GROUNDBREAKING DISCOVERIES IN SCIENCE AND TECHNOLOGY

DNA – The Essence of Life

Magnus Merscher
2. Semester Bachelor Wirtschaftsingenieurwesen/Maschinenbau SoSe2010
s762342
magnus.merscher@googlemail.com

All parts of this work have been edited only by myself. The text has not been published or used otherwise in any way. All foreign trains of thought and sources have been cited. Please note that all images without citation are the work of the author.
Content

1. Introduction 1

2. DNA in a nutshell 1
   1. Appearance and basic functionality in procreation 1
   2. Structure and cell functionality 2

3. The discovery of DNA and the deciphering of its code 3

4. Popular applications 5
   1. Hereditary research 5
   2. Genetic engineering in farming 6
   3. An example of a medical treatment relying on DNA research 8
   4. Organism design and human genetic engineering 9

5. Exotic applications 10
   1. DNA computation 10
   2. Data preservation 11

6. Conclusion 12

7. References 13

8. Additional helpful references 16
1. Introduction

DNA is one of the most important discoveries of the 20th century and fundamentally changed the way we perceive ourselves and the world we live in. The ability to understand the mechanisms of life itself enables us to find new solutions to problems such as heredity, overpopulation, disease, or even computation. Apart from the technical implementation of the technology, the sheer knowledge that we gained through DNA research took us to a new level of self understanding.

This paper is designed to give an overall impression of what DNA is, how it was discovered, and to showcase some of the various applications of our present day, popular and exotic, highlighting the importance of this technology. New technology rarely comes to fruition without opposition, so a short glance is taken on the ethical implications to see whether the benefits outweigh the downsides. Recent developments have been taken into account up to 18 August 2010.

2. DNA in a nutshell

2.1 Appearance and basic functionality in procreation

Deoxyribonucleic Acid (DNA) is a large organic molecule that can be found in every biological organism known today. It functions as heredity information storage inside cells and is therefore critical in the process of procreation. All contained information, called the genome, is passed from one generation of an organism to the next where it leads to a specific physical development, called the phenotype, that does not necessarily match the originating cell. This alternate phenotype may be caused by an alternate interpretation of the genetic code, called the genotype, due to genetic dominance effects (see Figure 1) or by mutation of the code itself. With every shift of generations, both dominant (Y) and recessive (y) genomes are passed on, while only the dominant variant becomes the phenotype when different variants of the parental generations are combined in one offspring; e.g. a Yy genotype becomes a Y phenotype (Orel 1996, 105).

The molecule contains all the genetic information in one large molecule at a time and can be found primarily in two locations inside a cell (if present): the nucleus, the core of the cell that controls most operations, and mitochondria, organelles responsible for producing ATP – a molecule...
functioning as the primary energy source for most cellular processes (Starr and McMillan 2008, 48-52). While mitochondrial DNA (or mtDNA) in humans only passes to the offspring from the female side, nucleic DNA may originate both from the male or female line, offering the opportunity to conclude about genetic connections between separate populations (Bandelt et al. 2006). Its crucial function in every organism suggests the danger of alteration of the genetic code. While it is imperative for species to adapt to new environments to survive, the alteration of certain genes may have disastrous effects on the individual when serious illnesses occur.

2.2 Structure and cell functionality

Technically, a DNA molecule is a single strand of a sugar-phosphate backbone containing one of four nitrogen bases: guanine, cytosine, adenine and thymine. These are the letters of our genetic code: G-C-A-T. They are the same for all life forms. The sugar-phosphate-base compound is the basic unit of all nucleic acids and called nucleotide. When occurring inside a cell, two DNA strands pair together to form the characteristic double helix (which then folds to form a chromosome), where the direction of the bonds in the backbone is opposite in both strands: they run anti-parallel. The carbon backbones coil around each other, while each base has a complement that it is paired with on the other strand. In between the base pairs, hydrogen bonds hold this structure together (Watson 1953, 737). Guanine always pairs with Cytosine (G-C), and Adenine with Thymine (A-T), which is important for the duplication process of DNA and RNA (see below).

