ambiguous. On the one hand, early evidence for *H. sapiens* in Africa (about 130,000 years ago) and the rapid disappearance of older *H. erectus* (e.g. Neanderthal) traits favors the replacement model. On the other hand, there are a few intermediate fossils, suggesting interbreeding, although “Out-of-Africa” proponents explain these as natural within-species variation in *Homo sapiens*.

Ancient DNA sequencing (the sequencing of DNA extracted from ancient bones) suggests a combination of the two models, with a greater contribution of the replacement model. Modern humans arose in Africa and replaced other *Homo* species. However, there is evidence for at least some genetic exchange in Europe and Asia. Modern European and Asian genomes contain a small proportion (1–3%) of Neanderthal DNA. Similarly, Oceanic genomes (Melanesian and indigenous Australian) contain a small proportion (4–6%) of Denisovan DNA, which was sequenced from an ancient finger bone found in Denisova cave in Siberia and is from a previously-unknown *Homo* lineage.

3. Mitochondrial Eve

DNA sequence data can be used to estimate “coalescent times” to the most recent common ancestor of parts of the human genome. Coalescent times are estimated from the number of differences among DNA sequence in a population sample, assuming neutral evolution. For a pair of sequences,

$$T_{\text{MRC}} = S / 2u$$

where S is the number of differences and u is the mutation rate. The mutation rate u can be estimated from the rate of substitutions k , usually from divergence data between human and chimp, using the relationship $u = k$, with k being measured in terms of substitutions per site per year. When the number of sequences is greater than two, more refined mathematical models are used to estimate the shape and branch lengths of the coalescent tree.

The most famous (and best studied) case is mitochondrial DNA. There are several advantages to using mtDNA:

- i) mtDNA is a single, non-recombining, maternally inherited locus.
- ii) mtDNA evolves rapidly (estimates from human–chimp comparison suggest $u \approx 1.7 \times 10^{-8}$ mutations per site per year (more than 10 times the rate of nuclear genes).

iii) The small mtDNA genome (16,500 bp) is easy to work with, amplify, etc...

There are two main findings:

- Worldwide samples of human mtDNA coalesce in a woman that lived about 200,000 years ago, known as “mitochondrial Eve”.
- All trees are rooted in Africa, meaning that African mtDNA coalesces closest to the root of the tree: mitochondrial Eve was African.

This does *not* mean:

- “Eve” was the only woman alive at her time.
- “Eve” was the only woman of her time to have children.
- “Eve” was special in terms of her fitness (the model is based on neutral evolution).
- “Eve” is the common ancestor of the rest of our genes (nuclear genome).

We *can* conclude that all humans are all closely related to each other, with a common mitochondrial ancestor who lived in Africa about 200,000 years ago.

By dating the nodes of the tree, we obtain the following picture of migration (which fits reasonable well to fossil data):

Leaving Africa: 100,000 years ago

Presence in SE Asia/Australia: 40,000–60,000 years ago

Presence in Europe: 40,000 years ago

Presence in North America: 20,000 years ago

Presence in South America: 13,000 years ago

This result favors the “Out-of-Africa” model. There is similar evidence for other parts of the genome. For example, “Y-Adam”, the common ancestor of the non-recombining part of human Y chromosome lived about 50,000 to 100,000 years ago, also in Africa.

For mitochondrial DNA, there is independent evidence against interbreeding of modern humans with Neanderthals: comparison of ancient Neanderthal mtDNA with mtDNA of modern humans reveals more differences than the typical variation observed among modern humans. However, a small proportion (<5%) of the nuclear genome of Europeans and Asians appears to be derived from a Neanderthal ancestor.

Overall, the genetic studies provide clear evidence for a second “Out-of-Africa” migration. Nevertheless, it does not appear that there has been complete replacement of the older *Homo erectus* species without any interbreeding. Even if local *H. erectus* populations were not completely replaced, but instead managed to contribute to the modern gene pool, there is little doubt about the disproportionately large contribution of the ancestral African population to our genetic make-up.

4. Genetic variation and the effective population size of humans

DNA sequence polymorphisms in worldwide samples of humans can also be used to estimate the effective population size (N_e). For nuclear genes, the nucleotide variability is $\pi \approx 0.001$ (for noncoding sites). According to the neutral theory

$$\pi \approx 4N_e u$$

where u is the neutral mutation rate per site per generation. For nuclear data the neutral mutation rate is approximately $u = 2 \times 10^{-8}$ per site per generation (estimated from comparison

of pseudogenes between human and chimpanzee). Thus, the effective size is about:

$$N_e \approx \pi / 4u \approx 0.001/8 \times 10^{-8} \approx 12,500$$

This is an incredibly low effective population size relative to the current census population size of >6 billion. This is due to the historically small population size of humans who, up to about 40,000 years ago, existed in groups of very small sizes. In fact, there is a signature of a population size expansion in the molecular data around 35,000–40,000 years ago. The signature of population size expansion consists of most polymorphisms occurring at a low frequency. That is, most mutations are “young” and did not have sufficient time to increase in frequency by drift. A second major expansion of the human population occurred about 10,000 years ago, when agriculture began to spread. Genetic diversity is smaller in non-African populations than in African populations, probably reflecting a population bottleneck when a relatively small group left Africa about 100,000 years ago. When population size fluctuates over time, N_e can be estimated from the harmonic mean of the population sizes. Thus, small population sizes have a disproportionately large influence. A simple example calculation demonstrates the effect of historically small population sizes on N_e : Assuming a founding population of 10,000 individuals 100,000 years ago and an exponential increase to a population size of 6 billion today (generation time 20 years): one obtains an effective population size of $N_e = 18,648$.

5. Geographic differentiation in humans

Natural populations do not exist as a single, large, randomly mating population. Instead, they are more like a set of sub-populations that are connected by gene flow (incorporation of alleles from one population into another). Gene flow by migration or dispersal of pollen homogenizes the genetic variability among the sub-populations and, in particular, maintains the total population as a single species. If there is high gene flow, the genetic variation in all sub-populations will be similar. If there is low gene flow, there will be greater variation between the sub-populations than within each sub-population. In the extreme, this can lead to genetic incompatibility of sub-populations and speciation. A standard measure of the degree of genetic differentiation is Wright’s F_{ST} , which can be defined (on the nucleotide level) as follows:

$$F_{ST} = (\pi_T - \pi_S) / \pi_T$$

where π_T is the nucleotide variability in the total population and π_S is the average variability within subpopulations. F_{ST} assumes values between 0 (sequences within and among subpopulations have the same average number of differences, $\pi_T = \pi_S$) and 1 (all differences are between sub-populations).

In humans, observed F_{ST} ranges from 0.05 to 0.15. This is a very low value and means that 85–95% of all human variation can be found in any regional population. This is compatible with high levels of gene flow or, alternatively, the recent origin from a single African population that spread throughout the world.

With enough data (i.e., genome-wide polymorphism data from many individuals of many populations), it is possible to distinguish populations, even within regions. For example, using information about the frequencies of variants at many loci in different European populations, the ancestry of an individual can be inferred from his/her genome sequence.