

What is 'molecular evidence'?

All living organisms share fundamental molecules essential for life: DNA (deoxyribonucleic acid), RNA (ribonucleic acid), and proteins. In order to understand the evolutionary history of organisms, scientists compare these molecules between life forms, and also study patterns of change.

DNA analysis

Structure of DNA is conserved in all life forms: Archaea, Bacteria and Eukaryotes, providing strong evidence for evolution of all life from a single common ancestor.

Genetic information is encoded in DNA as a sequence of nucleotides (base pairs): adenine, thymine, cytosine and guanine (represented by the letters A, T, C, G). Over time, mutation results in nucleotide changes to the DNA code.

By comparing DNA of different organisms it's possible to document genetic change over time. Counting the number of nucleotide differences between species, in a segment of DNA, provides information on how long ago these species diverged from a common ancestor. The longer two species have evolved separately, the more nucleotide differences will be observed.

What type of DNA?

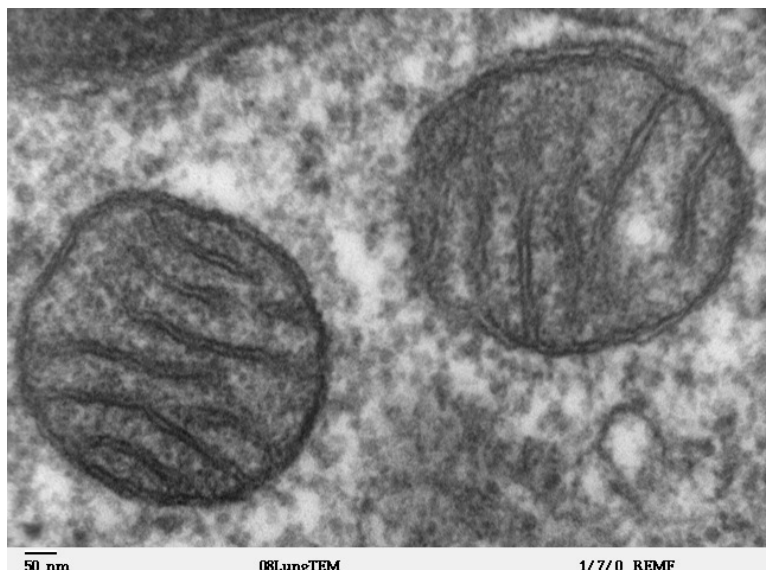
DNA found in the nucleus of cells is called nuclear or nDNA. DNA is also found in mitochondria where it's called mitochondrial or mtDNA. Both are used in evolutionary studies.

	nDNA	mtDNA
protein-coding genes	approx 20 000	13
structure	packaged in chromosomes	circular
location	nucleus	mitochondria
mutation rate	low	high

Mitochondria are cellular organelles involved in respiration. Each cell has approximately 100 mitochondria, which means there is a lot of mtDNA in each cell. For this reason mtDNA is frequently used in forensic or archaeological studies.

Choice between nDNA and mtDNA in evolutionary studies depends on the genes that are being studied. nDNA codes for many more genes than mtDNA so a greater choice of candidate genes is available. mtDNA has been used in much evolutionary research because while there is only a small amount of DNA in each mitochondrion, the large number of mitochondria means there is lots of mtDNA in a cell.

Differing mutation rates between the two types of DNA, and the fact that mtDNA is inherited from the maternal line only, may lead to apparently contradictory results when studies are compared.



TEM image of mammalian lung tissue showing mitochondria, by Louisa Howard public domain, remf.dartmouth.edu/images/mammalianLungTEM/source/1.html

Protein analysis

Amino acid sequences of proteins are compared to determine species' evolutionary histories. For instance, analysis of the amino acid sequence for beta-globin, a subunit of the protein haemoglobin, shows a single difference between humans and gorillas, but over twenty amino acid differences between humans and horses. From this we can infer that humans are more closely related to gorillas than horses.

Table 2: This comparison of beta-globin amino acid sequences for human, gorilla and horse highlights differences from the human sequence. Amino acids are indicated by their single-letter symbols.

BETA-GLOBIN AMINO ACID SEQUENCE		
human	gorilla	horse
MVHLTPEEKS AVTALWGKVN	MVHLTPEEKS AVTALWGKVN	VQLSGEEKAA AVLALWDKVN
VDEVGGEALG RLLVVYPWTQ	VDEVGGEALG RLLVVYPWTQ	DEEVGGEALG RLLVVYPWTQ
RFESFGDLS TPDVAMGNPK	RFESFGDLS TPDVAMGNPK	RFFDSFGDLS NPGAVMGNPK
VKAHGKKVLG AFSDDLHL	VKAHGKKVLG AFSDDLHL	VKAHGKKVLH SFEGGVHLD
NLKGTFATLS ELHCDKLHVD	NLKGTFATLS ELHCDKLHVD	NLKGTFATLS ELHCDKLHVD
PENFRLLGNV LVCVLAHHFG	PENFRLLGNV LVCVLAHHFG	PENFRLLGNV LVVLAHHFG
KEFTPPVQAA YQKVVAGVAN	KEFTPPVQAA YQKVVAGVAN	KDFTPELQAS YQKVVAGVAN
ALAHKYH	ALAHKYH	ALAHKYH

Which genes are best for analysis?

Sequencing the full genome of an organism provides the most comprehensive information about evolutionary history, but this is time consuming and expensive. Instead, scientists decide which genes they'll study. Criteria used when selecting a specific gene include biological function and rate of evolutionary change.

Genes involved in vital cellular processes, such as respiration, are often used in evolutionary analysis because they're present in most organisms. These are commonly termed housekeeping genes. Scientists may also target genes with a known biological function. This approach is particularly useful in studies of disease.

Why cytochrome *b*?

The cytochrome *b* gene codes for a mitochondrial protein functional in cellular respiration and found in mtDNA. It's a housekeeping gene, present in many organisms where it performs a vital cellular function. This makes it an excellent candidate gene for analysis.

Also, due to its sequence variability, cytochrome *b* is commonly used in evolutionary studies. It displays a high rate of mutation, and is considered most useful in resolving evolutionary relationships at the family or genus level.

Another advantage of cytochrome *b* is that data for many species are available, particularly mammals, enabling a large number of comparisons.



Image of cytochrome bc1 complex from the RCSB PDB (www.pdb.org) of PDB ID 1BCC (Zhang Z., Huang L., Shulmeister V., Chi Y., Kim K., Hung L., Crofts A., Berry E. and Kim S. (1998). Electron transfer by domain movement in cytochrome bc1. *Nature* 392, 677-684)