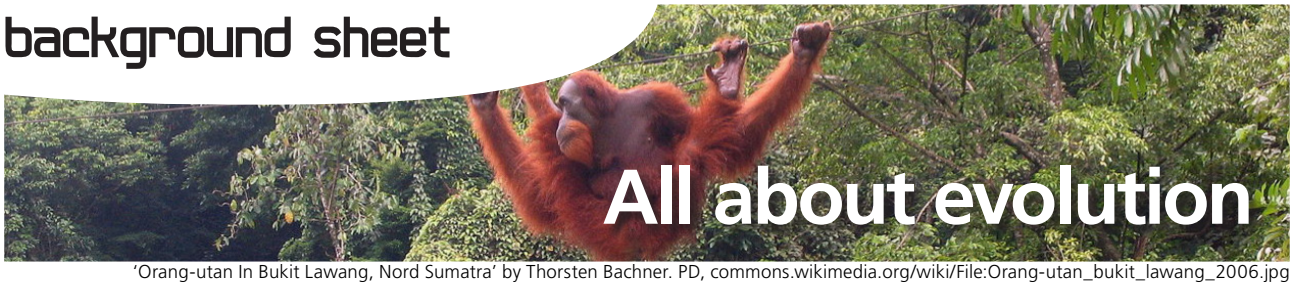
background sheet



All about evolution

'Orang-utan In Bukit Lawang, Nord Sumatra' by Thorsten Bachner. PD, commons.wikimedia.org/wiki/File:Orang-utan_bukit_lawang_2006.jpg

Evolution has long been observed from the basis of fossils, homologous and vestigial structures and embryology. However, the power of DNA technologies now enables us to see evolution in action.

'The DNA sequence of each species is a complete record of the present … [it] is also a window into the recent and deep past.' (Carroll, 2006, p33)

'Every evolutionary change is recorded in the DNA.' (op cit, p14)

Evolution

Evolution is based on three main concepts: variation, mutation and natural selection.

VARIATION	MUTATION	NATURAL SELECTION
Within a species there's always variation between individuals. We aren't exact clones of each other. As genes are passed from parents to offspring, recombination, or errors in the replication process, lead to variation in genes. As individuals are unique, variation or genetic difference allows, for example, forensic identification.	Variations arise as a result of mutations, which are errors in DNA replication. Mutations are random and are not directed; they occur by chance. Mutations or changes in DNA sequence may be harmful, beneficial or have no effect. If a mutation has no impact on survival or reproduction it won't exert evolutionary pressure. (Gregory, 2009) *'Mutation generates random variation, selection sorts out the winners and losers.' (Carroll, 2006, p57)*	Organisms generally have to struggle or compete to survive. Some organisms are more likely to succeed and produce offspring as their genetic variations (traits) make them more suited to a particular environment. While generation of mutations, or genetic variation, is random, selection is not. Fitter traits, or genes, increase in proportion, over time. However, if the environment changes, other individuals with different genetic variations may be more successful, and be 'selected'.

Table 1: comparison of variation, mutation and natural selection

Using DNA

A DNA sequence can be considered the 'fundamental text for evolution.' (Carroll, 2006, p26) Every evolutionary step is a result of random changes in DNA sequence. These changes are expressed in the phenotype that's selected. With new technologies these genetic changes can be traced, and differences and similarities in DNA sequences of different organisms, plotted.

Scrutiny of DNA sequences reveals insights into relatedness of organisms. For example, there are approximately 500 genes that all organisms must have to function. These genes are necessary for key cellular processes. Therefore, any mutation that renders a key gene inoperable is not selected, and not passed on to offspring because they'd have zero reproductive success.

What type of DNA?

Much work in relation to evolution is a comparison of DNA between different species. There are potentially many thousands of genes to compare. Thus, the question arises: which type of DNA and which gene or sequence should be used for comparison?

Mitochondrial DNA and nuclear DNA

Full genome sequencing requires significant time, cost, effort and computing power. For example, the full gorilla sequence contains 3 041 976 159 base pairs, with 20 962 protein-coding genes and 6701 RNA genes. (Scally et al, 2012)

Evolutionary studies require the full gorilla sequence to be aligned with sequences of human, chimpanzee and orang-utan, and differences and similarities plotted. Serious computing!

