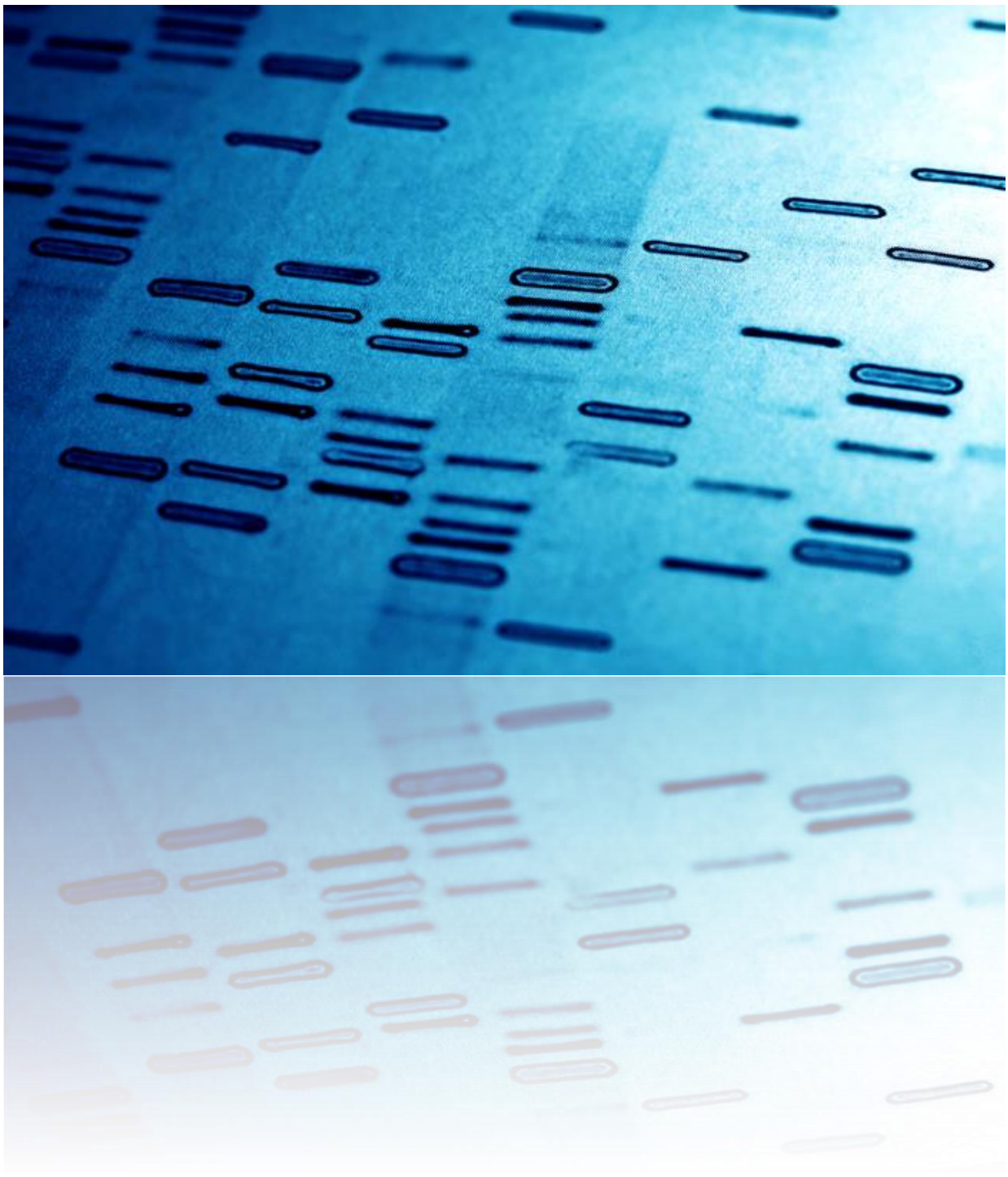
DNA Profiling



2. Mai 2016

DNA Profiling

Index:

1. Preface

We were always curious how come that one could prove a suspects guilt or innocence by just analyzing a stain or swab and how that technique works. But analyzing DNA is not only used in forensic crime investigation but also to look for single point mutations that can cause conditions like cystic fibrosis, sickle-cell anemia or color blindness however that is usually referred to as full genome sequencing. Whereas the term „DNA Analysis "refers to a forensic procedure used for person identification.

