# DNA profiling of single sperm cells and single skin flakes in forensics using micromanipulation and whole genome amplification.

THEUNISSEN, G.M.G.

2020

# DNA profiling of single sperm cells and single skin flakes in forensics using micromanipulation and whole genome amplification

Glenn M.G. Theunissen

PhD 2020

# DNA profiling of single sperm cells and single skin flakes in forensics using micromanipulation and whole genome amplification

Glenn M.G. Theunissen

A thesis submitted in partial fulfilment of the
requirements of the Robert Gordon University
for the degree of Doctor of Philosophy

May 2020

# Declaration

I hereby declare that this thesis has been written entirely by myself and has not been presented previously for any degree award of this university or other academic institute. The research was completed at the partner university Hochschule Bonn-Rhein-Sieg University of Applied Sciences in Rheinbach, Germany. All information and work, other than my own, are acknowledged and cited by reference.

# Acknowledgements

I hereby would like to thank the following people that played a very big important role during the time being a PhD student and working on the research described in this thesis. The research performed and described in this thesis was not possible without them.

**Prof. Paul Kong Thoo Lin and Prof. Richard Jäger**
Thank you so much for supervising me. You both supported me with great effort in many aspects of my PhD. You both helped me to stay positive, even during the hard times. It was a great honor to have you as my supervisors. I have learned a lot from you.

**Dr. Andrew Gibb**
Thank you very much for supporting me with your forensic expertise and your highly appreciated help on recruiting volunteers for sample donation. It was a great honor to have you as one of my supervisors.

**Dr. Colin Thompson, Prof. Susan Duthie and Dr. Andrew Lamb**
Thank you for supporting me and helping finding a solution whenever I had any problems.

**Carol Rogers, Scottish Police Authority – Forensic Services**
Thank you for the support on recruiting volunteers and sample preparation

**Jakob Waldhausen**
Thank you for your work on staining and DNA extraction of single skin flakes.

**Dr. Melanie Grabmüller, Dr. Julia Sabirova, Dr. Claudia Till, Sarah Heß, Jan-Philipp Biermann, Prof. Ralf Thiele**
Thank you for your help and support on the NGS analysis

**Dr. Burkhard Rolf, Eurofins Medigenomics Forensics**
Thank you for supporting me by offering your facilities to perform X- and Y-STR analysis.

**Dr. Harreld Schneider, Hessian state criminal police office**
Thank you for teaching me how to isolate single skin flakes from tape liftings.

**Prof. Hans Weiher, Dr. Barbara Roitzheim Esther van Dorp, Klaus Peter Schmidt, Lara Kehret**
Thank you for offering your lab facilities used in many conducted experiments.

**All voluntary donors**
Thank you for donating the biological material which was used in the experiments. These experiments were not possible without your help.

**Angela, Pathology, Uniklinik Bonn**
Thank you for the training on micromanipulation.

**Prof. Klaus Olek, Dr. Sophia Forat**
Thank you for offering your laboratory facilities.

**German Federal Ministry of Education and Research (BMBF)**
For the financial support of the research project

**The Graduate Institute and the Faculty of Applied Sciences of Hochschule Bonn-Rhein-Sieg University of Applied Sciences**
For financially supporting me with a scholarship

**ISFG**
For supporting me financially with a conference travel fellowship

**Friends and family**

# Abstract

In forensic DNA profiling the occurrence of complex mixed profiles is currently a common issue. Cases involving intimate swabs or skin flake tape liftings are prone to mixed profiles, because of more than one donor contributing to a DNA sample. By DNA profiling of single spermatozoa and skin flakes, problems associated with mixed profile could ideally be overcome. However, PCR is not a sensitive enough method to generate DNA profiles by STRs on single cells. Moreover, high quality intact DNA is required, but is not always available in skin flakes due to degradation. Additionally, single skin flakes are difficult to discriminate from other similar looking particles on tape liftings used to secure DNA samples from evidence.

The main purpose of this study was to develop a method which enables DNA profiling of single sperm cells and skin flakes. By studying multiple whole genome amplification (WGA) protocols, REPLI-g Single Cell WGA was selected due to its suitability in the pre-amplification step of template DNA. Micromanipulation was used to isolate single spermatozoa. Furthermore, micromanipulation in combination with REPLI-g Single Cell WGA resulted in successful DNA profiling of single spermatozoa by using autosomal STRs as well as X- and Y-chromosomal STRs. The single spermatozoa DNA profiling method described in this thesis was successfully used to identify male contributors from mock intimate swabs with a mixture of semen from multiple male contributors.

Different dyes were analysed to develop a staining method to discriminate skin flakes from other particles including particles such as those from hair cosmetic products. From all dyes tested, Orange G was the only dye which successfully discriminated skin flakes from hair product particles. Also, an alkaline based lysis protocol was developed that allowed PCR to be carried out directly on the lysates of single skin flakes. Furthermore, REPLI-g Single Cell WGA was tested on single skin flakes. In contrast to the single spermatozoa, REPLI-g Single Cell WGA was not successful in DNA profiling of single skin flakes. The single skin flake DNA profiling method described in this thesis was successfully used in correctly identifying contributors from mock mixed DNA evidence. Additionally, a small amplicon based NGS method was tested on single skin flakes. Compared to the PCR and CE approach, the small amplicon based NGS method improved DNA profiling of single skin flakes, giving a significant increase in allele recovery.

In conclusion, this study shows circumventing mixtures is possible by DNA profiling of single spermatozoa using micromanipulation and WGA. Furthermore, DNA profiling of single skin flakes has been improved by the staining of tape liftings methodology with Orange G, alkaline lysis, direct-PCR and a small amplicon based NGS approach. Nonetheless, future work is required to assess the performance of the single spermatozoa method on mock swabs with more diluted semen. Also, commercially available NGS kits should be tested with single skin flakes and compared with the in-house NGS method.



**Keywords:**
Forensic, DNA, single cell, sperm, skin, whole genome amplification, WGA, micromanipulation, next generation sequencing, NGS.

# Table of Contents

# List of figures

# List of tables

# Abbreviations

| Abbreviation | Meaning |
|---|---|
| ADO | allele drop-out |
| ADI | allele drop-in |
| bp | basepair |
| BSA | bovine serum albumin |
| *Bst* | *Bacillus stearothermophilus* |
| CTS | Christmas Tree Staining |
| DAPI | 4',6-diamidino-2-phenylindole |
| DI | degradation index |
| DNA | deoxyribonucleic acid |
| dsDNA | double stranded DNA |
| FACS | fluorescence-activated cell sorting |
| FISH | fluorescent in situ hybridization |
| gDNA | genomic DNA |
| ht | height |
| Kb | kilo base pairs |
| LCM | laser capture microdissection |
| LR | likelihood ration |
| LV-PCR | low volume PCR |
| MALBAC | multiple annealing and looping based amplification cycles |
| MDA | multiple strand displacement amplification |
| miPEP | modified improved primer extension preamplfication |
| PCR | polymerase chain reaction |
| Phi29 | bacteriophage Φ29 |
| PI | paternity index |
| qPCR | quantitative real-time PCR |
| RNA | ribonucleic acid |
| rpm | rounds per minute |
| SNP | single-nucleotide polymorphism |

| | |
|---|---|
| ssDNA | single stranded DNA |
| STR | short tandem repeat |
| *Taq* | *Thermus aquaticus* |
| *Tth*PrimPol | primer synthesizing polymerase from *Thermus thermophilus* HB27 |
| WGA | whole genome amplification |

# Chapter 1

# Introduction

# Chapter 1. Introduction

In criminal investigations the analysis of DNA samples is of great importance. Typically, DNA evidence provides the only link between the donor and the object or location where the DNA was found. In addition, a link between various crime scenes and objects can be based on a common trace of DNA, even when the person who left the DNA is unknown. Therefore, forensic DNA analyses are crucial to civil security.

## 1.1. Forensic DNA profiling

Forensic DNA profiling is based on the analysis of highly variable locations (loci) which are scattered over all chromosomes. These loci have been termed short tandem repeats (STRs) and consist of tandemly repeated DNA sequences that are 2 to 7 base pairs (bp) long. The STRs that are used for DNA profiling contain repeating DNA sequences which are 3bp or 4bp long. These DNA loci are also referred to as microsatellites. Every STR genotype consists of two alleles: one maternally inherited allele and one paternally inherited allele. The number of repeats at these loci vary naturally in the population and display a high degree of polymorphism meaning that they are well-suited as a tool for discriminating between people. A person's DNA profile is the full list of genotype combinations for any particular set of loci looked at, typically the loci targeted by the DNA profiling kit used. By utilizing large number of loci, an extremely high number of genotypic combinations are possible which greatly reduces the chance of adventitious matches.

The varying numbers of repeat units result in differences in length between the respective alleles that can be analysed by using polymerase chain reaction (PCR) in combination with capillary electrophoresis. A multiplex PCR makes many copies of multiple STR loci of the recovered DNA thus ensuring high sensitivity of the analysis. Many multiplex PCR kits for forensic STR analysis are available [1].

These include not only STR loci, but also another commonly used locus named Amelogenin. This locus is present on the X- and Y-chromosomes and the length in bp of each allele depends on which chromosome the allele is present [2]. A female person contains two X-chromosomes, which results in both alleles having the same length. However, a male person contains one X- and one Y-chromosome which results into two Amelogenin alleles differing in length. By relating the length of the alleles to the corresponding sex-chromosome, Amelogenin can be used to determine the genetic sex of the DNA's donor.

The PCR products generated with multiplex kits are nowadays analysed by capillary electrophoresis which allows the different Amelogenin alleles, STR alleles and loci to be distinguished based on size as well as fluorescent-tags incorporated during the PCR process. As a result, the PCR products from each allele are displayed as peaks based on relative fluorescent units (RFU) in a graphical output known as an electropherogram (see figure 1.1.1, page 25).

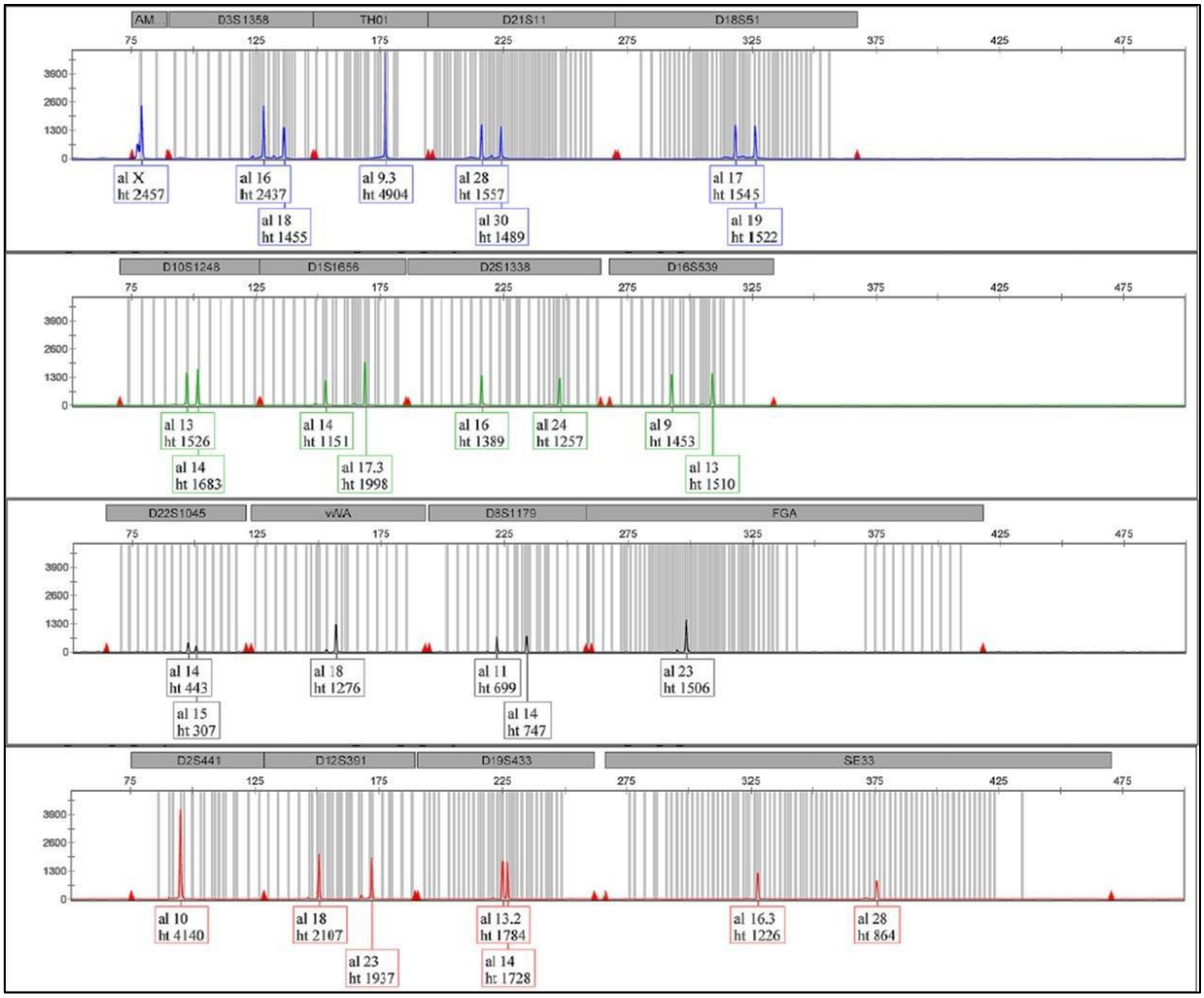


**Figure 1.1.1. Electropherogram showing the alleles from different forensic STR loci of a DNA sample from a female person. Peaks represent the amplicons of different alleles amplified using multiplex PCR and detected by capillary electrophoresis. The different allele amplicons were detected and shown using 4 colour channels (blue, green, black and red). The relative fluorescent units (RFU) are set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. On top, the X-axis represents the fragment size in base pairs (bp) with the allele corresponding STR locus name above in the grey horizontal bars. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. The DNA profile was generated with the multiplex PCR kit PowerPlex ESX17 (Promega, Madison, USA).**

The size of the PCR products corresponds to the respective number of repeat units within the targeted region. Therefore, a DNA profile contains purely numerical information about the number of repeating units in each allele. Thus, a DNA profile of a female person who is homozygous for STR locus TH01 could for example look like this:

*Amelogenin*

Allele 1: X

Allele 2: X

*TH01*

Allele 1: 8

Allele 2: 8

A DNA profile of a male person which is heterozygous for STR locus TH01 could look like the following:

*Amelogenin*

Allele 1: X

Allele 2: **Y**

*TH01*

Allele 1: 8

Allele 2: **9**

The DNA profile based on Amelogenin and multiple STR loci form the basis of the national criminal DNA databases which are used by federal and state criminal police in many countries for storing and comparing DNA profiles of offenders and DNA during investigations. The set of different STR loci used, also called core loci, differs between countries. In the UK core loci include 10 STR loci (FGA, TH01, VWA, D2S1338, D3S1358, D8S1179, D16S539, D18S51, D19S433, D21S11) and one locus to determine gender (Amelogenin) [1]. Whereas in Germany 8 STR loci (FGA, TH01, SE33, VWA, D3S1358, D8S1179, D18S51, D21S11) and one gender locus form the core loci (Amelogenin) [1]. A full database DNA profile includes the core loci plus additional STR loci. For example, in Germany a full database DNA profile consists of 16 STR loci (D18S51, D21S11, TH01, D3S1358, D16S539, D2S1338, D1S1656, D10S1248, FGA, D8S1179, vWA, D22S1045, D19S433, D12S391, D2S441 and SE33) and Amelogenin. In order to add a DNA profile to the criminal DNA database the core loci are essential and therefore at least these

should be determined. Thus, in each country multiplex PCR kits are used to enable amplification of the country's specific core loci plus additional STR loci.

Other than autosomal STR multiplex kits, multiplex kits exist that are used to amplify X- or Y-chromosomal STRs [1,2]. In contrast to autosomal STRs, X- and Y-chromosomal STRs contain loci located only on the X- or Y-chromosome. In forensic DNA analysis, Y-STR kits can be used to analyse DNA from sexual assault cases with female victims [2]. As the Y-STR kit only works on male DNA, a DNA profile can be established from DNA left behind from the male perpetrator even when DNA is mixed with DNA of the victim. In contrast to autosomal STRs, Y-STRs are completely paternally inherited which results in many males containing the same Y-STR profile. In other words, the Y-STR profile is less unique than an autosomal STR profile. This makes the power of using Y-STR profiling for matching the DNA profile to a person's DNA profile much lower than when using autosomal STRs.

Whenever a suspect's DNA profile matches the DNA profile of the DNA sample taken from the evidence, the strength of the match, or DNA evidence, needs to be determined. This is necessary to enable the court to evaluate the evidence. To determine the strength of the DNA evidence, a calculation is done which states how likely it is obtaining a matching DNA if the DNA is from the suspect rather than if it is from an unknown unrelated individual. This so-called likelihood ratio (LR) calculation is based on the DNA profile's frequency ($P$) in a population [3]. STR allele frequencies have been determined in many different populations e.g. European, African and South Asian populations. These allele frequencies ($P_{allele}$) are used to calculate the genotype frequency for each STR locus. The locus frequency ($P_{locus}$) of a homozygous locus is calculated by $P_{allele\ 1} \times P_{allele\ 2}$. In case where the locus is heterozygous $P_{locus} = 2 \times (P_{allele\ 1} \times P_{allele\ 2})$. Then the frequency of locus 1 is multiplied with the frequency of locus 2 and with the frequency of locus 3 and so on. When the frequency of the complete DNA profile is determined, the LR can be calculated by 1 divided by the DNA profile's frequency. So a profile frequency of 0.0002 will give a LR of 5000. This means it is 5000

times more likely to see the recovered alleles if the suspect is the source rather than if they were randomly chosen from another male. If more STR loci are determined, then more STR loci could be matched, hence the higher the LR could be. A perfect match of 16 STR loci can often lead to a LR of more than $10^9$.

After years of optimization, many forensic PCR kits can now amplify template DNA at very low concentration e.g. less than 100pg [4]. This sensitivity allows for generating a full profile even though small amount of DNA as little as 16 somatic cells is available [5]. However, intact DNA is still required. Due to environmental factors like UV-light and heat or the type of tissue like hairs, DNA from forensic evidence often is degraded. The latter implies that the DNA may not be intact at every locus anymore resulting in allele drop outs (ADO) and hence leading to unbalanced profiles. If both alleles of an STR locus drop out, then the whole STR locus drops out. Since the LR depends on the allele frequency of the recovered alleles, the power of the LR increases with the number of STR loci analysed. ADO can massively decrease the LR.

Also, stochastic effects can lead to unbalanced profiles and ADO. Stochastic effects are random variations arising from the PCR and predominantly occur due to sampling error. Usually for PCR a fraction of the whole extracted DNA sample is used as template DNA in the reaction. The number of copies from every locus is thus equal for all loci. However, when the extracted amount of DNA is too low, the proportion of extracted DNA which is used as template DNA in the PCR does not represent the whole extracted DNA sample anymore. As a result, ADO or unbalanced allele peak heights can occur [6]. Furthermore, the high sensitivity of currently available forensic PCR kits increases the chance of contamination which could lead to allele drop in (ADI). If present, contaminating DNA makes up a tiny fraction of the total DNA. As a consequence, the contaminating part is underrepresented during PCR and subsequently allele peaks will stay underneath the analytical and/or stochastic threshold. However, when trace amounts of DNA template are used the amount of DNA from the main component can equal or almost equal the amount of the contaminating part which can lead to a mixture and thus can interfere with the interpretation of the DNA profile. To prevent ADI

from contaminating DNA a few precautions are commonly taken in forensic DNA investigations. The pre-PCR rooms (evidence examination rooms, DNA extraction rooms and PCR set up/pipetting facilities) are separated from post-PCR rooms (rooms for PCR, DNA quantification and capillary electrophoresis) as well as separating the victim's evidence from the accused's evidence. Furthermore, work surfaces are cleaned with DNA decontaminating agents e.g. bleach and forensic examiners wear gloves and surgical masks to prevent depositing the examiner's DNA onto the evidence. Also negative control samples are used to check for any DNA contaminated reagents e.g. DNA extraction buffers and PCR mastermixes. Additionally, the forensic examiner's DNA (laboratory staff as well as forensic examiners in the field) is profiled. If ADIs pop-up, even though taken all necessary precautions, the specific ADIs can be compared with examiners' DNA profiles to check for any contamination that happened during securing or examining the evidence.

## 1.2. Strategies to isolate single sourced DNA

Ideally, single sourced DNA is used as template for PCR. Template DNA from more than one person complicates the interpretation of DNA profiles. Due to the high sensitivity of current PCR systems, very often minor quantities of DNA from a second donor will be amplified alongside the perpetrator's DNA resulting in mixed profiles. There are, however, situations where mixed profiles are hard to avoid. For example, this is the case when DNA is extracted from DNA samples which are composed of biological material left by more than one person. Intimate swabs from sexual assault cases and tape-liftings of skin flakes are such DNA samples that often lead to mixed profiles [7]. Often in these samples the DNA of the perpetrator is even in minority. However, mixture interpretation is possible and has advanced with the use of statistical software packages e.g. TrueAllele® and STRmix™ [8][9]. Though, often mixed profiles can be uninterpretable especially with trace amounts of DNA overlapping. This means evidence to solve cases is lacking. Therefore, the ability of isolating and subsequently analysing

DNA from different individuals separately would likely improve interpreting forensic casework. This can be achieved by analysing individual cells or other biological material that can be left only from one individual, such as single skin flakes.

### 1.2.1. Isolation of single sperm cells

Regarding sexual assault cases, an abundance of epithelial cells from the victims is often found on their own intimate swabs together with the sperm cells from the semen left by the perpetrator. The extracted DNA from these mixed samples will contain DNA from the victim as well as DNA from the perpetrator [10,11]. After PCR of those samples, the resulting electropherograms will display alleles from the victim next to the alleles of the perpetrator. When only a few sperm cells could be recovered from intimate swabs, the DNA from the perpetrator will be in minority compared to the victim's DNA. The abundance of DNA from the victim will outcompete the DNA originating from the perpetrator during PCR, resulting in the masking of the perpetrator's profile by the victim's alleles.

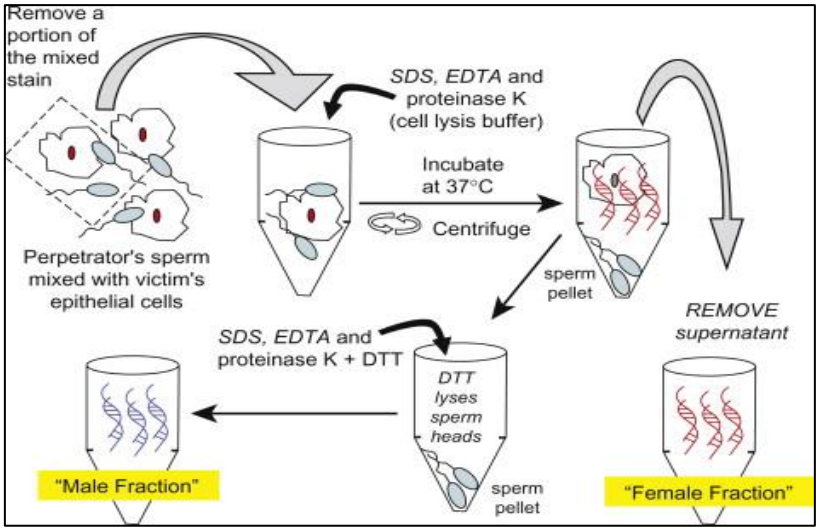


**Figure 1.2.1.1. Schematic view of the differential extraction which is used to separate the male fraction from the female fraction from DNA samples like intimate swabs. DNA samples containing a mix of the victim's epithelial cells and semen from the perpetrator is first incubated in a lysis buffer which lyses only the epithelial cells. Afterwards, the still intact spermatozoa are pelleted by centrifugation and the supernatant with the victim's DNA (the female fraction) is removed. Then the sperm pellet is lysed by using a lysis buffer with DTT to release the perpetrator's DNA (male fraction). (Figure taken from Butler 2012** [12]**).**

To overcome this problem, differential lysis was introduced as an extension to the normal DNA extraction method especially for analysing vaginal swabs. As shown in figure 1.2.1.1, the protocol is based on separating the DNA extraction of epithelial cells from the sperm cells [13,14]. Despite using differential lysis, DNA profiling from the DNA of the perpetrator is still problematic due to cell-free DNA from the victim present in the lysates [15]. Furthermore, with trace levels of sperm, the sperm fraction is often "contaminated" by epithelial cell which is caused by carry-over of epithelial cells which have not been completely lysed [16]. This results in the victim's epithelial DNA masking the seminal DNA. This problem commonly arises if intimate swabs are taken long after the sexual intercourse as the number of spermatozoa in the vaginal track tends to reduce with time.

Moreover, sperm cells from more than one donor can lead to problematic DNA profiling. This can happen in rape cases when, for example, sperm from the victim's partner is present on an intimate swab together with the perpetrator's sperm or in cases when sperm from more than one perpetrator is found on DNA evidence such as intimate swabs, pants, interchanged/reused condoms or in multiple rape cases. In such cases, differential lysis does not solve the situation where more than one male contributes to the sperm sample. While DNA mixture interpretation is possible, when dealing with trace amounts of sperm it might not be feasible, and so ideally DNA profiling should be performed on single isolated sperm cells, since one cell can only originate from one donor. Although in principle the sensitivity of PCR is high enough to amplify DNA from single cells [17], forensic multiplex PCR analysis of single cells is usually compromised by the prevalence of stochastic effects [17]. The sensitivity limit of current protocols is about 30pg of DNA, which corresponds to 4-5 diploid cells [18]. Thus, sensitivity is not yet high enough to reliably amplify DNA from a single cell. However, there are methods (see chapter 1.3) that can be applied to the single cells prior to any PCR attempt and which can improve the amplification of DNA content.

For any attempt at single cell isolation, a method is first required for cell detection, and there are several such methods available to detect the cell of interest. The particular characteristics of sperm cells have allowed for the development of methods that specifically target their visualization, often using specific cellular staining techniques. Christmas Tree Staining (CTS) is probably the most known stain to type sperm cells and has been used for many years. CTS is based on two staining components, nuclear fast red and picroindigocarmine [19]. Nuclear fast red stains the sperms' heads red and the acrosomal cap pink. Picroindigocarmine stains the neck and tail of the sperm cell green and blue. Thus, CTS is based on cell morphology to distinguish sperm cells from other cell types. The staining process is fast and can be used with a normal light microscope. This makes CTS a rather specific staining for sperm cells.

On the other hand, fluorescence microscopy has now been used for many years, and a wide variety of fluorescence stains are commercially available which can be used to specifically detect many different cell types. Also, fluorescence in-situ hybridization (FISH) which works *via* fluorescently tagged nucleic acid probes can be used to detect cells that differ in chromosomes or alleles [20,21]. FISH is therefore able to distinguish diploid from haploid cells and male from female cells. This feature makes it not only possible to distinguish haploid sperm cells from other diploid cells, but also to distinguish male epithelial cells from female epithelial cells in cases when no sperm cells are found on intimate swabs after sexual assault. A previous study reported successful DNA profiling by using FISH to detect male epithelial cells on a vaginal swab from a sexual assault case in which a sterile perpetrator raped a woman [22]. However, FISH is time consuming, laborious and therefore difficult to incorporate into a busy case working laboratory environment. Another alternative fluorescent staining for the detection of sperm which has successfully been applied in a few forensic laboratories is called Sperm HY-LITER™ (Independent Forensics) which uses a human sperm specific antibody [7,11,23,24]. This makes it a highly specific sperm detection method when compared to CTS. Furthermore, Sperm HY-LITER™

is fast and does not require many laborious steps which make it readily amenable with the procedures in an active forensic laboratory.

Once visualized and in order to isolate single cells, various techniques can be used, e.g. fluorescence-activated cell sorting (FACS), micromanipulation, laser capture microdissection (LCM), optical tweezers, microfluidics and di-electrophoresis [11,25,26]. Depending on the application, each technique comes with its advantages and disadvantages.

FACS is fast and can be used for high throughput, but requires many cells which in forensics are not often available [27]. In addition, the high shear forces involved with FACS can lead to cellular disruption and releases of DNA prior to any profiling attempt.

The use of LCM has increased over the years due to its speed and simplicity. It has also proven its applicability in forensic cases [7,22,28,29]. However, very often tissue or DNA from cells surrounding the selected specimen is cut out together with the cell or tissue of interest and this can lead to mixed profiles.

Another technique that has been increasingly used for the last years is microfluidics [27,30,31]. The strengths of using microfluidics for isolating cells include its high throughput capability, closed environment system which eliminates contamination, and the possibility to integrate amplification directly on one platform. However, specialized devices are required for setup and the microfluidic chips are not reusable. This means chips need to be bought or fabricated in house of which the latter is laborious and costly.

Another prominent cell sorting system available is based on di-electrophoresis. A commercially available device, DEP array (Silicon Biosystems) has recently been released and has proven its success in isolating single tumor cells, epithelial cells, white blood cells and sperm cells [32]. The unique feature of this technology is its ability to visibly follow a cell towards the place of deposition. This property

ensures that a cell is always present wherever it is transported without any post deposition control step. Unfortunately, this technology is very expensive since the chip required for this device is not reusable.

In the current study micromanipulation was chosen for cell isolation, because the micromanipulation platform was already available at the place of research. Other than that, micromanipulation allows very precise isolation of single cells including full visual control and enough throughput, since the isolation of hundreds of cells is not required. Also, since micromanipulation does not require many cells, it is a better option than FACS. Furthermore, micromanipulation has proven to be successful in a wide variety of applications including the isolation of sperm cells in forensic samples [10,33]. In contrast to LCM, contamination with neighbouring cells is unlikely with micromanipulation. In addition, micromanipulation is cheaper than LCM. However, micromanipulation is not suitable for high throughput cell isolation. Furthermore, micromanipulation is a technically and manually demanding technique. Although automated micromanipulations platform has been proven successful, automated cell isolation however is not available, hence making micromanipulation of single sperm cells labour intensive. Despite the low throughput and the required labour of micromanipulation, these would not be a large disadvantage in a forensic biology laboratory environment. This is because the number of cases this technique would be applied to is very low, compared to cases which can be analysed with conventional techniques. Thus, without the need of large amounts of expensive equipment or consumables and with a simple laboratory set up, micromanipulation can outperform other available techniques based on its simplicity and costs. Finally, the applicability/integration of using micromanipulation in a forensic laboratory does not have to be tested, since a complete set up is often already present in many forensic laboratories.

### 1.2.2. Isolation of single skin flakes

Using skin flakes as DNA evidence has been suggested in literature by only a few studies and is currently applied in some specialized forensic laboratories [34–36]. Every contact leaves a trace which can contain DNA. Other than DNA, fibres can be transferred from one person/object to another person/object. That's why decades ago when DNA was not used as evidence, investigators already searched for fibres on body parts, clothes or furniture. Pieces of tape were placed on the object's surface. Existing fibres then stuck to the adhesive side of the tape. By lifting the tape, fibres were collected from the object's surface and could be used for analysis. Other than fibres, also skin flakes can be collected along with fibres by tape lifting as seen on figure 1.2.2.1 (page 36). As a skin flake contains several cells which are shed or can be cast off by force and then left behind, they are considered as a biological trace containing DNA. Nowadays tape lifting is also used to secure DNA. A theoretical case example in which tape lifting can be used is for example, when a woman wearing a t-shirt is strangled by a man, the latter can deposit skin flakes from his arms onto the neck and breast area of the woman's t-shirt. By tape lifting or swabbing the neck and breast area of the t-shirt, DNA of the man can be collected and be used for DNA extraction and subsequent PCR. However, DNA from the victim wearing the t-shirt, DNA from the victim's friend she possibly hugged before being strangled and DNA from the person the t-shirt was borrowed from can all be present on the t-shirt along with the perpetrator's DNA. Since very often complete tapes in its whole are used for DNA extraction, the extract may contain DNA from multiple sources which then leads to mixed profiles. On the other hand, the victim is often the wearer of the piece of clothing and therefore continuously touching and leaving DNA on the surface of the clothing while wearing. In this case the vast majority of the skin flakes or of cell free DNA present on clothing is from the victim itself. This results in the masking of the perpetrator's DNA.