While the bonds in the backbones are covalent and therefore relatively strong, hydrogen bonds can be easily broken, which proves to be essential in its function. Molecules inside a cell break the hydrogen bond, read out the genetic code and rewrite it to a related molecule, the ribonucleic acid (RNA). The RNA assembles on one strand of DNA and the amino acids are assembled on the RNA, forming a protein.

![Figure 2: DNA double helix](image)

![Figure 3: The encoding process of a protein from DNA.](image)
the DNA helix and is lining up as its complement. This is executed by molecules called RNA polymerase, and replaces what would be thymine in a complementary DNA strand with uracil in the RNA strand. After RNA synthesis in completed, it is detached from the DNA strand, which coils up to the double helix again. The RNA strand is then interpreted into amino acids in steps of three base pairs, called codons. Finally, those amino acids are chained together to form protein molecules. Since there are 4 base pairs, there are $4^3 = 64$ possible combinations for a codon, or three-letter word, that make up the code necessary for the creation of a protein. In this context, the term 'gene' refers to the entire stretch of code responsible for the creation of a certain type of protein. (Alberts 1998, 184-185)

The old question of the chicken and the egg, or rather DNA and the cell, may be solved as there are indications that RNA may synthesize spontaneously under certain conditions (Szostak 2001), making the forming of the first rudimentary 'proto-cell' a mere result of (complicated) coincidental biochemical processes in the evolution of life.

3. The discovery of DNA and the deciphering of its code

The discovery of DNA is a story of hard and tedious work by not one, but many enthusiastic men and women that can fill scores of books.

The first scientific experiments concerning parentage were conducted by Czech monk Gregor Mendel. He mixed differently colored pea plants and in 1865 discovered that the heredity of an attribute followed distinct rules (see section 2.1, Figure 1). He coined the terms genotype and phenotype to distinguish between hereditary information and outer appearance of an organism, laying the cornerstone of modern genetics (Orel, 1996).

Independently, in 1868 the Swiss physician Friedrich Miescher isolated a new substance from the cells, at first not realizing it was a novelty. With instructions of biochemist Felix Hoppe-Seyler he was analyzing the chemical composition of cells using leukocytes (white blood cells). When adding acid to an extract of those cells, a white substance was deposited in the test tube, which could in turn be dissolved when adding a base, proving it was not a protein. Miescher speculated that it originated from the nucleus of the cells and went to work on a proof for this hypothesis. After long hours of gradually dissolving membranes and cytoplasm with hydrochloric acid and ether in a low temperature environment to prevent decomposition, he chose a new name for his finding: nuclein. In a following composite analysis he isolated carbon, hydrogen, nitrogen and phosphorus, but no sulfur, relieving himself of the last doubts that his discovery was genuine (Lagerkvist, 1998).

In 1929, Phoebus Levine identified the building blocks of DNA including the four bases adenine,
cytosine, guanine and thymine. His discovery of the presence of two types of sugar in nucleic acids, ribose and deoxyribose, led to the modern names or ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). The term 'nucleotide', the basic unit of all nucleic acids, was his creation to describe a three part construct: a phosphate group, a sugar, and a nitrogen base. (Newton, 2009, 156)

American scientist Oswald Avery connected DNA and heredity in 1944 when he and his team managed to transfer the sickening features of one kind of bacteria to another, which has then been able to pass this on to the next generation, and proved that the transferred features had been stored inside the DNA. In 1928, Frederick Griffin had already conducted a similar experiment, but was unable to pinpoint the source of the 'transforming principle' (Shmaefsky 2006, 150).

Having no other way to observe a DNA molecule directly, a group of scientists made use of the available but highly complicated X-ray diffraction photography to gain hints at the possible structure of the molecule. The legendary “Photograph 51” finally proved a helical structure, and working with plastic models, the group around J. D. Watson and F. H. C. Crick (1953, 737 and 2001, 202) deduced a double helix in the molecule inside cells, as opposed to a suggested triple helix. It should be noted that shortly before Watson and his team published their discovery, Rosalind Franklin and Maurice Wilkins achieved the same results of X-ray diffraction, but wanted to further confirm their findings before publication. Watson and Crick saw the results and used them to deduct the final step in their research (Newton 2009, 7).

James D. Watson, Francis Crick and Maurice Wilkins were awarded the Nobel Prize for Medicine and physiology in 1962 for these discoveries.