Our questions:

- How do you analyze a person's DNA?
- How reliable is such an analysis?
- Can this technique be used elsewhere than Forensic Genetics?

2. Introduction

Over the history people have tried to gain information from physical evidence of crime scenes, especially information about the people that were involved in them. Though methods like comparing fingerprints have existed and been used by the police since the late 19th centuries, many controversies have risen us about their validity and accuracy. The reason for that is simple. The act of comparing fingerprints is a very subjective process that can be easily misled by external information. It is for this reason that detectives and scientists alike sought for a method that is scientifically validated and resistant to personal bias to link traces found at crime scenes with possible subjects.

The method of using and comparing DNA was made possible by two major breakthroughs in genetic research on different sides of the Atlantic around 1983-1984. In the United States biochemist Kary Mullis has developed a process called PCR (Polymerase Chain Reaction) that can take a few copies of a strand of DNA and make multiple thousand to million copies of it. A year later Sir Alec Jeffreys in the University of Leicester, UK discovered after years of looking into genetic code of humans that it was possible to tell a lot about a person by looking at its DNA at specific regions. Thereby, he developed the first method of "DNA fingerprinting" called RFLP (Restriction Fragment Length Polymorphism) which would take certain areas of different lengths, which are unique for every person, and compare them to each other.

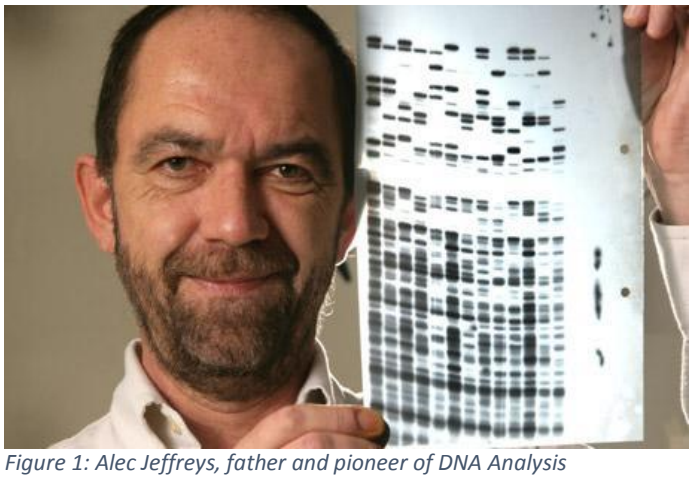


Figure 1: Alec Jeffreys, father and pioneer of DNA Analysis

Using this method it was then possible in 1986 to link two cases of rape/murder by finding a perfect match in the genetic code form the semen found at the scenes.

This result not only helped to convict the perpetrator but also to free an innocent man who was thought to have committed the crime, marking the first application of DNA analysis in law enforcement. To this date scientist have refined and improved upon this process and are nowadays able to get a result form just a few cells within a day and the fields of application for this analysis has expanded to many areas like .

The Innocence Project is an organization founded to free wrongfully convicted people with these DNA tests, many of which were prosecuted due to outdated and unreliable forensic methods like fingerprints or lie detectors. Another application of DNA analysis is paternity testing in which the genetic profile of a child can be used to find or verify its biological parents, mostly the fathers. This can be taken further to finding discovering a person's family tree or even some methods nowadays allow us to study human migration dating back thousands of years ago.

Today DNA analysis is scientifically validated and an essential procedure that has proven itself in accuracy and versatility especially in the field of forensic science and has swept away older methods of identifying people, like hair analysis or fingerprinting as these have proven to be subjective with little to no scientific background.

3. Process and Methods of Profiling

In almost every cell of our body, from the roots of our hair to the cells of our skin, there are long chains of a negatively charged acid called DNA that dictates the structure of our body. Though most of which within humans is almost identically the same, many parts with its combinations are unique to the individual. We know that not all of the genetic code is relevant and expressed as proteins that our bodies produce. In fact, this is only a very small fraction of our DNA. The so-called "non-coding" area therefore most easily prone to mutations and such and are highly specific in a specimen, be it human or any other living creature. We can use this uniqueness to find characteristics of a person and establish a genetic profile that is not same to any other person on this world (with the exception that they are monozygotic twins).

There are many methods of analyzing and making a profile with strands of DNA that have been developed over time. But before we can start to analyze it we have to do a few pre-procedures in order for the DNA to be processed from then on it can take several different ways. We will only cover the most commonly used methods.

3.1 Extraction, Purification and Multiplication of DNA

When human cells are found at the crime scenes, in for example cigarette buds or skin cells underneath a victim's nails. But first it is necessary to extract the DNA from the cell. This can be done by various procedures and machines and differs slightly for the types of samples like buccal swab, blood or hair. The sample is suspended in different solutions in order to separate the DNA from the rest of the cell and to purify the sample of inhibitors.

After the DNA has been extracted it will be quantified to determine whether there is enough DNA collected in order to proceed. In the past about a quarter of a droplet of blood was required, nowadays 15-20 cells are enough to build a sufficient profile of a person.

Figure 2: DNA extraction robot (Pomerga Maxwel 16)

The DNA will then be "amplified". This is where the PCR comes into play. In this process specific sequences will be multiplied by a machine that performs this task. The DNA sample will be inserted to the machine along with a few other materials necessary for the reactions. It will then be heated up to around 95C degrees so that the DNA two strands split up (Denaturation). The temperature will be lowered to 68C degrees and so-called

"primers" will be attached specifically to the 3' end of the target sequence (Annealing). Then, at 72C degrees, the "taq polymerase" will begin to copy the DNA strand starting from the primer with free nucleotides from the surroundings (Elongation).

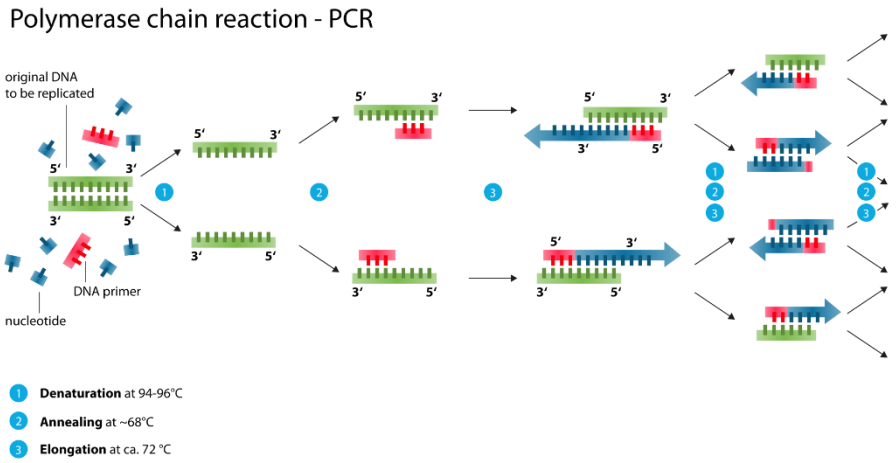


Figure 3: PCR (Polymerase Chain Reaction)

Once this is finished, the composition will be heated up to 95C to repeat the process again. After the third cycle we will get two copies only of the target sequence and from that point each additional cycle will increase the number of copies exponentially and after just a few hours we will get millions of copies of the desired DNA part. From this point on there are multiple methods that utilize different attributes of DNA sequences to analyze and create a profile.

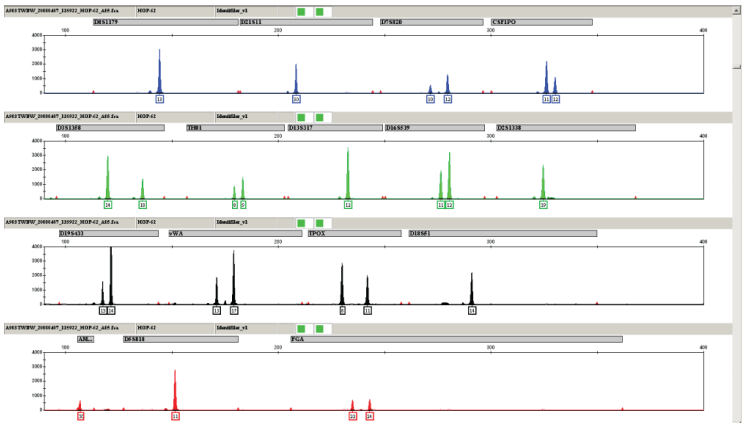
3.2 Restriction Fragment Length Polymorphism (RFLP)

This method is the first method of analyzing DNA created by Alec Jeffreys in 1984. The process begins by cutting the DNA with a restriction enzyme. Restriction enzyme are found in bacteria and archaea as a defense mechanism against viruses that cuts away foreign DNA. This process cuts the DNA only after a specific bases sequence occurs. Those fragments that are being produced will have different lengths for every individual because rarely does somebody have the same amount of these sequences and the same distance between them. Then the fragments will be sorted by size by a process called "electrophoresis". The negatively charged DNA sample will be inserted into a jelly-like substance called "agarose gel" and a voltage will be applied. Smaller pieces will travel faster down through the gel than larger ones. This will create a unique pattern and can be compared with other ones for example found at crime scene. In short **Restriction** enzymes cut DNA into **Fragments** of different **Length** that exhibit **Polymorphism**.

RFLP is nowadays still used for Paternity test as closer relatives will have similar patterns in their profile. This is especially true when examining the Y chromosome of a male child to compare it with a supposed father. This tool is also a vital tool for the location of genetic diseases as family members that are effected by them will probably have the same fragment at that certain location where the disease is located at. Its use in the field of forensics today is less practical as the process slower and less accurate for generating genetic profiles compared to new methods like STR.

3.3 Short Tandem Repeats analysis (STR)

Short Tandem Repeat analysis was introduced in the 1990s and became very popular in the field of forensics. STR's or Microsatellites are base sequences known to be found at certain regions of the non-coding part of DNA that have a specific base sequence usually the length between 2-5 base pairs that are repeated typically 5-50 times (for example a sequence of AATG can be found at certain region (TPOX) in the first chromosome). These regions have a very high mutation range (a single mutation would interrupt the repeat cycle) and have, thus, a high variation among people making them suitable for comparison.



[1]*Figure 4: Computer result of a STR analysis*

In this process restriction enzymes cut the DNA at the specific region where the STRs can be found, then they will be amplified with PCR and then also go through the process of electrophoresis. These results can then be analyzed by the computer. Although, STRs are polymorphic, the range of different kinds is small. The chance that two people have the same STR is around 5%- 20%. The strength of that method lies at comparing multiple STRs simultaneously. The US government uses 13 different loci making the chance of two having the same profile at 1: 1 quintillion (10^{18}).

On the right you can see how a paternity test result looks like. These results show that the alleged fathers DNA profile matches the child's profile in these five markers. In total it is tested for 16 markers whether they match or not. From this the probability of a blood relationship is determined by multiplying all the probabilities of a match of each marker.

DNA Marker	Mother	Child	Alleged father
D21S11	28, 30	28, 31	29, 31
D7S820	9, 10	10, 11	11, 12
TH01	14, 15	14, 16	15, 16
D13S317	7, 8	7, 9	8, 9
D19S433	14, 16.2	14, 15	15, 17

Figure 5: STR Family Relationship Analysis

3.4 Mitochondrial analysis and other methods

In the case of highly degraded samples for which a complete analysis of 13 STRs would be impossible, forensic scientists have found a method analyze mitochondrial DNA (mtDNA). With this method it is even possible to extract information out of old bones and teeth. The issue is that it is inherited from the mother making traces back only possible from the side of the mother.

Over the years many other methods of risen up to analyze specific parts of the DNA with different approaches, for instance, AmpFLP or Y-chromosome analysis etc. though the most widely used are RFLP and STR analysis.

4. Interview

We contacted the forensic institute of Basel and they were happy to give us an interview. Unfortunately it wasn`t possible for us to visit the lab because of contamination danger. However Mr. Dion was so kind to

provide us with some pictures of Machines and Robots that are also used in their lab. The interview took place at the forensic institute of Basel, Mülhauserstrasse 113, with Mister Daniel Dion, Dipl. anthropol. , Deputy Head of Department, Forensic Genetics Basel:

What type of DNA Analysis do you use the most?

We have like a standardized DNA Analysis, what we do: We extract DNA from the cell and then we do a PCR, short tandem repeats, with this PCR-product we do a capillary electrophoresis, so we separate the fragments we get from the PCR and with the capillary electrophoresis and from that on we have a software which calculates our DNA profile.

There is another type of analysis with mitochondrial DNA. Do you use that technique here too?

Uh no. Mitochondrial DNA Analysis is interesting for lineage purposes, from the mother's side to the children. But mitochondrial DNA we don`t use that. We do standard PCR, that's what we use for stain and person identification and paternity testing, it is more specific than mitochondrial DNA testing and sufficient for our purposes.

What are the most common problem you encounter when analyzing a DNA sample?

Good question. The most common problems are, if you have a DNA sample taken at a crime scene which is not only biological sample but also inhibitors, so inhibitors could be for example dirt but also any kind of soil or stone. Which could all inhibit the polymerase. So inhibitors are the main factor. Plus degraded DNA for example so we have usually fragments that are 400 base pairs long but at some point you have degraded DNA so you can only analyze till 200 base pairs and after 200 base pairs you don`t get any results so degraded DNA samples are very much a problem. And also the amount of DNA can also be a problem because we need about 15-20 cells for a full DNA profile but sometimes at a crime scene you don`t get that many cells. For example skin cells you don`t get that many cells so if you don`t have enough DNA you sometimes get absolutely no DNA profile. And of course if you have multiple contributors at one stain then you cannot interpret the DNA profile anymore. So we say we can interpret a DNA profile with up to 2 contributors but more than 2 contributors it depends. Three to four contributors we can still do a so called complex profile and do a local comparison so we can visually compare the DNA profile of a suspect and tell if he`s in that complex profile or not.

Is there a control mechanism, which ensures that the DNA amplified by PCR is correct? So we want to know if I got a DNA profile how can I be sure that this is really the correct DNA profile?

Yes, so in Switzerland if you want to do DNA profiles from stain or people you need to have an accreditation. So every lab which gets stains and swabs from suspects needs to have an accreditation. That means you need to fulfill certain standards from how the stain arrives in the lab you need to protocol every single step you do in the lab. And all the methods like PCR or DNA extraction we use are validated. So you need to do a validation on all your instruments. And we also have positive and negative controls which allow us to ensure that the DNA which is amplified is correct.

How long does it take for you to analyze a sample?

There are two different procedures so if you have a swab from a person I would say we have about four to five hours. These methods of course have improved in in the last four five years, everything needs to be faster and quicker so four to five hours is quite fast. Before it took us about a day.

With the stains it is different because it involves a different extraction method and also different PCR-Kits. So for a full profile I'd say it still takes us about a day if we have very good samples. The usual process is one to two days now.

Which tasks during analysis can be given to machines to fulfill, which tasks still need humans to be done?

A lot of processes are automated already for example we are just in the course of changing the DNA extraction method, we have done that in the last few years manually and now we are switching to a robot not a big robot but a table robot. But first you need to do a cell lysis before you can extract the DNA. And this new method is with purification. We also use pipette robots. But in the end you always have to do some things manually in the lab… there`s not like in CSI where you can put a swab into a machine and you get a full DNA profile (laughs). Well at least not now maybe in 20 or 30 years this is possible. But especially when you work with stains you need to be very very careful because the amount of DNA is crucial and you sometimes don`t have much DNA in a sample; so that needs to be done very carefully by hand. So yes, there are tasks we can use machines and there still are processes which need to be done manually.

The machine you use for purification is it one which uses metal beads because we were at the "Tag der Biomedizin" and they used one with metal beads there?

Well it indeed has metal beads where the DNA is attracted to because DNA is negatively charged. So you can bind it to magnets and you can purify the whole sample as I said you can purify it from all the inhibitors. Which is quite important. So it will be more purified than other methods where you don't actually purify the DNA you just separate the DNA from the rest of the cell but you still have inhibitors in there. That's the main difference. But also this step you could do manually.

What would be a major breakthrough in DNA analysis in your opinion?

Well the best thing we could have would be if you could do a full DNA profile just from one cell. As easy as that. One cell and we`d get a full DNA profile. Because the major problem is really the amount of DNA so if we just have one cell… right now we would never get a full profile because we need our 15-20 cells. Would be nice if we'd get that.

What are the current research points your institution is working on?

Uhm… there are different projects in our department but these projects are more involved in validation of new methods… like new extraction methods. But one thing is for example the micro dissection of like DNA, sperms. If you have male and female cells you want to separate them. So you stain the sperm cells and you shoot them out with a laser under the microscope. Laser dissection. It`s not research exactly but it takes a lot of time and when one lab does it and it works doesn`t mean that it works in another lab. You need to do it again and again.

5. Discussion

The current technique used by forensic labs all over the world makes it possible to have a one in a quintillion probability of a random match of unrelated individuals. So the chances of a false match are

practically zero percent. Which is of course ideal for a forensic procedure. Full genome sequencing, which is really the continuation of DNA Profiling, has the potential to change our medical diagnostic system fundamentally when it becomes possible to exactly determine what mutations combined with which epigenetic and environmental factors cause which diseases.

5.1 Ethical concerns

So when it comes to DNA Analysis there`s no ethical discussion possible. Because it doesn`t reveal information about the individual itself.

However when it comes to full genome sequencing it is possible. Because full genome sequencing reveals much about a patient and diseases this person is likely or less likely to get and also the same information about close relatives of the patient.

5.2 Advantages and Disadvantages

Advantages:

- It makes it possible to test a person for a certain disease and treat it before it shows symptoms and also treat relatives which could have inherited the same mutation.
- It can be used to better understand which mutations combined cause which diseases.
- It could make diagnostic medicine a lot more individual and help to better understand diseases.

Dangers and problems:

- A patient doesn’t necessarily want to know what disease he/she is going to get because it would be extremely depressing to know that you are going to get Alzheimer’s for example and would change your entire life.
- A conflict of interests can occur when a patient refuses to share information on a diagnosis, which is not only preventable but also where there is a risk of a relative carrying the same disease mutation. That could lead to a problem with the patient-doctor confidentiality.
- Information about a person’s genome must be kept private under all circumstances otherwise it could be a potential discriminating factor when it comes to a applying for a job for example.
- It also raises legal questions concerning dispositional rights over a DNA sample and the information it contains.
- Also public scientific databases for genotypes must keep an individual’s phenotype secret in order not to make the individual possible to identify.
- Today it wouldn`t be possible to actually store that much information about so many individuals because the human genome is really huge.

6. Summary

Overall DNA profiling is a technique that take the advantage of the diversity of different aspects in a person's genetic code, be it the frequency of reoccurring patterns or the distance between certain base pairs, to distinguish and back trace microscopically small evidence at crime scenes to people or find hidden ancestors in the family tree. It is a method that has been made possible thanks to scientific breakthroughs of the 1980s and is continuingly improved in both accuracy and speed and has become a standard

procedures in not only forensics but also many other fields, like the analysis of diseases or bringing back families together. It will probably be replaced with full genome sequencing down the line once we have found a way to store all that information. However, even then privacy issues and the question of who will have the access and control over these profiles will remain as problems.

7. References

- http://biotechlearn.org.nz/focus_stories/forensics/dna_profiling
- http://www2.le.ac.uk/departments/emfpu/genetics/explained/profiling-history
- https://www.youtube.com/watch?v=8tC_x1-JuwI&list=PLdpycn66Td7WuIvQm9YNL7YjqUDNumgH-&index=4#t=507.201312
- http://www.innocenceproject.org/about/
- http://www.dnai.org/d/
- https://www.chetson.com/2010/12/how-forensic-dna-collected-extracted-profiled/
- https://upload.wikimedia.org/wikipedia/commons/thumb/9/96/Polymerase_chain_reaction.svg/2000px-Polymerase_chain_reaction.svg.png
- http://www.omicsonline.org/str-profiling-of-human-cell-lines-challenges-and-possible-solutions-to-the-growing-problem-2157-7145.S2-005.php?aid=2120
- http://www.omicsonline.org/articles-images/2157-7145-S2-005-g001.gif
- https://en.wikipedia.org/wiki/DNA_profiling, 7.35PM 3/24/2016
- https://en.wikipedia.org/wiki/Whole_genome_sequencing, 4.25PM 4/27/2016