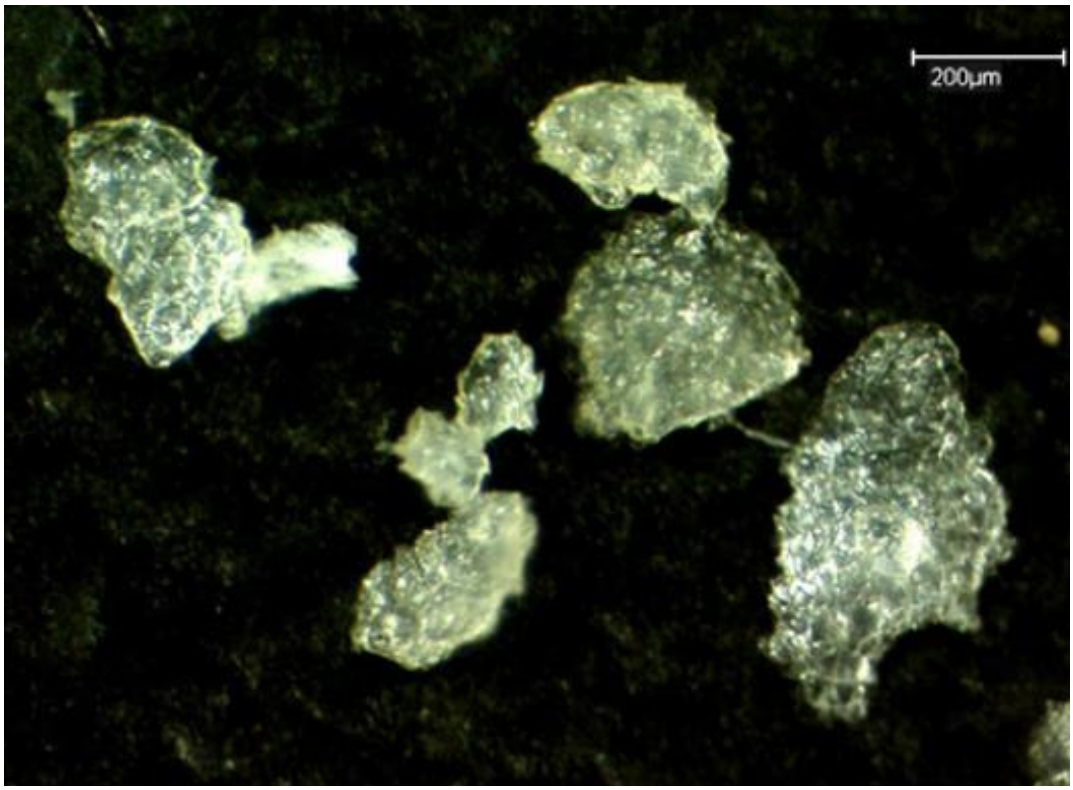


**Figure 1.2.2.1. Microscopic view of skin flakes secured on a tape lifting. (Figure taken from Schneider et al 2011** [37]**).**

Thus, the power from this type of biological trace is compromised by the mixtures derived from analysing a bunch of skin flakes instead of single ones separately. Thus, the strategy of isolating and DNA profiling of single cells could also overcome mixed profiles with DNA profiling from collected skin flakes. As with single sperm cells, single skin flakes can only originate from one donor. Furthermore, the possibility of DNA profiling of single skin flakes has been demonstrated in casework [37]. However, DNA profiling was successful only in 15% of single skin flakes. A third of the 15% gave full profiles and 10% yielded partial profiles [37].

PCR success relies on intact DNA from several cells. Skin flakes are derived from the epidermis, the outermost layer of the skin. The epidermis is predominantly

composed of keratinocytes, but also contains melanocytes, Merkel cells and dendritic cells. As keratinocytes travel from stratum basale, the inner part of the epidermis, to the outermost part, the stratum corneum, they undergo cornification as part of the cell differentiation to end up as corneocytes in the stratum corneum. Skin flakes consist of many cells and they contain little intact DNA due to DNA degradation during cornification. Furthermore, skin flakes are often difficult to distinguish from other particles present on collected samples e.g. cosmetic hair product particles and plant derived particles. For these reasons, DNA profiling of skin flakes often leads to profiles with many ADOs and exhibits a low success rate. A different approach should therefore be applied that combines the reliable identification of skin flakes with a sensitivity high enough to successfully profile their challenging DNA content.

Cell or tissue specific staining can be valuable to highlight skin flakes and to discriminate them from other similar looking particles on tape liftings. Methylene blue and trypan blue are dyes that might be suited to highlight skin flakes on tape liftings. Both dyes are used to discriminate live cells from dead cells [38]. Methylene blue is a cationic dye that includes a positively charged dimethylamino group as seen in figure 1.2.2.2. Methylene blue stains the nuclei dark blue by binding to the negatively charged organelles of the cell like DNA and RNA.

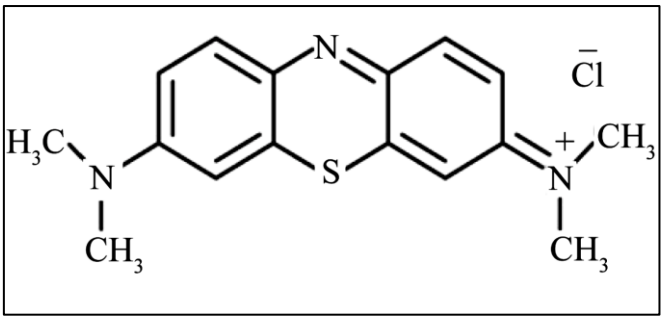


**Figure 1.2.2.2. Chemical structure of methylene blue. (Figure taken from Elmorsi 2011** [39]**).**

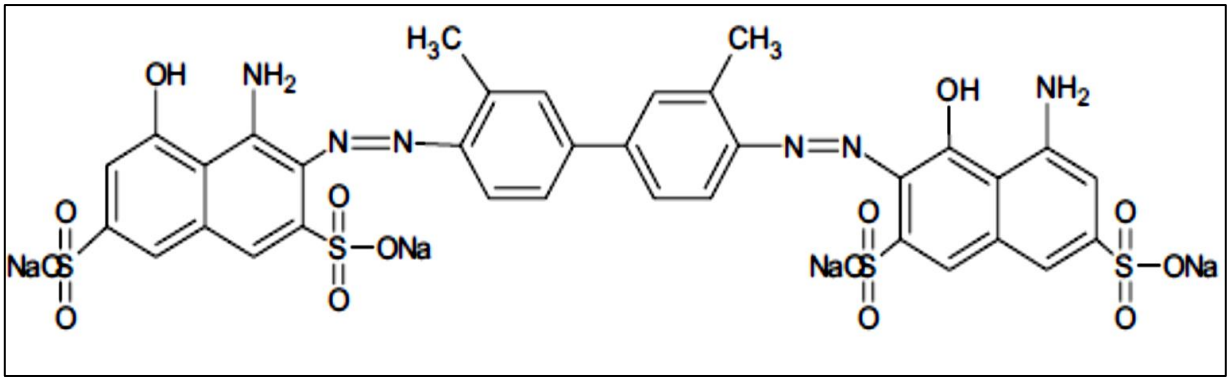


**Figure 1.2.2.3. Chemical structure of Trypan blue. (Figure adapted from Yathisha and Nayaka 2020** [40]**).**

Trypan blue is an anionic azo dye. Azo dyes are characterized by their functional group R-N=N-R' (*figure 1.2.2.3*) and are commonly used in the textile and food industry. As Trypan blue is an anionic azo dye it binds, through ion exchange, to cationic compounds e.g. proteins. Therefore, Trypan blue can stain complete cells and has been used to highlight biological particles including skin particles and loose epithelial cells to be used for forensic DNA profiling [33,41]. However, trypan blue stains all dead tissue or cells blue. Consequently, trypan blue does not discriminate skin tissue from other bio particles[42]. Moreover, trypan blue is a dye used in the staining of cotton fabrics. This feature can complicate the analyses of such fibres that are often found on tape liftings as well. A different anionic azo dye which is often used in combination with other dyes as a counter stain is Orange G (*figure 1.2.2.4)*.

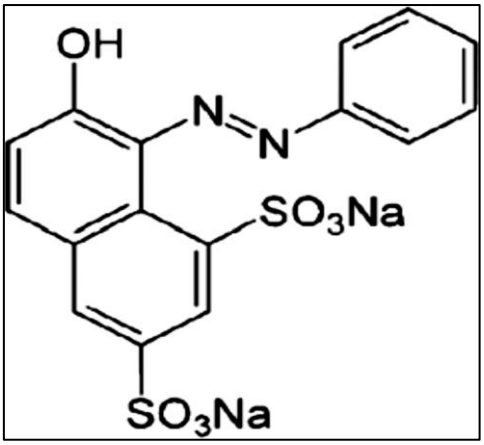


**Figure 1.2.2.4. Chemical structure of Orange G. (Figure taken from Mondal et al 2010** [43]**).**

This dye may be particularly well suited as it is used in the Papanicolaou stain to stain keratin [44] as well as to treat wool which contains keratin. Skin flakes are made up of keratinocytes which contain keratin. Keratin is only found in vertebrates, but not in plants. Accordingly, Orange G could be a potential stain to distinguish skin flakes from example plant derived particles. Furthermore, certain staining methods have the additional advantage of highlighting intact nuclei which might indicate less degraded nuclear DNA. Here, fluorescence microscopy can be used in combination with a DNA stain such as 4',6-Diamidin-2-phenylindol (DAPI) or Hoechst 33342. These fluorescent stains are widely used to detect nuclear DNA and nuclei in live or dead cells [45,46]. Both stains bind to dsDNA that is rich in adenine-thymine base pairs. DAPI has an affinity to bind to RNA as well [45]. That makes Hoechst 33342 more appropriate to stain skin flakes without false positives due to RNA fluorescence. Furthermore, Hoechst staining was used to investigate the correlation of allele recovery with the number of nuclei found in hairs [46]. On the other side, the intercalating or DNA-binding property of chemicals can cause changes in the DNA, also named as mutations. Consequently, DAPI and Hoechst 33342 could be less suitable for down-stream forensic DNA analysis compared to the non-intercalating dyes (e.g. methylene blue, Orange G and Trypan blue).

## 1.3. Strategies to analyse low copy number DNA

### 1.3.1. Classical approaches

Currently available forensic PCR kits use 28-30 cycles to amplify the set of forensic STR loci to be analysed. If low template DNA is used, these numbers of cycles are often not sufficient to analyse the PCR products. One classical approach to battle this problem is to increase the number of PCR cycles up to 33 cycles to increase the number of amplicon copies. This method can ideally result in less ADOs. However, more cycles also induce higher sensitivity and subsequently can lead to ADIs from contaminating DNA.

Another approach would be performing PCR in lower volumes. Previous studies had shown increased sensitivity when PCR is performed in reduced reaction volumes [4,28,47–50]. This technique is called low volume PCR (LV-PCR) and have been applied in several studies to increase allele recovery of LCN DNA samples. By adding template DNA in a lower reaction volume (<25µl to several pl), the concentration of template DNA in the total reaction volume increases compared to the same amount of template DNA added into standard PCR reaction volumes (25µl). Despite the increased sensitivity, LV-PCR has its limitations in a normal set up. Evaporation limits the reduction of reaction volume to a minimum volume of 5µl. However, lower volumes methodology can be used in combination with oil, the latter which seals the reaction volume and eliminates evaporation. Previous studies have used this approach in amplifying gDNA and single sperm cells in reaction droplets of 1.0µl – 2.5µl by using the AmpliGrid (Advalytix) slide as shown in figure 1.3.1.1 [48,50,51].

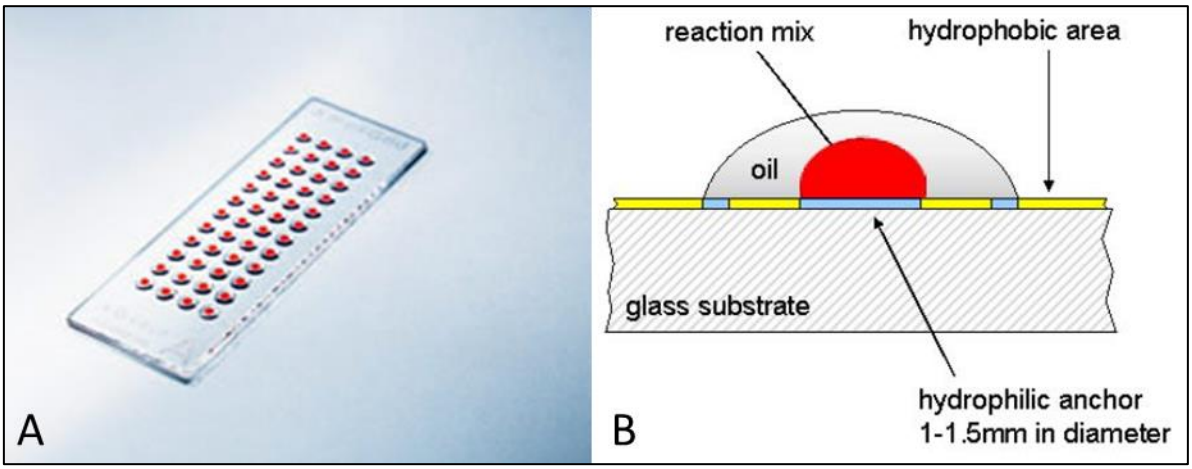


**Figure 1.3.1.1. The AmpliGrid system. An AmpliGrid slide after loading samples, reaction mix and Sealing Solution (A). The glass slides are designed with a structure of hydrophilic spots and hydrophic/hydrophylic and rings to hold the aqueous sample and oil cover in place(B).(Figure taken from** [52,53]**).**

Instead of using a PCR tube, the PCR reaction is performed in a transparent glass slide (called AmpliGrid) that contains hydrophilic spots and hydrophobic rings that keep the reaction droplet in place and prevents the oily sealing solution from

running away. gDNA or a single cell can be pipetted onto the hydrophilic spot which is then followed by pipetting the reaction mix and sealing with sealing solution to prevent evaporation. These slides allow easy work flow, since isolated cells can be directly deposited onto the slide under the microscope without any loss in transfer. Additionally, the slide can be put on a thermocycler for immediate PCR. Despite these advantages, the AmpliGrid slide is unfortunately no longer available on the market.

However, PCR in droplet form is still used but in smaller reaction volumes, e.g. in the format of emulsion PCR or of digital PCR. Emulsion PCR is based on emulsifying the reaction mix into oil which leads to small droplets that act as individual reaction chambers surrounded by oil [54]. Emulsifying can be performed by microfluidic chips, ultrasound or simply vortexing. This technique has been successfully applied for template enrichment by beads, single cell PCR, real time PCR and single cell sequencing. Digital PCR is also based on amplification in droplets which can be generated by in oil emulsifying as well as droplet in oil printing [55–58]. PCR in droplets assure an increased sensitivity as well as shorter run times and is cost effective with the use of fewer reagents.

A previous study showed the use of compartmentalization of single cells into agarose droplets for forensic DNA profiling on single cells [59]. Single cells were mixed with STR primer bond beads into agarose droplets and dispersed into an oil phase on microfluidic chip. The solidified agarose keeps the DNA inside the droplet during cell lysis and the beads are used to amplify the STRs. However, this method has several drawbacks. For example, every droplet produced theoretically contains the right combination of one cell and one bead but can in practice also contain only one of two components or none of them. Additionally, primer binding to beads is problematic since not every bead is covered with an even amount of each primer. This can lead to underrepresentation of STR loci and ADOs. Also, the need of statistical dilution of the beads to divide them into a micro well plate for PCR results in a high number of samples without amplification resulting in high costs due to loss of reagents. The concept of amplification in

nanoliter (nl) droplets is still a promising method for the future, but this approach requires optimization. The use of oil has its advantage in that it conducts heat better than an aqueous phase. However, oil can make a PCR less efficient since the oil-aqueous interface of the droplet is prone to polymerase sticking. When polymerase sticks on the interface it is not able to participate in the amplification enzymatic reaction anymore. This can often be prevented by adding bovine serum albumin (BSA) to the reaction mix [60]. BSA has also been used as a PCR improver which means it will not have a negative influence on the reaction. Additionally, a previous study used BSA for hydrophilic spots as an alternative to the laborious fabricating of PDMA chips which are used in droplet printing PCR [61]. Another study showed that the nonionic surfactant Brij L4 decreases the adsorption of polymerase on the interface when Brij L4 was mixed into the oil [60]. Therefore, Brij L4 could also be used as an alternative to or combined with BSA to increase PCR efficiency.

### 1.3.2. Whole genome amplification

The STR loci used in forensic DNA profiling are amplified by using a multiplex PCR. The multiplex PCR only amplifies the loci of interest and not the rest of the genome which makes PCR a specific amplification reaction. In contrast to PCR (whether a monoplex or multiplex PCR) another amplification reaction, known as the whole genome amplification (WGA), is a rather unspecific amplification reaction since as it amplifies the whole genome available in the sample. Different WGA methods are discussed below. Furthermore, WGA is used on single cells and has been tested in combination with forensic DNA profiling. Also, why WGA could be a potential tool to be used in forensic DNA typing of single spermatozoa and skin flakes is described further on.

### *1.3.2.1. WGA methods*

Despite the availability of many techniques for cell isolation, to the best of my knowledge there has been no published study that describes the successful DNA profiling on single cells with currently available forensic STR kits, with quality and reliable results for routine forensic analysis. To date, only next generation DNA sequencing has been successfully applied on isolated single cells [62,63]. To ensure that enough DNA is available to be sequenced, the genomic DNA can be first pre-amplified by whole genome amplification (WGA) protocols. The latter can be performed by various protocols, some of which are based on multiple displacement amplification (MDA), whereas others are PCR based, such as modified improved primer extension pre-amplification (miPEP) or by a combination of the two, such as multiple annealing and looping based amplification cycles (MALBAC) [55,62,64–69].

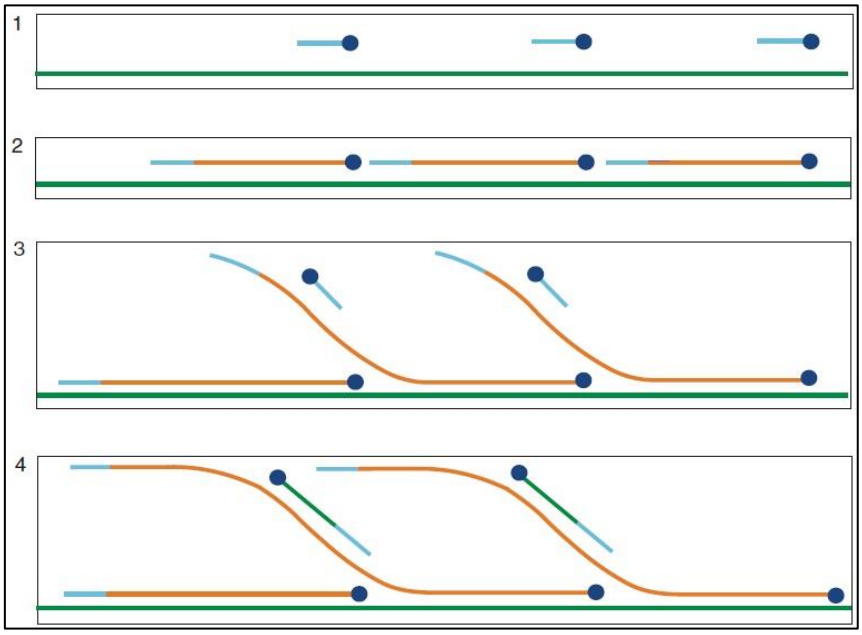


**Figure 1.3.2.1.1. Schematic representation of the multiple displacement amplification MDA. At first, random hexamers (blue lines) bind to denatured DNA (green line) (1). Secondly, phi29 DNA polymerase (blue circles) extends primers until it reaches newly synthesized dsDNA (orange line) (2). Afterwards, the enzyme displaces the strand and pursues polymerization, while primers bind to newly synthesized DNA (3). At last, DNA synthesis starts on the new strands. (Taken from** [68]**).**

In MDA, random hexamer primers and the isothermal phi29 DNA polymerase are used to randomly amplify DNA along the genome (*figure 1.3.2.1.1*). During DNA synthesis, phi29 displaces double stranded DNA (dsDNA) thus generating new single stranded DNA (ssDNA) strands. By again binding random primers to new single strands, DNA synthesis subsequently continues thus generating an arborized pattern of amplified DNA. A modified version MDA uses a second polymerase, called TthPrimPol, instead of the random primers [70]. TthPrimPol randomly synthesizes primers on ssDNA. To be effective, MDA requires long stretches of intact DNA. Several protocols (such as RCA-RCA or BL-WGA) have adapted the MDA principle to enable amplification of degraded genomic DNA samples as are often encountered in forensic context [71,72].

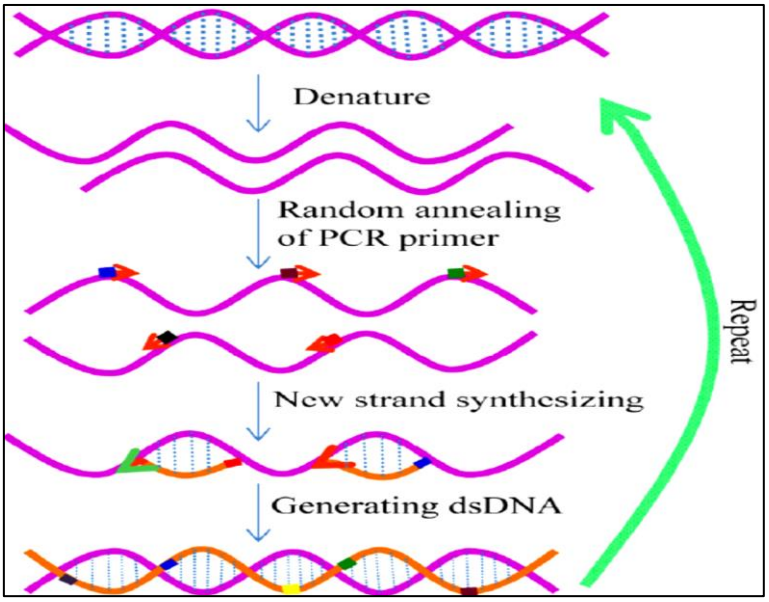


**Figure 1.3.2.1.2. Schematic representation of the modified improved primer extension preamplfication (miPEP). In each cycle, primers randomly anneal on ssDNA after denature of the DNA template. Then new DNA strands are synthesized by a mix of *Taq* DNA polymerase and a high fidelity DNA polymerase. Afterwards, a new cycle starts again. (Figure taken from** [73]**).**

miPEP uses random 15 bases primers (*figure 1.3.2.1.2*) which are combined with a high fidelity DNA polymerase mix and PCR with low annealing temperatures (37°C) and long extension steps (4 min on 55 °C).

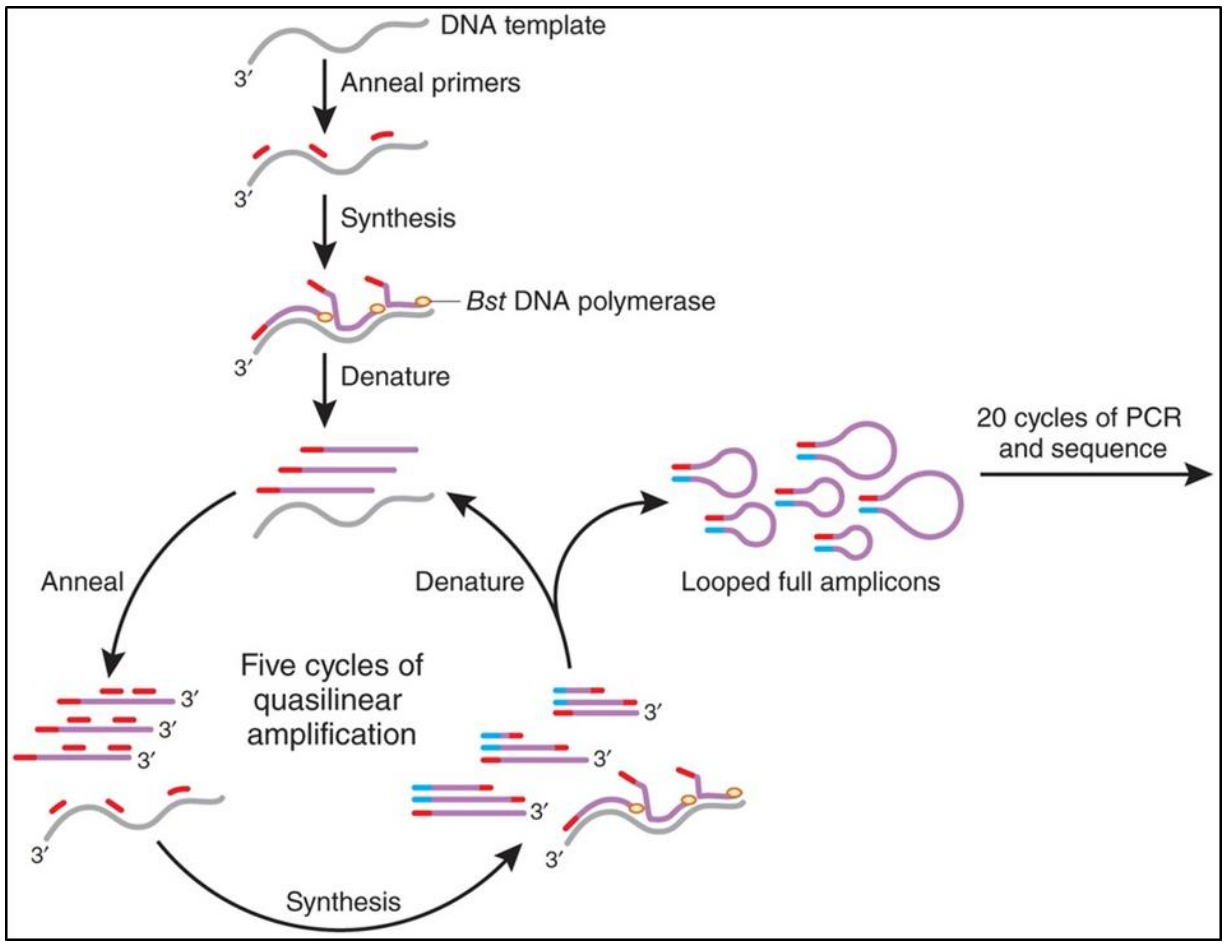


**Figure 1.3.2.1.3. Schematic representation of the multiple annealing and looping based amplification cycles (MALBAC) reaction** [74]**. First, MALBAC primers (red) anneal to the DNA template. Strand-displacement synthesis by *Bst* polymerase synthesizes partial amplicons. These amplicons are then denatured from the template at 94 °C. Again priming is followed to new positions on the DNA template and synthesizes more partial amplicons. Priming and extension on partial amplicons produces complete amplicons having the MALBAC primer sequence at the 5′ end (red) and its complementary sequence at the 3′ end (blue). By subsequently denaturation at 94 °C, the original template regenerates a now larger and more diverse pool of partial amplicons. The complete amplicons form loops, which are resistant to further amplification or hybridization. After five cycles, the complete amplicons are exponentially amplified by 20 cycles of PCR using primers which are complement to the common regions of the MALBAC primers.(Figure taken from** [74]**).**

MALBAC combines the displacement feature from MDA with a PCR. By using special MALBAC primers and *Bst* polymerase, DNA is randomly synthesized and displaced (*figure 1.3.2.1.3*). After displacement the dsDNA is denatured into ssDNA which enables not only primer binding to the newly synthesized DNA but

also to the template. By priming and extension on partial amplicons, complete amplicons are generated which by subsequent denaturation form loops. These amplicon loops prevent further amplification by the displacement reaction thus ensuring quasi-linear amplification. After several cycles, the amplicon loops are exponentially amplified by PCR with primers complement to the MALBAC primers.

#### *1.3.2.2. WGA of single cells*

Currently a variety of WGA kits for single cells are commercially available: REPLI-g (Qiagen, MDA based), TruePrime (Sygnis, MDA based), Genomiphi (GE Healthcare, MDA based), Genomeplex (Sigma Aldrich, PCR based), MALBAC (Yikon Genomics, MALBAC based) and Ampli1™(Silicon Biosystems, PCR based) [64,75–77]. Many of the single cell WGA methods available have been successfully used on single cells of human origin, but also on single bacterial cells [25,78–80]. In research focussing on the genetics of single cells, DNA sequencing is used to analyse the DNA sequence of the genes or other loci of interest. In order to sequence these loci, a PCR is used to amplify enough copies to be analysed with a DNA sequencing method e.g. sanger sequencing. As previously described, PCR requires enough template DNA to produce sufficient copies of the loci of interest. Since single cells do not contain enough DNA for PCR, WGA is used to pre-amplify the whole genome of the cell first to produce enough DNA. Thereafter, the WGA product is then used as template DNA for the PCR to sufficiently produce enough copies which can be used for DNA sequencing.

#### *1.3.2.3. WGA in forensics*

Up to now several WGA methods have been tested on their potential applicability in forensics as pre-amplification to enable DNA profiling on low template DNA [67,68,81,82]. Unfortunately, none of them were successful. Despite many reaction modifications, no WGA has been shown to increase the sensitivity in DNA profiling to the level of one cell. Strikingly, one study showed that the sensitivity of the regular DNA profiling, PCR gave higher sensitivity than when a WGA approach as pre-amplification was executed [67]. This could be due to multiple

reasons. Firstly, the use of random primers could lead to primer binding within an STR locus instead of outside the STR locus. This leads to incomplete amplification of the STR locus and PCR primer binding regions. When using the WGA product containing an incomplete amplified locus, PCR is not able to amplify the whole locus and can therefore not be detected with capillary electrophoresis. Secondly, stochastic effects could cause an uneven amplification of the genome and randomly underrepresenting specific regions. Thirdly, fragmentation-based WGAs such as Genomeplex may cut within an STR locus which prevents amplification of complete STR loci. These factors put together can lead to major ADOs or ADIs. In particular Bst polymerase, used in MALBAC, is prone to allele slippage and could lead to false alleles or ADIs [83].

#### *1.3.2.4. Could WGA still be an option for forensic DNA profiling of single spermatozoa and single skin flakes?*

One study showed the ability of genotyping of single spermatozoa by using REPLI-g in combination with a multiplex PCR based on 5 STR makers [82]. Although the study was not performed in a forensic context, the REPLI-g based WGA was described to work well with multiple STR markers. This gives great potential to use WGA like REPLI-g as a possible tool to increase the sensitivity of forensic DNA profiling to the level of single spermatozoa. The previous study did show some errors as sometimes two alleles were determined instead of one allele. This was explained as pipetting error. Here, micromanipulation to isolate single spermatozoa was not applied. Instead, samples with single spermatozoa were prepared by pipetting of a strongly diluted spermatozoa suspension. By this, each tube would consist of one spermatozoon theoretically. This means that some tubes could contain more than one spermatozoon which could result in obtaining more than allele from a specific STR locus. In the case of the current study, micromanipulation of single spermatozoa will be used which could exactly assure not more than one spermatozoon will be added to the WGA and subsequently not more than one allele should be obtained.

Taken together these positive and negative features, the applicability of WGA in forensic DNA profiling of single spermatozoa and single skin flakes is still

questionable. Furthermore, recent WGA methods, especially the WGA kits specifically designed for single cells have yet to be tested in forensic DNA profiling applications. Therefore, it is necessary to further test WGA including newer WGA and single cell WGA kits as a potential tool to provide forensic DNA profiling of single spermatozoa and single skin flakes.

## 1.4. Improved analysis of degraded DNA using Next generation DNA sequencing

STR analysis using CE is based on the detection of different fragments (PCR products) separated by length. This is also called fragment analysis. The so identified STR alleles differ in size. However, for several forensic additional variations in sequence have been reported that will not lead to differences in length. For example, at STR locus D2S441 person A and person B are both homozygously containing allele 8. Though, at position 93bp in the DNA sequence of allele 8, person A contains a "G" (guanine) in contrast to person B which contains an "A"(adenine). Thus, even though allele 8 contains the same fragment size, two different "versions" of alleles can occur/exist based on the SNP at position 93bp of the DNA sequence. These two different "versions" of the same allele are called isoalleles. As fragment analysis using CE is only based on size, these isoalleles cannot be determined. Thus identification of isoalleles requires DNA sequencing. Classical DNA sequencing by using Sanger sequencing, however, is not suited for forensic STR typing, as it does not allow for sequencing multiple STR loci or alleles in parallel, i.e. does not permit a multiplexed analysis.

A sequencing technique capable of multiplex analysis is next generation DNA sequencing (NGS) which is often called massively parallel sequencing (MPS), second generation sequencing or high-throughput sequencing (HTS) [84]. Even the STR alleles of DNA mixtures can be analysed, since NGS is based on sequencing of single DNA molecules. As the sequence information at the same

time bears the information of fragment size, NGS is well suited for STR typing and for the identification of isoalleles.