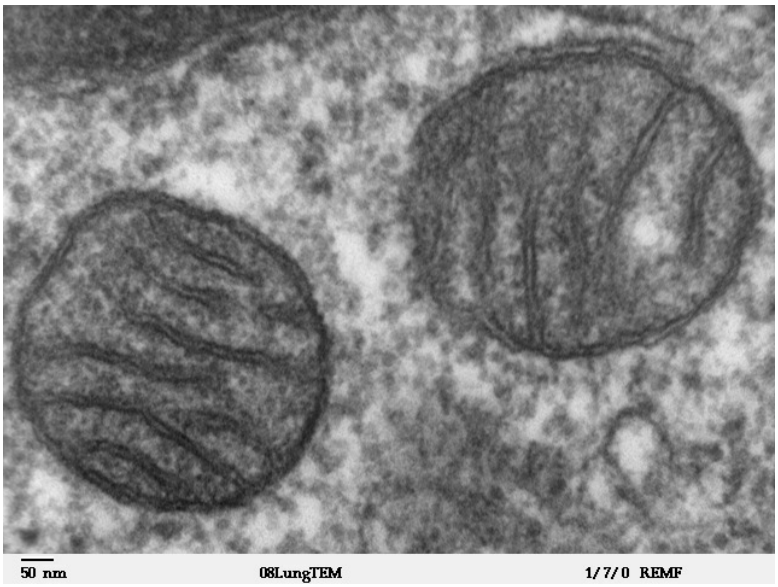
Mitochondrial DNA (mtDNA), as opposed to nuclear DNA, has been used in much evolutionary research because, while there is a smaller amount of DNA in mitochondria to manipulate, there's lots of it in a cell. Each cell has approximately 100 mitochondria and each mitochondrion has 5–15 copies of the mtDNA genome. (Ingman & Gyllensten, 2001)

mtDNA has: a smaller number of genes; a uniparental mode of inheritance (maternal); and both rapid and slow evolving regions. In addition, there's a 5–10 times higher mutation rate in mtDNA compared to nuclear DNA. (Melnick & Hoelzer, 1993)

This is attributed to lack of mitochondrial histones (packaging molecules which are protective of DNA), plus a high concentration of oxidative radicals. (Ingman & Gyllensten, 2001)

Melnick & Hoelzer (1993) found the higher mutation rate in mtDNA causes greater genetic variation than is found in nuclear DNA, so evolutionary change can be seen more clearly. The mtDNA genome is circular and encodes for 13 proteins, 22 tRNA genes and 2 rRNA genes. (Ingman & Gyllensten, 2001)

As there's a lot of mtDNA in each cell it can be copied from many different tissue types, including quite degraded material, making it suitable for forensic and archaeological studies.



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

There are limitations in using mtDNA compared to nuclear DNA. Firstly, mtDNA can only ever reflect maternal trends. This may potentially produce a skewed distribution of genetic variation if, for example, females stay in social groups while males leave. (Melnick & Hoelzer, 1993)

Indeed, the high rate of genetic variation found in mtDNA may poorly reflect nuclear DNA variation. If used in evolutionary studies this may distort the construction of phylogenies.

Candidate genes

Another approach is to select a specific gene to scan, then compare. The candidate gene's selection is usually based on what 'variation of phenotype' (Zhu & Zhao, 2007) is under investigation. For example, disease genes or genes associated with economic traits, such as wool strength. The potential candidate for research would be a gene (or genes) with known biological function (direct or indirect action). The disadvantage is obvious: pre-existing knowledge is needed or the researcher must spend large amounts of time 'searching for a needle in a haystack'.

Housekeeping genes are genes usually involved in some form of cellular maintenance, or a key process, and are generally expressed at a constant level in most tissues. As these genes are usually highly conserved, differences in their sequences can be used in evolutionary studies. For example, in the learning object, *Building evolutionary trees*, a mtDNA gene, cytochrome *b*, is analysed. Cytochrome *b* (gene) codes for a mitochondrial protein with a functional role in respiration (electron transport chain).