The next significant milestone in DNA research was reached in 1966, when the genetic code was cracked. From now on, every possible combination of letters in a codon of RNA led to a known amino acid (Lagerkvist, 1998).

Polymerase chain reaction, or PCR, was devised in 1983 by Kary Mullis, enabling scientists to duplicate only certain sections of a DNA molecule. This process is much easier and faster in handling than the formerly used restriction fragment length
polymorphism, or RFLP, speeding up deciphering processes substantially (Newton 2009, 37). From 1990 until completion in 2003, the human genome project set the goal of sequencing the entire DNA sequence of a single human being, working out the blueprints for a more complete DNA comparison. Alongside the sequencing of DNA, its interpretation of what functions certain parts of it fulfill within a cell or organism was the most crucial step in transferring the new knowledge from research to development of applications (Newton 2009, 35).

4. Popular applications

4.1. Hereditary research

The ability to decipher genetic sequences and compare their codes is the foundation of heredity research, mapping the paths of evolutionary progress within species. One branch of this research focuses on humans, our species, the homo sapiens. As Green (2010) and his team show, comprehensive information may be obtained from the comparison of present-day humans and Neandertals, which vanished from the genetic pool some 30,000 years ago. Retrieving DNA for sequencing from ancient bones has proven to be a challenging task, having to patch holes in the code, differentiate between Neandertal genes and those of organisms decomposing the bodies of the deceased, and devise strategies of data interpretation. The genetic code of present-day humans and chimpanzees was helpful in identifying those sequences belonging to the Neandertal by comparison. Contrary to a long-standing belief, modern man is not a descendant of the Neandertal, but rather a distant cousin with a common ancestor within the last 500,000 years. The time frame can be estimated by dividing the number of genetic differences by the natural rate of mutation within a population, given a certain size of one generation in years. There is some speculation about interbreeding of the two lines of heritage, but no complete assessment can be made with the data available today except that both species have met and were able to interact in their time. The key finding in Green's work was another:

We show that Neandertals shared more genetic variants with present-day humans in Eurasia than with present-day humans in sub-Saharan Africa, suggesting that gene flow from Neandertals into the ancestors of non-Africans occurred before the divergence of Eurasian groups from each other. (Green 2010)

Using this technique, a complete bloodline of all known species can be drawn up, showing more or less precisely where the lines have forked, marking the last common ancestor even among species that separated hundreds of millions of years ago.
4.2. Genetic engineering in farming

DNA technologies have taken traditional forms of breeding to the next level, from mixing hereditary traits of two individual organisms in a long and time consuming line of succession, as Mendel did with his pea plants in the 1800s, to directly accessing and modifying hereditary information within the cells themselves to create an optimized organism for farming. One famous example is Monsanto, a US-based company selling seeds of genetically modified food plants such as soy, corn, cotton, canola and many more which are aiming for resistance and yield enhancement. Monsanto made use of brand development and genetic patenting, offering farmers an easy to use farming product with a quick overview of what to expect in comparison to regular, unaltered plants (see Figure 6).

The Genuity™ brand describes the family of traits that enables farmers to do what they do best even better. Genuity™ represents numerous traits such as drought tolerance, cold tolerance, nitrogen use efficiency, yield enhancement and more. Genuity™ stands as the gold standard in trait technology. (Monsanto 2010)

The term 'genetically modified organism' (GMO) has been introduced to classify these altered species. The earliest emergence of commercial use were in the early 1990s by China, introducing virus-resistant tobacco and tomatoes, followed by a US brand of delayed-ripening tomato called 'Flavr Savr' by Calgene in 1994. Since then, the number and volume of GMO products and crop areas have increased exponentially and can be found in North America, Mexico, Argentina and South Africa. This success was not unopposed, as concerns about consumer food safety, the environment, corporate control of agriculture and ethics are very strong internationally and have led to great restrictions on GMO farming in Europe (Nelson 2001).

In 2000, the genetically altered corn type 'StarLink' was accidentally released from its government-approved animal feed use into unapproved human food products. Upon releasing information about this incident, the US Food and Drug Administration (FDA) and the US Environmental Protection Agency (EPA) both received massive reports on allergic reactions on foods that may have been trace-contaminated with StarLink corn. Subsequently, in 2001 all foods bearing traces of StarLink were withdrawn from stores nationwide. After further investigation, the altered corn was found not to be responsible for the reactions. However, the fear of unforeseen consequences of genetic engineering in the general population of the USA became apparent (TransGen 2002).