Further advantage of NGS-based STR analysis over CE-based analysis consists in its improved ability to analyse degraded DNA samples [84]. The STR markers, which in forensic DNA profiling are amplified by PCR, differ in fragment size (e.g. 50bp – 350bp). The correct assignment of loci requires that amplicon size ranges must not overlap within one colour channel. Inevitably, inclusion of more loci results in bigger size ranges of the amplicons within one colour channel. Degraded DNA is fragmented DNA containing short fragments compared to intact high molecular DNA. As a result, long amplicons cannot be amplified by PCR. Therefore, short amplicons are required for STR analysis of degraded DNA. Mini-STRs are more suitable smaller amplicons, but color detection channels of capillary electrophoresis are limited. Consequently the number of STR loci not overlapping in size is limited. As a result, no mini-STR multiplex PCR system exists that completely cover all loci from a conventional STR multiplex PCR system. This is where next generation DNA sequencing (NGS) could become an alternative to capillary electrophoresis. With NGS size overlap is not an issue, because identification is based on sequence, not on size. Furthermore, with NGS the design of primers for small amplicons is possible for all loci.

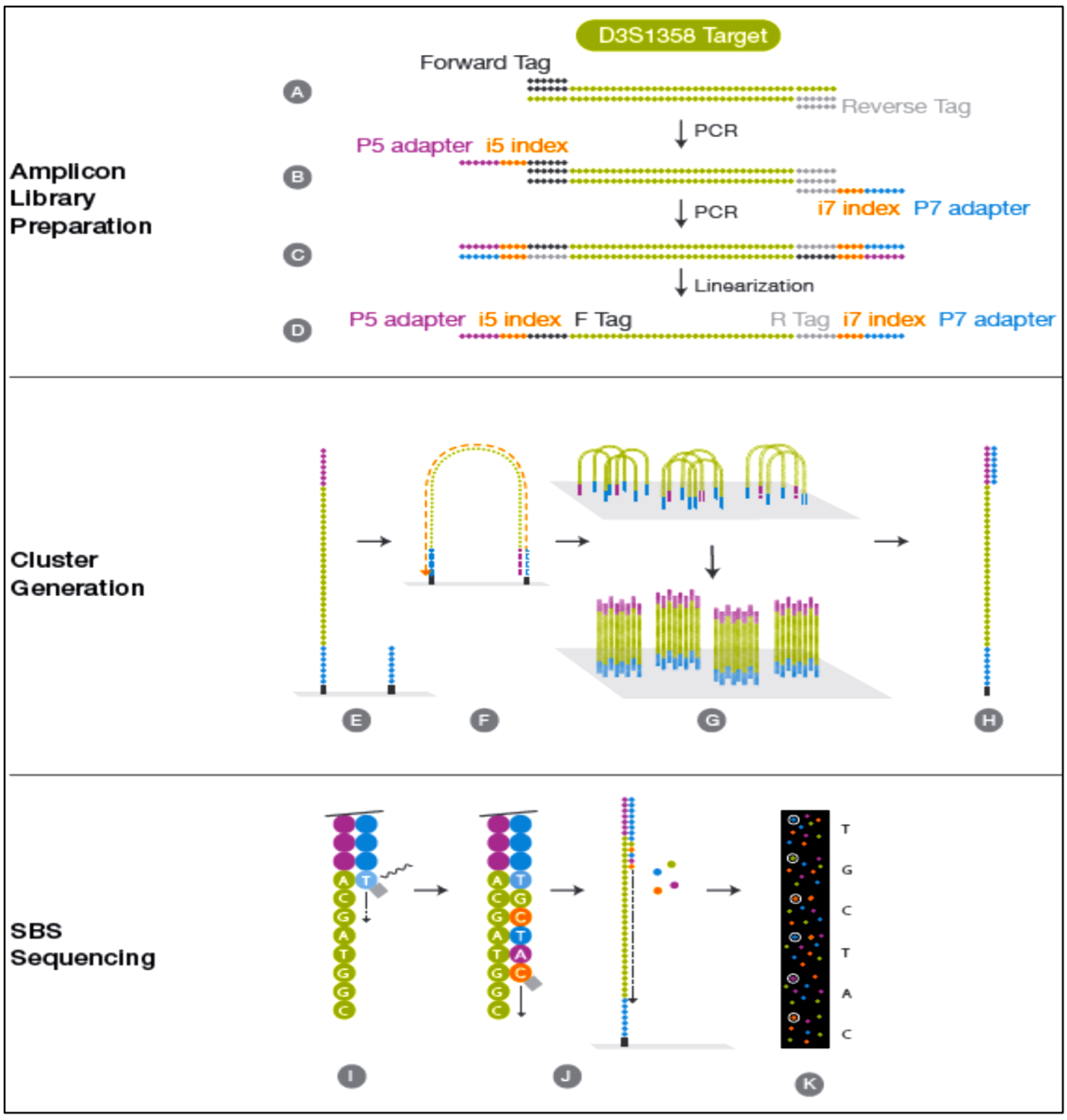


**Figure 1.4.1. Schematic overview of NGS using the Illumina platform. NGS starts with an amplicon library preparation. The amplicon, for example D3S1358, is first amplified using a PCR and tag including forward and reverse primers (A). Afterwards, indices and adapter sequences are added (B). The amplified products including the amplicon+tags+indices+adapters are purified, normalised and pooled (C). The pooled library is then linearised prior to be loaded onto the flow cell (D). Olligo's complementary to the adapters are bound to the surface of the flow cell and capture the linearised fragments (E). By bridge amplification, spatially separated clusters of many copies of identical molecules are generated (F, G). Following cluster formation, the sequencing primer is added and anneals to the adapter (H). Sequencing of the fragments is accomplished by sequencing-by-synthesis (SBS) which includes sequential incorporation of a fluorescently labelled dNTP which is followed by the cleavage of a terminator plus the fluorophore to allow for incorporation and detection of the next complementary base in the template's sequence (I, J). Detection of the incorporation of the fluorescently labelled bases is established by using a high resolution camera (K). (Figure taken from www.illumina.com**[85]**).**

With NGS millions of PCR amplicons are read on a base-by-base procedure. The principle of the NGS technique goes as follows. At first, individual target DNA molecules are spatially separated and amplified [86]. Second, the separated single DNA molecules are amplified. Afterwards, the sequence of each of these spatially separated, amplified copies is determined individually in spatially separated sequencing reactions.

The IonTorrent is one of the NGS platforms currently available. IonTorrent makes use of a bead-based emulsion PCR of individual template molecules and a complementary metal–oxide–semiconductor (CMOS) chip [87]. The beads on which the amplified single DNA molecules are bound are spatially separated by distributing them into individual cavities of the CMOS chip. By sequentially exposing the amplicon's DNA to individual nucleotides, hydrogen ions are released upon nucleotide incorporation and can be detected by the CMOS chip.

The NGS platforms from Illumina are based on the sequencing-by-synthesis (SBS) process and use a flow cell to separate the DNA molecules to be sequenced (see figure 1.4.1.) [85,86]. A high resolution camera is used to detect sequential incorporation of fluorescently labeled dNTPs that contain reversibly bound terminators. Incorporation is followed by the chemical cleavage of both fluorophore and terminator to allow incorporation of the next complementary base in an amplicon's DNA sequence. Although NGS has not been used as excessively as CE, the ability to analyze STRs without many errors makes SBS based NGS platforms such as the MiSeq Forensic Genomics (FGxTM) System the current leader in forensic DNA profiling. In contrast to CE-based analysis, signal strengths are not recorded as RFU's but as numbers of reads, i.e. how often a given sequence has been detected.

The advantages of NGS is that each single fragment can be analyzed simultaneously regardless of fragment size. This makes it possible to combine autosomal STR markers with lineage markers such as Y-STRs or X-STRs, identity SNPs as well as phenotypic and ancestry markers [88]. Currently, a few multiplex

PCR systems are available for NGS which enables the analyses of all of these markers at once. Moreover, an in-house (laboratory of Prof. Richard Jäger, Hochschule Bonn-Rhein-Sieg University of Applied Sciences) developed mini-STR multiplex PCR assay for SBS based NGS has currently been developed which covers the 16 STR markers included in the German system. The use of this mini-STR-NGS assay enables improved DNA profiling of degraded DNA. However, the use of this system has not yet been tested on many different types of DNA samples containing degraded DNA, including skin flakes.

## 1.5. Research aims

DNA profiling in forensic casework struggles with complex mixed profiles as a consequence of increased sensitivity of the analysis as well as the presence of contaminating cells from another contributor or DNA resulting from sampling the evidence. DNA profiling of single cells and single skin flakes could circumvent this problem. However, no specific method is yet available that describes successful DNA profiling of single sperm cells or of single skin flakes. Therefore, the general aim of this study is to develop a method which enables DNA profiling of single sperm cells or of single skin flakes, respectively, in mixed DNA samples.

Of the variety of techniques for cell isolation, cell staining and DNA amplification described above, the most appropriate ones will be selected for sperm cells as well as for skin flakes. These selected techniques will be integrated in a workflow which will be assessed with respect to its efficiency on mock sexual assault swabs spiked with sperm as well as on tape liftings taken from pillows with skin flakes from multiple persons.

Specific objectives to achieve the aim of the project:

1) Single sperm cells

i. Assessing and comparing the impact of different WGA assays on forensic DNA profiling of low copy number DNA based on the allele recovery and number of ADI.
ii. Optimize micromanipulation of single sperm cells in order to be compatible with the selected amplification method.
iii. Assessing the performance of the selected amplification method on single sperm cells from different persons based on allele recovery and ADI of autosomal, X- and Y-STR systems.
iv. DNA profiling of single sperm cells from mock sexual assault swabs with a single or multiple male contributors.

2) Single skin flakes

v. Assessing different staining assays to detect skin flakes on tape liftings based on discrimination power and PCR inhibition.
vi. Assessing the performance of DNA extraction methods on skin flakes based on rate of dissolving and PCR inhibition.
vii. Assessing the impact of the selected WGA method on single skin flakes.
viii. DNA profiling of single skin flakes on tape liftings taken from mocked up mixed DNA evidence.
ix. Determining and comparing DNA quantity, DNA quality and allele recovery of single skin flakes from different persons.
x. Assessing the performance of a mini-STR approach with NGS on single skin flakes.

# Chapter 2

# Materials & Methods

# Chapter 2. Materials and Methods

## 2.1. General laboratory conditions

To prevent DNA contamination all work surfaces were made DNA-free by applying the bleach "Sagrotan Schimmelfrei" (Reckitt Benckiser Deutschland GmbH, Germany) and rinsing with 70% ethanol. Also all pre-PCR tools such as pipettes, pencils, tweezers, tube racks, vortexers, centrifuges and microscopes were wiped with Sagrotan Schimmelfrei and 70% ethanol. All mastermixes for PCR and WGA were pipetted inside a UV-hood. All pre-amplification work was done while wearing hair nets and mouth masks in a PCR-free laboratory room. Negative controls of the reagents used were performed and the DNA profile of the person conducting the experiment was generated to check for any contamination incidences.

## 2.2. Biological material and mock sample preparation

In order to perform the experiments described in this research project, people were asked to volunteer by anonymously donating their semen, buccal, skin flakes or participating in a mock scene. Furthermore each volunteer donated one buccal mouth swab for DNA extraction as reference sample. Each volunteer gave their consent for their participation by signing a consent form. Examples of these consent forms can be found in appendix VI and VII.
Ethical approval for conducting the research described in this thesis was granted by the Robert Gordon University's Ethics Review Panel. A signed copy of the letter of ethical permission can be found in appendix V.

### 2.2.1. Sperm

Collected semen was stored at -20°C. All mouth swabs were dried and stored at -20°C.

Neat semen from two donors was used for testing with WGA and autosomal STRs as proof of principle. Furthermore, neat semen was used for testing with WGA combined with X-, and Y-STRs.

### 2.2.2. Mock intimate swab preparation

Mock intimate swabs were created by spiking mouth swabs from female donors with 100µl 1:50 diluted semen from one donor. For mock intimate swabs with two contributing men, female mouth swabs were spiked with 100µl 1:50 diluted semen from each male donor. Spiked mock intimate swabs were air dried at room temperature and stored at -20°C. All mock intimate swabs were created at the Scottish Police Authority – forensic services.

### 2.2.3. Skin flakes and hair product particles

Skin flakes were collected from one man by rubbing and shaking his head above a piece of paper and were stored in a light protected container at room temperature. Those skin flakes were used in bulk lysis experiments.

For staining experiments, skin flakes were collected from two men by rubbing and shaking their heads above a clean black cotton t-shirt. Tape liftings were used to recover the skin flakes from the t-shirt. Hair product particles were created by using Barbie dolls and three different hair products: Isana flexible finish hairspray (Rossmann, Germany), taft matt look hair gel (Schwarzkopf, Germany), got2be beach matt (Schwarzkopf, Germany) (see *figures 2.2.3.1 and 2.2.3.2, page 57*). Before applying the hair products, dolls were washed with dishwashing detergent, bleach and rinsed with tap water. After drying, each hair product was applied on the hair of one doll and kept at room temperature to dry for 2 hours. Hair product particles were released from each doll by combing its hair with a clean dry comb (also cleaned with dishwashing, Sagrotan and rinsed with tap water) above a clean black cotton t-shirt. Lifting tape (Neschen Filmolux 609 transparent,

Neschen, Bückeburg, Germany) was cut into 5cmx3cm pieces and were used to recover the hair product particles from the t-shirt and stored in a petri dish protected from light at room temperature. Tape liftings with skin flakes mixed with hair product particles were created by dropping both skin flakes and hair product particles from the dolls at the same spot on the t-shirt. Isolated single skin flakes on glass slides, tape liftings with either skin flakes, hair product particles or both were used to test the performance of the dyes Methylene Blue, Orange G and Trypan Blue.


**Figure 2.2.3.1. Hair products used to create solidified/dried hair product particles. Hair spray (Haarspray, ISANA), hair gel (taft LOOKS MATT LOOK, Schwarzkopf) and beach matt (got2b Strand matte, Schwarzkopf) where used.**


**Figure 2.2.3.2. Method to create hair gel particles by using a doll. (A) A Barbie doll which was washed and dried before applying hair products onto the doll's synthetic hair. (B) Combed hairs with dried hair gel.**

### 2.2.4. Mock skin flake DNA samples preparation

For validation studies, two different mock skin flake DNA samples were created at Hochschule Bonn-Rhein-Sieg University of Applied Sciences: mock samples with a single contributor and mixed mock samples. For the mock samples with a single contributor, 3 male volunteers (1 bald, 1 with short hair, 1 with medium long hair) and 1 woman (long hair) were asked to lay down, place their head on a clean pillow and relax while moving as natural as possible for 30 minutes. New bought pillows in vacuum cleaned 40cmx40cm new pillowcases (fabric: 100% polyester) were used. Then three 5cmx3cm tape liftings (Neschen Filmolux 609, Neschen, Bückeburg, Germany) per pillow were taken to recover the skin flakes left by each person. For the mixed mock samples, 2 or more volunteers were asked to lie down on one clean pillow after each other. Then 3 tape liftings were taken to recover the skin flakes left on the pillows.

### 2.2.5. Staining of tape liftings

Tape liftings with skin flakes, hair gel particles and beach matt particles were first separately tested by staining with different concentrations of three different dyes: Orange G (0.5 - 5%), Methylene blue (0.1 – 1,5%) and Trypan blue (0.4%). Staining was performed by placing tape liftings on a DNA free surface first. Then the staining was sprayed onto the tape lifting from a minimal distance of 10cm using a spray bottle (*figure 2.2.5.1.*). Stained tape liftings were incubated at room temperature for 5 minutes. Then the tape liftings were rinsed with DNA free water using a squeeze bottle or pipette. Tape liftings were then put into petri dishes and fixed with tape. The petri dishes containing the tape liftings were then put into a convection oven for 30 minutes at 25°C with the fan for air circulation switched on. The petri dishes were then closed with their lids and stored protected from light at room temperature

All tape liftings used for DNA profiling of single skin flakes were stained with a 3% Orange G solution. Tape liftings were put on a DNA free surface. Then the staining was sprayed onto the tape lifting from a minimal distance of 10cm using a spray bottle. Stained tape liftings were incubated on room temperature for 5 minutes. Then the tape liftings were rinsed with DNA free water using a squeeze bottle or pipette. Tape liftings were then put into petri dishes and fixed with tape. The petri dishes containing the tape liftings were then put into a convection oven for 30 minutes at 25°C with the fan for air circulation on. The petri dishes were then closed with their lid on and stored protected from light at room temperature until further analysis.

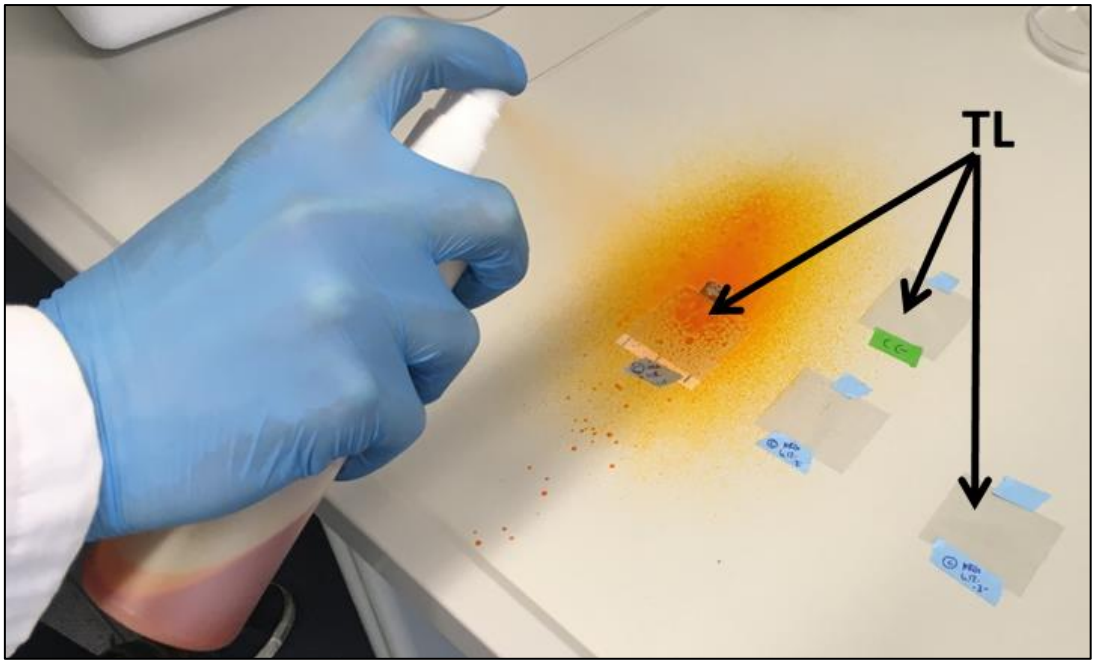


**Figure 2.2.5.1. Staining of skin flakes on 4 pieces of tape liftings (TL) by spraying with Orange G.**

## 2.3. Extraction of spermatozoa from swabs

In order to release spermatozoa from the cotton swabs, frozen mock swabs were thawed at room temperature for 30 minutes. Then the cotton part was cut off from the swab's end on a DNA free glass petri dish by using a scalpel and split in half. Each cotton half was put in a 1.5ml microcentrifuge tube by using a pair of tweezers. Then 500µl of buffer ATL (Qiagen, Germany) and 25µl of Proteinase K (Qiagen, Germany) was added to each tube. Afterwards, the tubes were vortexed for 10 seconds before incubating them in a thermoshaker for 30 minutes at 56°C with at least 900rpm shaking/agitation to get enough liquid moving around the cotton parts. After 30 minutes, the tubes were vortexed again for 10 seconds and incubated again for 30 minutes at 56°C with shaking/agitation. Tubes were then vortexed for 10 seconds and briefly spun down. Next, each cotton part was put into a MiniStrainer (pluriSelect, Leipzig, Germany) with 10µm pore size that was placed back into the tubes and centrifuged at 14,000g for 5 minutes to release

remaining liquid with spermatozoa left on the cotton parts. The MiniStrainer containing the cotton part and almost all of the supernatant were discarded without disturbing the pellet. Then 500µl 1x PBS was added to the tubes and pellets were re-suspended by tapping and pulse-vortexing. Afterwards, the cells were pelleted again by centrifuging at 14,000g for 5 minutes and almost all of the supernatant was discarded without disturbing the pellet. The pellet was resuspended in 300µl of DNA free water by tapping and pulse-vortexing. Then, tubes were spun down briefly in order to remove any liquid on the top of the tube and lid. Afterwards, 20µl of cell suspension was placed on a microscope slide and analysed with a light microscope at 200x magnification. In case the suspension was highly concentrated with spermatozoa, the suspension was diluted with DNA free water to sufficiently separate the spermatozoa.

## 2.4. Micromanipulation

### 2.4.1. Micromanipulation of spermatozoa

Single spermatozoa were isolated by using micromanipulation. Since spermatozoa stuck to the surface of glass microscope slides, the surface of different material was tested on non-stick properties first. The following materials were tested: polyethylene terephthalate, polycarbonate, polyethylene, polymethyl methacrylate, tesafilm® transparent adhesive tape (tesa SE, Germany), polystyrene petri dishes and non-adherent cell culture flasks, glass microscopy slides coated with WD-40® (WD-40, USA) or coated with Repel-Silane ES and petri dishes filled with 1% agarose gel.

The 1% agarose gel was selected for micromanipulation. This agarose gel was made by mixing 0.25g of agarose with 25ml DNA free water, followed by heating

in a microwave oven. First, petri dishes were filled with liquid 1% agarose just enough until the bottom plate was fully covered and incubated at room temperature until the agarose gel solidified. Then 100µl-300µl of cell suspension was placed on the gel, spread out by gently swirling the petri dish and left to dry for at least 30 minutes or until no liquid was visible. In case of diluted suspensions, an agarose gel petri dish for each dilution was prepared. Micromanipulation was performed by placing the agarose gel petri dish containing the spermatozoa underneath an inverted microscope (Axiovert S 100, Zeiss, Oberkochen, Germany) equipped with a Transferman micromanipulation system (Eppendorf, Germany) and the search for spermatozoa was carried out at 200x magnification. A small amount of adhesive was collected from a piece of 3M tape (3M, USA) with a tungsten needle with a tip diameter of 5µm (PLANO GmbH, Germany). The adhesive on the tip of the needle was used to pick up a single spermatozoa from the agarose gel surface and then transferred into a 200µl PCR tube containing 4µl DNA free water by swirling the tip of the needle in the water around for 10 seconds. Tubes containing isolated single sperm cells were then ready to be used for whole genome amplification with REPLI-g Single Cell WGA kit (Qiagen, Germany).

### 2.4.2. Micromanipulation of skin flakes

Single skin flakes from tape liftings were isolated by first placing tape liftings underneath a StemiDV4 stereomicroscope (Zeiss, Germany). By using a pair of Inox08 tweezers (Dumont, Switzerland) single skin flakes were picked from the tape lifting and placed at the bottom of a 200µl PCR tube. Skin flakes that looked like clumps were first cut into smaller pieces and mashed on a glass slide with a scalpel before placing in a 200µl PCR tube.

## 2.5. DNA extraction

### 2.5.1. DNA extraction of mouth swabs

Genomic DNA (gDNA) from buccal cells taken with cotton swabs was extracted using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). Swabs were incubated at 56 °C for 10 min in a mixture of PBS (400 µl), proteinase K (20 µl) and Buffer AL (400 µl). After incubation, ethanol (400 µl, 99%) was added to the sample and vortexed immediately. From that mixture, 700µl was added to a QIAamp Mini spin column and centrifuged at 8000rpm for 1 min. After discarding the flow through, the remaining volume of the mixture was loaded onto the same spin column and centrifuged under the same conditions followed by discarding the flow through. Afterwards, the spin column was washed with AW2 buffer (500 µl), centrifuged at 14,000 rpm for 3 min and the flow through was discarded. The spin column was centrifuged again at 14,000 rpm for 1 min to dry the column. Finally, the DNA bound to the column was eluted with distilled water (150 µl) followed by centrifugation at 8000 rpm for 1 min. The extracted DNA was stored at 4 °C.

### 2.5.2. DNA extraction of skin flakes

For DNA extraction from skin flakes 3 different methods were compared with skin flakes in bulk by assessing the lysis ability on solubility. For each of the protocols an 8mm³ pinch of skin flakes was added to a 1.5ml microcentrifuge tube and incubated for 3h. Lysis buffer and incubation temperatures were as follows:

- Chelex based protocol using 100µl Chelex (Biorad, USA) and incubated at 56°C. This is a slightly modified protocol described by Lorente et al. (1998) [35].
- Chelex based protocol using 100µl Chelex (Biorad, USA), 10µl Prot K (Qiagen, Germany) and 4µl 10% Tween20 and incubated at 56°C. This is a slightly modified protocol described by Schneider et al. (2011) [37].

- Alkaline based protocol using 100µl of buffer containing 0.4M NaOH and 8% Tween20 incubated at 60°C. This protocol was based on a protocol described by Sinkiewicz et al.(2017) [89].

For single skin flakes 4µl of lysis buffer containing 0,3M NaOH and 8% Tween20 was added to the 200µl PCR tubes containing the skin flakes at the bottom. Tubes were spun down briefly to make sure the skin flakes were submerged in lysis buffer. If necessary, the tubes were vortexed and subsequently spun down again to make sure the skin flake was covered by the solution. Tubes with single skin flakes were incubated in an ABI 9700 thermocycler (ThermoFisher, Germany) for 3h at 60°C with a heated lid at 103°C. Skin flake clumps were incubated for 12h at 60°C. After lysis, 4µl of 0.3M HCl was added to neutralize the lysates. Crude lysates were stored at 4°C and were directly used for PCR.

## 2.6. Whole genome amplification

Several WGA methods were tested with 1ng, 100pg or 30pg of HeLa gDNA (New England Biolabs, Ipswich, USA). The modified improved primer extension pre-amplification (mi-PEP) was performed as described. TruePrime single cell WGA kit (Sygnis AG, Heidelberg, Germany), MALBAC single cell WGA kit (Yikon genomics, Taizhou, China), illustra Single Cell GenomiPhi DNA Amplification Kit (GE Healthcare, Chicago, USA), GenomePlex® Single Cell Whole Genome Amplification Kit (Sigma-Aldrich, St. Louis, USA), REPLI-g single cell WGA kit (Qiagen, Hilden, Germany) and Ampli1™ WGA kit (Menarini Silicon Biosystems, Bologna, Italy) were used following manufacture's protocols. WGA products were cleaned up with innuPREP PCRpure Kit (Analytik Jena, Jena, Germany) by adjusting the manufacture's protocol by using 450µl Binding Buffer to the WGA product instead of 500µl and repeating the first centrifugation step with centrifugation at 12,000g for 2 minutes.

## 2.7. DNA quantitation and qualification using qPCR

DNA concentration and DNA integrity of gDNA from skin flakes was measured by real time PCR using the qTOWER 3 G (Analytik Jena, Jena, Germany) and following a slightly modified protocol described by Opel et al. (2008) [90]. ALU forward primer, 82bp reverse primer and the 201bp reverse primer described in Opel et al. (2008) were used. PCR was performed in a 20µl reaction volume including 1µl crude lysate or gDNA, 1µl ALU forward primer (20µM), 1µl 82bp/201bp reverse primer (20µM), 4µl AmpSolution (Promega, Madison, USA), 10µl Maxima SYBR Green/Fluorescein qPCR Master Mix (2X) (Thermo Fisher, Waltham, USA) and 3µl nuclease free $H_2O$. Additionally, a standard row of 5 different standard concentrations (0.1ng/µl, 0.5ng/µl, 1ng/µl, 2.5ng/µl HeLa gDNA) was used as references to calculate the DNA concentration of each sample. For reactions including the 82bp fragment, amplification was performed using the following cycling conditions: 1 cycle of 95 °C for 4 min, 45 cycles of 95 °C for 15 sec, 59 °C for 20 sec and 72 °C for 20 sec. For reactions including the 201bp fragment, amplification was performed using the following cycling conditions: 1 cycle of 95 °C for 4 min, 45 cycles of 95 °C for 15 sec, 63 °C for 20 sec and 72 °C for 20 sec. After each run, DNA concentrations where calculated by using the software package qPCRsoft (Analytik Jena, Jena, Germany). Determination of the DNA degradation index (DI) was calculated by dividing the concentration measured with the 82bp amplicon (gDNA small amplicon) through the concentration measured with the 201bp amplicon (gDNA big amplicon):

$$\mathrm{DI} = \frac{gDNA\ small\ amplicon}{gDNA\ big\ amplicon}$$

## 2.8. Amplification of STRs by using PCR

### 2.8.1. Autosomal STRs

At the beginning of this study 16 autosomal STRs and Amelogenin were amplified using MPX-5 ESS Multiplex PCR kit (Serac GmbH, Bammental, Germany). Later on the same markers were amplified using PowerPlex ESX17 kit (Promega, Madison, USA). Both multiplex STR kits included the following loci: D18S51, D21S11, TH01, D3S1358, D16S539, D2S1338, D1S1656, D10S1248, FGA, D8S1179, vWA, D22S1045, D19S433, D12S391, D2S44, SE33 and Amelogenin Reactions were carried out in 5µl volume and amplified with 30 PCR cycles using a GeneAmp PCR System 9700 thermocycler (Thermo Fisher, Waltham, USA). For PCR on WGA products, 1µl of 1:50 diluted WGA product was added as template DNA to the reaction. For PCR on gDNA of skin flakes, 1µl of AmpSolution (Promega, Madison, USA) was included in the mastermix and 1µl of crude lysate was added to the reaction as template DNA.

### 2.8.2. X-chromosomal STRs

X-STRs were amplified using Investigator Argus X12 QS kit (Qiagen, Hilden, Germany). Reactions were carried out in 5µl total reaction volume and amplified with 1 cycle of 94 °C for 4 minutes, 5 cycles of 96 °C for 30 sec, 63 °C for 120 sec and 72 °C for 75 sec, 25 cycles of 94 °C for 30 sec, 60 °C for 120 sec and 72 °C for 75 sec, 1 cycle of 68 °C using a GeneAmp PCR System 9700 thermocycler (Thermo Fisher, Waltham, USA). 1µl of 1ng gDNA or WGA product was added as template DNA to the reaction.

### 2.8.3. Y-chromosomal STRs

Y-STRs were amplified using PowerPlex Y23 (Promega, Madison, USA). Reactions were carried out in 10µl total reaction volume and amplified with 1 cycle of 96°C for 2 min, 30 cycles of 94°C for 10 sec 61°C for 1 min 72°C for 30 sec and 1 cycle of 60°C for 20 min using a GeneAmp PCR System 9700 thermocycler (ThermoFisher, Waltham, USA). 2µl of 2.5ng gDNA or WGA product was added as template DNA to the reaction.

## 2.9. Fragment analysis by using capillary electrophoresis

PCR products were analyzed by capillary electrophoresis on an ABI Prism 310 or ABI Prism 3130 Genetic Analyzer (ThermoFisher, Waltham, USA). A volume of 1 µl product was denatured in 12 µl deionized HiDi™ formamide (ThermoFisher, Waltham, USA) and 0.5 µl DL500 ORN size standard (Serac GmbH, Bammental, Germany), WEN ILS 500 (Promega, Madison, USA) or Biotype SST-BTO 60-500bp (Biotype, Dresden, Germany) at 95 °C for 4 min. A volume of 1 µl of 1:50 diluted autosomal STR PCR products from WGA products as well as Y-STR PCR products were denatured as described above. A volume of 1 µl of 1:100 diluted X-STR PCR products was denatured as described above. Denatured samples of single skin flakes were injected at 3 kv for 20 sec. All other denatured samples were injected at 3kv for 3 sec. Data was genotyped with GeneMapper™ v3.0 (Thermo Fisher, Waltham, USA). Allele recovery was calculated by dividing the number of alleles recovered in the sample (*N* sample) through the total number of alleles from the reference sample (*N* reference) multiplied by 100% according to the following formula:

$$\text{Allele recovery} = \frac{N\ \text{sample}}{N\ \text{reference}} * \ 100\%$$

When autosomal haplotypes of sperm were compared with the reference genotype the calculation was as follows:

$$\text{Allele recovery} = \frac{N \text{ sample}}{17} * 100\%$$

## 2.10. NGS analysis

An in-house developed NGS protocol (laboratory of Prof. Dr. Jäger, Hochschule Bonn-Rhein-Sieg University of Applied Sciences; manuscript in preparation) was used for DNA profiling of single skin flakes as an alternative method to the method with the PowerPlex ESX17 kit (Promega, Madison, USA) and CE described before.