Genome sequencing

Full genomic sequencing analyses the entire genome of an organism and is currently the gold standard in terms of evolutionary analysis. The 'whole genome shotgun approach' (WGS) involves fragmenting or cutting up DNA, cloning fragments, reading fragment sequences, and then reassembling fragments into the correct order. While this sounds straightforward there are many challenges. Full genome sequencing is costly, both in terms of time and computer power.

Visualising data

Following sequencing, whether a fragment, a gene or full genome, sequences are aligned and related regions assessed. Alignment aims to arrange sequences, usually represented as a series of rows, so that differences or similarities between sequences can be identified. (Roy, 2009)

Alignment programs use various mathematical algorithms that allow sequence data to be imported and aligned. A computer program 'slides' sequences past each other until sections with a good match are found.

Figure 2 shows two similar sequences. However there are two extra nucleotides in sequence 1. Alignment of sequence 1 and 2 requires two 'gaps' or 'indels' (insertions or deletions) to be inserted into sequence 2.

1	TGACTCCAAAAGGCTGGCGGCTTA
2	TGACTCAAAAGGCTGCGGCTTA

(a)

1	TGACTCCAAAAGGCTGGCGGCTTA
2	TGACTC AAAAGGCTG CGGCTTA

(b)

Figure 2: comparison (a) and alignment (b) of two DNA sequences

An alignment program such as BLAST (basic local alignment search tool) finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases, and calculates the statistical significance of matches. The relatedness of sequences are then visualised through construction of a phylogenetic tree. If sequences are homologous (they have the same pattern) it is assumed they are related through a common ancestor.

Molecular clocks

The 'molecular clock' is a concept that assumes a roughly constant mutation rate, and hence a constant rate of evolution. The idea was first proposed in 1962, when Zuckerkandl and Pauling attempted to date origins of different haemoglobin proteins. (Kumar, 2005)

From the earliest days of this concept there were concerns regarding accuracy and general applicability. Fossil records are often used to compare molecular clock estimates.

While the concept can provide time estimates there are obvious problems in comparing different organisms. For example, species with longer generational time go through fewer replications per unit of time, and would therefore expect to accumulate fewer mutations. Humans take twice as long to reach sexual maturity as chimpanzees (Elango et al, 2005), and have an approximately 20% slower mutation rate than other Hominidae. (Kumar, 2005)

Despite limitations, the molecular clock concept provides a framework for estimating species divergence. In addition it's now well known that different genes and types of DNA evolve at different rates. Mitochondrial DNA mutates at a much faster rate than nuclear DNA, while some highly conserved genes (eg housekeeping proteins) show slow mutation rates. Therefore, analysis of different types of sequences, accompanied by appropriate background information regarding mutation rate, can offer insights into diverse aspects such as closely related species, or earlier divergence of life. (Kumar, 2005)

Misconceptions about evolution

Common ancestor

A common misconception is that 'humans evolved from chimpanzees'. If this were the case, there'd be no chimpanzees, as they would have evolved into humans. Chimpanzees are considered our closest relatives and, at some stage, estimated to be 6 million years ago, chimpanzees and humans shared a common ancestor. Four million years before that, gorillas shared a common ancestor with the human-chimpanzee common ancestor. (Scally et al, 2012)

Teleology

Teleology refers to purpose-based explanations of natural phenomena, that children use, to help make sense of their world. Kelemen and Rosset (2009) however, discuss the possibility that teleological explanation is not just part of childhood, but may persist throughout life.

Teleological explanations for natural phenomena are based on a function or purpose. That is, for a teleological thinker, change occurs in response to a need. This relates to the misconception that individuals evolve in response to environmental challenges. However, changes based on a perceived need, experience or internal desire, don't occur.

Gregory (2009) provides an example of a poor explanation of evolution (a teleological explanation) by a trusted authority, National Institutes of Health, that perpetuates this misconception:

As microbes evolve, they adapt to their environment. If something stops them from growing and spreading — such as an antimicrobial — they evolve new mechanisms to resist the antimicrobials by changing their genetic structure. Changing the genetic structure ensures that the offspring of the resistant microbes are also resistant.

The above text implies that microbes decide they need to change their genetic structure so they'll maintain antibiotic resistance. However, as a result of random mutation, microbes have a changing genetic structure. If environmental conditions change, some individual microbes will be favoured, and some will be disadvantaged. Over time, genetic variations that provide reproductive advantage will be selected.

Evolutionary change is based on genetic mutation that occurs by chance.

Evolution is based on use, or disuse, of body parts

This misconception stems from a 'use it or lose it' view of evolution. Students think that if a particular body part is not used, it will disappear through the process of evolution. For example, students may believe that animals living in darkness no longer have eyes because they stopped using them. This is not the case.

Allied to the misconception 'evolution is based on the use, or disuse, of body parts', is the idea that characteristics acquired through life, for example strength through training, are passed onto offspring.

> Random mutation of DNA sequences in 'eye' gene(s) produces genetic variation. If having 'eyes' results in reproductive success, the eye phenotype would be advantageous and therefore selected. However, if you live in darkness and you have eyes it's not necessarily advantageous.
>
> If over time there was a random mutation that resulted in animals with no eyes, it may be that having no eyes would not affect reproductive success. If all animals in the cave have no eyes, then the no-eye phenotype is clearly successful and has been selected.

This is the Lamarckian view of evolution.

Language

The language of evolution is problematic in that some scientific terms, such as: 'fitness', 'adaptation' and 'evolve', have different everyday meanings.

The term 'survival of the fittest' appears to refer to the survival of the strongest rather than survival of those organisms that are best suited to the environment (due to their genetic variations). However, fitness, in an evolutionary sense, refers to 'a measure of the total reproductive output of an organism with a particular genotype.' (Gregory, 2009)

Put more simply, fitness means reproductive success. In a scientific context, fitness does not refer to strength or physical conditioning.

'Adaptation' is a word in everyday language that usually refers to an ability to change to an environment. We speak of adapting to a situation or environment, by choosing or deciding to do something.

'Evolve' is also a word in common use, that is, used in the context of changing or developing, often from simple to complex.

When used in evolutionary terms, 'adaptation' refers to a process where populations (not individuals) evolve a particular phenotype, through natural selection, that helps them survive in their environment.

Hominidae (great ape) evolution

Taxonomic classification traditionally placed humans in their own family, Hominidae, with other apes in the Pongidae family. (Simpson, 1945; Goodman et al, 1994)

However, phylogenetic analysis, made possible through advances in DNA technology, has established a different evolutionary relationship between humans and living apes.

It is now firmly established that humans (genus *Homo*) are most closely related to chimpanzees and bonobos (both belong to genus *Pan*). *Pan*'s relationship to genus *Homo* is closer than it is to genus *Gorilla*, who is their next closest relative. As a consequence, *Homo*, *Pan* and *Gorilla* are now grouped together in the sub-family Homininae. Orang-utans are the sole member of Pongidae.

Names of the groups are extremely confusing and can vary, depending on views of different researchers. For example, Table 2 shows how Goodman and colleagues (1994) classified members of Hominidae, based on the sequence of non-coding regions of β-globin gene cluster. It also shows a more recent classification by Wood and Harrison (2011) that used analysis of mitochondrial and nuclear genes.