Popular fear of GMO crops has a scientific basis, discussed by Young (2004, 18). A few of the stated hazards are ecological stability and horizontal genetic transfer. While the implemented trait of a GMO crop may raise its yield, for example, it is possible that other parts of the genome may be
affected as well, since our understanding of DNA today is that one gene may control many traits, not just one. A mustard plant, engineered for herbicide resistance, has been found to be 20 times more fertile than its originating strain, posing an imbalance risk to the ecosystem if a wild release should occur. Horizontal genetic transfer means that genes are not only exchanged between plants of the same species, bearing the risk of uncontrolled breeding of GMO plants, but also within the same ecosystem. Introducing new sets of genetic material then poses the risk of altering other plants in the ecosystem with unforeseen consequences.

A genuine threat to farmers arose when GMO crops, namely Monsanto's 'Roundup Ready' canola strains, were planted in open fields, from which they spread to neighboring farms. The Canadian Supreme Court Case of 'Monsanto vs Schmeiser' (CanLII, 2004) concluded this instance of confrontation between a Canadian farmer – Percy Schmeiser, who had his own breed of seeds contaminated with patented Monsanto genetic material by natural pollination – and the Canadian Monsanto branch. Monsanto demanded he either hand over all his seeds or pay C$15 in licensing fees per acre for use of its invention. After a long and expensive legal battle – Schmeiser paid around C$400,000 for legal representation – the court ruled in favor of Monsanto: In general, all profits gained from the use of patented GMOs, willingly or not, were to be handed over to the patent owner. But Schmeiser won a partial victory: since he did not have an advantage from the presence of the gene in question, no damages were to be paid to Monsanto, including legal costs. The crucial point in this case was that the crop in question holds a gene that makes it resistant to Monsanto's 'Roundup' herbicide, effectively cleansing the fields of all weeds other than the desired crop. Schmeiser had not used the herbicide, thereby not gaining any profit from the presence of the gene.

In an out of court settlement, Monsanto agreed to pay for all costs arising from the removal of its patented seeds from Schmeiser's fields. Schmeiser believes that

this precedent setting agreement ensures that farmers will be entitled to reimbursement when their fields become contaminated with unwanted Roundup Ready canola or any other unwanted GMO plants. (Schmeiser, 2008)

It remains to be seen whether this is true or not, but the incident shows that small farmers who cannot finance long-term and expensive court cases would be doomed in a confrontation with the global players of genetic engineering. Several other countries struggle with similar cases – see 'Monsanto vs Geerston Seed Farms' in the USA – while in some regions, farmers may slowly slide into dependancy on biotech companies, unable to take on a fight. As farmers shift from traditional to GMO crops, their income situation changes drastically. GMO crop seeds can be obtained at low prices, while any crop yield from these seeds carries a licensing fee to be paid to the GMO company. At the same time, the seeds may only be obtained from that one company, effectively
building a monopoly on seeding. Should the company decide to raise licensing fees, many farmers may not have a choice to go back to traditional crops or even other brands of GMO crops (Nelson, 2001).

The behavior displayed by Monsanto and all the imminent threats to the environment and the food markets emphasize the need for appropriate risk assessment and regulation before releasing any project into the wild. So far, inexperience has led to unwanted contaminations which cannot be reversed. It is up to the scientists behind this technology to invent safe methods of implementation, such as infertile crops that are unable to spread by themselves, or larger ranges of testing before a crop is allowed to be seeded for use. At the current state of affairs, GMO crops are an incalculable risk for anyone and anything they touch.

4.3. An example of a medical treatment relying on DNA research

One serious illness that cannot be cured as of today is multiple sclerosis, a nerve-degenerating disease that destroys fatty myelin sheets around axons of the brain and spinal cord, ultimately leading to asphyxiation and cardiac arrest by nervous dysfunction. Even DNA research has not led to any significant progress towards a cure, but at least some of the symptoms can be treated as of today. Baker (2000) points out that tetrahydrocannabinol (THC), the intoxicating substance found in cannabis, and related molecules may ease distressing symptoms such as tremor, spasms on walking, leg weakness, facial pain, impaired balance, anxiety and depression.