The NGS protocol is particularly suited for degraded DNA as it is based on small amplicon sizes (mini-STR amplicons). Furthermore, the mini-STR NGS protocol has a sensitivity of 100 pg input DNA and DNA profiles are fully concordant with CE-based profiles (Prof. Dr. Jäger, personal communication). The NGS protocol was adapted from a previous study described by Kim et al. (2016) [84]. The in-house developed protocol included D18S51, D21S11, TH01, D3S1358, D16S539, D2S1338, D1S1656, D10S1248, FGA, D8S1179, vWA, D22S1045, D19S433, D12S391, D2S441, SE33 and Amelogenin. The amplicon size range was 76 – 296 bp of which amplicons from 13 loci were <200 bp. Amplicons were first amplified by using one multiplex PCR including 1µl of crude single skin flake lysate as template DNA and 5µl of AmpSolution (Promega, Germany) was included in the PCR mastermix. Afterwards, the PCR products where purified by a post-PCR clean-up using Agencourt AmPure XP Beads (Beckman Coulter, Brea, California). Then, PCR indices (Nextera XT Index kit, Illumina, San Diego, USA) were introduced during a second PCR after which a second post-PCR clean-up followed. Afterwards, the cleaned up PCR products from the index PCR were quantitated using the KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems, Wilmington, USA). Sequencing was performed on a standard v3 flow

cell using paired-end sequencing (600 cycles) with dual index reads using the MiSeq Reagent Kit v3 (Illumina, San Diego, USA) on the Illumina MiSeq FGx device (Illumina, San Diego, USA). Data acquisition and data processing was performed through a proprietary bioinformatics pipeline by Jan Biermann and Prof. Ralf Thiele at the informatics department of Hochschule Bonn-Rhein-Sieg University of Applied Sciences. In this pipeline, identification of loci is based on the DNA sequence, and allele calling is based on comparison to sequence of reference alleles. Signal strength is recorded as number of reads. Thresholds for data analysis were set as follows:

- The analytical threshold was set at 10 reads.
- Calling threshold was set at 0.2% of the total number of reads of the quality control sample.
- Stutter threshold was set at 25% of the number of reads of the parent allele.

## 2.11. Statistical analysis

The statistical package from the data analysis tool in Microsoft Excel 2010 (Microsoft, Redmond, USA) was used to calculate:

- The correlation between allele recovery with gDNA concentration and gDNA integrity.
- Significant difference in allele recovery of skin flakes between different persons by using the analysis of variance (one factor)
- Significant difference in allele recovery between skin flakes analysed by using PowerPlexESX17 with CE and the same skin flakes analysed by using the in-house developed mini-STR based NGS method, by using a paired two sample t-Test

# Chapter 3

# Results & Discussion

# Chapter 3. Results and Discussion

## 3.1. Sperm

### 3.1.1. Introduction

In order to develop a method which enables forensic DNA profiling of single spermatozoa multiple different experiments were performed.
First, different WGA assays were assessed and compared based on their impact on forensic DNA profiling of low copy number DNA (chapter 3.1.2.). The best WGA assay was then selected and used for subsequent experiments on single spermatozoa.
Second, the isolation of single spermatozoa by using micromanipulation was optimized to be compatible with the selected WGA assay (chapter 3.1.3.).
Afterwards, the performance of the selected WGA method on single spermatozoa was assessed based on allele recovery of autosomal, X-chromosomal and Y-chromosomal STR systems (chapter 3.1.4.).
At the end, mock sexual assault swabs were used to validate the single spermatozoa DNA typing approach in mixtures (chapter 3.1.5.). These mock swabs included epithelial cells from 1 female contributor mixed with semen from 1 or 2 male contributors.

### 3.1.2. Performance of various WGA protocols on diluted genomic DNA

In order to amplify enough DNA from a single sperm cell, a WGA or LV-PCR was selected to be suitable for STR analysis. Therefore, different WGA methods and LV-PCR were tested with different amounts of HeLa gDNA used as template DNA and assessed in terms of allele recovery and of ADIs.

**Table 3.1.2.1. Performance of WGA methods on different amounts of template DNA.**

| WGA method | Template DNA (pg) | Allele recovery (%) | ADI (*N*) |
|---|---|---|---|
| mi-PEP | 1000 | 100 | 0 |
| | 100 | N/A | N/A |
| | 30 | 21 | 4 |
| TruePrime | 1000 | 88 | 0 |
| | 100 | N/A | N/A |
| | 30 | 0 | 0 |
| MALBAC | 1000 | 53 | 4 |
| | 100 | N/A | N/A |
| | 30 | N/A | N/A |
| GenomiPhi | 1000 | 93 | 0 |
| | 100 | 93 | 0 |
| | 30 | 30 | 0 |
| GenomePlex | 1000 | 89 | 0 |
| | 100 | 65 | 3 |
| | 30 | 24 | 0 |
| REPLI-g | 1000 | 100 | 0 |
| | 100 | 100 | 0 |
| | 30 | 59 | 0 |
| Ampli1 | 1000 | 54 | 0 |
| | 100 | 18 | 0 |
| | 30 | 12 | 0 |
| No WGA | 1000 | 100 | 0 |
| | 100 | 100 | 0 |
| | 30 | 38 | 0 |

Of all methods tested, REPLI-g turned out as the best performing WGA method with an allele recovery higher than the conventional method (*table 3.1.2.1*). In addition, REPLI-g as well as TruePrime did not produce any ADIs.

As mentioned in chapter 1.3.2.3. ADIs can make interpreting forensic DNA-profiles difficult. In general, ADIs are often seen with PCR on LCN-DNA in combination with many PCR cycles. Amplicons present in only minute amounts in the environment can then be amplified to some extent enough to be detected by capillary electrophoresis. ADIs can also arise from allele slippage caused by the polymerase used in the specific reaction. Allele slippage can cause stutters. However, when stutters are much higher than the stutter threshold, stutter peaks could be interpreted as real alleles which results in ADIs. From previous studies

it is known that ADIs often occur at PCR based WGA methods due to the generation of non-specific products from amplification errors[91]. This matches the results shown here. mi-PEP as well as GenomePlex are both predominantly PCR based WGA methods and both showed ADIs. Furthermore, the semi-PCR based MALBAC showed ADIs already when 1000pg (equivalent to approximately 160 somatic cells) of template DNA was used. Although PCR is part of MALBAC, the ADIs are probably caused at the part where MALBAC is based on MDA like REPLI-g or TruePrime. Of note, MALBAC uses a different strand displacement polymerase than the latter two methods. This particular polymerase called *Bst* polymerase is prone to allele slippage during amplification and might therefore be the cause of the observed ADIs [83].

Previous research showed WGA methods were successfully used in single cell genome sequencing [25,62,63,78,92–94]. For example multiple studies showed successfully single cell analysis by using MALBAC in combination with DNA sequencing [62,95]. However, many ADI were shown with MALBAC in the present study which made MALBAC unsuitable for STR typing. Therefore, it could be that some WGA methods struggle with loci which include many repetitive units like the STR loci used in forensic DNA analysis, but can easily amplify non-repetitive loci. As this goes beyond the scope objectives of this thesis, more research is necessary to find out the exact cause why certain WGA methods generate more ADI than others.

In conclusion, not all WGA methods tested here were suitable for STR analysis on LCN-DNA. ADI were seen with some PCR based and semi-PCR based WGA methods, but not with MDA based methods which used the Phi-29 polymerase. Furthermore, REPLI-g was shown to be the WGA method with the highest allele recovery (59%) obtained with the lowest amount of gDNA test, 30pg (approximately 5 somatic cells), and did not show any ADIs. Therefore, REPLI-g was selected as a suitable WGA method for forensic DNA typing and was chosen as WGA method to be used in experiments on single sperm cells as well as single skin flakes shown in the following chapters (chapter 3.1.4 and chapter 3.2.5).

### 3.1.3. Optimisation of micromanipulation of spermatozoa

In order to isolate single spermatozoa, micromanipulation was used. At first the isolation and transfer of single spermatozoa on microscopy slides was carried out using glass capillary based micromanipulation. Glass capillaries are generally used in *in vitro* fertilisation (IVF). Unlike IVF, in the present study glass capillaries were used to transfer the spermatozoa from the medium/petri dish into a tube. However, glass capillaries proved rather difficult for this purpose as these were very fragile and broke during transfer into PCR tubes.

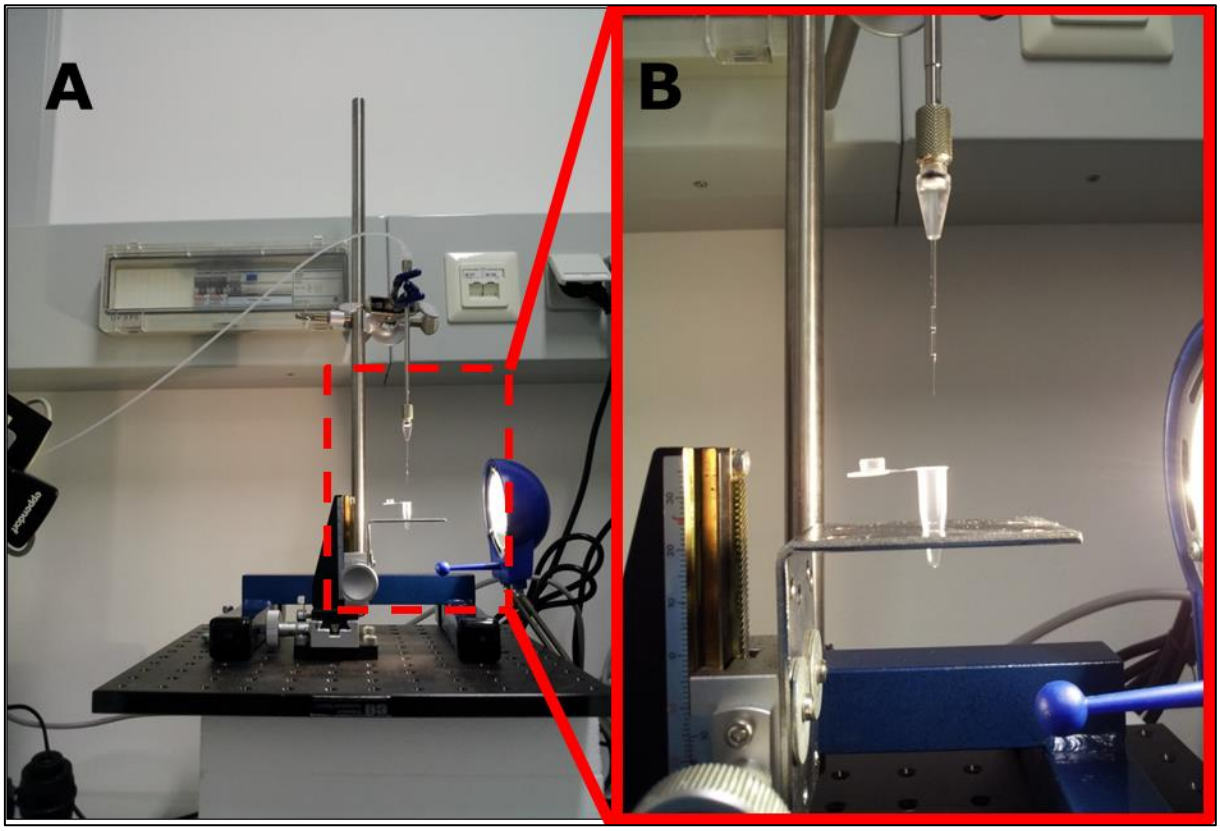


**Figure 3.1.3.1. Cell transport platform. (A) Platform used for transport by glass capillary based micromanipulation isolated spermatozoa into a PCR tube. (B) Close-up where glass capillary is hanging above the PCR tube shortly before the tip of the capillary was moved into the PCR tube.**

To prevent breakage of the capillary during transfer of spermatozoa and ensure successful transfer into PCR tubes, a transfer platform/station was built up (*figure*

*3.1.3.1.*). Here the capillary holder including the capillary could be fastened. The station included an X-Y table screw table in which the PCR tube with the liquid was placed. The position of the X-Y table could be finely adjusted which enabled the precise positioning of the PCR tube in order to insert the tip of the capillary into the liquid without touching the tube's walls.

Nevertheless it turned out that the cells were not always transferred properly. This was mostly due to the fact that cells often were not released out of the capillary, but rather stayed in the capillary by sticking to the walls of the capillary, disappeared by liquid spontaneously flowing out of the tip on its way to the station/PCR tube or cells got stuck on the outside of the tip.

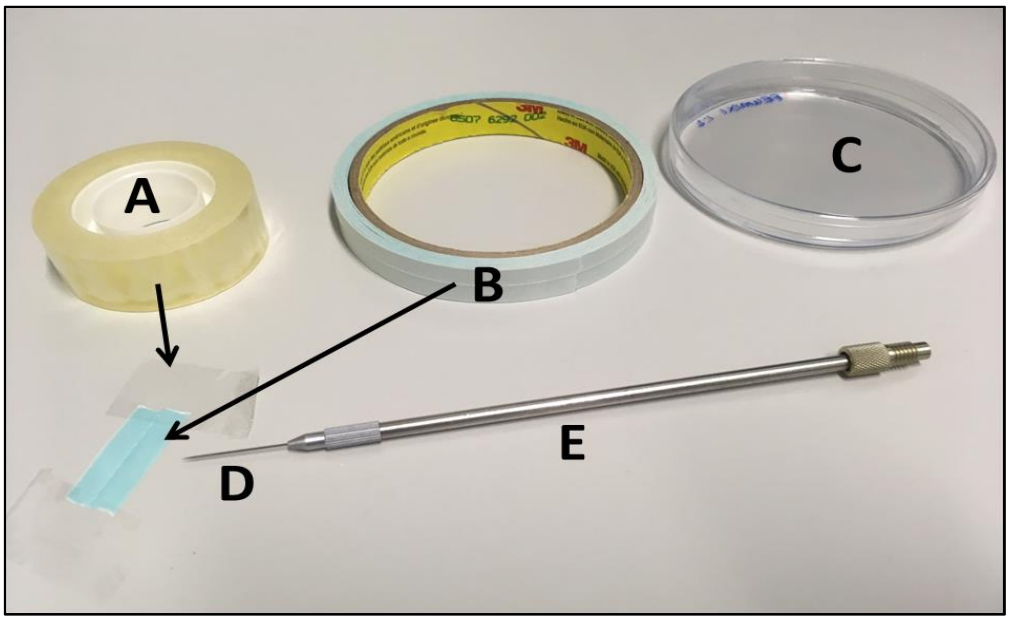


**Figure 3.1.3.2. Components used for tungsten needle based micromanipulation of spermatozoa. Multipurpose transparent adhesive tape (TESA film) (A) was used to fasten a piece of the 3M water soluble adhesive tape (B) to a clean work surface. A tungsten needle (D) was inserted in a capillary/needle holder (E). 3M soluble adhesive (B) on the tip of the tungsten needle (D) was used to isolate single spermatozoa from the surface of an agarose gel filled petri dish (C). Adding 3M soluble adhesive onto the needle (D) was accomplished by scraping the tip of the needle (D) on the piece of fastened tape.**

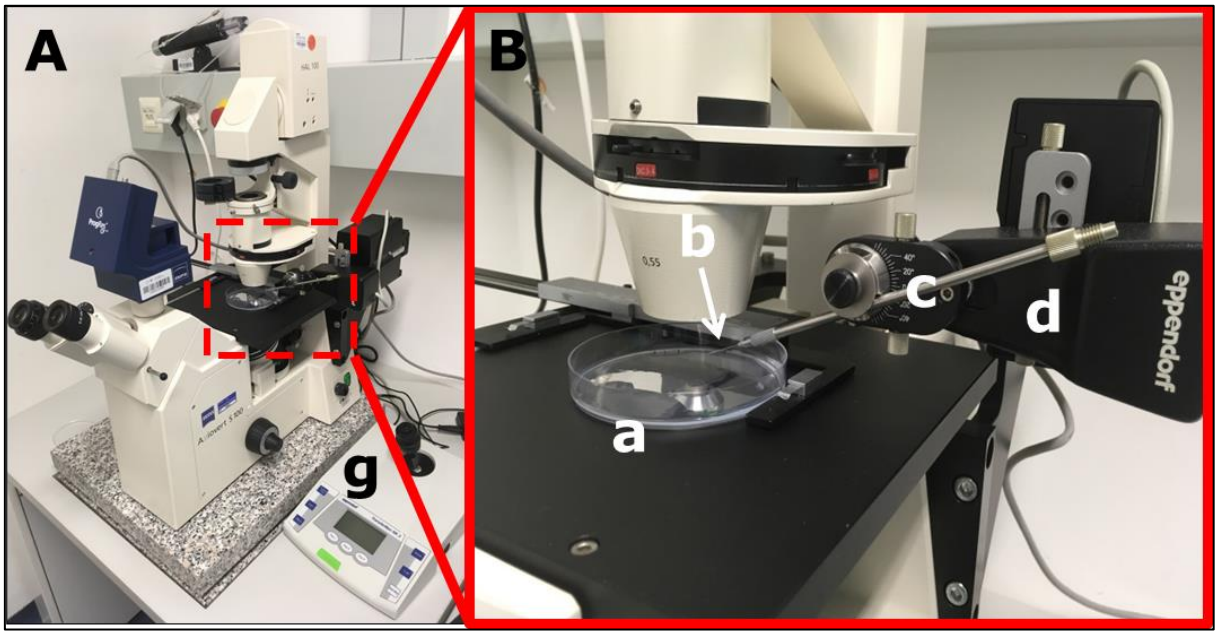


**Figure 3.1.3.3. Needle and agarose gel based micromanipulation set-up. (A) Fully equipped microscope with additional joystick (g) to control the arm (d) of the needle holder (c) movements. (B) Close up of the agarose filled petri dish (a) ready to perform micromanipulation of spermatozoa by using a tungsten needle (b).**

To allow for better control of the transfer of cells into PCR tubes, micromanipulation using a tungsten needle in combination with water soluble adhesive was used *(figure 3.1.3.2. and figure 3.1.3.3.)*. This method has been described [33,41]. Compared to glass capillaries, tungsten needles are very robust and won't break if touched by the PCR tube's wall. By applying a tiny amount of adhesive on the needle's tip, cells can be picked up and securely moved into for example, a buffer containing PCR. As the adhesive is water soluble, the secured cell can be released into the buffer inside the tube. To allow picking up cells by the water soluble adhesive, a dry surface of the substrate is required. On the other hand, if cells are in a liquid medium (cell suspension), the adhesive will be released from the needle's tip and dissolve into the cell medium. For this reason, spermatozoa were dried on a glass microscopy slide at first. Yet, spermatozoa stuck to the surface and did not detach from the slide when performing micromanipulation with the adhesive. Therefore, an alternative material to the glass microscopy slide was required.

**Table 3.1.3.1. Surfaces of different materials tested for non-stick properties of spermatozoa.**

| Surface material tested | Non-stick? |
|---|---|
| Polyethylene terephthalate | No |
| Polycarbonate | No |
| Polyethylene | No |
| Polymethyl methacrylate | No |
| tesafilm® | No |
| Polystyrene petri dishes | No |
| Non-adherent cell culture flasks | No |
| WD-40® coated microscopy slide | No |
| Repel-Silane ES coated microscopy slides | No |
| 1% agarose gel | Yes |
| 0.5% agarose gel | Yes |

Different materials were tested on non-stick properties with sufficient transparency to clearly recognize spermatozoa on the surface of the specific material tested (*table 3.1.3.1.*). Of all materials tested, a 1% agarose gel was both transparent enough to use for microscopy and at the same time was capable to release spermatozoa when picking them up with the adhesive (*figure 3.1.3.4, page 82*). The 0.5% agarose gel was also suitable for microscopy because of its high transparency, suitable for microscopy as well. However, due to the higher water content it did not keep spermatozoa in place on the gel's surface and spermatozoa did not stick to the adhesive. Instead, spermatozoa were pushed away. Although all the other surfaces tested (polyethylene terephthalate, polycarbonate, polyethylene polymethyl methacrylate, tesafilm®, Polystyrene petri dishes, non-adherent cell culture flasks, WD-40® coated microscopy slide and Repel-Silane ES coated microscopy slides) were capable of keeping spermatozoa in place, they however did not release sperm during micromanipulation. This could be due to the spermatozoa's membrane surface structure/proteins binding to surface structure of the materials tested when drying. The same effect happens when spermatozoa are dried on swabs or fabric. For this reason, it has been described that proteinase K is essential for efficient

sperm release from swabs by digesting the proteins binding to the surface of the cotton fibers [96,97].

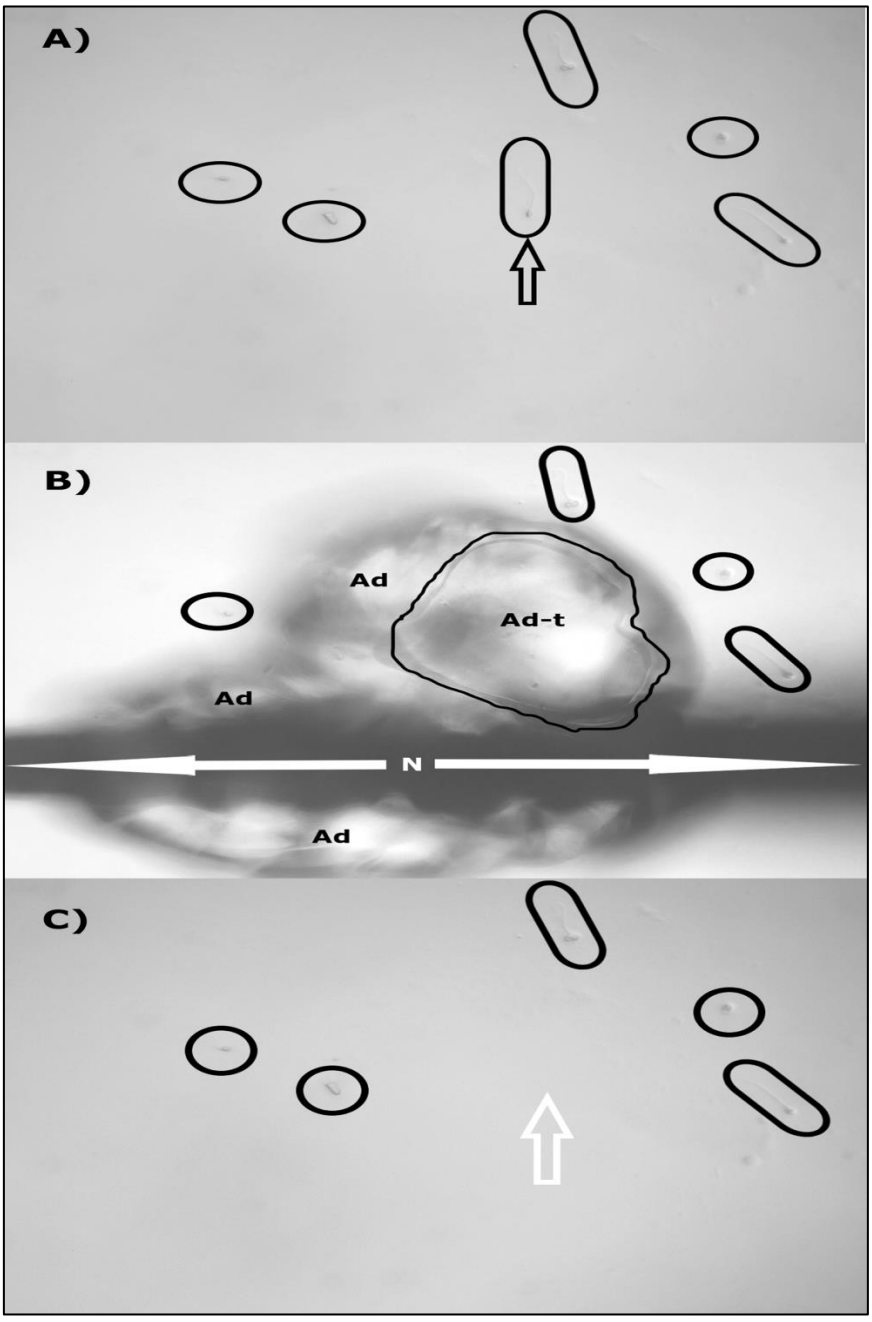


**Figure 3.1.3.4. Microscopic view of isolation of single spermatozoa by using the tungsten needle and agarose gel based micromanipulation (200x magnification). A) Multiple single spermatozoa laying on a 1% agarose gel surface and one spermatozoon (arrow) was selected to be isolated. B) Applying water soluble adhesive (Ad) by using a tungsten needle (N) on top of the selected spermatozoon with only a part of the adhesive touching (Ad-t) the surface of the spermatozoon and the surrounding agarose gel surface. After the spermatozoon stuck to the adhesive, the spermatozoon was isolated by lifting the needle upwards away from the gel surface. C) View after the spermatozoon was isolated.**

Micromanipulation of single sperm cells by using the capillary method has been used for decades. Also, micromanipulation of particles, tissue or cells carried out by the combination of a tungsten needle and water-soluble adhesive has also been used for forensic trace analysis [33,41]. Nevertheless, no agarose gel has thus far been included as substrate surface. A different method which uses agarose gels for sperm analysis is the TUNEL assay [98]. However, the spermatozoa are not lying on top of an agarose gel surface. Instead, sperm cells are embedded between two layers of agarose gel. Furthermore, the purpose of the TUNEL assay is analysis of the quality of the spermatozoa's DNA and not the isolation of spermatozoa discussed in the present work. So far, no other publicly released study describes isolation of sperm cells by using the agarose gel method. This makes it a new and unique micromanipulation method to isolate single sperm cells. Furthermore, it does not require large amounts of expensive equipment like with laser-based micromanipulation platforms, microfluidics and di-electrophoresis platforms [11,20,25–27,99]. Also the set-up/parts (control arm, capillary/needle holder, joystick) for this method, can be attached on multiple different microscopes. They also do not take up much space and do not require any software, thus making it easy to integrate in any forensic biology laboratory.

### 3.1.4. Performance of WGA on single spermatozoa

In order to test the performance of WGA on single spermatozoa, the REPLI-g protocol was applied on single spermatozoa from two persons which were isolated from neat semen by micromanipulation. WGA products generated were used for autosomal STR typing for both X- and Y-STR. At best, autosomal STR typing of each spermatozoon would show a haploid DNA profile. In case of X- and Y-STR typing, only one of the two profiles would be obtained since each spermatozoon will only carry an X-chromosome or a Y-chromosome.

The PCR product from most samples was diluted by 1:50 in order to not oversaturate the capillary electrophoresis. A few samples had to be further diluted to 1:100 because the amount of PCR product of some alleles was still too high to be analysed. That can also lead to ADO from alleles which could be adequately analysed with a dilution of 1:50. Therefore, both 1:50 and 1:100 dilutions of the PCR product per sample were analysed and allele recovery was then defined as the consensus recovered peaks from the two dilutions. Autosomal STR allele recovery from WGA products ranged from 18% to 100% (*table 3.1.4.1.*). WGA products were positive for Amelogenin and showing only AMELX or AMELY as expected. However, this doesn't necessarily indicate that only one spermatozoon was typed since multiple spermatozoa might have been typed but all with the same X or Y chromosome segregation. The autosomal STR electropherograms from successfully amplified STRs showed no evidence of multiple spermatozoa being typed as only one of the two possible alleles at each STR marker was present (*figure 3.1.4.1.*). This is exactly what was expected from one spermatozoon as it only contains one allele per locus, a so called haploid profile. This confirms the efficacy of the presented isolation/micromanipulation method.

**Table 3.1.4.1. Autosomal STR allele recovery, ADO rate and X/Y-chromosome determination of WGA products from single spermatozoa from two donors (donor A and donor B).**

| Donor | WGA product | Amelogenin | | ADO (n) | Allele recovery (%) |
|---|---|---|---|---|---|
| A | **gDNA no WGA** | **X** | **Y** | **0** | **100** |
| | Spermatozoon 1 | | Y | 14 | 18 |
| | Spermatozoon 2 | | Y | 2 | 88 |
| | Spermatozoon 3 | | Y | 2 | 88 |
| | Spermatozoon 4 | | Y | 3 | 82 |
| | Spermatozoon 5 | | Y | 12 | 29 |
| | Spermatozoon 6 | | Y | 3 | 82 |
| | Spermatozoon 7 | X | | 0 | 100 |
| | Spermatozoon 8 | X | | 3 | 82 |
| | Spermatozoon 9 | X | | 1 | 94 |
| | Spermatozoon 10 | X | | 0 | 100 |
| | Spermatozoon 11 | X | | 3 | 82 |
| B | **gDNA no WGA** | **X** | **Y** | **0** | **100** |
| | Spermatozoon 1 | | Y | 1 | 94 |
| | Spermatozoon 2 | | Y | 1 | 94 |
| | Spermatozoon 3 | | Y | 2 | 88 |
| | Spermatozoon 4 | | Y | 6 | 65 |
| | Spermatozoon 5 | | Y | 0 | 100 |
| | Spermatozoon 6 | X | | 1 | 94 |
| | Spermatozoon 7 | X | | 4 | 76 |

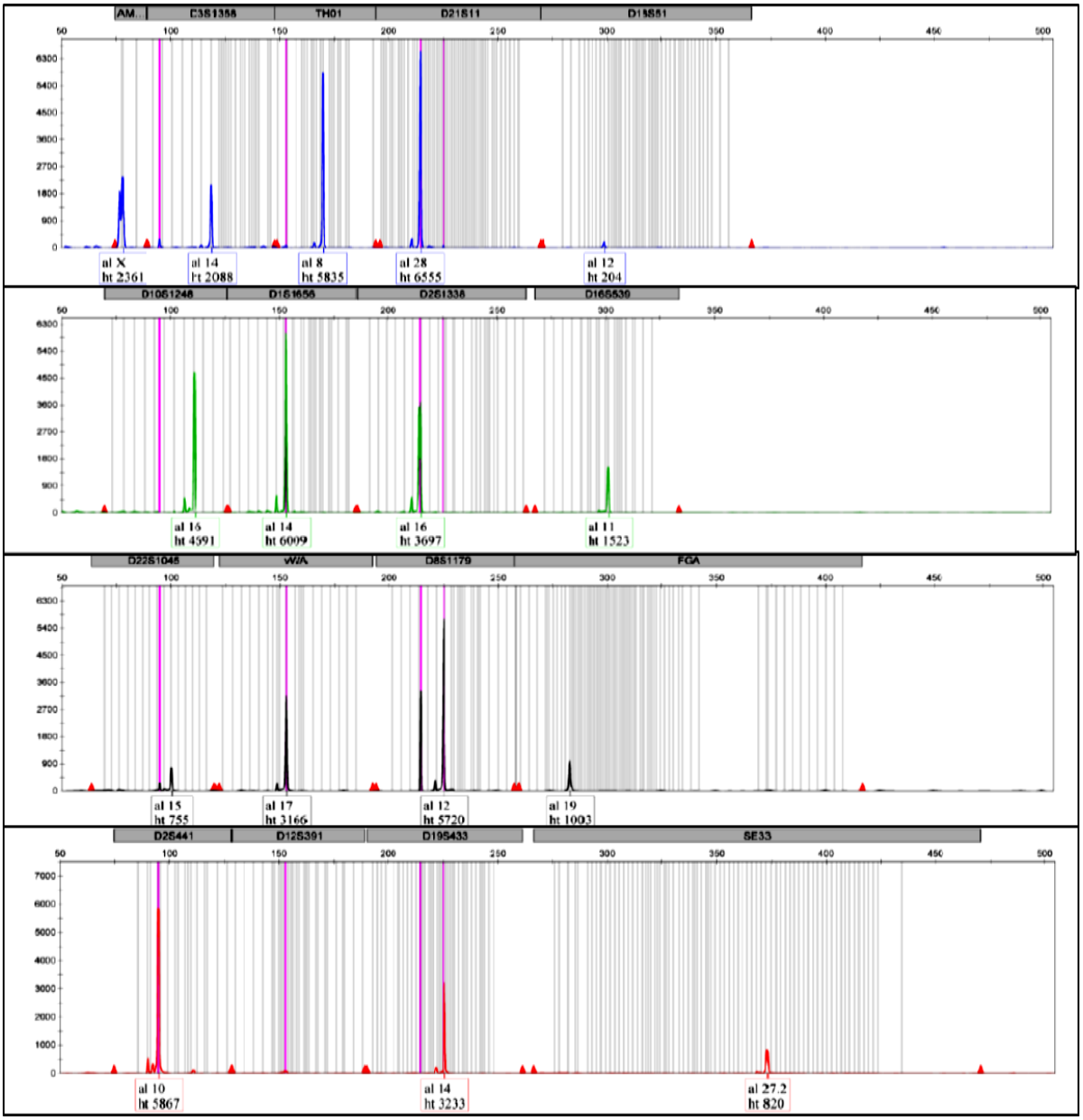


**Figure 3.1.4.1. Electropherogram of a haplotype autosomal STR profile from a single spermatozoon analysed with REPLI-g based WGA and PowerPlex ESX17. The relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. On top, the X-axis represents the fragment size in base pairs (bp) with the allele corresponding STR locus name above in the grey horizontal bars. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. The purple lines show the locations of peaks not representing a detected allele, but are rather due to fragments of alleles which were in such abundance the corresponding fluorescent was high enough to create false peaks in other detection channels (blue, green, black or red) too. These false peaks were therefore annotated with a purple line. False alleles were not called and therefore no label box is present underneath a false allele peak.**

To further confirm that only one spermatozoon has been through the whole process, X- and Y-STR profiles were both generated from the same WGA products. All WGA products from single spermatozoa positive for AMELY, but negative for AMELX, showed only Y-STR alleles and no X-STR alleles were detected (*table 3.1.4.1*). A typical example of a single spermatozoon positive for Y-STRs is shown in *figure 3.1.4.2.* In addition a STR profile from a spermatozoon without a Y chromosome is shown in *figure 3.1.4.3*.

For WGA products from which an X-allele at Amelogenin was detected X-STRs were found, but were all negative for Y-STR alleles. An example of X-STR electropherogram from a spermatozoon negative for X, but positive for Y is shown in *figure 3.1.4.4*. Also, an X-STR electropherogram from a spermatozoon positive for X, but negative for Y is given in *figure 3.1.4.5*.

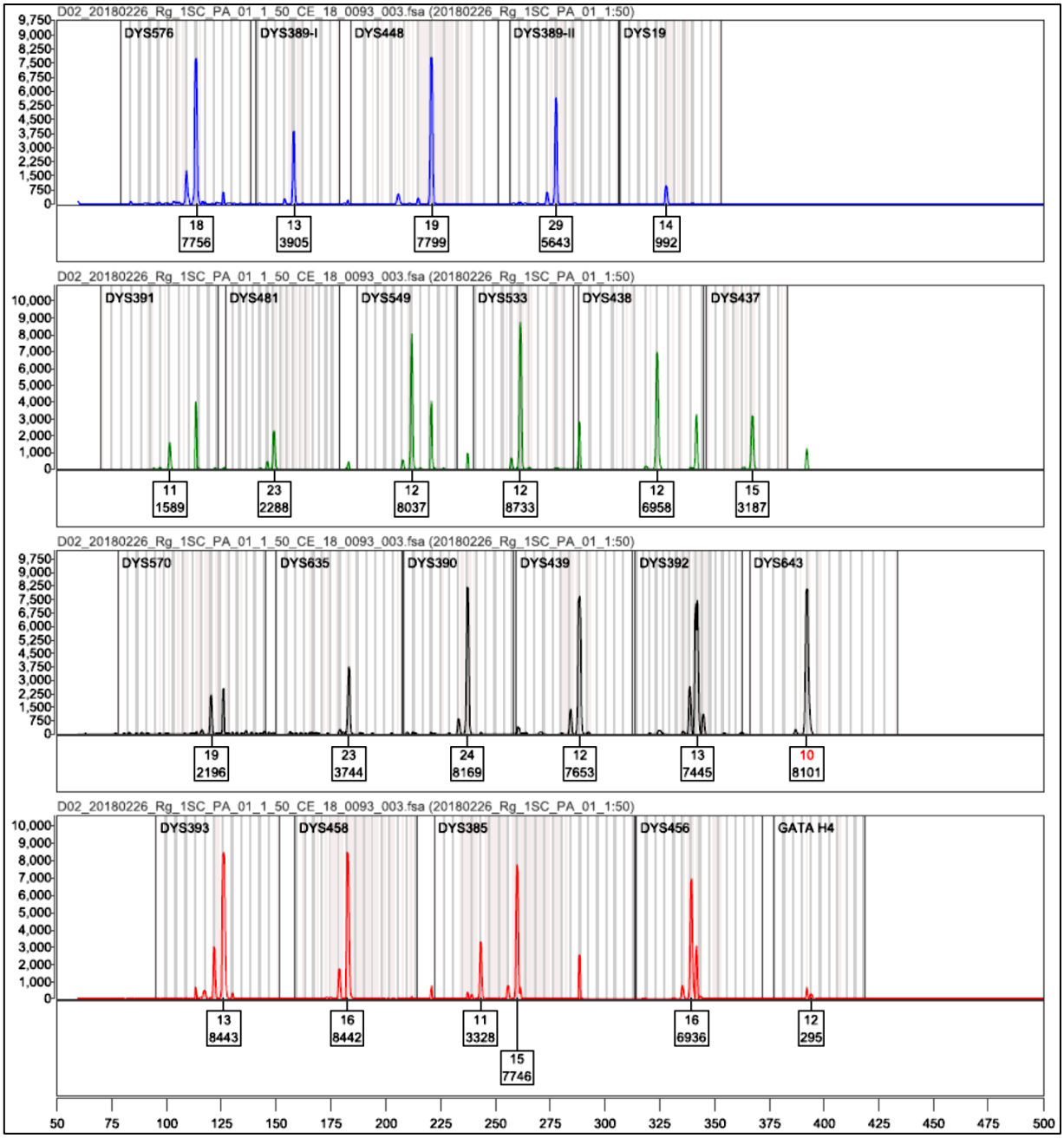


**Figure 3.1.4.2. Electropherogram Y-chromosomal STR profile from a single spermatozoon analysed with WGA and PowerPlex Y23. The spermatozoon used here was positive for Y-chromosomal STRs, but negative for X-chromosomal STRs. The relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. Below the red channel (with red allele peaks), the X-axis represents the fragment sizes for all detection channels in base pairs (bp). The STR locus name is shown on top in each detection channel above the allele peaks. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. False alleles were not called and therefore no label box is present underneath a false allele peak.**

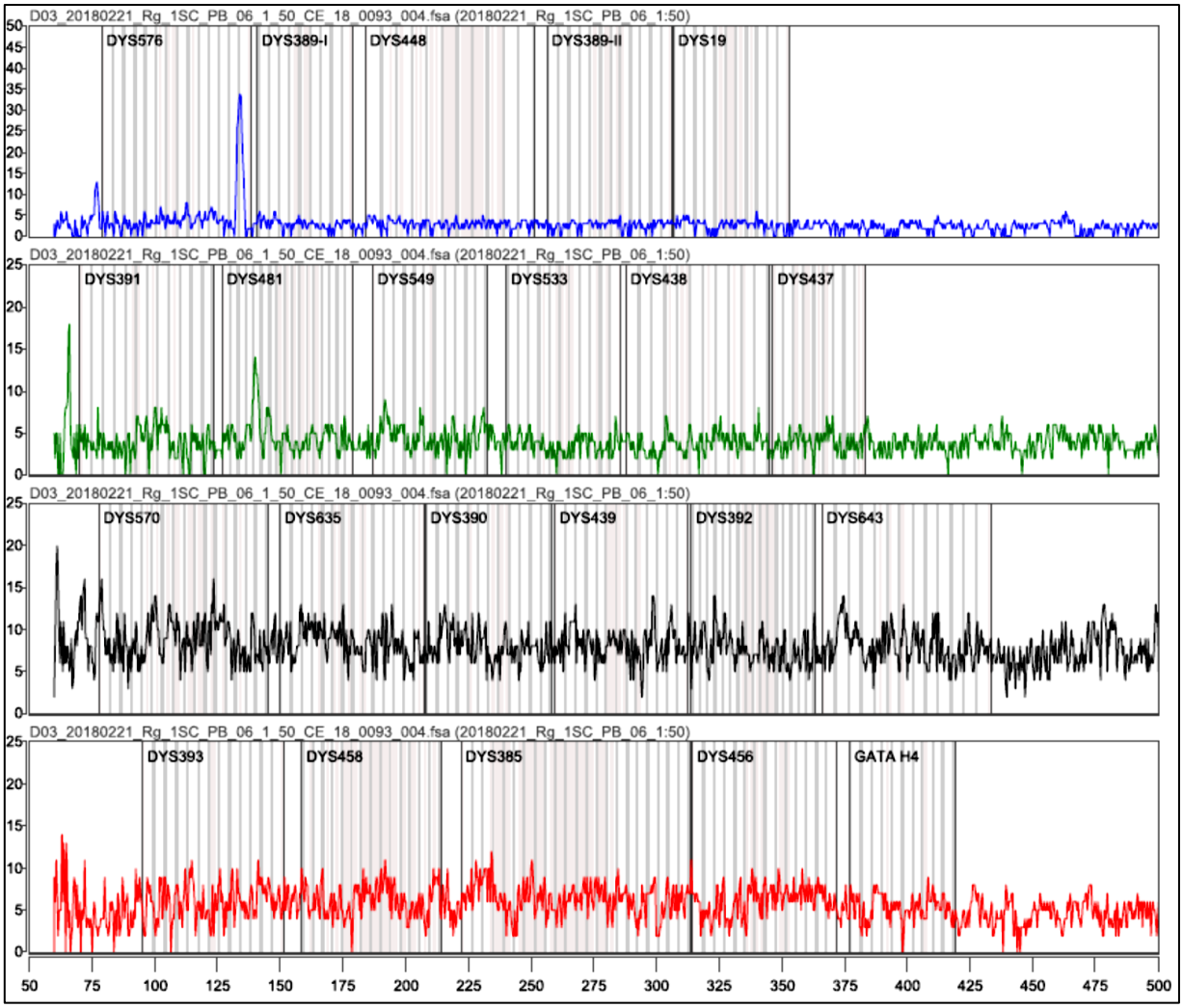


**Figure 3.1.4.3. Electropherogram Y-chromosomal STR profile from a single spermatozoon analysed with WGA and PowerPlex Y23. The spermatozoon used here was negative for Y-chromosomal STRs, but positive for X-chromosomal STRs. The relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. Below the red channel (with red allele peaks), the X-axis represents the fragment sizes for all detection channels in base pairs (bp). The STR locus name is shown on top in each detection channel above the allele peaks. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. False alleles were not called and therefore no label box is present underneath a false allele peak.**

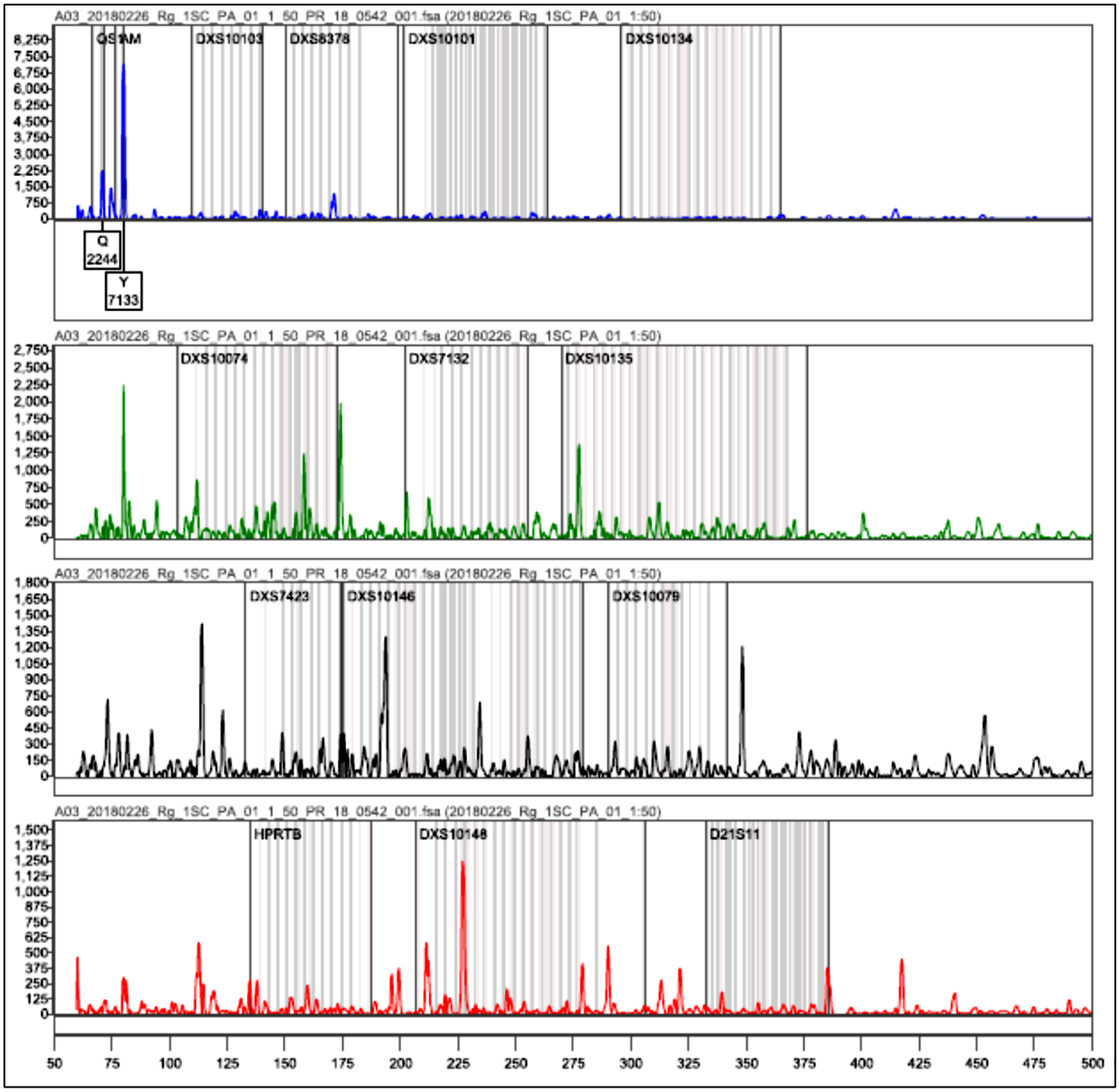


**Figure 3.1.4.4. Electropherogram X-chromosomal STR profile from a single spermatozoon analysed with WGA and Argus X-12 QS. The spermatozoon used here was negative for X-chromosomal STRs, but positive for Y-chromosomal STRs. The relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. Below the red channel (with red allele peaks), the X-axis represents the fragment sizes for all detection channels in base pairs (bp). The STR locus name is shown on top in each detection channel above the allele peaks. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. False alleles were not called and therefore no label box is present underneath a false allele peak.**

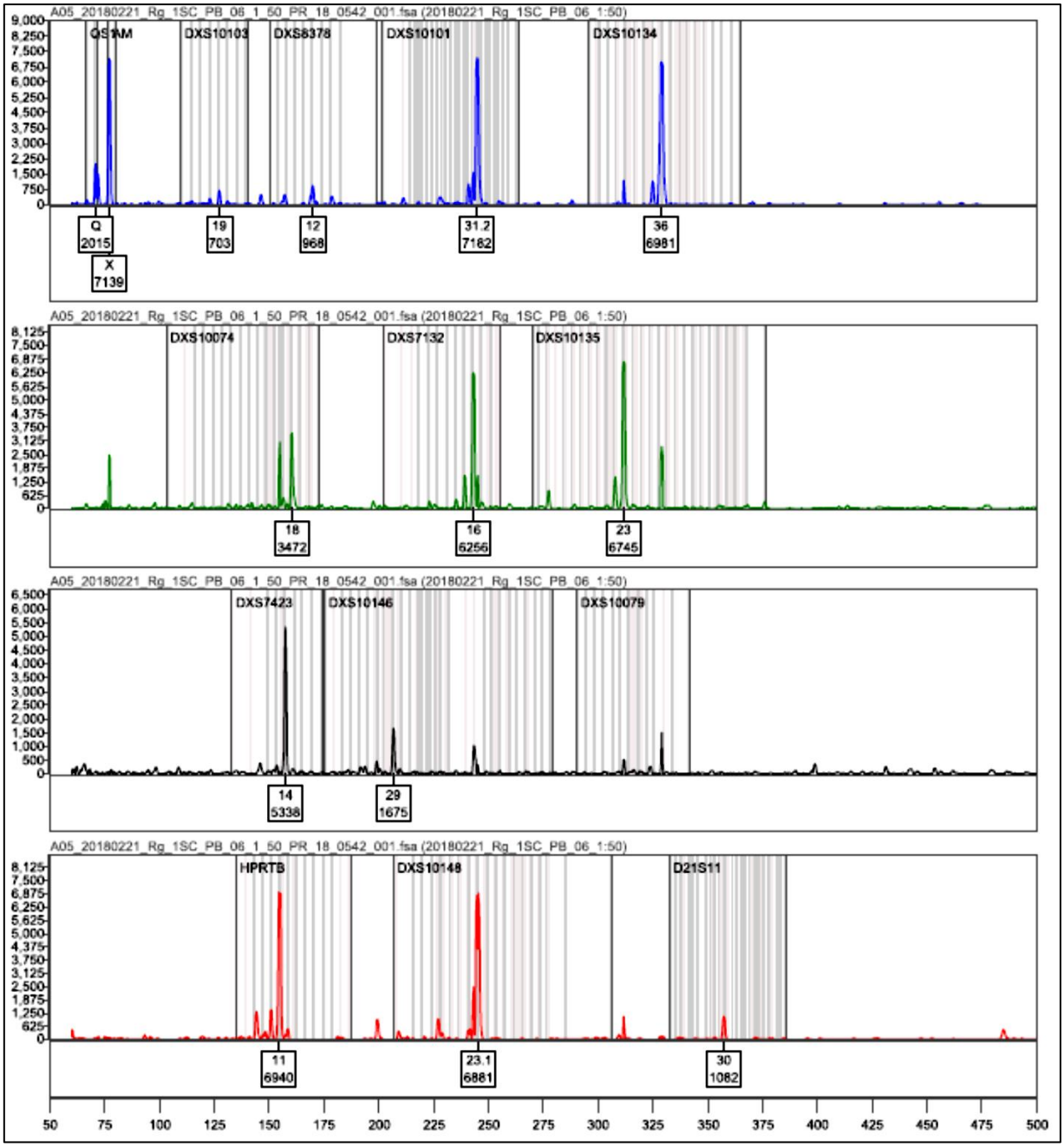


**Figure 3.1.4.5. Electropherogram Y-chromosomal STR profile from a single spermatozoon analysed with WGA and Argus X-12 QS. The spermatozoon used here was positive for X-chromosomal STRs, but negative for Y-chromosomal STRs. The relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. Below the red channel (with red allele peaks), the X-axis represents the fragment sizes for all detection channels in base pairs (bp). The STR locus name is shown on top in each detection channel above the allele peaks. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. False alleles were not called and therefore no label box is present underneath a false allele peak.**

Differences in Y-STR profiles between the two donors were determined (as expected as they had no known relatedness). The markers of which alleles were different between the two donors were therefore called diagnostic markers (*table 3.1.4.2*). Since all alleles which were detected were in concordance to the donors' profiles, no differences in alleles were detected in between single spermatozoon WGA products from the same donor for Y-STR typing.

**Table 3.1.4.2. Y-STR allele recovery of WGA products from single spermatozoa from two donors.**

| Donor | Sample | Alleles of diagnostic markers | | | | | | Allele recovery (%) | |
|---|---|---|---|---|---|---|---|---|---|
| | | DYS576 | DYS481 | DYS570 | DYS439 | DYS385 | | Total | Diag-nostic |
| A | **gDNA no WGA** | **18** | **23** | **19** | **12** | **15** | | **100** | **100** |
| | Spermatozoon 1 | 18 | 23 | 19 | 12 | 15 | | 100 | 100 |
| | Spermatozoon 2 | 18 | 23 | - | 12 | 15 | | 91 | 80 |
| | Spermatozoon 3 | 18 | 23 | 19 | 12 | 15 | | 95 | 100 |
| | Spermatozoon 4 | 18 | 23 | 19 | 12 | 15 | | 82 | 100 |
| | Spermatozoon 5 | 18 | 23 | - | 12 | - | | 68 | 60 |
| | Spermatozoon 6 | 18 | 23 | 19 | - | 15 | | 68 | 80 |
| | Spermatozoon 7 | - | - | - | - | - | | 0 | 0 |
| | Spermatozoon 8 | - | - | - | - | - | | 0 | 0 |
| | Spermatozoon 9 | - | - | - | - | - | | 0 | 0 |
| | Spermatozoon 10 | - | - | - | - | - | | 0 | 0 |
| | Spermatozoon 11 | - | - | - | - | - | | 0 | 0 |
| B | **gDNA no WGA** | **19** | **22** | **17** | **11** | **11** | **14** | **100** | **100** |
| | Spermatozoon 1 | - | 22 | 17 | 11 | 11 | 14 | 91 | 80 |
| | Spermatozoon 2 | - | - | 17 | 11 | 11 | 14 | 91 | 60 |
| | Spermatozoon 3 | 19 | 22 | - | - | 11 | 14 | 91 | 60 |
| | Spermatozoon 4 | 19 | 22 | 17 | 11 | - | 14 | 68 | 80 |
| | Spermatozoon 5 | 19 | 22 | 17 | 11 | 11 | 14 | 100 | 100 |
| | Spermatozoon 6 | - | - | - | - | - | - | 0 | 0 |
| | Spermatozoon 7 | - | - | - | - | - | - | 0 | 0 |

Also differences in X-STR profiles between donors were determined, again as expected (*table* 3.1.4.3). Similar to the Y-STR results, X-STR alleles were consistent with donor genotypes except for the WGA product of sperm 9 from donor A. In this sample allele 18 at marker DXS10074 was not in concordance with allele 19 from the control sample (*table 3.1.4.3 and appendix I*). Allele mutation rates are very high in sperm cells and are mostly due to natural cell replication slippage, leading to a 1 repeat deletion or insertion [100,101]. For DXS10074 a paternal mutation rate of 0.0087 has been reported [102]. Due to the fact that inconsistencies have been observed at only one locus and the deletion corresponded to 1 repeat unit, the observed allele is probably due to a mutation and not a result of a contamination. To support this, many negative controls are necessary during the WGA phase. By determining the ADI rate of the negative controls and comparing those with the paternal mutation rate of a certain locus, it can be concluded whether the ADI (the allele observed not matching one of the two alleles of the diploid reference profile) is due to environmental contaminating DNA or to a mutation.

The observed allele mutation sheds further light on a very important aspect. If only one cell out of many cells contains the mutated allele, the latter probably would not show up in the electropherogram, since copies of the wild type/original allele is present in abundance. The same would be the case with spermatozoa. Semen consists of, among others, many spermatozoa. When DNA typing semen in which only 1 or a few spermatozoa out of the many spermatozoa contain a mutated allele, this allele does not show up in an electropherogram. However, the sperm bearing the mutated allele can be seen when single spermatozoa are analysed separately as could have been the case in the results shown in this study.

**Table 3.1.4.3. X-STR allele recovery of WGA products from single spermatozoa from two donors.**

| Donor | Sample | Alleles of Diagnostic markers | | | | | | | | | | | | Allele recovery (%) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | DXS8378 | DXS10101 | DXS10134 | DXS10074 | DXS7132 | DXS10135 | DXS7423 | DXS10146 | HPRTB | DXS10148 | D21S11 | | Total | Diagnostic |
| A | **gDNA** | **11** | **29.2** | **37** | **19** | **13** | **29** | **15** | **30** | **12** | **25.1** | **28** | **32.2** | **100** | **100** |
| | Sperm 1 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 2 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 3 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 4 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 5 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 6 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 7 | 11 | 29.2 | 37 | 19 | 13 | 29 | 15 | - | 12 | 25.1 | 28 | - | 93 | 91 |
| | Sperm 8 | - | 29.2 | 37 | 19 | - | 29 | 15 | 30 | 12 | 25.1 | 28 | - | 86 | 82 |
| | Sperm 9 | - | - | 37 | 18 | 13 | - | 15 | 30 | 12 | 25.1 | 28 | - | 79 | 73 |
| | Sperm 10 | - | 29.2 | 37 | 19 | 13 | - | - | 30 | - | 25.1 | 28 | - | 71 | 64 |
| | Sperm 11 | 11 | 29.2 | 37 | - | 13 | 14.1 | 15 | 30 | 12 | 25.1 | 28 | - | 93 | 91 |
| B | **gDNA** | **12** | **31.2** | **36** | **18** | **16** | **23** | **14** | **29** | **11** | **23.1** | **29** | **30** | **100** | **100** |
| | Sperm 1 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 2 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 3 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 4 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 5 | - | - | - | - | - | - | - | - | - | - | - | - | 0 | 0 |
| | Sperm 6 | 12 | 31.2 | 36 | 18 | 16 | 23 | 14 | 29 | 11 | 23.1 | - | 30 | 93 | 100 |
| | Sperm 7 | - | 31.2 | 36 | 18 | 16 | 23 | 14 | 29 | 11 | 23.1 | - | 30 | 93 | 91 |

Besides mutated alleles, a contamination had probably occurred at sperm 11 from donor A. Here, allele 14.1 at marker DXS10135 was found and is not in concordance with allele 29 from the control sample (*see table 3.1.4.3. and appendix II*). Since the theory of -1 or +1 repeat mutation does not fit here, contamination may be the cause of the obtained result cannot be ruled out. On the other side, a mutation can neither be ruled out. Although a mutation of +3.1 would be rarer than a -1 or +1 mutation, it is not ruled out. ADI caused by contamination are easily seen here when two alleles per marker are obtained. The 14.1 allele was not an additional allele, but the only allele of marker

DX10135. Therefore the true cause of obtaining the in concordant allele cannot be confirmed.

DNA contamination is a very important aspect in forensic DNA analysis. In a forensic lab the main concern is sample:sample or staff:sample contamination as these really can ruin an DNA profile interpretation. Hence making case work extremely problematic. ADIs are a very specific form of DNA contamination and are caused by amplification of only minute amounts of amplicons present in the environment. This will result into no or little detection of alleles from the contaminating part. Generally, ADIs are easily recognized since they are rare and usually observed in electropherograms under 90 rfu (personal communication A. Gibb). Therefore, ADIs will not harm interpretation of DNA profiles. Although, the lower the DNA concentration to be analysed is, the higher the contaminating fraction gets. This could result in allele peaks as high as the main allele peaks which then can become problematic. As single spermatozoa contain extremely low gDNA, contaminating DNA could easily be detected in electropherograms next to the spermatozoa's allele peaks. Furthermore, using a method as sensitive as REPLI-g single cell WGA, an environmental contamination could easily be pre-amplified enough to produce detectable PCR product when used as a template. Furthermore, WGA could result in variation in allele peak heights which is caused by over- and underrepresented amplification of parts from the template DNA during WGA [103]. This means that WGA could also increase the height of ADI peaks to look like real allele peaks. This indeed warrants for DNA contamination precautions such as DNA decontamination of all surfaces to be used, performing sample preparation and amplification in separate areas, using DNA free water, wearing protective clothing like mouth caps and DNA free long sleeve gloves and the like. Such DNA contamination precautions were taken while executing the presented experiments as good as possible. As a possible contamination (as described above with sperm 11) occurred only once, the contamination rate of DNA typing of single spermatozoa performed as described in this thesis was still low. On the other side, the number of samples analysed within this study was also very low in comparison with the high number of samples analysed in a

standard forensic lab. Therefore, a study including a high number of samples which are analysed within a forensic lab environment could be executed to better determine the contamination rate of the method used here.

Since single sperm cells can only produce a haploid profile only one allele per marker can be compared to the suspect's full/genotype DNA profile. In case of a mixture of single spermatozoa derived from multiple donors the donor from the generated haploid DNA profiles are not known. This could make comparing and finding a match more difficult (not impossible). By assigning all autosomal STR profiles from single spermatozoa WGA products with the same X- or Y-STR profile to one group, a consensus DNA profile of the autosomal STR alleles could theoretically be established. Here, a consensus DNA profile means the accumulated profile from multiple DNA profiles from spermatozoa from the same donor. It is basically an established diploid autosomal STR DNA profile based on multiple haploid autosomal DNA profiles.

For example, 11 spermatozoa from Person A and 6 spermatozoa from Person B where typed. 6 spermatozoa from person A (spermatozoa 1 – 6) and 5 spermatozoa from person B (spermatozoa 1 – 5) showed a Y-STR profile. All 6 Y-STR profiles from person A matched each other and the reference profile as did all 5 Y-STR profiles and the reference profile from person B. Though Y-STR profiles from person A did not match with Y-STR profiles from person B as expected. At the 6 spermatozoa from Person A showing a Y-STR profile (spermatozoa 1 – 6), allele 8 was observed at autosomal STR locus TH01 in 5 out of 6 spermatozoa (spermatozoon 5 did not show an allele). No other allele was observed at TH01. At autosomal STR locus D3S1358, allele 14 was observed in 3 out of 6 spermatozoa and allele 18 in 2 out of 6 spermatozoa (spermatozoon 1 did not show an allele). The consensus of TH01 resulted in allele 8 homozygous and for D10S1248 alleles 14 and 15. This was done for all markers to establish a consensus profile including all 16 STR markers. Additionally, a consensus profile was established from haploid autosomal STR profiles with the same X-STR profile. The consensus profile of the Y-STR group and the consensus profile of the X-STR

group were compared and matched each other, as expected. A final consensus profile was then established based on all 11 haplotypes. The final consensus profile matched the reference profile. By this strategy, a full diploid autosomal STR profile could be generated for person A (*table 3.1.4.4*). For person B an autosomal STR profile could be generated with an allele recovery of 96% (*table 3.1.4.5*). The difference in allele recovery can probably be explained by the lower number of single spermatozoa typed from person B. Altogether, results have shown it is possible to establish a diploid autosomal STR profile by taken the consensus from all typed spermatozoa's autosomal STR profiles from the same donor based on their Y- and X-STR profile. As a full consensus profile includes twice as much alleles than a full haploid profile, the power of match probability or likelihood ratio calculations increases when using the consensus profile instead of a haploid profile.