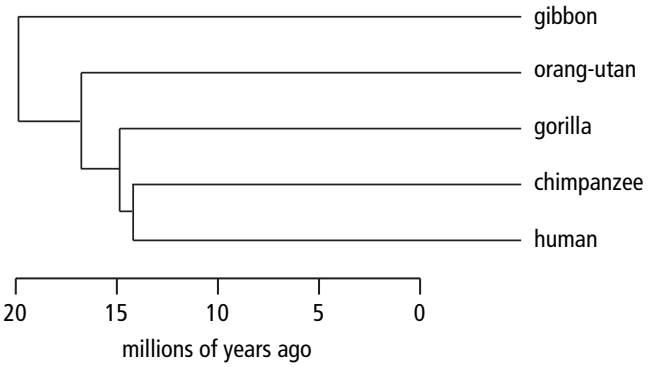


Figure 4: Hominoidea evolutionary tree

	1994	2011
Superfamily		
	Hominidae	Hominoidea
Family		
		Hylobatidae (gibbon)
		Hominidae (hominids includes all modern and extinct apes)
Subfamily		
	Hylobatinae (gibbon)	Ponginae (orang-utan)
	Homininae	Homininae (human, chimpanzee, gorilla)
Tribe		
	Pongini (orang-utan)	Gorillini — *Gorilla*
	Hominini	Panini — *Pan* (chimpanzee, bonobo)
		Hominini — *Homo* (modern and extinct humans)
Subtribe		
	Gorillina	
	Hominina	
	Pan	
	Homo	

Table 2: Hominidae/Hominoidea classification

Three of our closest relatives (chimpanzee, gorilla and orang-utan) have had their full genome sequenced, enabling systematic identification of similarities and differences. Varki and Altheide (2005) found a difference between chimpanzee and human genomes of approximately 4%, rather than the oft-quoted 1%, an estimate that was made using shorter DNA fragments. Proteins, however, were found to be similar between humans and chimpanzees, with 30% identical in structure and most proteins only differing in one or two amino acids.

Differences in DNA sequence, between genomes of different animals, don't provide any understanding about evolutionary change; basically you need something with which to compare. Orang-utans are close relatives of chimpanzees and humans. Locke et al (2011) recently estimated that Homininae diverged from Hominidae about 12–16 million years ago. This can provide a basis for determining what lineage (chimpanzee or human) experienced the change.

The full genome of the gorilla has also recently been sequenced by Scally et al (2012). Speciation is estimated to have occurred 10 million years ago. Interestingly, the researchers found that 15% of human genes are more like gorilla genes than chimpanzee genes, raising yet more questions regarding great ape evolution.

References

Carroll, S.B. (2006). *The making of the fittest: DNA and the ultimate forensic record of evolution*. W.W. Norton & Co: New York.

Elango, N., Thomas, J.W., NISC Comparative Sequencing Program and Soojin, V.Yi. (2006). Variable molecular clocks in hominoids. *Proceedings National Academy of Sciences, 103*(5), 1370-1375.

Goodman, Bailey, Hayasaka, Stanhope, Slighton and Czelusniak. (1994). Molecular evidence on primate phylogeny from DNA sequences. *American Journal of Physical Anthropology 94*:3-24.

Gregory, T.R. (2009). Understanding natural selection: essential concepts and common misconceptions. *Evolution Education and Outreach, 2*, 156-175.

Ingman, M. & Gyllensten, U. (2001). Analysis of the complete human mtDNA genome: methodology and inferences for human evolution. *The Journal of Heredity, 92*(6), 454-461.

Kelemen, D. & Rosset, E. (2009). The human function compunction: teleological explanation in adults. *Cognition, 111*, 138-143.

Kumar, S. (2005). Molecular clocks: four decades of evolution. *Nature Reviews Genetics, 6*, 654-662.

Melnick, D.J. & Hoelzer, G.A. (1993). What is mtDNA good for in the study of primate evolution? *Evolutionary Anthropology, 2*(1), 2-10.

Roy, D. (2009). *Bioinformatics*. Alpha Science International Ltd: Oxford.

Scally, A., Dutheil, J.Y., Hillier, L.W. et al (2012). Insights into hominid evolution from the gorilla genome sequence. *Nature, 483*,169-175.

Varki, A., & Altheide, T. K. (2005). Comparing the human and chimpanzee genomes: searching for needles in a haystack. *Genome Research, 15*(12), 1746-1758.

Zhu, M. & Zhao, S. (2007). Candidate gene identification approach: progress and challenges. *International Journal of Biological Sciences, 3*(7), 420-427.

'chimpanzee at Whipsnade Zoo' by William Warby. CC-BY-2.0, www.flickr.com/photos/wwarby/5026590985/in/photostream/

'gorilla tales 6' by William Beem. CC-BY-NC-2.0, www.flickr.com/photos/wbeem/469121925/

'orangutan time 1' by Darren Deans, www.darrendeans.com, used by permission