

What we may gain from a Fly – Diseases, Genomes and Genome Projects

In April this year researchers of the private firm Celera announced having finished the first step towards sequencing a complete human genome. In the years to come, nearly the complete genetic information of a human being will be available to some researchers. Disputes are going on with respect to granting patents and making the information available to the public, but here we shall not focus on legal issues. We will discuss possible scientific developments based on the knowledge of the human genome within the next years and see the impact a small fruit fly might have in this area.

It happened widely unnoticed by the public in March this year that researchers of the private firm Celera and the Berkeley Drosophila Genome Project jointly published the complete genome of the fruit flyⁱ. The fruit fly is one of the best-studied organisms to date and having elucidated the complete DNA sequence of *Drosophila melanogaster* is the successful end of more than 80 years of *Drosophila* research. This genome might have some impact on human disease research and could be, for the next years, even more important than the human's own genome.

The problem in using information human DNA provides

This problem lies in the incomplete understanding of the expression of proteins from the information contained in DNA. The information DNA conveys is a four-letter code of

bases, of adenine (A), thymine (T), cytosine (C) and guanine (G). Combinations of these bases encode amino acids – the sequence GGC for example codes for the smallest amino acid glycine. Those amino acids in turn are used to form bigger compounds - between a few dozens and many hundreds of those amino acids are linked to form a protein molecule. Proteins fulfil a large variety of functions in an organism; examples for those proteins are the family of haemoglobins, which are involved in human oxygen binding and transport, or the insulins, which regulate the level of blood sugar with the consequence of diabetes if the body-internal synthesis is not working correctly.

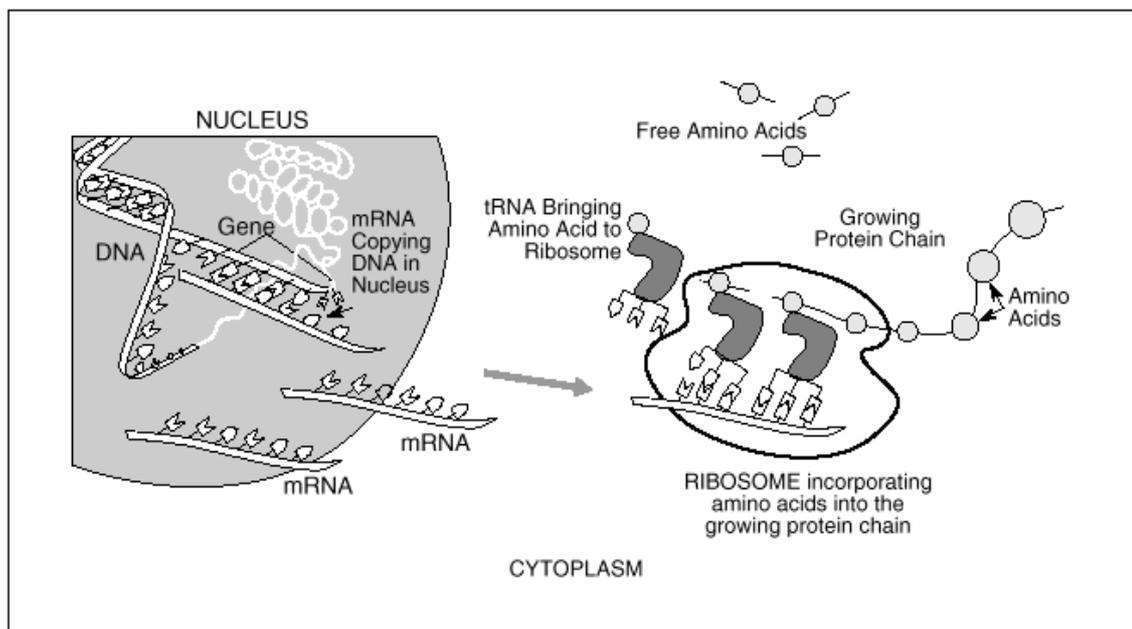


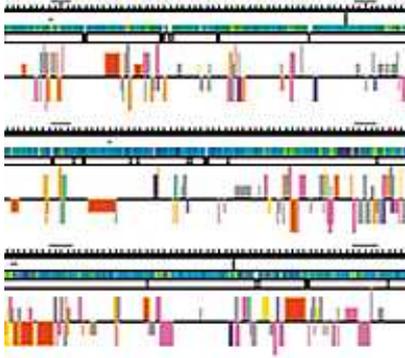
Fig. 5. Gene Expression. When genes are expressed, the genetic information (base sequence) on DNA is first transcribed (copied) to a molecule of messenger RNA in a process similar to DNA replication. The mRNA molecules then leave the cell nucleus and enter the cytoplasm, where triplets of bases (codons) forming the genetic code specify the particular amino acids that make up an individual protein. This process, called translation, is accomplished by ribosomes (cellular components composed of proteins and another class of RNA) that read the genetic code from the mRNA, and transfer RNAs (tRNAs) that transport amino acids to the ribosomes for attachment to the growing protein. (Source: see Fig. 4.)

But which part of DNA contains the information for which gene and therefore accounts for which functions within an organism is widely unknown. Parts of the DNA are *non-coding* and hence unessential “heterochromatin”, they will never be expressed as proteins– but some genes are hidden within those areas. The question is: Where? Another problem is that only some parts of the DNA within coding regions - the so-called “exons”- will be used to build up proteins, other parts will be removed in a process called “splicing”. But which parts? If those problems are overcome, one will recognize which parts of the DNA will form proteins in the living cell. But the next question is already looming – which functions do these proteins fulfil?

Common approaches are based on the idea that DNA sequences are conserved among similar organisms, that means for example that the haemoglobin of apes closely related to the humans should have a DNA sequence similar to the DNA encoding human haemoglobin. Computer algorithms try to align DNA sequences in order to find those similarities, and the more similar sequences are the more likely is it that they have similar functions. This “sequence aligning” approach proved to be successful to a certain degree, especially due to increasing computational power. But still the functions of most sequences remain unknown.

I am sure that most of these problems will be overcome within the next decades, that DNA sequences, structures of the encoded proteins and their functions will finally be known. But that is – for the vast majority of genes – not where we are today.

The advantages of the fruit fly



A part of the DNA of *Drosophila* – the whole genome is freely available to the public at <http://www.fruitfly.org>

“In terms of evolutionary conservation of sequence similarity, *Drosophila* is the closest of the invertebrate model organisms to humans. Moreover, in terms of morphological, physiological, and behavioural complexity *Drosophila* is by far the closest to humans of these model organisms, yet its genome is not substantially bigger than the least complex metazoans (= multicellular animals with specialized cell types).”ⁱⁱ

Similarities between humans and *Drosophila* range from Aging and Cancer to Alzheimer’s and Parkinson’s

I want to focus on similarities in aging, cancer and neurological diseases, but that it is just an extract of the stunning degree of relatedness between *Drosophila*’s and the human genome.

In a healthy body there exists a balance between cell proliferation and cell death maintained by apoptosis, the so-called “cell suicide”. An oversupply of apoptosis is presumed in case of HIV infections, where T-cells, a part of the human immune system,

might commit suicide because of a message sent by the HIV infected cells. Also an insufficient degree of apoptosis is dangerous and might lead to an uncontrolled cell growth, because hazardous cells no longer receive the order to terminate their own existence. This system of apoptosis, the suicide for the body's dangerous cells, is found in humans as in *Drosophila*. Both humans and the fruit fly also have in common a tendency to develop cancer as a result of mutations of genes controlling the cell cycle (although many of those mutations are repaired or the cells commit suicide, see above). It is believed that 67 percent of known human cancer genes have related genes in the fly (the overall gene conservation is estimated at 60 percent).

One of the most advanced research areas is the field of neurological diseases such as Alzheimer and Parkinson'sⁱⁱⁱ, a field where the similarities between human and *Drosophila* are stunning. The fly shows the same symptoms of Parkinson's as humans - tremors and loss of motor coordination. Mel B. Feany and Welcome W. Bender of Harvard Medical School report that flies bred with the human gene "alpha-synuclein," show symptoms of brain damage that not only mimic those in humans but also appear to be closer in the fly than those used in mouse models of Parkinson's. Therefore, flies could become an ideal model for testing new anti-Parkinson's therapy quickly and effectively, especially considering the fly's short 60-day lifespan. Mel B. Feany has already discovered some genes that suppress the onset of the disease by about 10 days. Also other genes that are known to prevent the formation of "Lewy bodies", abnormal lumps inside nerve cells in Parkinson patients in humans as in *Drosophila*, have been

discovered. The crucial questions are now, how to influence these genes and the transferability to human beings.

A possible Explanation for the Similarities – the Frugality of Nature

Human beings, without doubt, have more genes than *Drosophila* (about 80,000 compared to 13,000) but Gerald M. Rubin, head of the Berkeley *Drosophila* Genome Project, adds: "About 60 percent or more genes are conserved between fly and human. Complexity does not come from the number of genes but from the way in which they are used. Humans may have four copies of a gene where the fly has one, but if you look at the core proteome - the core set of parts - they're not that different. The fly has learning and memory and behavior, it's just not as complex as ours." He likes the analogy to supercomputers and desktop PCs - the human is a supercomputer to the PC fly. "It's an organizational issue. The parts are basically the same."

Therefore, the fruit fly offers a powerful simulation tool to disease researchers. Let's see what we will gain from it!

ⁱ Adams, M.D. et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185-2195
"complete" in this sense means that 99.7 percent of known genes are accounted for in the published data

ⁱⁱ Rubin, G.M. (1996) Around the genomes: The *Drosophila* genome project *Genome Research* 6, 71-79
freely available for the public at <http://www.fruitfly.org/methods/>, the website of the Berkeley *Drosophila* Genome Project

ⁱⁱⁱ Feany, M.B., Bender, W.W. (2000) A *Drosophila* model of Parkinson's disease. *Nature* 404, 394-398

Written by Andreas K P Bender for the International Science Writer Award 2000.

Contact: mail@andreasbender.de or bendera@tcd.ie.