Since THC was outlawed in most of the countries in Europe and in the USA, this imposed the question of legality both on the patients and their treating doctors who strive for proper medication. Some countries, such as Germany, have legalized the medical use of marijuana to some extent, but are reluctant to alleviate the ban on cultivation of marijuana plants and the regulations on extraction. So far, THC is only to be extracted from the fibers of the plant, which yield only up to 0.2% of THC as compared to up to 25% in other parts. This tedious process provides only a total supply of about 20kg of substance per year while demand is at a soaring high of one metric ton per year (TU-Dortmund 2010).

German scientist Oliver Kayser found a way to transfer parts of the cannabis genome into a certain strain of bacterium, animating it to produce THC. Bypassing the conventional supply problems, he provided the medical community with the base for a fresh and reliable source of a much needed active agent for medication (TU-Dortmund 2010).
4.4. Organism design and human genetic engineering

One weakness of applications relying on genetic transfer, as demonstrated by Kayser (see section 4.3.), is that the desired genetic sequences need to be available before any action may be taken to utilize them. This constraint may soon be a concern of the past, as recent development shows. Venter (2010) and his team have a history of pushing forward the boundary of progress in biotechnology. In 1995, they were the first who were able to sequence the complete genome of a self-replicating bacterium. In the same year, they set themselves the ambitious goal of recreating a bacterium from scratch with only its essential genome, without any inactive 'junk' sequences. A highly complicated process of genetic elimination was developed, until two strands of 'highly accurate' genetic sequences were put together in yeast cells, which was then sequenced and taken as a model for the process of synthesis. An external manufacturer synthesized several 'cassettes' full of small DNA sequences which were then put together in a very delicate operation comprised of countless steps to reproduce the desired, final sequence within a yeast cell, checking for errors along the way.

The final step of the organism design was the transfer of the genome from the yeast nucleus into a prepared empty nucleus of an 'M capriolum' bacterium. The cell with the translated genome began replicating itself and resumed normal cell functions of the original donor cell, proving to be the first synthesized organism in the history of mankind. This process took not one try, but many over the course of months until it succeeded (Venter 2010). Venter himself said: 'We created a new cell. It's alive. But we didn't create life from scratch.' (CNN 2010)

Most important in this development is the ability to design DNA sequences to manufacture substances without relying on natural occurrence. Complicated active agents in medication, raw material for construction purposes or even breathable air may all be supplied by designed strains of bacteria, optimized for a single task and providing a new kind of strategic access to resources.

Genetic engineering technology is the cause for a lot of controversy. In 2010, after Venter and his team created their synthetic organism, the Vatican made a statement about DNA technology and the synthetic organism in particular (CNN 2010). It praised the work as 'important research,' and 'the work of high-quality genetic engineering'. This stance over genetic engineering acknowledges the potential benefits for disease research, but also urges to cautiousness. In general, the Vatican does not oppose genetic engineering, as long as no embryonic stem cells, cloning or any other experimentation with human cells are involved.

Many dangers of DNA technology are not a matter of unsafe methods, but of abusive execution. The above-mentioned mistrust in experimentation with human genes goes back to a long-standing discussion about the boundaries of the acceptable. The analyzing of the genome seems like a good
idea for the individual, but can quickly turn sour when information about possible diseases leaks to unwanted third parties, such as health insurance companies or employers, who would then try to avert possible damages to their business by ending collaboration. Also, the prenatal selection of embryos, the 'designed child', poses a great risk of unwanted selection processes in our society: there is no excuse for choosing one child over another, or for manipulating one's genome, simply following the latest fashion in genetic configuration.

The boundary of genetic engineering of human beings – and others – needs to be drawn today. We cannot allow ourselves to create beings that simply derived from our own species, and end up mixed with traits from all kinds of organisms, dooming it to alienation and suffering. The old and seemingly innocent dream of immortality might be within reach via DNA technology, but this brings to life a startling idea that could originate from a science fiction movie: eternal subjugation by an immortal tyrant and his army of clones.

Forming this boundary has hardly started and will probably be going on for a very long time, if ever coming to a conclusion with new problems emerging and old ones being reassessed.

