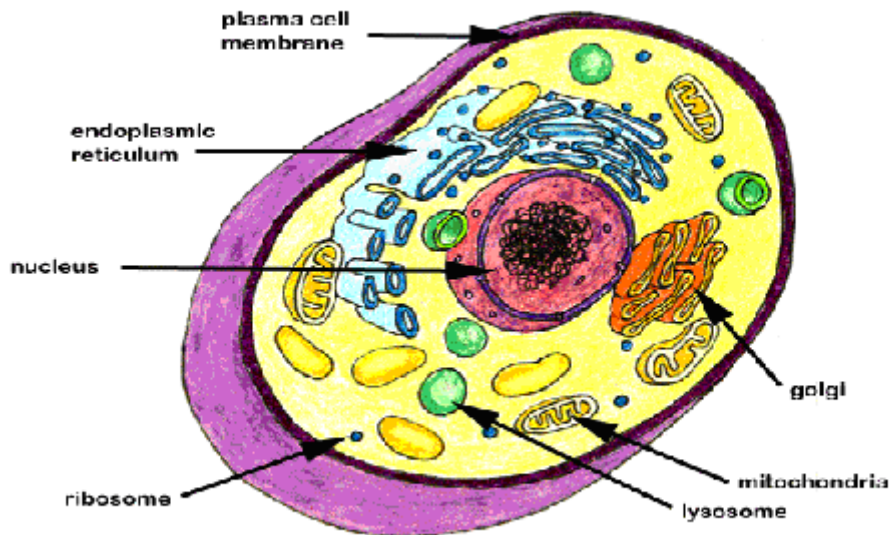


Analysis of mitochondrial DNA (mtDNA)

The human cell consists of a core surrounded by a fluid (cytoplasm) in which many cells and organelles are located. The liquid is surrounded by a membrane. An important organelle is the mitochondrion. In each cell are certainly hundreds and possibly thousands of mitochondria present. The mitochondria provide the energy of the human body.



The nucleus of every cell contains the DNA, the genetic material of humans.

DNA (desoxyribo nucleic acid) is the large collection of genetic information in the cell. A DNA molecule consists of two strands in the form of a double helix linked by weak base pairs of nucleotides, each two of four bases: adenine (A), cytosine (C), guanine (G) and thymine (T). The sequence of bases in DNA is the code for specific information about the genetic blueprint of the individual.

Part of a strand looks like this: **gatcacaggctatcacctattaaccactcacg**

DNA is found not only in the core of the cell but also in a mitochondrial. We cut it down to mtDNA.

