

## Molecular Phylogeny

1. Evolution
2. Extinction
3. Molecular Clock - The clock ticks
4. Human evolution - fossil & molecular phylogeny
5. Methods of molecular phylogeny
6. Evolutionary Biology & Bioinformatics:
  - Paralogs and Orthologs
  - COGS

# 1. Evolution:

- Biological changes of living organisms over time
- Evidence supports existence of evolution
- Theory of evolution (assumptions) answers questions on diversity, similarity between organisms etc
- Conclusions from these studies may differ between investigators
- Two levels of evolution studied:
  - (A) Microevolution- small changes over few generations
    - \*4 forces: (i) natural selection, (ii) mutation, (iii) gene flow & (iv) genetic drift
    - \*Makes alleles in a population more or less common over time

## (i) Natural selection- 4 major principles

- More offsprings produced which environment can not support
- Offspring with an inherited advantage survives
- Individuals pass on these advantages to their offsprings
- No two individuals are identical

Eg Peppered moths live on gray lichen (color protects against birds); black moths get eaten more frequently BUT when the lichens blacken due to pollution the reverse happens

(ii) Mutation- genetic source of variation

(iii) Gene flow (migration)- Allele sharing, eg mixed marriages

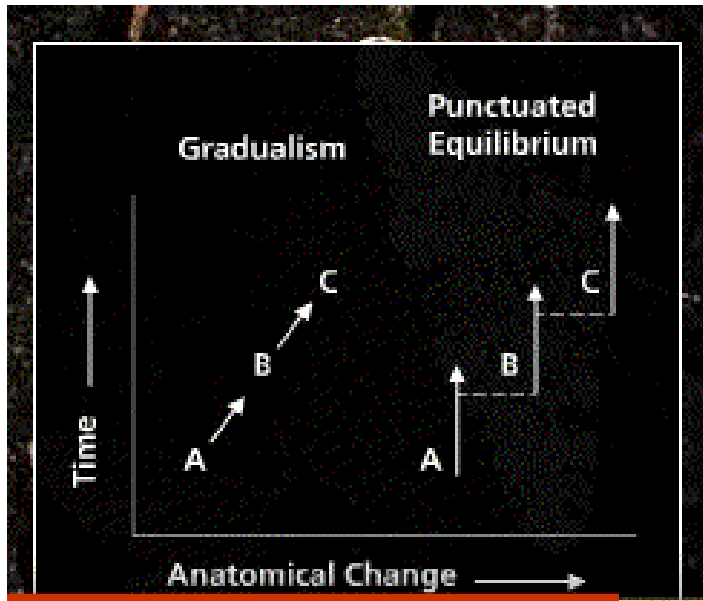
(iv) Genetic Drift- Change due to random event, eg volcano eruption kills people with type A blood allele in the vicinity. The allele is lost until gene flow or mutation returns it to the population.

Two special cases of genetic drift:

**Founder Effect**- A small population establishes a new colony & marries within itself; alleles that exist in the parents increase in the new ones. The rare Ellis-Van Creveld syndrome frequency is higher in the Lancaster County Amish (1:14) than in the parent population (1:400)

**Bottlenecking**- This occurs when a large number of population dies and the population rebuilds from a much smaller genetic base than before; the new population will show very little genetic variation. The Siberian tiger faces extinction due to this reason.

**(B) Macroevolution-broad pattern changes over 1000s of generations - 2 models used**



**Gradualism- slow steady change**

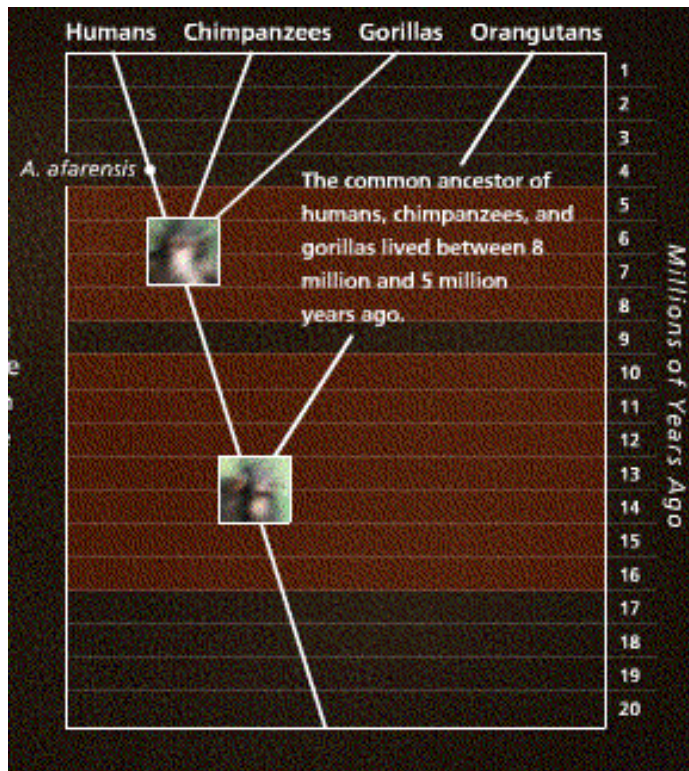
**Punctuated- short periods of rapid change interrupted with long periods of little change**

## 2. Extinction:

- Normal part of the evolutionary process
- Many species, including our hominoid ancestors, that existed before are extinct
- Healthy ecosystem balances species formation & extinction- dinosaur extinction created opportunities for mammalian ancestors to thrive.
- Is our ecosystem out of balance? Vast human population which was never seen before populates earth, uses extensive resources, creates pollution- humans have caused extinction of a large numbers of species. What is the future?

### 3. Molecular Clock:

- Relationship between 2 species estimated using DNA- all life shares the same bases
- Species that are genetically similar share a common ancestor
- Eg Comparison of Orangutans, African apes, humans and chimpanzees.



**Table of Distance matrix\***

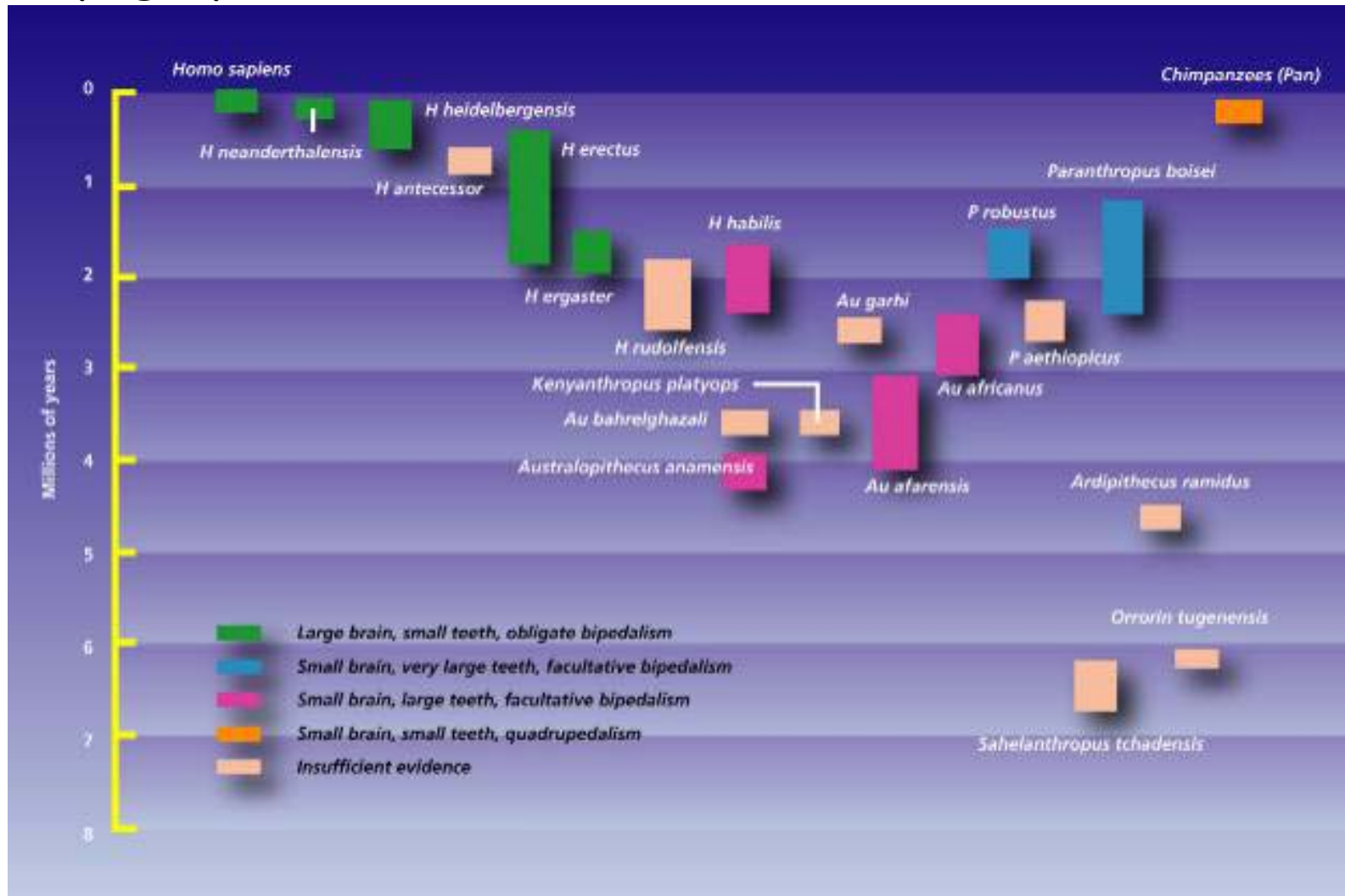
	<u>Human</u>	<u>Chimpanzee</u>	<u>Gorilla</u>	<u>Orangutan</u>
<b>Human</b>	100			
<b>Chimpanzee</b>	98.55	100		
<b>Gorilla</b>	98.49	98.43	100	
<b>Orangutan</b>	97.02	92.45	96.96	100