5. Exotic applications

5.1. DNA computation

An intriguing conjunction of mathematics and molecular biology emerged when Leonard Adleman (1994) demonstrated the feasibility of parallel computation via DNA strands. He took advantage of similarities in mathematical problems, such as the Hamiltonian Path Problem, and PCR (polymerase chain reaction), the chemical process of multiplying certain sequences of DNA in vitro.

The Hamiltonian Path Problem, in short, is the effort of finding one or more solutions for the connection of a given set of nodes in such a way that the nodes are interconnected by straight lines, forming one single path. This single line touches each node exactly once and forms a loop.

An algorithm able to solve such a problem must follow three types of steps: 1. generate random paths, 2. only keep those paths which fulfill a certain requirement, and 3. read out Hamiltonian paths, if any remain.
Step one can be accomplished with a certain logic of encoding every node in a small strand of DNA and then letting it interact with other strands to form small single or double strands representing a possible path between two nodes. Step two can be completed via PCR (duplication), electrophoresis (length separation) and affinity purification (selection by specific code fragments). Step three, the solution of the problem, is obtained by graduated PCR, where the remaining strands are duplicated and compared with the initial building blocks representing the nodes via electrophoresis to determine the sequence of the nodes on the Hamiltonian paths.

This process took Adleman seven days of laboratory work, in which he recreated the biological equivalence of a memoryless filter with the simple operations of 'detect', 'separate' and 'extract', each represented by a certain interaction of test tubes via DNA manipulation techniques. Adleman specifically demonstrates that every action in the laboratory can be translated to a command in a programming language and vice versa, proving his newly devised method to be a programmable computation device much like the computers we use today.

Subsequently he constructed algorithms that would generate operations like 'merge', 'copy', 'discard' and many more, leading to the solution of combinatorial problems, determining the number of possible ways of arrangement and selection.

More development can be expected in this field as molecular computers offer two crucial benefits over traditional, silicon and circuit board models, at least from today's point of view: the sheer mass of molecules that can partake in a calculation and the density of gained information per step. Even if not many types of problems or real-time applications may be realized by the method described above, highly specialized machines with efficient outputs are within reach.

As a reminder: Konrad Zuse invented his first digital computer 'Z3' in 1941, and it has taken several decades for this technology to develop into personal computers and the Internet.

5.2. Data preservation

We live in a time with such immense knowledge and fast-pacing advancement that hardly anyone seems to waste a thought on long-term archiving. While we all know the stone buildings and clay tablets of civilizations past, we have nothing reliable to show for ourselves apart from outdated magnetic or optical storage volumes.

Canadian experimental poet Christian Bök has made the illustrious attempt to have one of his poems coded into the genetic sequence of a bacterium in 2007. Regardless of the outcome, DNA data storage and preservation is worth a thought. If we don't build obvious monuments, we could as
well start inscribing our signatures into life itself, which is going to be here long before we are gone ourselves.

Such a long-term data preservation brings up a problem: with every natural procreation of the original organism, its DNA may mutate or mix with non-altered specimens, ultimately obscuring the stored information. In order to prevent this loss, the information needs to be stored with high redundancy, either enlarging the genome of the single organism, or spreading it out over several organisms. The manipulation of an organism's DNA, in turn, may pose a threat to the organism or its ecosystem (see section 4.2).

6. Conclusion

Taking all the before-mentioned examples into account, DNA technology is a fast evolving field of research with a lot of current applications which already touch the lives of millions of people today. Surely the future holds more surprises for us, and we need to be careful about applying this advancement without an adequate risk-assessment – which in most cases still needs to be developed, as the case of Monsanto vs Schmeiser shows. The biggest issue today is the use of GMO crops with all its benefits and downsides.

As long as we keep our focus on biosafety and bioethics along the way, the revolutionary discovery of DNA will be remembered as a groundbreaking discovery with a positive twist.
7. References


8. Additional helpful references

CAFFEINE. $C_8H_{10}N_4O_2$.


GOOGLE BOOK SEARCH. Available from: http://books.google.com


Several issues of 'Scientific American' and its German counterpart 'Spektrum der Wissenschaft'.

The Human Genome Project. Available from: