f23 Tree analysis method of genetic variation in a species

< mitochondrial Eve >

> But nature *is* the unity. We do not have to invent that. It exists out there, visibly, as its own proof.
>
> No. There is no correct description of nature. Nature is more subtle, more deeply intertwined and more strangely integrated than any of our pictures of her—than any of our errors. It is not merely that our pictures are not full enough; each of our pictures in the end turns out to be so basically mistaken that the marvel is that it worked at all.
>
> —dialog fragments from *The Abacus and the Rose* by J. Bronowski.[1]

> A genome, now no more than a set of chemical structures of the organic molecules involved in inheritance, will need natural history to give it value. —Steven A. Benner.[2]

Given our understanding of the mechanism of mutation, the genome cannot avoid becoming changed. Turning this around, the genetic distance between species should be a measure of how much time has passed since they had a common ancestor (***Footnote f23.1***).

A genealogical tree can be built for species. Using what is called the *parsimony method*, the tree's branching network is constructed in a way that minimizes the number of mutations required to relate the types. To convert the resulting network into a tree, the ancestor or root must be placed, which requires additional information or assumptions. The closest available species to the group being investigated is assumed, by default, to be one that diverged at the root. The genetic composition of this closest related taxon, called the "outgroup," is used as the genetic composition of the common ancestor. The method is not perfect as root choice will significantly affect how the internal structure of an evolutionary tree is resolved and so, then, models of the possible evolution of various traits.

Mitochondrial DNA (mtDNA) is easier to isolate and sequence than DNA in the nucleus of a cell. mtDNA is useful for tracing the matriarchal family tree because it and not the father's mtDNA is passed on by the mother's egg. mtDNA from a sperm's midpiece (and nonexistent in the sperm's head or tail) are about dozen that release energy needed for the journey to the egg.[3] These persist among the 100,000 or so egg mitochondria when the sperm fertilizes an egg. However, during the sperm's creation, its mitochondria acquire ubiquitin. This protein tags cell-molecules and organelles for recycling. Gerald Schatten and Peter Sutovsky found in 2000 that while the tag is ignored during the journey, the sperm's mtDNA are destroyed within the first few cell divisions of a fertilized egg.[4]

In Europe, farming began in the Neolithic Period 5,000-10,000 years ago. Language and archaeology indicate that farming spread gradually westward from its inception in the Middle East. But did the farmers themselves? The canonical hypothesis has been that Middle Eastern populations spread through Europe and "swamped" the less technologically advanced people they found as they went. To test this, Bryan Sykes obtained a swab of cells from inside the cheek from several thousand individuals across Europe. 15-20 per cent of the population sampled have mtDNA clearly related to mtDNA from Middle Eastern populations and these are also of young lineages. The remainder have mtDNA that places their ancestors in Europe tens of thousands of years *before* the arrival of farming. Sykes concludes that rather few farmers left the Middle East, and that the in-place Europeans learned from their example and adopted the farming way of life.[5]

In 1987, Allan C. Wilson (1934-1991) with Rebeca L. Cann and Mark Stoneking presented an initial finding (their 'Eve' paper) that of all mtDNA types in contemporary Europeans, Asians, and American blacks, stem from a common female ancestor present in an African population some 120,000 to 150,000 (with uncertainty that ranges from 50,000 to 500,000!) years ago.[6]

As reported by Linda Vigilant in 1991, Wilson's 1987 study continues and has been extended to include many sub-Saharan peoples (!Kung of Botswana, Eastern Pygmies of Zaire, Western pygmies of Central African Republic, Yorubans of Nigeria).[7]

The outgroup method was used preferentially as it does not rely on the assumption that the rate of evolution is the same in all lineages. To place the human mtDNA ancestor on the network, a sequence from another species (the "outgroup"), in this application that of a chimpanzee, was used. The outgroup attaches to the network relating the human mtDNA types at the position that minimizes the total number of mutations in the tree. The point of attachment is thus the position of the human mtDNA ancestor on the tree.

The genealogical tree relating the 135 mtDNA types found among the 189 individuals was studied. Significantly, of the people studied, identical mtDNA types are shared within but not between populations.

This work is preliminary. For the (rightly) cautious, in reconstructing detailed evolutionary trees from DNA sequences, Gavin J. P. Naylor has found that the amount of sequence is not as important as a preparatory selection of sequence portions that does reconstruct the gross phylogeny known from the fossil record.[8] For the (seriously) curious, valuable references are provided by Fredrik Ronquist who reviews in *Science*,[9] Joseph Felsenstein's *Inferring phylogenies*, 2004.[10] ❑

Footnote f23.1 **Molecular clocks**

Jacob and Monod showed that genes are not restricted to chromosomes. They found free-floating genetic elements, called *episomes* and *plasmids*, in bacteria; other scientists soon found these elements in higher organisms as well.

Mitochondria, the cell's power plants, and chloroplasts in green plant cells were later found to have their own genes, inherited independently from those on the chromosomes. —Nathaniel C. Comfort.[11]

... preservation of favourable variations and the rejection of injurious variations, I call natural selection. Variations neither useful nor injurious would not be affected by natural selection, and would be left a fluctuating element ... —Darwin.[12]

Natural mutations occur at a small but non-negligible rate. The result is that DNA sequences over time change. The rate at which different biological sequences within an organism do change can be governed by natural selection. For example, the amino acid sequence of a protein encoded by a gene changes more slowly than the DNA sequence of the underlying gene. This is because amino acid changes can alter function, and natural selection eliminates the majority that under ordinary circumstances turn out to be disadvantageous. By contrast, a significant proportion of DNA changes may be selectively neutral because they create a synonymous codon (one that specifies the same amino acid). So at the molecular level evolutionary changes are caused not by Darwinian selection but by random drift of selectively neutral mutants (as presciently recognized by Darwin). This is the "Neutral Theory of Molecular Evolution" advanced by Motoo Kimura in 1968 to explain the unexpectedly high rate of evolutionary change and very large amount of intraspecific variability uncovered by DNA sequencing.[13]

Evolutionary genetic studies of distant species are now, to quote Simon Tavaré: "often carried out by examining amino acid sequences of proteins, while evolutionary comparisons among more closely related species are better done by examining DNA sequences within or between genes." Most rapid, is the rate of mutation observed in the small circular chromosome of mitochondria. As mitochondria occur outside the nucleus, they are not subject to certain DNA repair mechanisms.

In 1981, the human mitochondrial genome DNA sequence, S. A. Anderson could report, "has been completely determined."[14] It consists of 16,569 base pairs. Using a "molecular clock" (dubious though its ticking), **Vincent M. Sarich** updated the time, in million years ago, that humans and chimpanzees shared a common ancestor to about 5 rather than the 15 to 20 that "bone" paleontologists had decided.[15]