\*5.3kb noncoding DNA of two separate regions  $\eta$ -globin locus (2.2kb) &  $\eta$ - $\delta$  globin intergenic region (3.1kb) were aligned, pairwise distances generated & a tree inferred

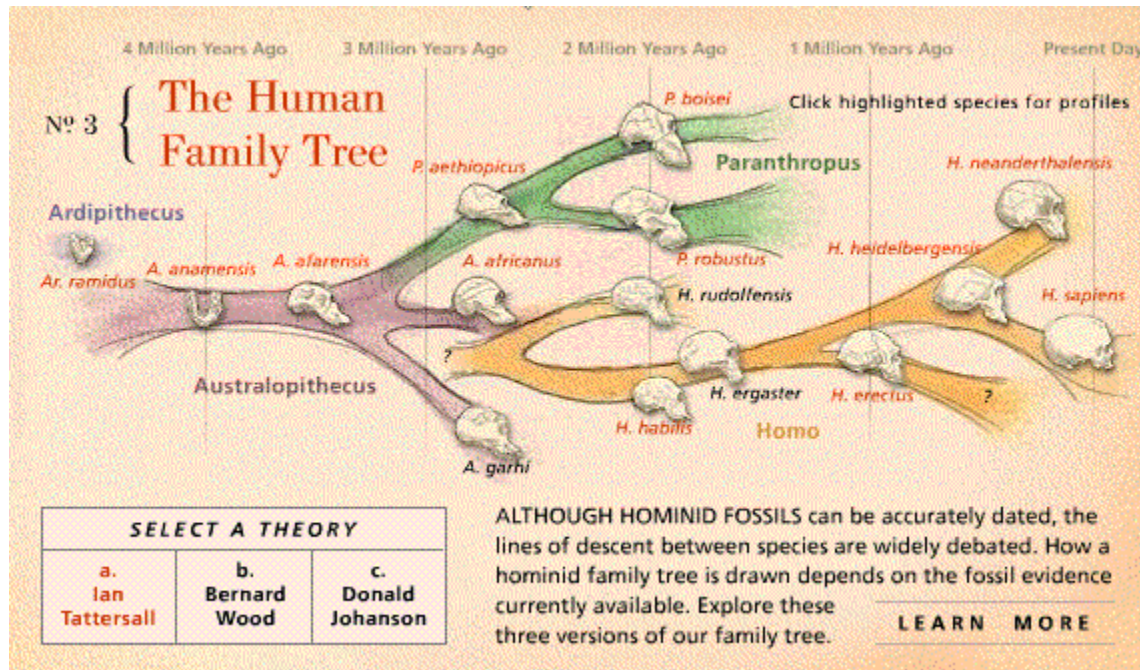
## 4 Human evolution - fossil & molecular phylogeny

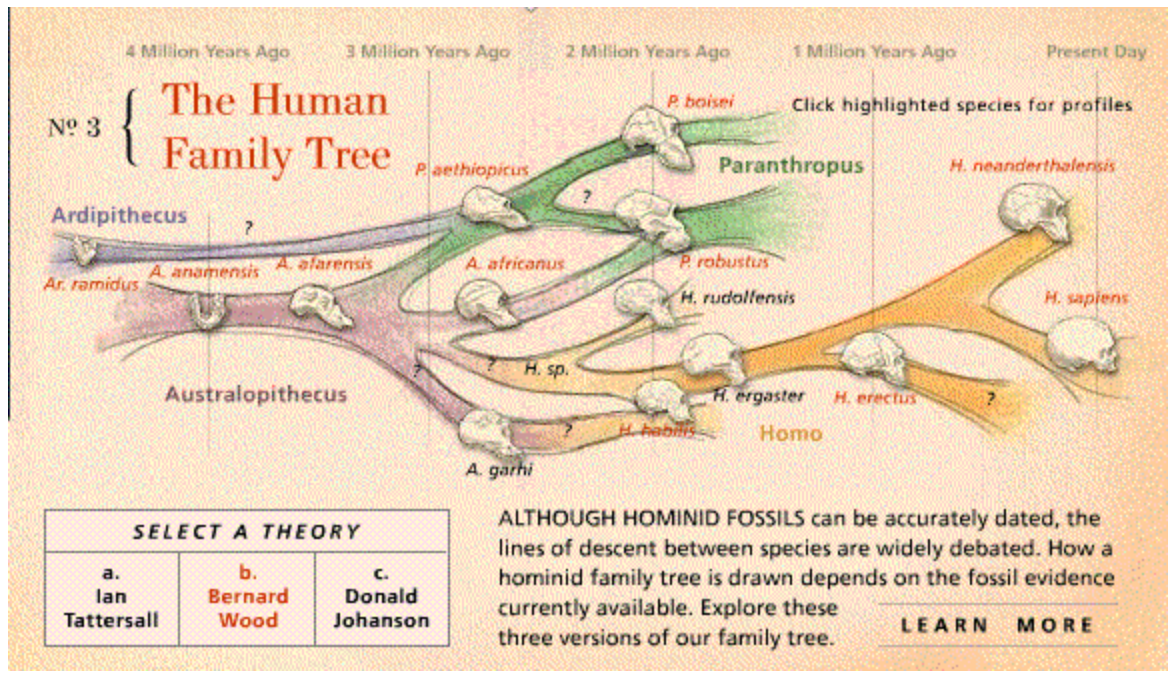
- Until the complete sequencing of the human genome, studies used fossils & molecules
- Different techniques (Ag-Ab), DNA-DNA hybridisation, gene comparisons) gave almost similar (but not the same) results on evolutionary relationship:
  - Humans (Hominidae) are related to apes (Pongidae)
  - Chimps (*Pan*) & gorilla (*Gorilla*) are our closest relatives - origin in Africa
- New phylogeny tools & complete sequencing of the human genome:
  - Chimps (*Pan troglodytes*) should be reclassified with humans as *Homo troglodytes* (humans = remodelled chimpanzees)
  - Controversial; acceptable - time will only tell; look out for comparative genomics of the human and ape genomes.

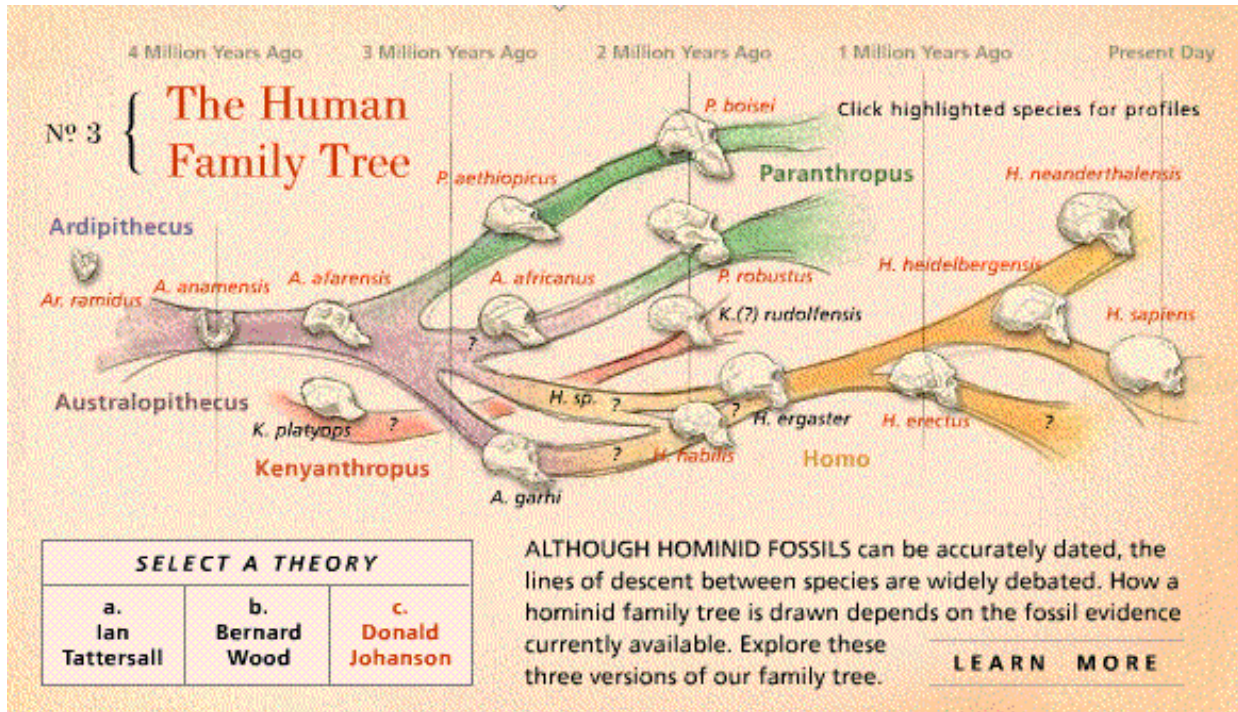
# Phylogeny based on fossil evidence- size of brain, teeth etc



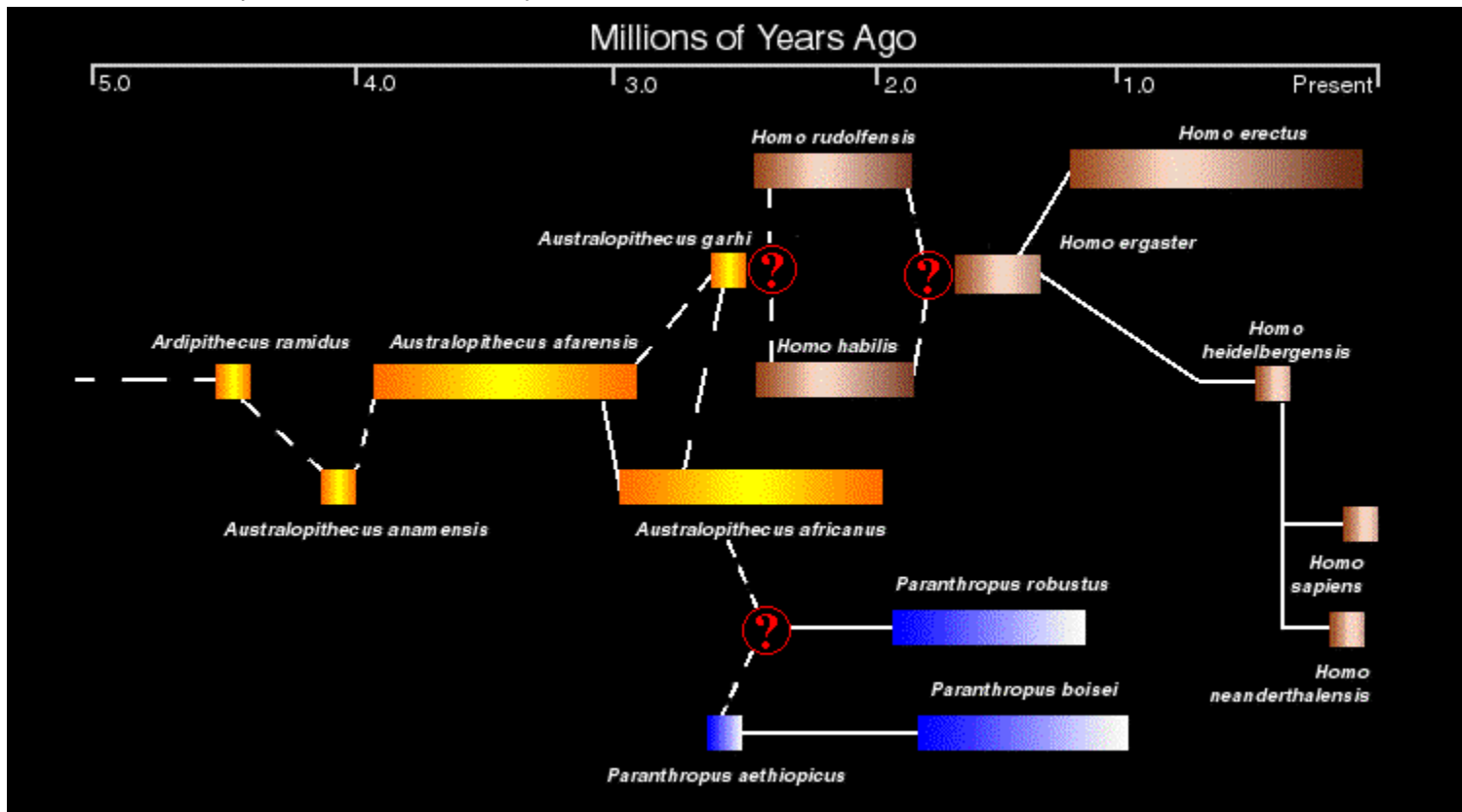






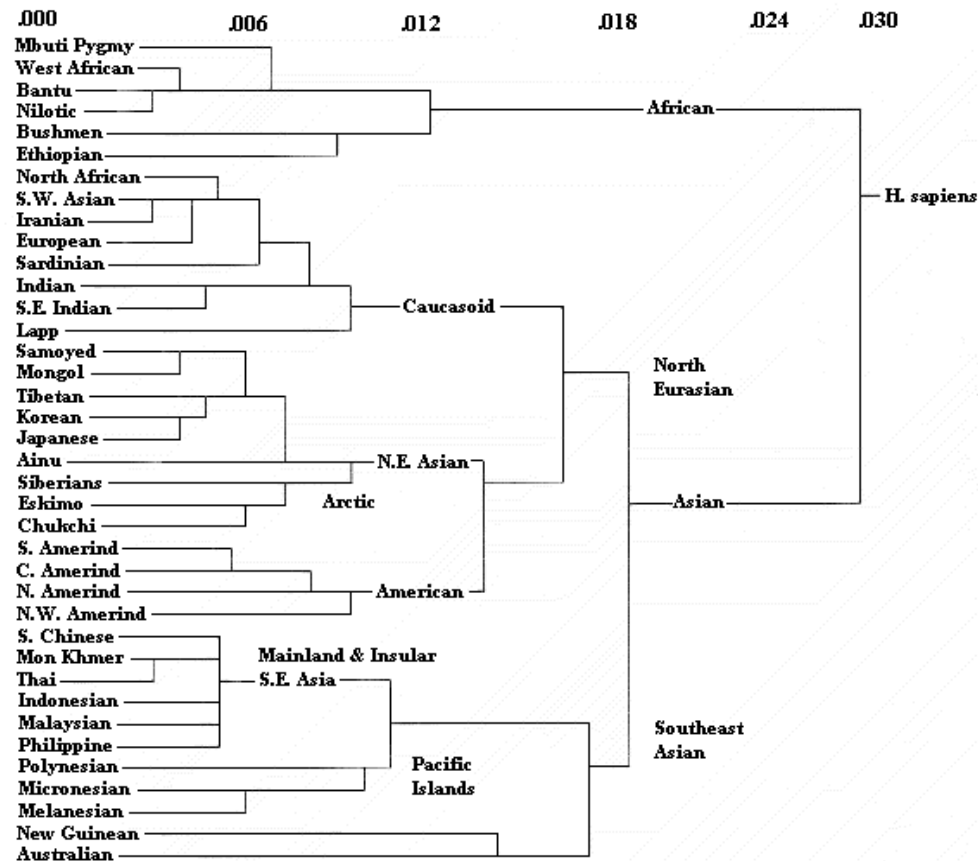


# Popular theory of human evolution until 2002



## Genetic Distance Between Autochthonous Human Populations

"Genetic drift" is the phenomenon by which small mutations in DNA (or RNA, and the proteins that are coded by them) add random variations over time to genetic material, resulting in differences between isolated groups of animals, whether of different species or of the same species that are not inter-breeding. Thus there is about a 5% difference in DNA between humans and our closest Primate relations, Chimpanzees. Between human populations, there is always much less than 0.1% difference. What these differences are and how much they vary can be used to construct a tree showing the relationships between human populations. The following tree is the result of such research, reported by Luigi Luca Cavalli-Sforza. The numbers across the top show the percentage difference in DNA, which is thus no more than 0.03% for all human beings. The most dramatic characteristic of the tree is the division between populations in sub-Saharan Africa and the rest of humanity. This is usually interpreted to mean that modern humans originated in Africa and that the population from which the rest of humanity descended left Africa somewhat less than 300,000 years ago, ultimately replacing earlier humans, like the Pithecanthropines (*Homo erectus*, like *Peking Man*, etc.), who had also evolved in Africa but left many thousands of years earlier.



Part of this research was the theory of "Eve," a single female in Africa, around 200,000 years ago, from whom every living human being is now descended. This does not mean that there were not other human females -- there were -- or that we are not descended from them too -- we are. The theory is based on the circumstance that some human genetic material is contained in the mitochondria, little organs in a cell outside the nucleus (where most genetic material is contained). Sperm cells do not pass on their mitochondria to a fertilized egg and so all human mitochondrial DNA is inherited from the mother. If a woman has only sons (a highly desirable result in many traditional cultures), then her own mitochondrial DNA is actually lost. Over time, this seems to have happened to all lines of descent of mitochondrial DNA, except one, the line from "Eve." Another interesting feature of the chart is the closeness of American Indians to modern population across Europe, the Middle East, and northern East Asia. Thus, curiously, Europeans are more closely related to American Indians than to Polynesians. Finally, it is noteworthy that skin color is not at all helpful in providing clues to genetic affinity. The darkest colored people on earth, in Africa, India, Melanesia, and Australia, are scattered between groups that are only distantly related. Dark skin color is certainly a function of living under the equatorial sun for many generations, but all human populations have the genetic wherewithal to make that adaptation.



A little girl hugging a chimpanzee demonstrates how different humans and chimps look physically, but they are between 95 and 98.5 percent genetically identical.

# Concept of Paralogous & Orthologous Genes

Initially defined by evolutionary biologists

Now used in genome annotation by biologists

Structural origin vs biological function

Different definitions and meaning. Definitions are important. Why?

Lateral & horizontal gene transfer, fused genes, gene duplication events - relationship to & impact on the whole organism & its genome

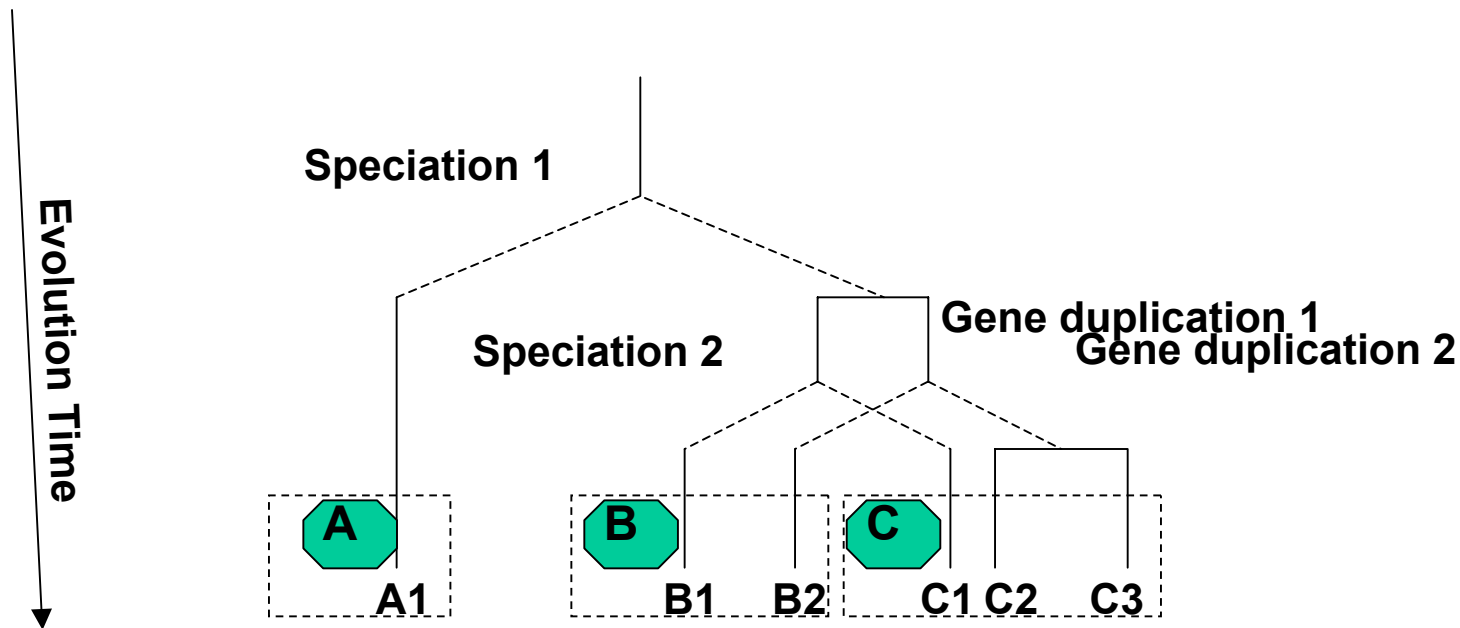
Flexibility vs specialisation ie new function, loss of old functions, modification of some functions- this is evolution

## Biologists definition:

- Orthologs- homologs in different species that catalyse the same reaction
- Paralogs- homologs in the same species that do not catalyse the same reaction



# Definition of evolutionary biologists



Four events of evolutionary divergence are shown. 2 are events of speciation and 2 are events of gene duplication yielding 6 contemporary genes in 3 organisms (A, B & C). Tracking any pair of genes back to where they join shows either orthology (inverted “Y”) or paralogy (horizontal bar). Thus A1 has 3 orthologs in species C but only C1 is an ortholog of B1. On the other hand, B2 has two orthologs in species C (C2 & C3) whereas B2 & C1 are paralogs. Orthology and paralogy differs in that one proceeds from speciation and the other from gene duplication.

## Cluster of Orthologous Groups of Proteins (COG)

- Phylogenetic classification of proteins from complete genomes
- Database of orthologous proteins connected through vertical evolutionary descent ie evolutionary biologists definition
- Includes lineage specific gene duplications, one to many and many to many relationships (therefore termed COG)
- COG & KOG are annotated for functions and genome evolution.
- 4873 COGs consists of 138,458 predicted proteins encoded in 66 genomes of unicellular organisms.
- 4852 KOGs consists of 59,838 predicted proteins from 7 eukaryotic genomes: three animals (the nematode *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster* and *Homo sapiens*), one plant, *Arabidopsis thaliana*, two fungi (*S. cerevisiae* and *S. pombe*), and the intracellular microsporidian parasite *Encephalitozoon cuniculi*.

# Cluster of Orthologous Groups of Proteins (COG)

Old COG URL- <http://www.ncbi.nlm.nih.gov/COG/>

New COG URL- <http://www.ncbi.nlm.nih.gov/COG/new/>