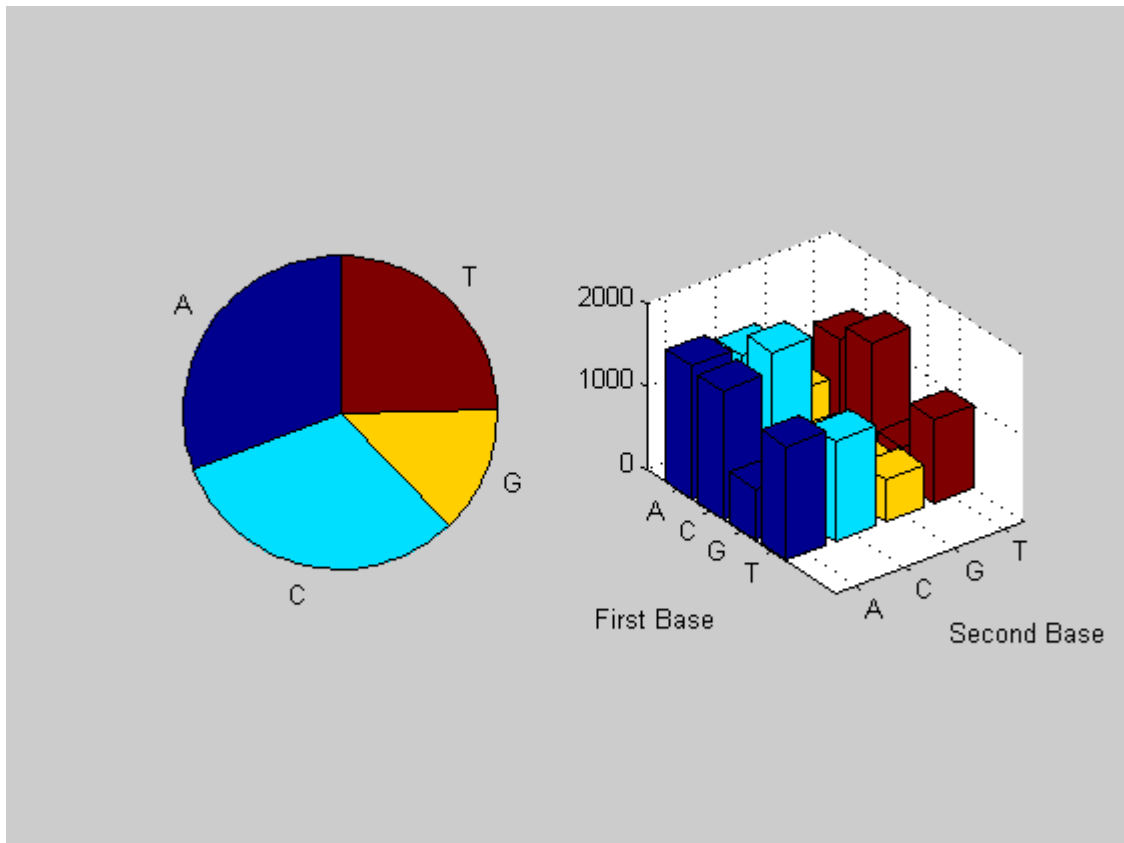
See: <http://www.mitomap.org>

The 'Human Mitochondrial Genome' consists of 16,568 nucleotides (base pairs).

An analysis of the total quantity provides the following breakdown:

A (5124 = 30.9%), C (5181 = 31.3%), G (2169 = 13.1%) and T (4094 = 24.7%).

Map of breakdown mtDNA (first and second Base)



With an equal distribution of the four bases every 4142 base (25%) would occur once. T is true that about. A and C are each approximately 6% more than average and G is almost 12% less.

The question is how these differences with an equal probability distribution can be explained.

The following analysis concerns the question what is the maximum number of successive equal bases and how often such a string is present. The maximum number of consecutive bases is equal 8. This occurs one time and concerns A.

The list is as follows.

	8 times	7 times	6 times
A (adenine)	1	7	12
C (cytosine)	0	1	10
G (guanine)	0	0	1
T (thymine)	0	1	4

Although the numbers of A and C are almost equal (C is even slightly more than A) it is particularly striking that a string of 7A much more frequently occurs (7 times) than a string of 7C (1).

The question is how these differences can be explained.

Finally, an analysis of the committed number of times a particular combination of four different bases occur.

With an equal distribution of A, C, G and T is the probability of 1 of 4 (ie the probability of A) is $\frac{1}{4}$. The chance of drawing eg ACGT will be: $\frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = 1 / 256$

Of a total of 16,567 bases, the number of times eg ACGT occurs on average $1 / 256 \times 16\,567 = 64$ times.

Four elements have 4! different combinations. $4! = 4 \times 3 \times 2 \times 1 = 24$

The frequency of reality is as follows.

ACGT	21	ATGC	37	AGTC	38	ACTG	39	ATCG	46	AGCT	63	Totaal	244
CGAT	23	CGTA	30	CATG	48	CAGT	49	CTGA	69	CTAG	97	Totaal	316
GATC	23	GTCA	32	GTAC	35	GACT	49	GCAT	56	GCTA	64	Totaal	259
TCGA	29	TACG	37	TGCA	37	TCAG	38	TGAC	54	TAGC	95	Totaal	290
	96		136		158		175		225		319		1109

Again the question is, how it is possible that such differences may arise in actual presence of a certain combination compared with the theoretical possibility.

Solution.

To all appearances, there is a Fractal distribution. In two articles the Fractal Analysis of mitochondrial DNA are discussed.

http://biocomplexity.indiana.edu/jglazier/docs/student_dissertations/cheryl_thesis.pdf

<http://www.fractal.org/Fractal-Research-and-Products/Fractal-Structure-mtDNA.pdf>

A fractale visualisatie of the mtDNA can be find:

<http://www.mathworks.com/matlabcentral/forums/21986/1/content/nmervisdemo.html>