**Table 3.1.4.4. Consensus autosomal STR profile donor A.**

| Donor | Sample | Amelogenin | D3S1358 | TH01 | D21S11 | D18S51 | D10S1248 | D1S1656 | D2S1338 | D16S539 | D22S1045 | vWA | D8S1179 | FGA | D2S441 | D12S391 | D19S433 | SE33 | Allele recovery (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | **gDNA** | X Y | 14 18 | 8 | 28 32.2 | 12 15 | 14 16 | 14 17.3 | 16 24 | 11 | 15 16 | 17 | 12 14 | 19 21 | 10 | 17 22 | 13 14 | 21.2 27.2 | **100** |
| | Sperm 1 | Y | | 8 | | | 14 | | | | | | | | | | | | 18 |
| | Sperm 2 | Y | 14 | 8 | 32.2 | | 16 | 14 | 24 | 11 | 15 | 17 | 12 | 21 | 10 | 17 | 13 | | 88 |
| | Sperm 3 | Y | 18 | 8 | 32.2 | 15 | 14 | 17.3 | 24 | 11 | 16 | 17 | 12 | | 10 | 17 | 14 | | 88 |
| | Sperm 4 | Y | 14 | 8 | 28 | 12 | 14 | | 16 | 11 | | 17 | 12 | 19 | | 22 | 14 | 27.2 | 82 |
| | Sperm 5 | Y | 14 | | | 15 | 14 | | | | | | 14 | | | | | | 29 |
| | Sperm 6 | Y | 18 | 8 | 32.2 | 12 | 14 | 14 | 16 | | 16 | 17 | 14 | 21 | | 22 | 13 | | 82 |
| | Sperm 7 | X | 14 | 8 | 28 | 12 | 14 | 17.3 | 16 | 11 | 16 | 17 | 12 | 21 | 10 | 22 | 14 | 27.2 | 100 |
| | Sperm 8 | X | 18 | 8 | 28 | 15 | 16 | 17.3 | 16 | | 16 | | 12 | | 10 | 17 | 14 | 21.2 | 82 |
| | Sperm 9 | X | 14 | 8 | 28 | 12 | 16 | 14 | 16 | 11 | 15 | 17 | 12 | 19 | 10 | | 14 | 27.2 | 94 |
| | Sperm 10 | X | 18 | 8 | 28 | 15 | 16 | 17.3 | 16 | 11 | | 17 | 12 | 21 | 10 | 17 | 14 | 21.2 | 100 |
| | Sperm 11 | X | 18 | | 28 | 12 | 14 | | 16 | 11 | 16 | 17 | 12 | 21 | 10 | 17 | | 27.2 | 82 |
| | **Consensus** | ***X Y*** | ***14 18*** | ***8*** | ***28 32.2*** | ***12 15*** | ***14 16*** | ***14 17.3*** | ***16 24*** | ***11*** | ***15 16*** | ***17*** | ***12 14*** | ***19 21*** | ***10*** | ***17 22*** | ***13 14*** | ***21.2 27.2*** | **100** |

**Table 3.1.4.5. Consensus autosomal STR profile donor B.**

| Donor | Sample | Amelogenin | | D3S1358 | | TH01 | | D21S11 | | D18S51 | | D10S1248 | | D1S1656 | | D2S1338 | | D16S539 | D22S1045 | | vWA | | D8S1179 | | FGA | | D2S441 | | D12S391 | | D19S433 | | SE33 | | Allele recovery (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| B | **gDNA** | X | Y | 14 | 18 | 6 | 7 | 29 | 30 | 12 | 16 | 13 | 12 | 17.3 | 20 | 24 | 12 | 14 | | 18 | 12 | 13 | 20 | 21 | 11.3 | 14 | 20 | | 12 | | 14 | 18 | **100** |
| | Sperm 1 | | Y | 14 | | | 7 | 29 | | | 16 | 13 | | | 20 | | 12 | 14 | | 18 | | 13 | 20 | | 11.3 | | 20 | | 12 | | | 18 | 94 |
| | Sperm 2 | | Y | 14 | 18 | 6 | | | 30 | | 16 | 13 | | 17.3 | | 24 | | 14 | | 18 | | 13 | 20 | | 11.3 | | 20 | | 12 | | | 18 | 94 |
| | Sperm 3 | | Y | | 18 | | | | 30 | 12 | | 13 | | 17.3 | 20 | | 12 | 14 | | 18 | | 13 | | | | 14 | 20 | | 12 | | 14 | | 88 |
| | Sperm 4 | | Y | 14 | | | | 29 | | 12 | | 13 | | 17.3 | 20 | | 12 | 14 | | | 12 | | | | | | 20 | | | | | | 65 |
| | Sperm 5 | | Y | 14 | | | 7 | 29 | | 12 | | 13 | 12 | | 20 | | 12 | 14 | | 18 | 12 | | 20 | | 11.3 | | 20 | | 12 | | 14 | | 100 |
| | Sperm 6 | X | | | 18 | | 7 | | 30 | 12 | | 13 | | 17.3 | | 24 | 12 | 14 | | 18 | | | 20 | | 11.3 | | 20 | | 12 | | 14 | | 94 |
| | Sperm 7 | X | | | 18 | 6 | | | | 12 | | 13 | | 17.3 | | | | 14 | | 18 | 12 | | 20 | | | 14 | | | 12 | | | 18 | 76 |
| | **Consensus** | ***X*** | ***Y*** | ***14*** | ***18*** | ***6*** | ***7*** | ***29*** | ***30*** | ***12*** | ***16*** | ***13*** | ***12*** | ***17.3*** | ***20*** | ***24*** | ***12*** | ***14*** | | ***18*** | ***12*** | ***13*** | ***20*** | ***21*** | ***11.3*** | ***14*** | ***20*** | | ***12*** | | ***14*** | ***18*** | **96** |

In all, results have shown that using micromanipulation in combination with REPLI-g single cell WGA full haploid autosomal STR DNA profiles as well as X-STR or Y-STR DNA profiles can be obtained from single sperm cells. It must be taken into account that the sperm cells used here were not mixed with epithelial cells or dried and extracted from swabs. To address these questions, further experiments were conducted of which results are described in the following paragraphs.

### 3.1.5. Performance of WGA on single spermatozoa from mock sexual assault swabs

In many sexual assault cases, enough sperm exists to be analysed using conventional DNA extraction including differential lysis. However, in some cases not enough spermatozoa are present which often results in not obtaining a DNA profile. Additionally, in some sexual assault cases the number of spermatzoa is not a problem, but a mixture of spermatozoa from different men; conventional DNA extraction with differential lysis conditions is not able to separate the spermatozoa fraction of one donor from the spermatozoa fraction of the other donor. Furthermore, in some cases can include known and unknown suspects.

A typical casework scenario in which only a few sperm might be present could consist of a teenage girl who fell asleep in a room at a house party and woke up to find a male having vaginal intercourse with her. Since she was scared of anyone finding out, she did not report the incident for a week. After 7 days she went for medical examination upon talking to a friend. Vaginal swabs were taken and were sent to the forensic laboratory. Studies have shown that the maximum expected detection of semen on vaginal swabs is 7 days [104]. Based on this information, the forensic scientist receiving this case rejected conventional DNA extraction and profiling of the intimate swabs. Instead, DNA profiling of single spermatozoa as described in this thesis could be performed. As mentioned before in this thesis, not every sexual assault case includes a known suspect. Therefore, in these cases no reference profile of a suspect is available for comparison. Instead, the DNA profile is run through the national DNA database in order to find a matching DNA profile.

Other case work could include a victim who was raped by two unknown males. In this case DNA extraction and profiling of the intimate swabs from the victim could give a mixed DNA profile. In order to perform a search from the national DNA database, a clean (not a mixed DNA profile) DNA profile is required. Thus, a

mixed DNA profile from two unknown suspects can prevent potential identification by a national database search.

For all of these above described cases DNA typing of single spermatozoa could allow DNA analysis and contribute to the identification of the suspect. In order to verify the method of single spermatozoa DNA typing by using WGA as presented in this thesis, mock sexual swabs spiked with semen from either one or two known men were used. Additionally, mock sexual swabs spiked with semen from either one or two unknown men were used for full validation of the method. A set of reference DNA profiles of the sperm donors was established from mouth swabs and saved as a mock DNA database. Autosomal DNA profiles obtained from the single spermatozoa were compared to the reference DNA profiles in the mock DNA database.

#### *3.1.5.1. Mock sexual swabs with single male contributor*

To investigate the performance of WGA on single spermatozoa from mock swabs including semen from a single male, four swabs, each containing a different known female donor and a different known male contributor, were used. Since a DNA profile of a single spermatozoon could not always be obtained, the first swab, "female A with male A", was used to estimate how many spermatozoa had to be analyzed to successfully obtain profiles from at least three spermatozoa. From swab "female A with male A", an autosomal STR profile was successfully obtained from 6 out of 10 treated spermatozoa (*table 3.1.5.1.1*). Upon this result, it was decided to isolate 7 spermatozoa from all mock swabs with a single male contributor and 14 spermatozoa for all mock swabs containing two male contributors.

**Table 3.1.5.1.1. Success rate of DNA typed single spermatozoa, average allele recovery as well as number of spermatozoa with ADI from mock sexual assault swabs which included a single male contributor.**

| Swab | | *N** Spermatozoa isolated (n**) | Successfully typed spermatozoa (n) | | | Average allele recovery (%) | | | Spermatozoa with ADI (n) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Female | Male | | Autos. | X | Y | Autos. | X | Y | Autos. | X | Y |
| A | A | 10 | 5 | 3 | 2 | 55 | 83 | 29 | 0 | 1 | 0 |
| B | B | 7 | 7 | 4 | 3 | 56 | 69 | 49 | 0 | 1 | 0 |
| C | C | 7 | 2 | 0 | 1 | 59 | - | 61 | 1 (6 ADIs) | - | 0 |
| D | D | 7 | 2 | 2 | 0 | 77 | 67 | - | 0 | 0 | - |
| A | Unknown | 7 | 1 | 0 | 1 | 59 | - | 52 | 0 | - | 0 |
| B | Unknown | 7 | 4 | 1 | 3 | 65 | 83 | 73 | 0 | 0 | 1 |

* *N* = number

** n = number

As shown in table *3.1.5.1.1*, all successfully typed spermatozoa contained alleles matching the male contributor's reference profile. The number of successfully typed spermatozoa differed between the swabs. The robustness of spermatozoa against environmental impacts e.g. freezing/thawing, drying or swab treatment during cell extraction as well as the quality of the spermatozoa's DNA is important for successful DNA typing. Therefore, these two traits do probably differ from person to person. Many studies have shown that sperm quality as well as sperm count can be affected by factors such as lifestyle which can differ from person to person [105]. As a result, the quality and number of spermatozoa in a male's semen can differ between individuals. This individual dependent sperm quality could possibly explain the difference in success rates between the swabs since each swab contained semen derived from a different person. On the other hand, the WGA itself could have caused these differences due to under and over representing loci of the spermatozoon's whole genome during amplification. This could lead to amplification of loci not relevant to forensic DNA profiling with no or not enough amplification of relevant STR loci. Also, the number of ADI per sperm differed between swabs and between spermatozoa from the same swab. WGA product from only one spermatozoon contained ADIs. As the latter matched the alleles from the female contributor, DNA left behind from the female fraction was possibly transferred and co-amplified along with the spermatozoa's DNA during WGA. This was not as problematic as ADIs usually are, since only 1 out of 7 typed spermatozoa showed ADIs in its profile, the 6 ADIs were in addition to the correct male alleles and the ADIs matched the female DNA profile, these alleles can easily be filtered out by establishing the autosomal haploid profile of the sperm.

The WGA products of all successfully autosomal typed spermatozoa were subjected to both X- and Y-STR typing. As to be expected, all WGA products which showed AMELX were positive for X-STR alleles, whereas no Y-STR alleles were obtained. *Vice versa*, all WGA products containing AMELY generated Y-STR alleles, but not X-STR alleles. Therefore, the results shown here by obtaining an autosomal haploid profile as well as only either an X-STR profile or a Y-STR

profile proves that only a single spermatozoon has been analyzed in each case. Thus, together with the fact that alleles from both autosomal STRs and Y- or X-STRs match the male contributor's reference profile, the applied strategy could make it possible to examine DNA evidence containing only a few spermatozoa which are usually too low in number for differential lysis and conventional PCR.

The technique used here could be useful in cases where intimate swabs were taken late after sexual intercourse as presented in the example case scenario of the girl who was raped at the party and went for medical examination after 7 days. Instead of using conventional DNA extraction, the intimate swabs could be subjected to single spermatozoa profiling by using micromanipulation and WGA. A possible result could be that two out of 7 spermatozoa derived WGA products would be positive in autosomal STR profiling and generated a partial haploid profile, both with AMELY. An ADI occurred at one marker from one spermatozoon derived WGA product and matched the victim's reference DNA profile. Y-STR profiling was positive, but negative for X-STRs. These findings altogether confirm only single spermatozoa were analysed.

Both autosomal haplotype profiles, as well as both Y-STR profiles matched the accused reference DNA profile. The two haploid autosomal DNA profiles could be used to calculate the match probability with the suspect's diploid DNA profile. Normally the matching probability is calculated by multiplying the genotype frequency for each STR locus. The frequency per locus depends on if the locus is homozygous or heterozygous. In case of partial diploid DNA profiles calculations are different, since heterozygous markers with one allele dropping out is not the same as a homozygous marker. A full haploid DNA profile, like the one from a spermatozoon, contains neither homozygous autosomal STR loci nor is a partial DNA profile. As each spermatozoon can theoretically produce a child (or is part of a child) the alleles of the diploid profile make up half of the diploid profile of a child. Therefore, it is suggested to use a paternity based match calculation. Here, the paternity index (PI) is used to determine how much more likely it is that the alleged father (AF) is the biological father of the child than a random

man [106]. The obligate alleles of the child's DNA profile are used by subtracting the maternal alleles of the child's DNA profile. The alleles obtained from a single spermatozoon are also the obligate alleles and can therefore be used in the same way to calculate the probability. The question is how much more likely a haploid profile is to be seen if the sperm was derived from the suspect (provided consistency between the haploid genotype of the sperm and the diploid genotype of the suspect) as compared to a random man from the population. The corresponding likelihood ratio can be calculated based on the following rationale: If for a locus in question the suspect's genotype is heterozygous, then the chance for him to transmit the allele detected in the sperm's DNA is 0.5 (since he can transmit either of two alleles); if the suspects genotype is homozygous then the chance of transmitting the sperms allele will be 1. The chance to generate the sperms genotype by a random man equals the population frequency of the respective allele. For several STR loci, these values have to be multiplied.
For example, if the haploid profile of the spermatocyte were D3S1358 allele 15 and TH01 allele 9.3 and the suspect's genotypes at these loci were D3S1358 15,18 and TH01 9.3, the likelihood ratio (LR) would be calculated as follows:

LR = 0.5 x 1 /(P D3S1358 allele 15 x P TH01 allele 9.3) .

The outcome of the LR could enable the court to understand the strength of the match.
Although the approach described in this study shows only a few spermatozoa are needed for successful DNA typing, the number of spermatozoa from each swab was high. Hence, additional experiments should be carried out, including mock swabs with lower dilution of semen as well as casework samples to determine the efficiency at lower spermatozoa count samples.

### *3.1.5.2. Mock sexual swabs with two male contributors*

Mock swabs containing sperm originating from two male donors were also used to test the described protocol of single spermatozoa DNA typing using

micromanipulation and WGA to be applicable to sexual assault cases in which the victim was sexually assaulted by two men. Ideally from each donor 7 spermatozoa were typed. At first one swab, with two known male contributors, male F and Male G, was tested. Here, 14 spermatozoa were isolated of which 12 generated an autosomal haploid profile (*table* 3.1.5.4.1). 4 autosomal haploid profiles matched the reference profile of male F and the other 8 matched male G. Furthermore, an X-STR profile was obtained from 6 spermatozoa. From the remaining 6 spermatozoa, a Y-STR could only be obtained from 5 spermatozoa. All obtained Y- and X-STR profiles matched the correct male reference profile. ADIs were only noticed in two of the X-STR profiles where in each profile a single ADI was observed. One of the ADIs was one repeat unit longer than the reference allele and could therefore possibly be a slippage mutation. The other ADI differed in 3 repeat units is less likely due to a mutation, though still possible. This ADI could of course be due to contamination. However, the chance of ADI due to environmental contamination during WGA needs to be investigated. This is to conclude whether an ADI is due to a mutation or contamination. It did not match the other male's reference profile. It was not possible to confirm whether this ADI was derived from the female's DNA, since the female's reference DNA was not tested for X-STRs.

Taken together, these results demonstrate successful DNA typing of spermatozoa in a mixture sample of two male contributors. By using the approach described here, X and Y-STR profiles can be used to assign each spermatozoon to the correct male contributor. In a previous study the use of Y-STRs has been used in mixtures to differentiate between single spermatozoa from different donors [107]. However, only Y-STRs were used (no autosomal STRs). Y-STR profiles are statistically less powerful than autosomal STR profiles for calculating match probabilities. The complete Y-chromosome, including the Y-STRs, can namely only be inherited paternally. This means that the Y-STR profile stays the same through many generations resulting in many men having the same Y-STR profile. Furthermore, using Y-STRs to differentiate spermatozoa from different male contributors cannot be applied when the two male contributors are paternally

related. As the gDNA of the spermatozoa in that study was used directly as template in the Y-STR specific PCR, the approach did not allow for other analysis like autosomal STRs. In contrast, the method described in this thesis does allow for autosomal as well as X-/Y-chromosomal STR analysis at the same time from each spermatozoon. Hence, this method could also be suitable in case the male contributors are paternally related to each other.
In conclusion, the results here showed that it is possible to avoid mixed DNA profiles no matter how many males there are by single DNA profiling using the method described in this study.

### *3.1.5.3. Mock sexual swabs with a single "unknown" male contributor*

In order to determine whether the single spermatozoa DNA profiling method facilitates the correct identifying unknown perpetrators, mock sexual swabs containing a single unknown male contributor were used. Multiple single sperm were typed. Reference profiles of all male and female donors contributed to this study were used as an internal database. After autosomal haplotypes as well as X- and Y-STR profiles were obtained, profiles were compared with the reference profiles in the database to find a possible match.

All successfully typed spermatozoa contained alleles that matched the male contributor's reference profile. From one swab, a profile could be obtained only from one spermatozoon out of seven. The autosomal haploid allele recovery of this spermatozoon was 59% (*table 3.1.5.1.1*). Despite of the low success rate, the allele recovery was high enough to match the correct reference profile within the internal database. Also with the other swab, with a success rate of 4 out of 7 typed spermatozoa, it was possible to match the correct reference profile. One of the 4 successfully typed spermatozoa showed one ADI in its Y-STR profile. Since this ADI was only 1 repeat different from the reference allele number, no second allele of the same locus was seen and no further ADI were noticed at the other 22 markers of this Y-STR profile, the ADI was presumed to be a mutation. However, one cannot rule out that the ADI may be due to an environmental contamination

since WGA is an extremely sensitive amplification technique and a study to determine environmental contamination rates was not carried out in this study.

Having successfully typed spermatozoa from swabs containing an unknown male donor, the above method could be applied for sexual assault cases when the offender is unknown.

#### *3.1.5.4. Mock sexual swabs with two "unknown" male contributors*

Just as assailants in sexual assault cases involving a single offender can be known or unknown, so can assailants in sexual assault cases in which multiple offenders are involved. Therefore, the single sperm DNA typing method presented in this thesis was tested on mock sexual assault swabs containing semen from two unknown male contributors. From one swab an autosomal STR haplotype profile from 12 of the 14 isolated spermatozoa was obtained (*table 3.1.5.4.1*). The success rate of the other swab was slightly lower, with 10 out of 14 successfully typed single spermatozoa. From both swabs autosomal STR profiles as well as X- and Y-STR profiles matched the correct reference profiles. Both male contributors of the second swab were identified by comparing the 10 autosomal haplotypes to the reference profiles in the "internal database". Two of the haplotypes matched with male A and the other 8 matched with male G. From the other swab all 12 obtained autosomal haplotypes matched one and the same reference profile in the database. Therefore, only one of the male contributors, male F, was identified. This could possibly be explained by person dependent sperm quality differences [105]. Also the concentration of spermatozoa in semen could differ from person to person which could have led to a ratio different from 1:1. The presence of spermatozoa from one person would then be more abundant than spermatozoa from the other contributor. This could have led to the isolation of more (or only) spermatozoa from one contributor. Another possible cause for the result obtained, result could be statistical sampling where by chance all 12 isolated spermatozoa were derived from the same donor. Although if the number of spermatozoa from each contributor are in equal

amounts, the chance that all 12 isolated spermatozoa are from the same contributor would be very low, namely $P=2.44 *10^4$ ($P= 1 / 2^{12}$).

The mock swabs containing unknown male contributors used here, demonstrated that it is possible to assign and identify male contributors by single spermatozoa DNA profiling without generating mixed profiles. DNA profiling of single spermatozoa as described in this study could therefore be applied to use similar samples such as intimate swabs from victims who had been subjected to a multiple-offender rape scenario.

In sexual assault case work that entails mixtures of two unknown suspects would normally complicate the search in the national DNA database. Although the results in this study have shown that mixtures can be avoided by using single spermatozoa DNA profiling, the method presented here enables for DNA database searches no matter how many unknown males contributed to the sample. Therefore, this method can greatly benefit forensic case work.

**Table 3.1.5.4.1. Success rate of DNA typed single spermatozoa, average allele recovery as well as number of spermatozoa with ADI from mock sexual assault swabs which included two male contributors.**

| Swab | | | *N** Spermatozoa isolated (n**) | *N* successfully typed spermatozoa (n) | | | Average allele recovery (%) | | | *N* Spermatozoa with ADI (n) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Female** | **Male 1** | **Male 2** | | **Autos.** | **X** | **Y** | **Autos.** | **X** | **Y** | **Autos.** | **X** | **Y** |
| **F** | **F** | **G** | 14 | 12 | 6 | 5 | 76 | 75 | 51 | 0 | 2 | 0 |
| **E** | **Unknown** | **Unknown** | 14 | 12 | 4 | 8 | 73 | 85 | 71 | 1 | 0 | 0 |
| **F** | **Unknown** | **Unknown** | 14 | 10 | 7 | 3 | 38 | 69 | 52 | 0 | 1 | 0 |

* *N* = number

** n = number

### *3.1.6. Conclusion*

Briefly, the micromanipulation method described here in combination with WGA enabled DNA profiling of single spermatozoa. The single spermatozoa DNA profiling method allows to analyse autosomal STRs as well as X- and Y-chromosomal STRs from the same spermatozoon. Therefore, the method described in this thesis facilitates the avoidance of mixtures of multiple males while producing single profiles, demonstrating great potential for casework.

.

## 3.2. Skin flakes

### 3.2.1. Introduction

In order to develop a method which enables successful forensic DNA profiling of single skin flakes multiple different experiments were performed. First different staining assays were assessed and compared based on their discrimination power and PCR compatibility/inhibition (chapter 3.2.2.). Second, multiple different DNA extraction methods on skin flakes were assessed based on rate of dissolving skin flakes and PCR inhibition (chapter 3.2.3.). Then, the selected DNA extraction method was tested with single skin flakes in combination with a direct PCR approach on the lysate and assessed based on allele recovery (chapter 3.2.4.). Afterwards, the WGA method, which was used for single spermatozoa in this thesis, was also applied to determine its impact on DNA profiling of single skin flakes based on allele recovery (chapter 3.2.5.).

The combination of the selected staining method and DNA extraction method was then tested on skin flakes secured from mock mixed DNA evidence as part of the validation process (chapter 3.2.6.). The mock DNA evidence consisted of pillows touched by a single or multiple persons. Additionally, the quantity and quality of DNA of single skin flakes was investigated by DNA profiling of single skin flakes from different persons using the method described in this thesis (chapter 3.2.7.). Finally, the performance of a mini-STR based NGS approach on single skin flakes was assessed based on allele recovery in comparison with the conventional PCR and CE approach (chapter 3.2.8.).

### 3.2.2. Staining of skin flakes on tape liftings

The following figures show a series of tape liftings on which only skin flakes (figure 3.2.2.1) or particles from hair products (figure 3.2.2.2 and figure 3.2.2.3) were secured. Morphologically skin flakes and the hair product particles are different from each other and could be distinguished from each other.

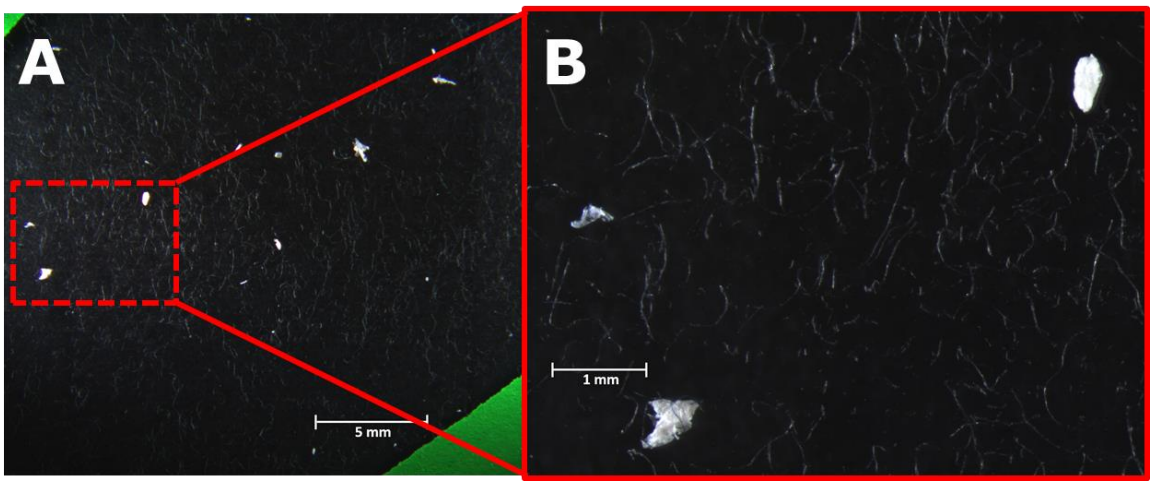


**Figure 3.2.2.1. Microscopic view with 6x magnification (A) as well as a close up with 25x magnification (B) of skin flakes on a tape lifting.**

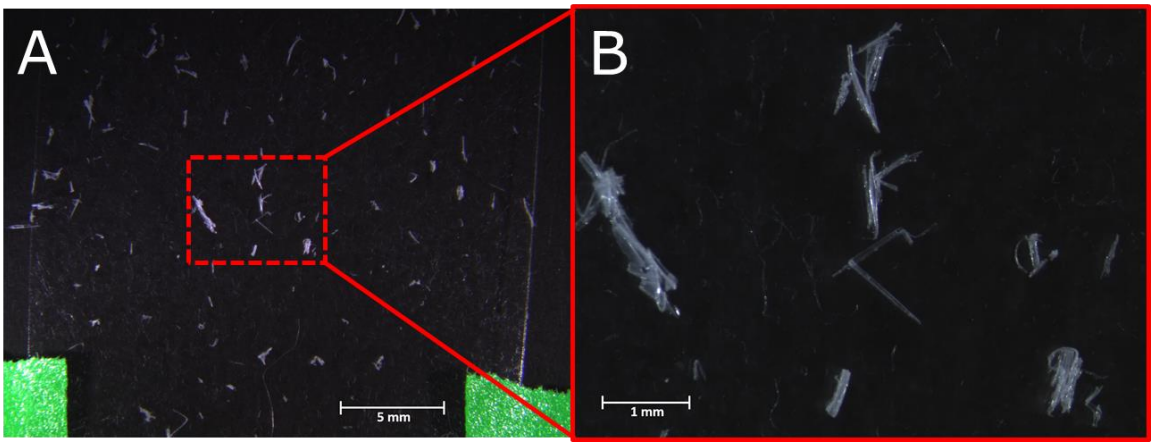


**Figure 3.2.2.2. Microscopic view with 6x magnification (A) as well as a close up with 25x magnification (B) of hair gel particles on a tape lifting.**

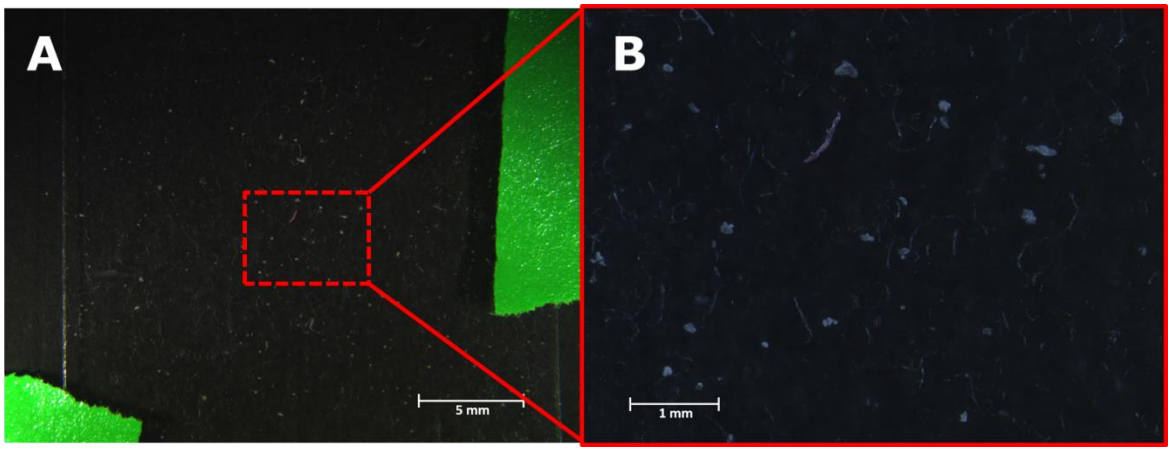


**Figure 3.2.2.3. Microscopic view with 6x magnification (A) as well as a close up with 25x magnification (B) of beach matt particles on a tape lifting.**

However, this can get complicated when tape liftings contain a mixture of skin flakes and hair product particles as seen in figure 3.2.2.4. In order to make skin flakes visible on tape liftings and to distinguish them from hair product particles, Orange G, methylene blue and trypan blue were used for staining. They were investigated to confirm whether a single or a combination of two dyes could highlight skin flakes while providing discrimination from hair product particles.

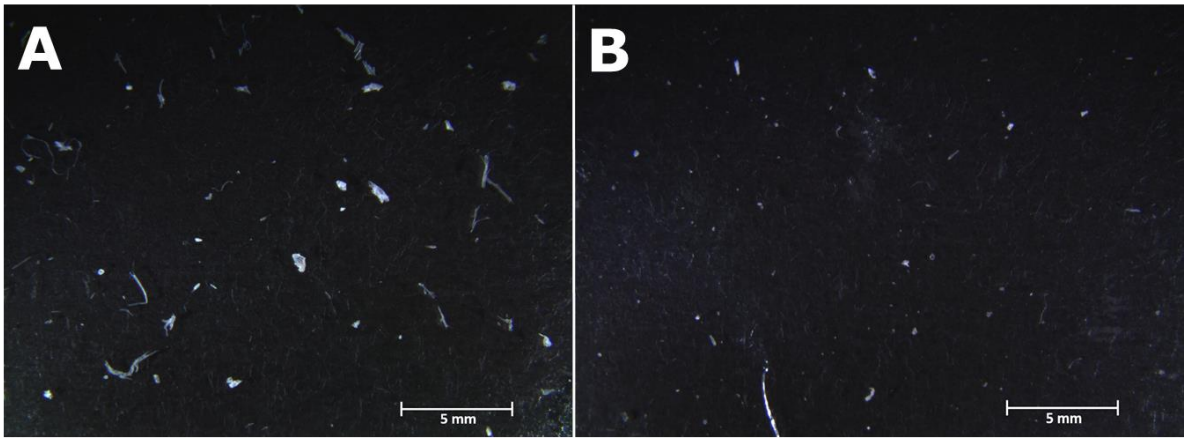


**Figure 3.2.2.4. Microscopic view (6x magnification) of skin flakes mixed with hair gel particles on a tape lifting (A) as well as skin flakes mixed with beach matt on a tape lifting.**

Among all the stains, Orange G performed the best. Skin flakes on tape liftings that were stained with Orange G turned bright orange (see *figure 3.2.2.5*). This was expected, since Orange G is used to selectively stain keratin [44]. However, the tape lifting adhesive remained unstained.

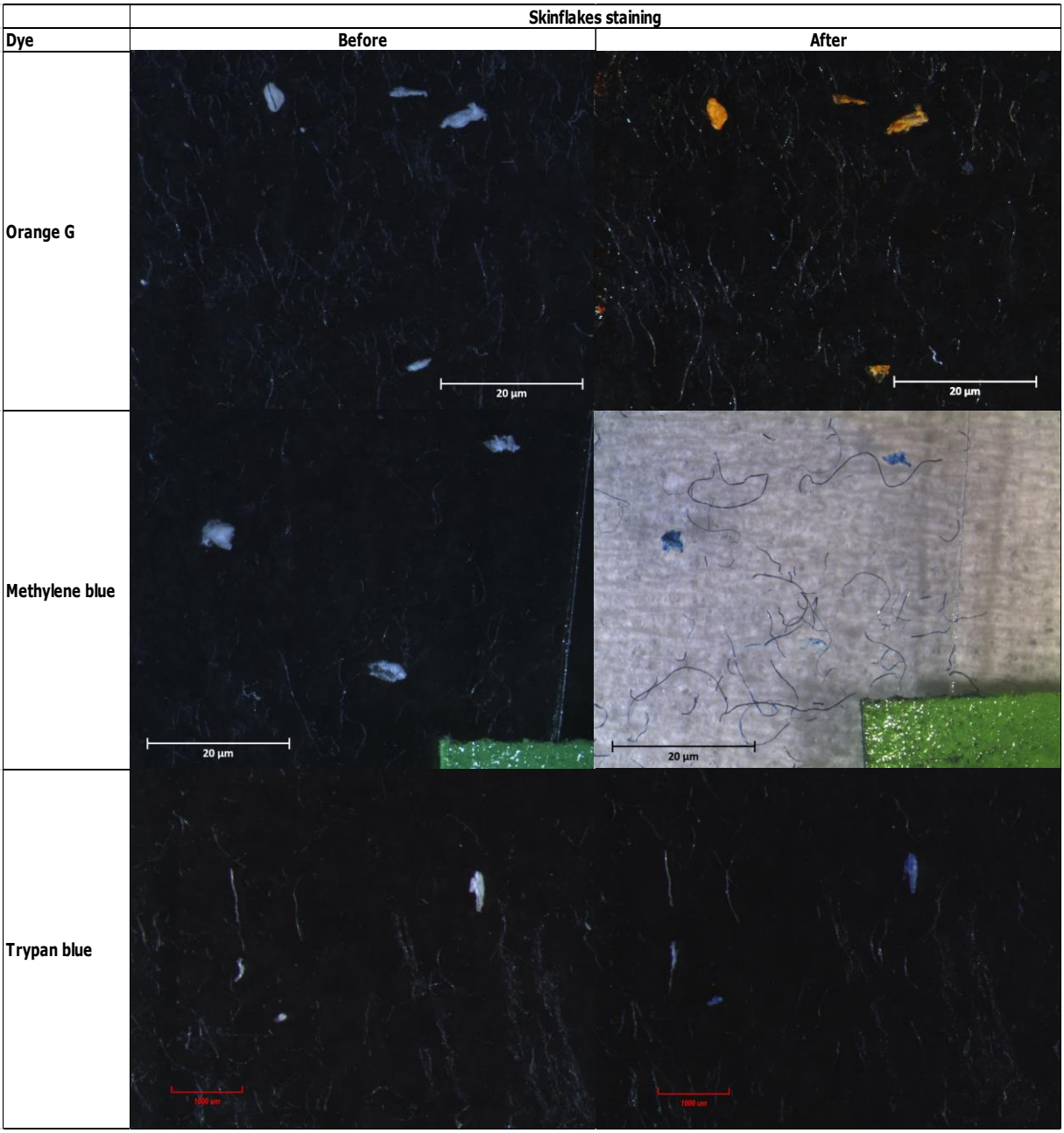


**Figure 3.2.2.5. Results on efficiency of different dyes on staining skin flakes on tape liftings. Skin flakes secured on tape liftings where stained with Orange G, methylene blue or trypan blue. Microscopic images where taken before and after each different staining.**

Furthermore, particles from both hair gel and beach matt remained unstained (see *figure 3.2.2.6* and *figure 3.2.2.7*). This makes Orange G a very good candidate dye to highlight skin flakes on tape liftings and discriminate them from other particles. This was clearly seen when Orange G was applied to tape liftings with skin flakes mixed with either hair gel particles or beach matt particles (see *figure 3.2.2.8).* Here, not all particles were highlighted after staining. From the tested concentrations (0.5% - 5%) a 3% Orange G staining was chosen for further experiments as this was the optimal concentration for specifically highlighting skin flakes on tape liftings.

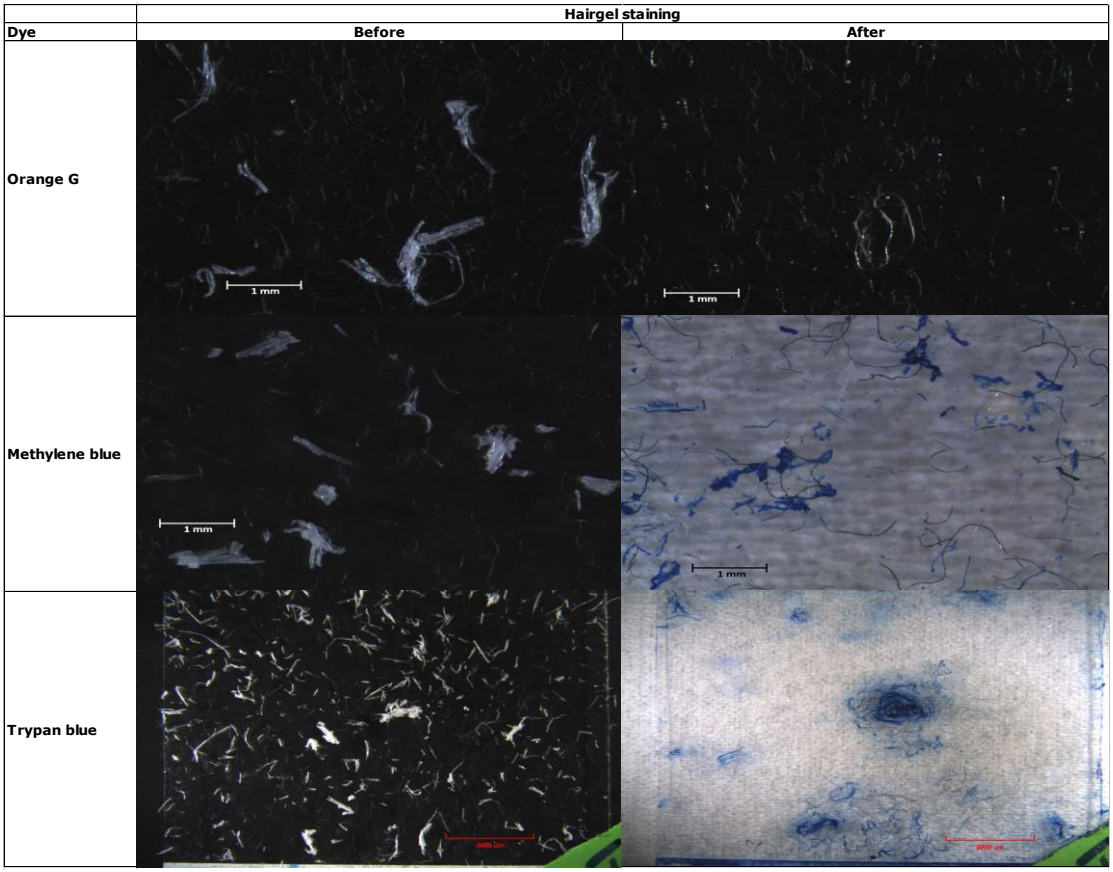


**Figure 3.2.2.6. Results on efficiency of different dyes on staining of hair gel particles on tape liftings. Hair gel particles secured on tape liftings where stained with Orange G, methylene blue or trypan blue. Microscopic images where taken before and after each different staining.**

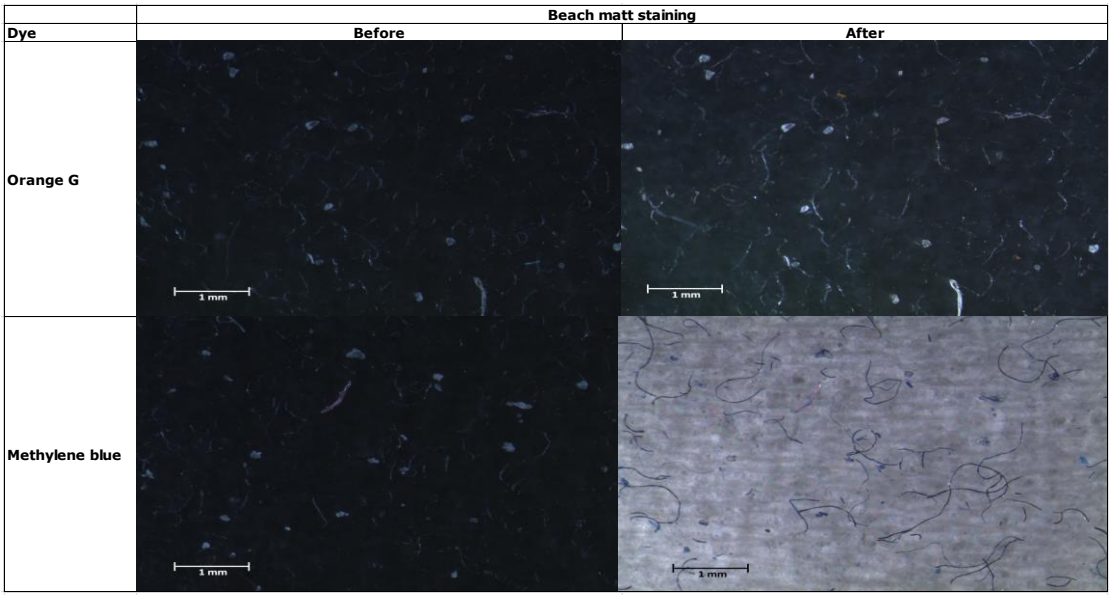


**Figure 3.2.2.7. Results on efficiency of different dyes on staining of beach matt particles on tape liftings. Beach matt particles secured on tape liftings where stained with Orange G and methylene blue. Microscopic images where taken before and after each different staining.**

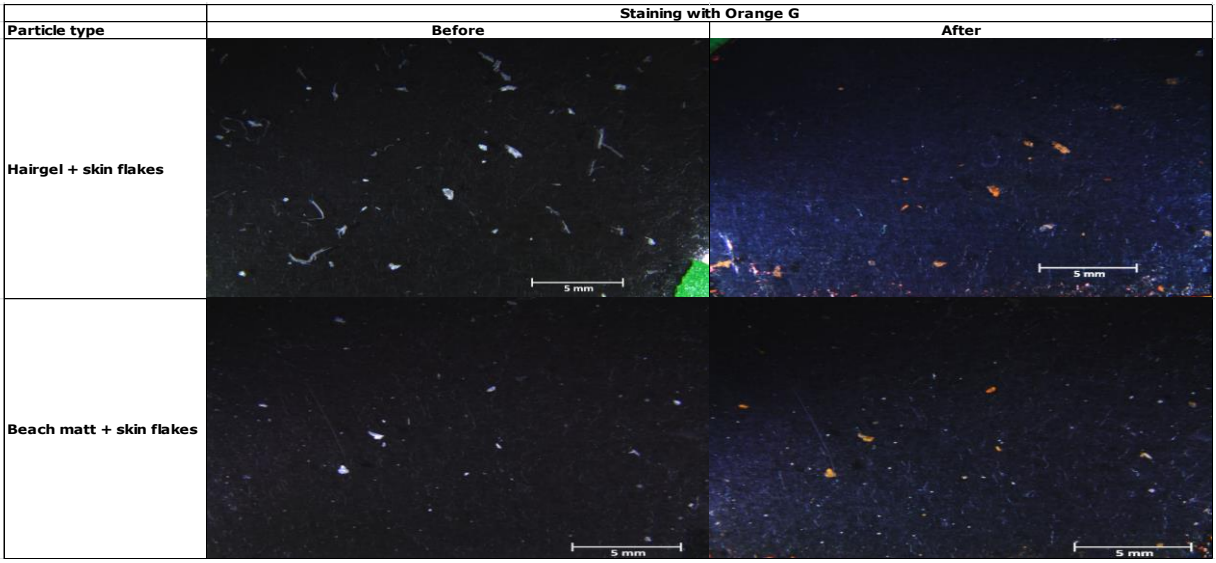


**Figure 3.2.2.8. Results on efficiency of Orange G on differentiating skin flakes from different hair product particles on tape liftings. Tape liftings with mixtures of skin flakes with hair gel or with beach matt particles where stained with Orange G. Microscopic images where taken before and after the staining.**

In contrast to Orange G, methylene blue turned out not to be a suitable candidate. Skin flakes turned blue after staining with methylene blue.

Furthermore, the adhesive surface also turned blue when clean tape liftings were stained with several concentrations (0.1 – 1.5%) of methylene blue. Even at the lowest concentration (0.1%) the skin flakes and the adhesive turned blue. In addition, particles from hair gel and beach matt turned blue when stained with methylene blue (see *figure 3.2.2.6* and *3.2.2.7*).

Results obtained with trypan blue were like those obtained with methylene blue. Trypan blue had been used in previous studies to stain tape liftings to highlight skin flakes and other bio particles [33]. Like in previous studies, trypan blue did not stain the adhesive surface of the tape liftings in the present study. However, trypan blue stained other particles from biological origin too and would therefore not be suitable as a selective dye to highlight skin flakes. It was unknown whether trypan blue would also stain hair product particles. Though, when tested on hair product particles, hair gel turned blue while with beach matt the blue colour was less intense. Therefore, trypan blue was not a suitable dye and its use on skin flakes was discontinued.

As the lysates from stained skin flakes will ideally be used directly for PCR, the staining with Orange G should not interfere with subsequent DNA analysis. Therefore, Orange G was tested on PCR inhibition by adding different amounts (2% - 5%) of Orange G together with gDNA to the PCR master mix and checked electropherograms on profile quality.

Experiments on PCR inhibition revealed no inhibition when PCR was performed with 1µl 2% Orange G was added to the master mix. This concentration is far higher than the amount of Orange G which is actually added when adding crude lysate from a stained and lysed single skin flake. This also makes Orange G a suitable dye in cases when no DNA purification of the lysed sample is preferred. However, 3% Orange G and higher exhibited strong PCR inhibition (*figure 3.2.2.9*).

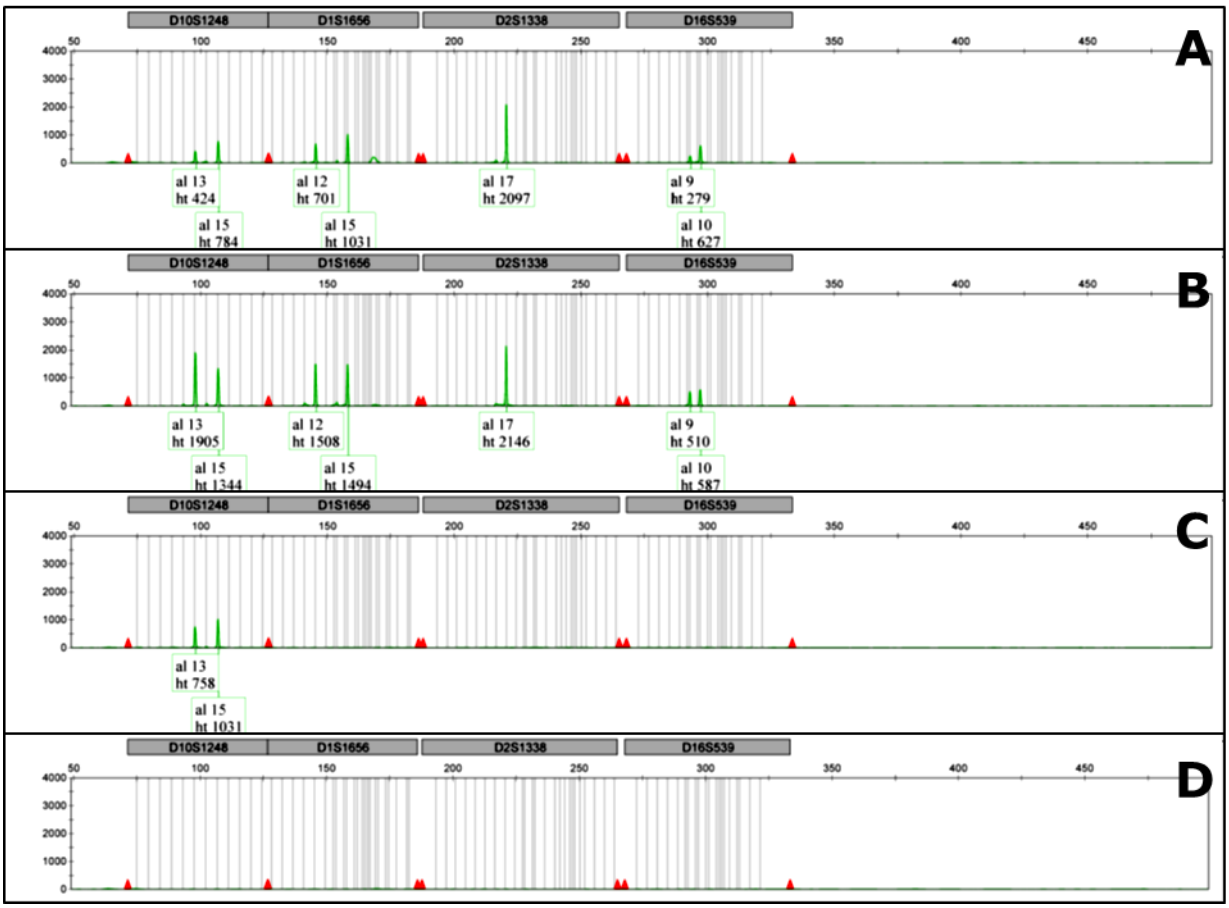


**Figure 3.2.2.9. Electropherograms that show inhibitory effects of 1µl 2% (A), 3% (B), 4% (C) and 5% (D) of Orange G on PCR. The relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. On top, the X-axis represents the fragment size in base pairs (bp) with the allele corresponding STR locus name above in the grey horizontal bars. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected.**

Altogether, the keratin selective staining property makes Orange G an excellent dye to highlight and discriminate skin flakes from other particles on tape liftings in forensic casework. This enables more specificity and less time-consuming isolation of skin flakes from tape liftings for downstream DNA analysis. Furthermore, skin flakes stained with Orange G do not require any DNA purification since Orange G does not inhibit PCR at the quantities required. This makes Orange G compatible with direct PCR assays of crude lysate. So far, no other dye has been reported in literature that had been used to highlight skin flakes on tape liftings. In contrast to many fluorescent dyes, Orange G is not expensive and does not require fluorescence which makes expensive fluorescent

microscopes unnecessary. In addition to its stability and the simple/fast staining protocol, makes Orange G a suitable dye easily implemented in a forensic biology laboratory. Tape liftings used in this study have been used for decades, which mean cold cases containing these tapes can also be stained and analysed with Orange G.

### 3.2.3. DNA extraction from skin flakes

Several lysis protocols were tested on bulk skin flakes in order to select one that is the most efficient in breaking down skin flakes to release their gDNA. The Chelex protocol described in a study by Lorente et al. (1998) [35] was the least efficient one (*table 3.2.3.1*). It was not able to dissolve most of the skin flakes after 3 hours. Even after 12 hours the number of undissolved skin flakes did not change noticeably. This was expected since Chelex is a metal ion binding resin and has no other lysing properties such as solubilizing lipids or degrading proteins.

**Table 3.2.3.1. Efficiency of solvation of skin flakes by different lysis protocols.**

| Lysis protocol | Level of skin flake solvation |
|---|---|
| Chelex | - |
| Chelex + Prot K + Tween20 | + |
| NaOH | -* / +++ |
| NaOH + Tween20 | +++ |

*In contrast to skin flakes donated during the colder months, skin flakes donated during the warmer months did not dissolve.

Also the Chelex based protocol described by Schneider et al. (2011) [37] was not able to break down skin flakes after 3 hours. However, extended incubation improved results since 12 hours incubation, did dissolve more intact skin flakes. Components that are often used in lysis buffers include proteinase K and Tween20 were not included in the protocol of Lorente et al. (1998). These components were probably responsible for the better results observed in the Chelex protocol of Schneider et al. (2011) [37] compared with the results observed with the protocol of Lorente et al. (1998).

However, both NaOH based protocols gave the best results. Even after a 3 hours incubation, all skin flakes were broken down and dissolved. Worth noting that these alkaline based protocols did not contain proteinase K or other protein digesting components and yet they perform best in digesting skin flakes. This is

probably due to the denaturing and hydrolytic potential of the alkaline buffer. Proteinases are often used to digest simple proteins. For more complex proteins including proteins with disulfide bonds, they are digested with proteinases in combination with a reducing agent such as DTT. The reducing agent opens the disulfide bonds and helps denaturing the protein [108]. Only then proteinases are able to efficiently digest the protein. The keratin which is present in skin flakes, hairs, feathers, nails and horns is a very complex protein and requires at least an agent which breaks up the disulfide bonds in order to denature and enable digestion by proteinases. In previous studies it was found that the use of an alkaline buffer was highly effective in dissolving chicken feathers and sheep wool [89,109]. This was in contrast to the use of proteinases such as proteinase K and reducing agents such as DTT which did not work as efficiently. Also, the ability to dissolve in liquid seems to be of importance for digestibility by proteinases, as undissolved solid proteins won't be as accessible [89,108,109]. After incubation, the pH of the alkaline buffer was neutralized, a white precipitate was observed. That white precipitate could be salts. However, as keratin is not soluble in neutral pH solutions, the white substance could also be explained by keratin precipitated out from the solution. These findings agree the results in a study where an alkaline buffer in combination with a neutralizing buffer was used to isolate keratin from chicken feathers [89].

However, the NaOH protocol without Tween20 did not work with skin flakes donated during the warmer months in contrast to skin flakes donated by the same persons during the winter months. Even after 12 hours, solvation did not occur. The human skin can be less fatty during the colder and drier winter months than it is during the warmer more humid time of the year [110]. Tween20 is a non-ionic detergent which solubilizes membrane proteins and emulsifies lipids. As only the Tween20 containing alkaline buffer was able to dissolve skin flakes donated in both seasons, the skin flakes donated during the warmer more humid months probably contained more fat. As the feathers in a previous study were first defatted prior to dissolve them by an alkaline solution, a

fatty layer surrounding the surface of the skin flake could protect it from dissolving by NaOH [89].

The Tween20 containing alkaline buffer was selected as a potential skin flake lysis buffer for further experiments in this project. In order to use the crude lysate directly as template DNA for PCR, the NaOH+Tween20 buffer after neutralisation with HCl was tested on PCR inhibition. Using a buffer containing 0,6M NaOH (neutralized with 0.6M HCl) PCR inhibition was noticed. Though, no inhibition was observed at 0.3M concentrations of NaOH (and HCl). Therefore, the neutralized crude lysate needed to be 1:2 diluted with DNA free water in order to be directly used as template for PCR. Since adding water meant also diluting the amount of gDNA residing in the crude lysate, lysis of skin flakes was tested with 0.3M NaOH instead of 0.6M NaOH. As 0.3M NaOH was as efficient as 0.6M NaOH in lysing skin flakes, that concentration was selected for further experiments including testing on single skin flakes.

The 0,3M NaOH + Tween20 buffer was able to dissolve single skin flakes for most of the skin flakes within 3 – 12 hours. However, this greatly depended on the morphology or type of skin flakes. Aside from the more abundant flat skin flakes, skin flakes looking like clumps were also seen (*figure 3.2.3.2.*).

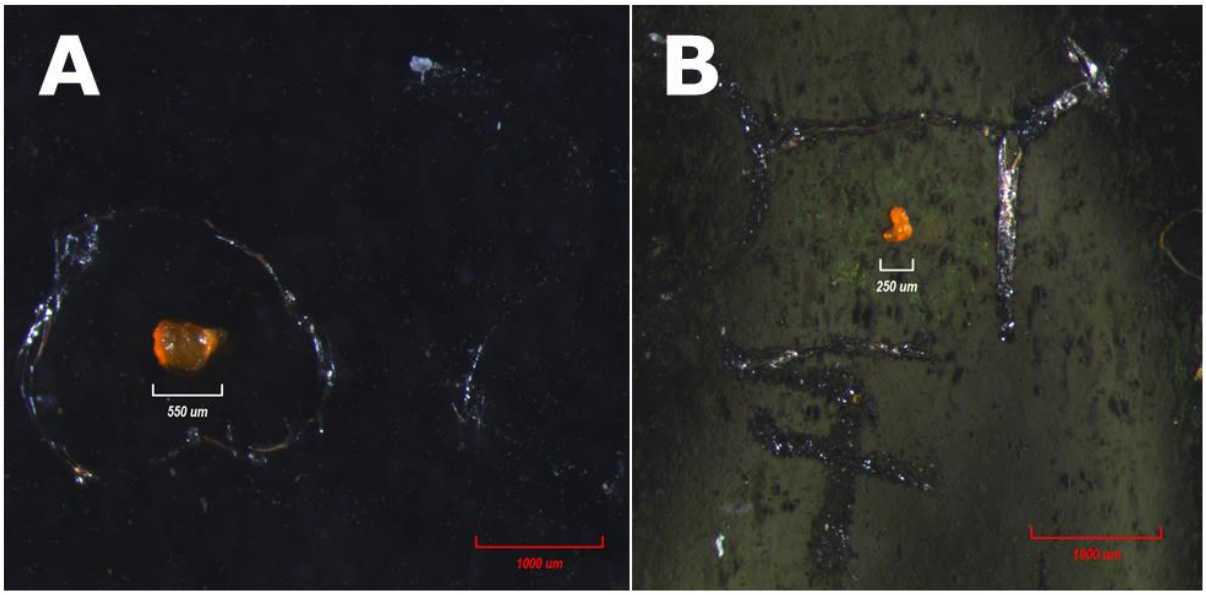


**Figure 3.2.3.2. Microscopic views of the clump morphotype of skin flakes with 25x magnification from person A (A) and from person B (B).**

Therefore, two types of skin flakes were characterized within this study. The flat skin flakes were almost all able to be lysed within 3 hours. However, the clumpy skin flakes were not yet completely lysed after 9-12 hours. To improve lysis of these clumps, mashing single skin flake clumps was performed using a scalpel on a clean microscopy slide before adding to the lysis buffer. This greatly improved lysis efficiency enabling complete lysis within 12 hours. The reason why the clump type of skin flakes was more difficult to lyse is probably the higher amount of keratin and/or fatty compounds. Results from longer incubation time and smashing as pre-lysis treatment support this hypothesis. Prolonged incubation enables the skin flake to interact more with the alkaline environment by penetrating deeper into the flake due to increased swelling. This would increase the molecular surface available which results into more dissolving of keratin. Also mashing the skin flake increases the surface area that interacts with the alkaline environment and leads to more dissolving of keratin.

Interestingly, the DNA content measured by the ALU based qPCR was dependent on both size of the skin flake and donor individual (*figure 3.2.3.3*). Although, DNA content of skin flakes increased with size when flakes were derived from person A, the average DNA content for small and big skin flakes was very low compared to skin flakes with the same size from person B. The average DNA content of clumps from person A was much higher when compared to clumps from person B. However, this was only due to one clump of which the crude lysate contained 7ng/µl DNA (84ng in total). The DNA concentration of the other clumps from person A was around 1ng/µl or lower. A possible explanation for the higher DNA content in the big skin flakes compared to small skin flakes could be the higher number of epithelial cells or intact nuclei present in big skin flakes. Although clumps are not always bigger than small or big skin flakes, among all three morphological types, clumps contained the most amount of DNA on average. A hypothesis for this observation could be that clumps consist of many epithelial cells packed tightly together. This could be a reason why these clumps were so hard to lyse in comparison with the flat skin flakes. The more and tighter keratin containing cells are packed together, the harder it would be for them to lyse.

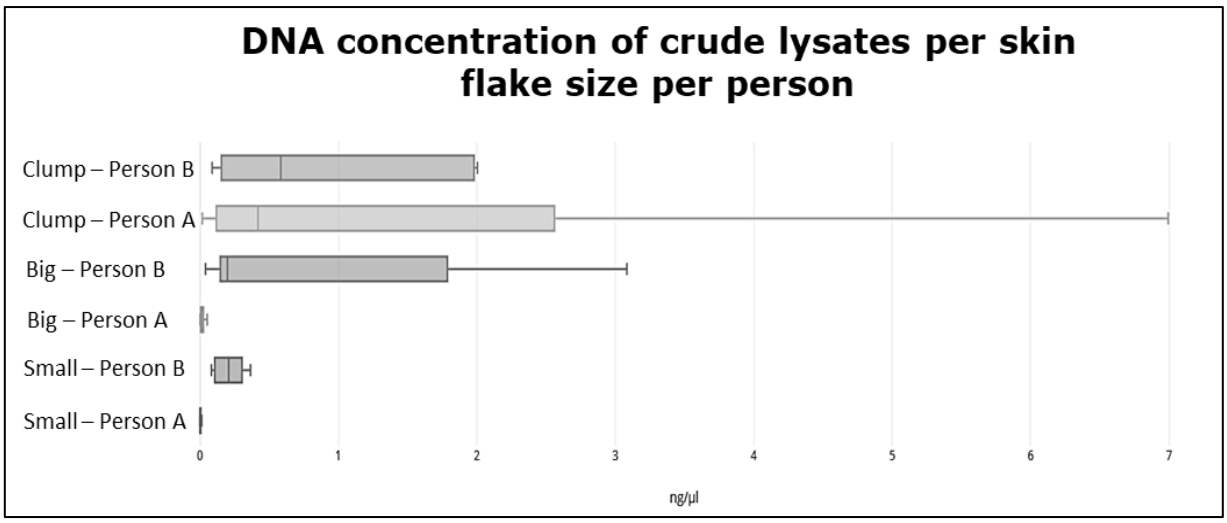


**Figure 3.2.3.3. Boxplot showing gDNA concentration in ng/µl of small, big and clump like skin flakes derived from two persons (Person A and Person B). For each size the gDNA concentration of five skin flakes was determined per person. The black line within each box represents the median of the gDNA concentration.**

In the skin flakes, the proteins such as keratin are broken down, denatured and subsequently dissolved. This in turn releases the DNA which was captured by the keratin matrix. As previously discussed, keratin precipitates after neutralisation. This means that proteins are still undigested and are present in the lysate. However, the presence of these proteins appears not to inhibit subsequent PCR, since the present results of measuring DNA concentration show DNA has been detected in many single skin flakes lysates.

### 3.2.4. Performance of direct PCR on single skin flakes

In order to prevent contamination and minimum loss of DNA, a direct PCR of the skin flakes' crude lysates was chosen rather than a DNA purification-based approach. PCR was carried out by adding 1µl of the skin flakes' crude lysates in a 5µl total reaction volume ESX17 multiplex PCR. 1µl PCR product was analysed by CE. Allele recovery was calculated by calculating the percentage of alleles recovered from the skin flakes in comparison with the reference profile obtained from the donor's extracted mouth swab gDNA.

Results on DNA typing of single skin flakes by direct PCR was found to be successful since DNA profiles were obtained. However, changes in allele recovery among skin flakes with different morphology were observed (*figure 3.2.4.1*). On average the bigger skin flakes resulted in higher allele recovery than the smaller skin flakes. However, the number of alleles that were recovered from clumpy skin flakes was even higher than from the bigger skin flakes.

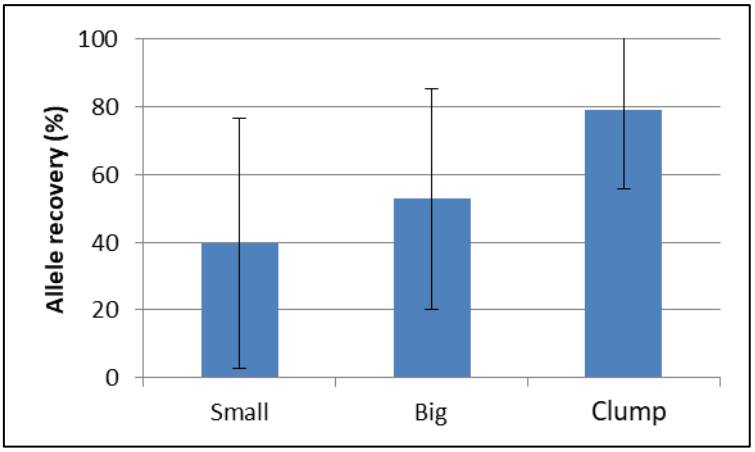


**Figure 3.2.4.1. Differences in allele recovery of autosomal STRs of different sizes/morphological types (small, big and clumps) of skin flakes. Allele recovery is presented in %. Bars represent the average allele recovery. Error bars indicate the standard deviation.**

The largest difference in allele recovery was seen between two donors (figure 3.2.4.2). The same trend was seen in between different sizes with person A. However, the average allele recovery of different sizes of skin flakes derived from person B were similar (around 80%).

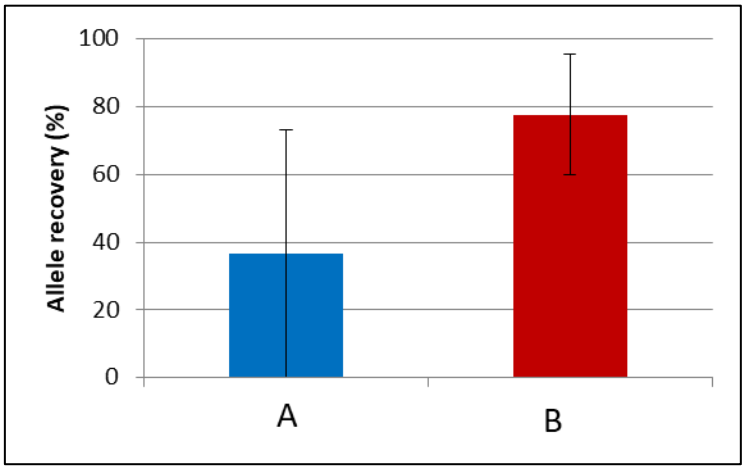


**Figure 3.2.4.2. The effect of person A (A) and person B (B) on the allele recovery of skin flakes. Allele recovery is presented in %. Bars represent the average allele recovery. Error bars indicate the standard deviation.**

The rate of ADIs was very low: One ADI was observed in only one out of all 60 skin flakes analysed.

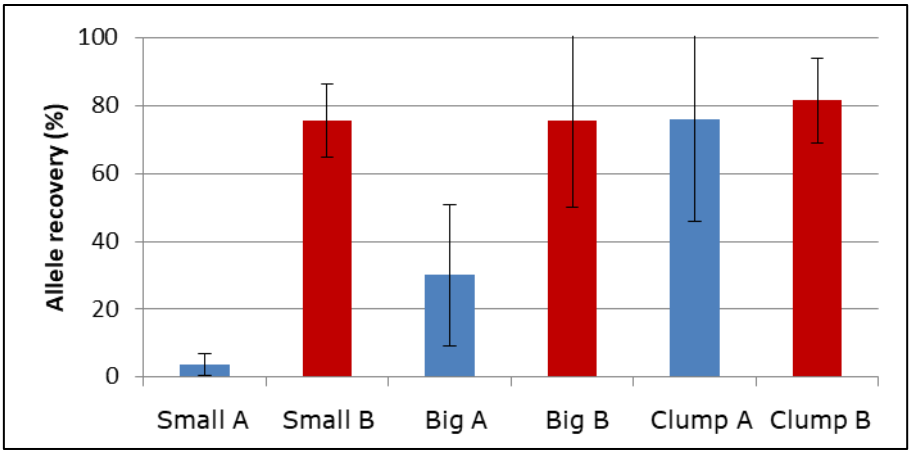


**Figure 3.2.4.3. The effect of two different persons (A = person A, B = Person B) and different sizes (small, big and clump) of skin flakes on allele recovery. Allele recovery is presented in %. Bars represent the average allele recovery. Error bars indicate the standard deviation.**

Altogether, results showed DNA profiling of single skin flakes was possible by using the method DNA extraction without DNA purification described in chapter 3.2.3. DNA can adhere to the walls of the tubes and pipette tips [16]. That is why pipetting steps and a step of transferring the extracted DNA to another tube (often included in DNA purification protocols) can cause a loss of DNA. The method presented here enables DNA profiling of single skin flakes without purification of the DNA from the lysate or any other steps of pipetting/transferring. This keeps the loss of DNA to a minimum.

### 3.2.5. Performance of WGA on single skin flakes

REPLI-g was applied on DNA from skin flakes to investigate whether WGA could improve the allele recovery of skin flakes that had not been successfully typed with the direct PCR method.

Initial results did not show any improvement at all. Strikingly, no allele peaks were displayed in the generated electropherograms. It was hypothesized the amount of salts present in the crude lysates might have led to an inhibition of the WGA reaction. According to the instruction's manual, the REPLI-g Single Cell WGA should not be performed with alkaline and/or neutralization buffers or DTT that are higher in concentration than the original protocol. Additionally, Tween20 was also present. However, Tween20 is often included in the PCR master mixes and so, its presence was not expected to inhibit the WGA reaction.

To determine whether the hypothesis of higher salt concentration to be inhibitory to the WGA reaction, an experiment was executed. gDNA with a concentration high enough, 100pg (the equivalent of approximately 16 somatic cells), as an effective template was added to the reaction tube. Also, the effect on the presence of the lysis buffer that came with the REPLI-g kit was tested. Although the controls were positive, no allele peaks were obtained in the electropherograms generated from experimental samples. These findings support the above hypothesis. Thus, it can be concluded that the higher salt concentrations from the lysis buffer is likely cause of the inhibition of the WGA reaction. Furthermore, WGA by using REPLI-g was not suitable with single skin flakes in the workflow used in this thesis. This does not rule out the possibility of successful WGA after removal of the salts and could be tested on single skin flakes in future research. A possible strategy might be to clean up the crude lysate from salts with a silica column or magnetic beads-based DNA purification. However, the added pipetting and washing steps may result in the loss of the already low DNA content. This can be expected when using silica spin columns which can result into more than 20% of DNA loss [111]. This can subsequently

lead to losing all copies of a certain allele or marker which results in ADOs following WGA.

### 3.2.6. DNA profiling of single skin flakes from mock DNA samples

Highlighting and DNA profiling of skin flakes on tape liftings, as the approach used in this thesis, was first performed on pillows at which DNA was deposited on (chapter 3.2.6.1). This type of DNA evidence was produced by having 4 persons, of which each had laid with their head down on a single pillow for 30 minutes as naturally as possible. Afterwards, tape liftings were taken from each pillow and stained to highlight any skin flakes recovered from the pillows. Highlighted single skin flakes were treated with the alkaline DNA extraction protocol and crude lysates where subsequently used in direct PCR for STR amplification. Allele recovery and ADI were determined by comparing the skin flakes profiles with the reference profiles of participating persons in the experiment.

As part of the validation process, the approach used above was tested on mock mixed DNA evidence (chapter 3.2.6.2.). For this, another set of 4 pillows with deposited DNA was used. In contrast to the first set of pillows, the second set included pillows where multiple persons had laid down on (see table 3.2.6.2.1) and was prepared on a different day as the first set was prepared on. Other than allele recovery and number of ADIs, the number of persons that had laid on the pillow was also determined by applying the single skin flake analysis approach.

### *3.2.6.1. Skin flakes from pillows used by one person*

Skin flakes were successfully highlighted with Orange G by staining the tape liftings taken from the pillows where one person had laid down on. (*figure 3.2.6.1.1*).

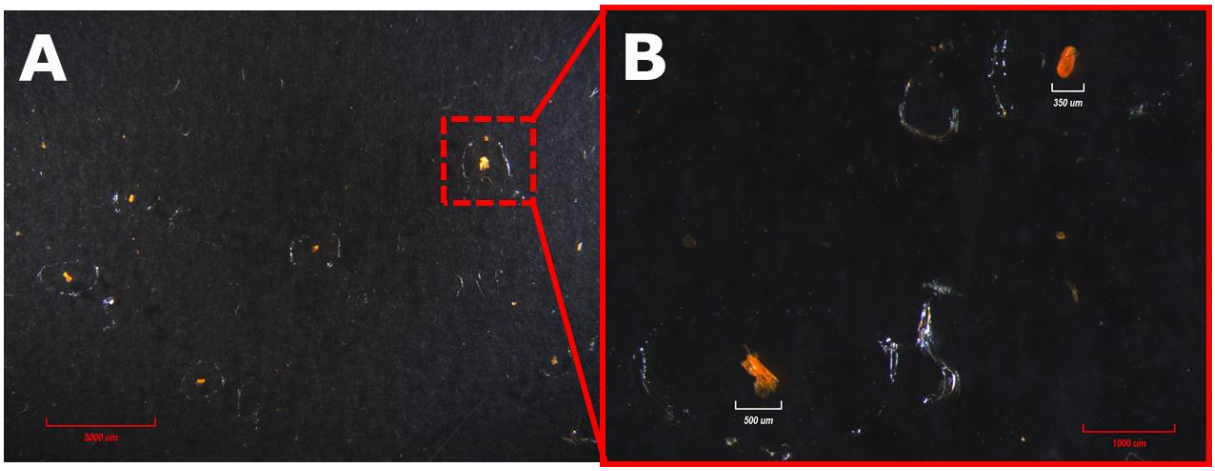


**Figure 3.2.6.1.1. Microscopic views of skin flakes on a tape lifting recovered from a pillow which person C had laid down on. (A) Multiple skin flakes could be highlighted after tape liftings were stained with Orange G (60x magnification). (B) Regular flat skin flakes as well as clump like skin flakes were found (25x magnification).**

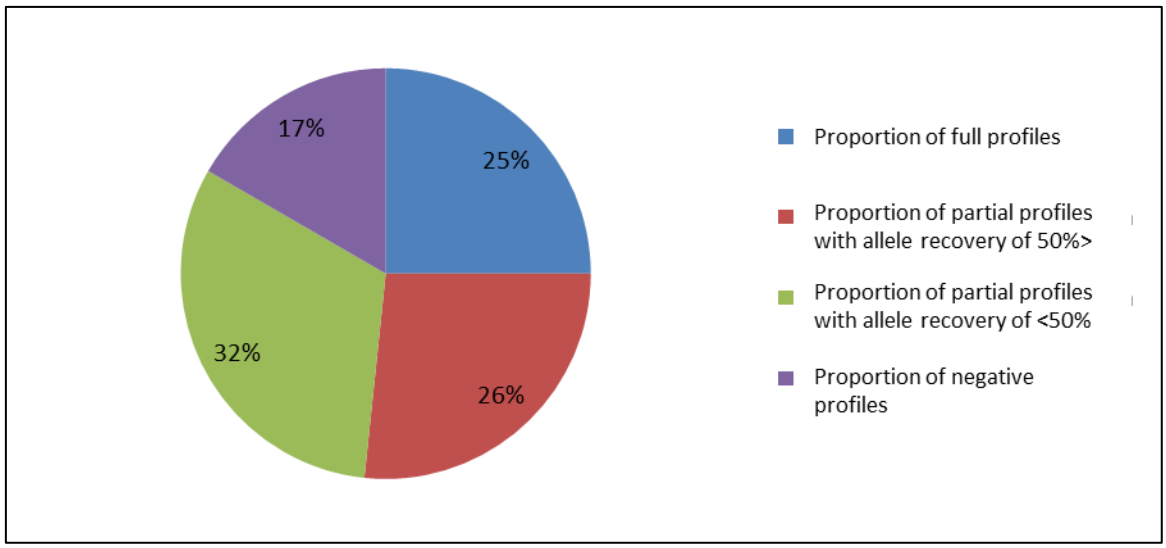


**Figure 3.2.6.1.2. Profile quality of single skin flakes recovered from pillows which were laid down on by a single person.**

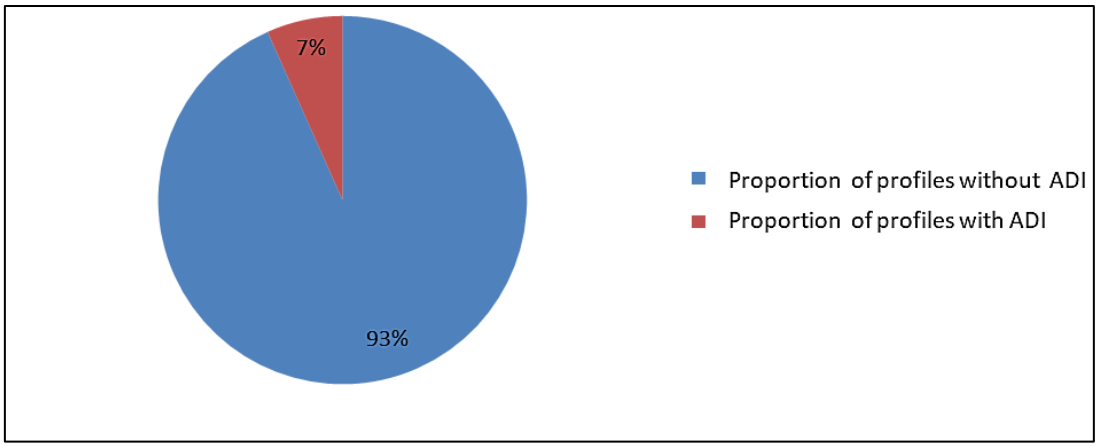


**Figure 3.2.6.1.3. Proportion of clean profiles and profiles with ADI generated from single skin flakes recovered from pillows which were laid down on by a single person.**

DNA typing of single skin flakes recovered from pillows was successfully performed. 58% of all skin flakes generated partial profiles, of which almost half still showed an allele recovery of more than 50% (*figure 3.2.6.1.2*). In 17% of all skin flakes no alleles were detected. ADIs were observed but were with only 7% of profiles containing ADI quite rare (*figure 3.2.6.1.3).*

In a previous study on DNA typing of single skin flakes a full profile success rate of around 5% was reported [37]. In contrast, in 25% of all skin flakes analyzed in the present work, full profiles were obtained (*figure 3.2.6.1.2*). This makes the protocol of single skin flakes analysis described here significantly more effective than current available methods. The skin flakes specific staining protocol used in this study decreases the chance of isolation of other particles that morphologically resemble skin flakes and subsequently yield negative profiles. Furthermore, currently available protocols include DNA extraction with many pipetting and washing steps hence increases the chance of losing DNA. Other protocols use a Chelex based protocol or include DNA extraction in the PCR mastermix [33,35,37,41]. This may work for many types of cells or tissue. However, the keratin in skin flakes makes them difficult to lyse. The protocol used in this work includes a simple lysis step specifically designed to release DNA

from high keratin containing samples and requires no pipetting/transferring or washing steps. Altogether, this decreases the chance of negative DNA profiles due to losing DNA and/or inefficient lysis of skin flakes.

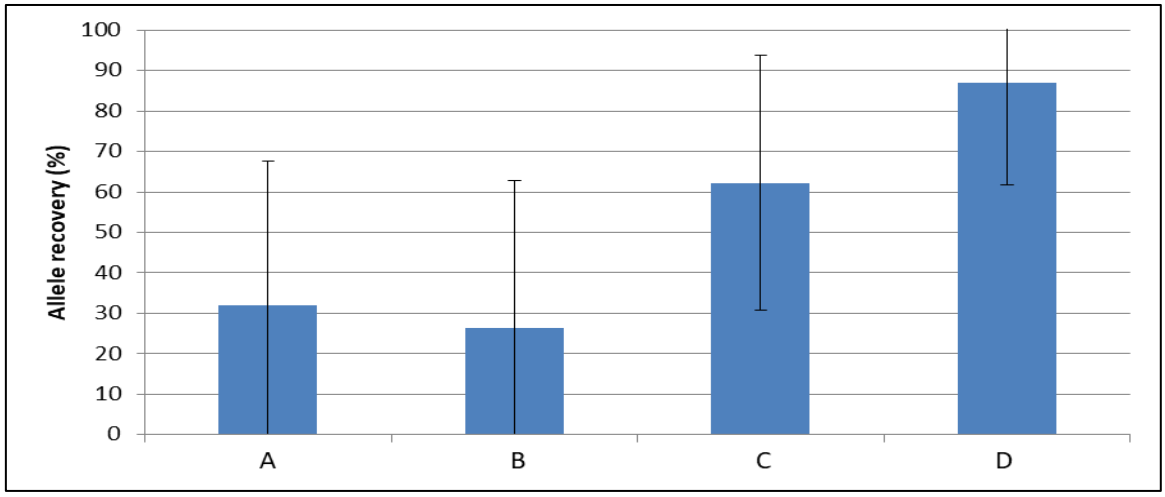


**Figure 3.2.6.1.4. Person specific average allele recovery of single skin flakes recovered from pillows which were each laid down on by a single persons (person A, person B, person C and person C). Allele recovery is presented in %. Bars represent the average allele recovery per person (A = person A, B = person B, C = person C, D = person D). Error bars indicate the standard deviation.**

As seen in previous experiments described in the present work, the average allele recovery of single skin flakes differed also within this experiment between donors (*figure 3.2.6.1.4*). Several skin flakes generated a full profile. These were all derived from Person D, whose skin flakes generated also the highest allele recovery on average of all four donors. A full DNA profile of a single skin flake from person D can be seen in appendix III. The cause of this difference is unclear. The only difference which correlates to the findings is age. The older the donors, the higher the average allele recovery was observed in this study. To truly conclude whether age is a significant factor in explaining the difference in allele recovery, skin flakes from more donors of different ages need to be analysed.

### *3.2.6.2. Skin flakes from pillows used by multiple persons*

Profiles were successfully obtained from 72% of all typed skin flakes derived from pillows which were touched by multiple persons. Of these positive profiles, only 15% were classified as mixed profiles. 35 of 57 typed skin flakes generated a clear single profile. (see figure 3.2.6.2.1 and figure 3.2.6.2.2). In total profiles showed on average 61% (SD 35%) allele recovery.

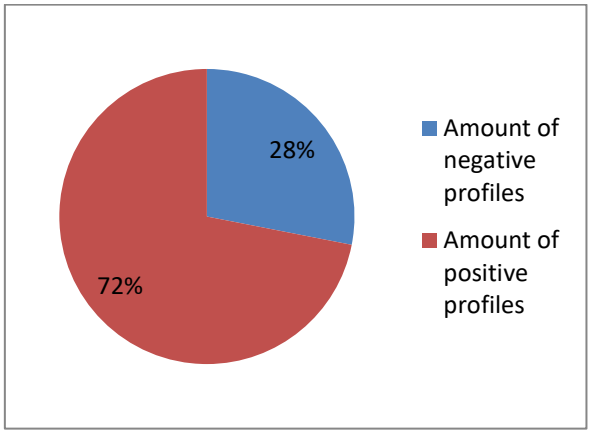


**Figure 3.2.6.2.1. Pie chart visualising the proportion of positive and negative DNA profiles obtained from single skin flakes recovered from 4 pillows at which multiple persons had laid down on.**

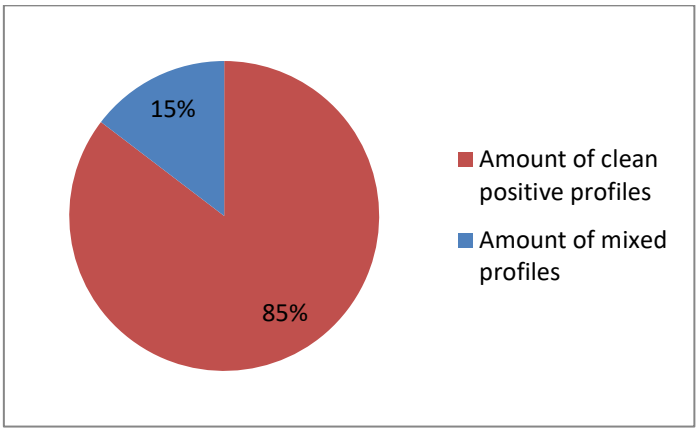


**Figure 3.2.6.2.2. Pie chart visualising the proportion of positive profiles (DNA profiles showing alleles and were generated from single skin flakes recovered from pillows where multiple people had laid down on) which were clean (no mixture or ADI) and the proportion of positive profiles which showed a mixture.**

As in the experiments with single donors (chapter 3.2.6.1), the average allele recovery between person A, B, C and D did significantly differ (P = 1.41E-05). Skin flakes from Person D generated profiles with on average a much higher allele recovery than all other persons within this study.

For all the pillows that were touched by different persons, all persons that touched the pillows were detected. The number of times each person was detected differed from pillow to pillow (*figure 3.2.6.2.3*). However, it was noticed that for each pillow the person who was detected the most was the last person who had touched the pillow (*table 3.2.6.2.1*). One explanation for this observation could be that DNA was picked up by subsequent contacts, thus lowering their number of skin flakes. This explanation is further supported by previously described study that reported friction to increase transfer [112].

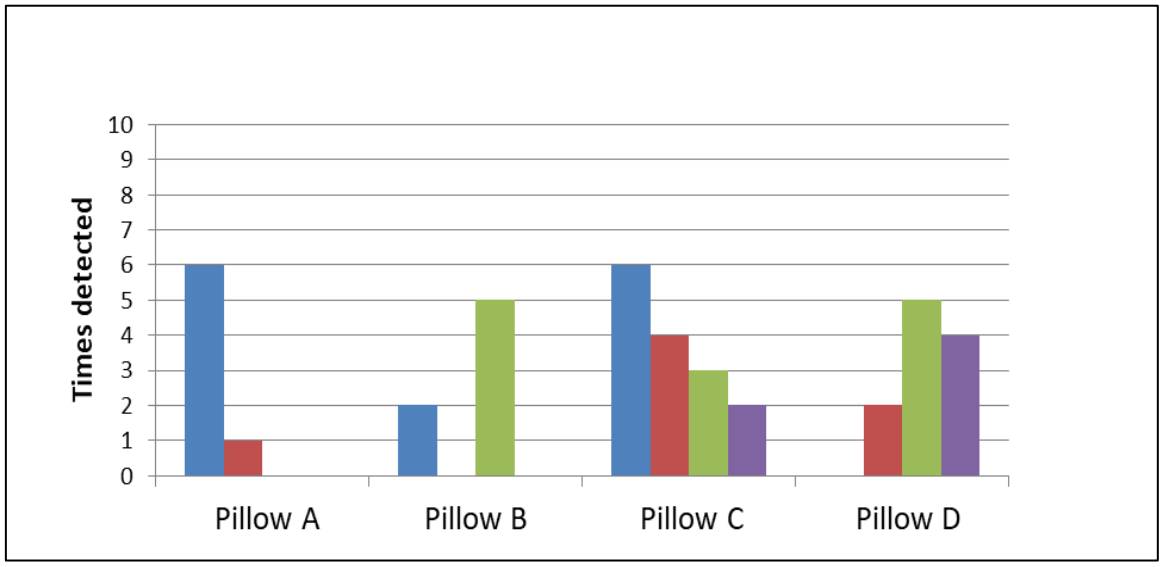


**Figure 3.2.6.2.3.Overview on the rate of detection of different persons on each pillow by DNA-typing of single skin flakes. Four pillows (pillow A, pillow B, pillow C, pillow D) at which 2-3 persons laid down on (person A, person B, person C, person D) were used to recover skin flakes. The different pillows are shown at the X-axis. The blue (person A), red (person B), green (person C) and purple bars (person D) represent the number of times the person's DNA profile was detected on each pillow. The number of times of detection is set on the Y-axis.**

**Table 3.2.6.2.1. Order by which different persons laid down after each other on each pillow.**

| Pillow | Persons and order of laying down |
|---|---|
| A | B -> A |
| B | A -> C |
| C | C -> D -> A |
| D | D -> B -> C |

Furthermore, a fourth person, person B, was detected on pillow C (*figure 3.2.6.2.3*). This was surprising, since person B did not lie down on pillow C (*table 3.2.6.2.1*). On the other hand, indirect transfer could have taken place which caused the detection of person B on pillow C. Skin flakes from person B could have been transported by person A from pillow A to pillow C. Such transportation of skin flakes would be a form of indirect transfer. Direct as well as indirect transfer of touch DNA or trace DNA has also been described in literature [113].

### 3.2.7. The effect of quantity and quality of DNA on allele recovery of single skin flakes

DNA typing of skin flakes has been successfully performed [35,37,114]. However, Schneider et al (2011) showed the rate of obtaining full profiles of single skin flakes is low [37]. In a previous study it was concluded that the low rate was mainly due to not enough gDNA present in skin flakes. However, the integrity/quality of the gDNA can also be contributory factor for the large number of incomplete profiles. This can be seen in figure 3.2.7.1 where a part of the electropherogram from a single skin flake is showing the effect of degraded DNA.

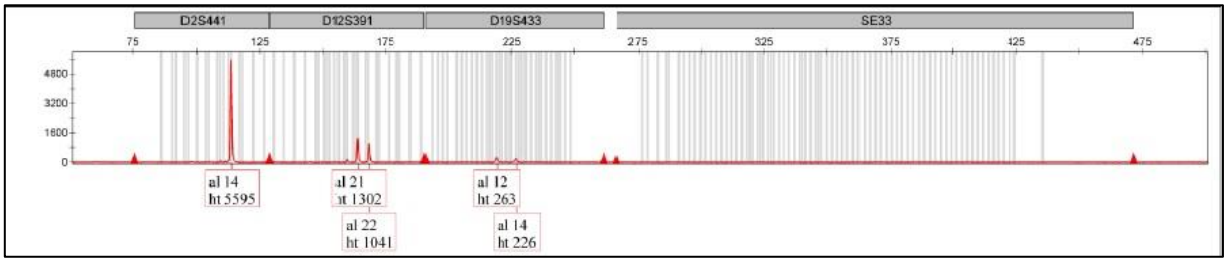


**Figure 3.2.7.1. Part of an electropherogram from a single skin flake showing the effect of DNA degradation on allele recovery. The height of the allele peaks drops as the size increases. Peaks represent the amplicons of different alleles amplified using multiplex PCR and detected by capillary electrophoresis. The relative fluorescent units (RFU) are set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. On top, the X-axis represents the fragment size in base pairs (bp) with the allele corresponding STR locus name above in the grey horizontal bars. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected.**

As degraded DNA contains less high molecular DNA the chance to amplify loci decreases with amplicon size. As a result the height of the peaks decreases with the fragment size and large alleles are starting to drop out as seen with both alleles from SE33. Although Schneider et al (2011) [37] did obtain more full profiles with the mini-STR multiplex PCR kit, the influence of the integrity of the DNA from skin flakes on the allele recovery was not investigated yet. To determine whether both the quantity or quality or a combination of both affected the allele recovery, the amount of gDNA and quality of gDNA of multiple single skin flakes was determined by using an ALU based qPCR method. Furthermore,

the allele recovery was determined from each skin flake by using PowerPlex ESX17 and capillary electrophoresis. Correlations were calculated to determine whether the amount of DNA and/or the level of DNA degradation had any influence on allele recovery.

A moderate positive correlation (Pearson's correlation = 0.53) was determined between the quantity of DNA and allele recovery. In figure 3.2.7.2 a logarithmic trend can be clearly seen; showing increased allele recovery with increasing amount of DNA. This observation agrees with similar results reported by Schneider et al. (2011) [37].

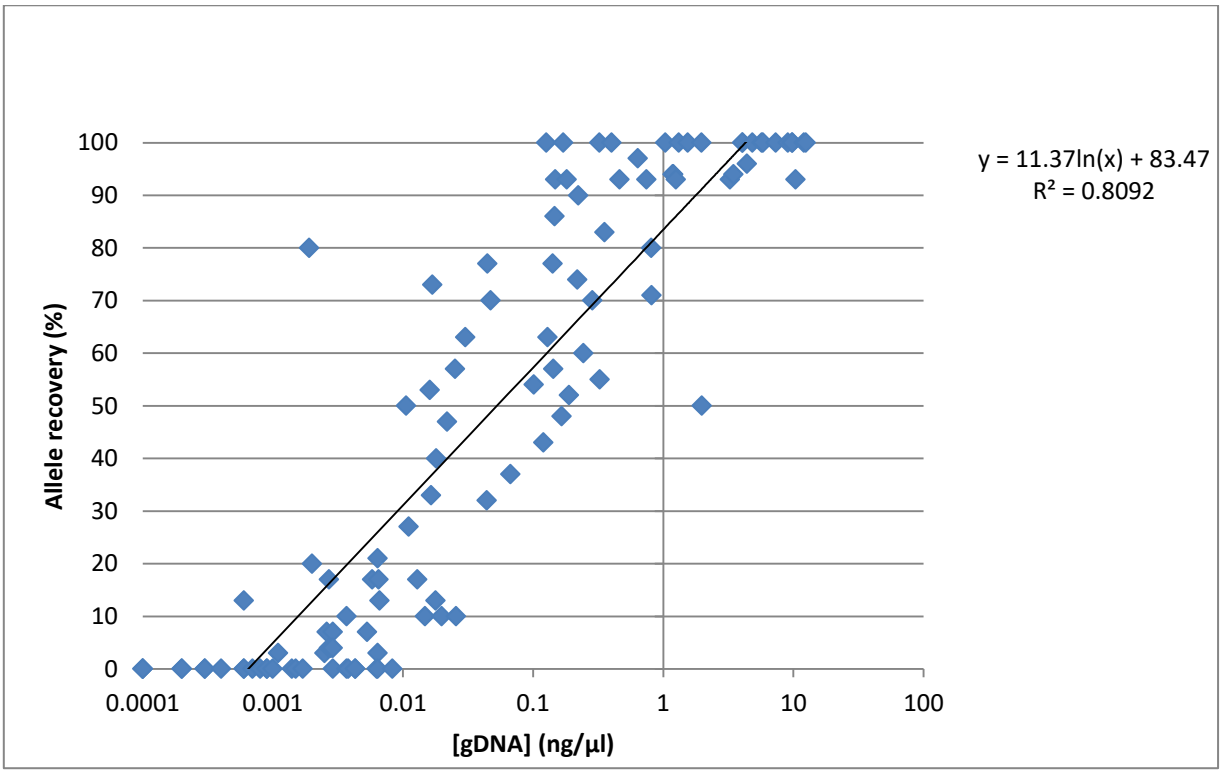


**Figure 3.2.7.2. Scatter plot of showing the relationship between the amount of gDNA in ng/µl (set on the X-axis) and allele recovery in % (set on the Y-axis) of the lysates of single skin flakes. Each data point (rectangles) represents a single skin flake.**

DNA degradation was also determined. The level of DNA degradation is specified by the degradation index and was calculated by dividing the DNA concentration measured with the small ALU amplicon through the DNA concentration measured

with the big ALU amplicon which gets degraded first because of its size. A degradation index of 1 means no degradation whereas a degradation index of 1› means degraded DNA. The higher the degradation index, the more the DNA is degraded. There was only a very weak positive correlation between the degree of degradation and the allele recovery (Pearson's correlation = 0.08).
As the level of DNA degradation gets more complicated with decreasing amounts of DNA, a second correlation calculation was performed by only including data from samples with a concentration of 0,100ng/µl (approximately 16 somatic cells per µl) or more (*table 3.2.7.1*). This showed a moderate negative correlation between degradation level and allele recovery (Pearson's correlation = -0.30). On the other hand, the correlation between concentration and allele recovery dropped slightly (Pearson's correlation = 0.45). This is probably due to at some concentration of DNA the allele recovery will not increase any further. For this reason many forensic multiplex STR kits recommend a template DNA amount of 0.5ng or more per PCR to guarantee full DNA profiles [115].

Because results had shown both quantity and quality of gDNA correlates to allele recovery, it can be concluded that both factors determined the quality of a DNA profile (not just the amount of gDNA). Thus, a sample with a high concentration of gDNA can give a full profile, but can also result in just a partial profile because of degradation which was for example seen in sample D7 (*table 3.2.7.1*). Here, 0.326ng of gDNA was used as template DNA. The PCR kit used requires a minimum amount of 0.050ng (approximately 8 somatic cells) of intact gDNA to generate a full DNA profile. Even though this minimum amount is exceeded 6-fold, the allele recovery was only 55% due to the high level of degradation (DI = 8.81).

The degradation level is even more important at lower levels of gDNA, since the chance of having at least a few copies of an intact locus is much lower compared to a sample with a higher gDNA concentration. This can be seen when samples R_1 and G_5 are compared (*table 3.2.7.1*). Both samples contain approximately 0.130ng/µl gDNA (approximately 20 somatic cells per µl), but differed in allele

recovery. Sample R_1 generated a degradation level of less than 1 a full profile. Although the degradation level of sample G_5 was not as high as seen in the previously described sample D_7, allele recovery was only 63%. Therefore, it is highly important to avoid losing any gDNA during DNA extraction steps.

**Table 3.2.7.1. Heat chart of single skin flakes with more than 0.100ng/µl in 8µl crude lysate with corresponding degradation index and sorted on gDNA concentration in ng/µl.**

| Sample | Conc (ng/µl) | Degradation-Index | Allele recovery (%) |
|---|---|---|---|
| J_14 | 0.102 | 3.05 | 54 |
| C_13 | 0.120 | 3.49 | 43 |
| R_1 | 0.126 | 0.89 | 100 |
| G_5 | 0.129 | 1.98 | 63 |
| G_11 | 0.141 | 1.32 | 77 |
| C_4 | 0.143 | 7.29 | 57 |
| J_1 | 0.147 | 2.53 | 86 |
| J_15 | 0.148 | 0.91 | 93 |
| R_7 | 0.166 | 9.44 | 48 |
| B_1 | 0.170 | 3.43 | 100 |
| D_13 | 0.181 | 2.43 | 93 |
| D_1 | 0.189 | 10.43 | 52 |
| R_10 | 0.220 | 2.30 | 74 |
| C_5 | 0.223 | 9.02 | 90 |
| D_14 | 0.244 | 11.18 | 60 |
| G_10 | 0.285 | 2.14 | 70 |
| N_15 | 0.324 | 3.20 | 100 |
| D_7 | 0.326 | 8.81 | 55 |
| B_4 | 0.355 | 1.33 | 83 |
| G_2 | 0.403 | 4.51 | 100 |
| N_2 | 0.462 | 0.41 | 93 |
| A_2 | 0.639 | 3.72 | 97 |
| D_4 | 0.746 | 2.24 | 93 |
| C_10 | 0.811 | 8.44 | 80 |
| A_3 | 0.814 | 8.33 | 71 |
| R_8 | 1.041 | 0.50 | 100 |
| D_15 | 1.185 | 1.82 | 94 |
| R_13 | 1.198 | 1.11 | 94 |
| G_9 | 1.255 | 2.01 | 93 |
| R_3 | 1.321 | 1.73 | 100 |
| J_7 | 1.547 | 1.42 | 100 |
| R_2 | 1.975 | 3.56 | 100 |
| D_5 | 1.987 | 5.19 | 50 |
| N_13 | 3.256 | 1.19 | 93 |
| D_6 | 3.476 | 13.83 | 94 |
| R_9 | 4.066 | 0.69 | 100 |
| N_12 | 4.076 | 2.14 | 100 |
| C_9 | 4.409 | 7.64 | 96 |
| C_3 | 4.845 | 22.75 | 100 |
| R_14 | 5.707 | 1.52 | 100 |
| R_11 | 5.838 | 1.36 | 100 |
| D_11 | 7.335 | 2.15 | 100 |
| C_6 | 9.107 | 2.95 | 100 |
| R_5 | 9.808 | 1.85 | 100 |
| R_12 | 9.817 | 1.27 | 100 |
| B_2 | 10.411 | 4.46 | 93 |
| N_14 | 12.087 | 0.74 | 100 |
| R_4 | 12.462 | 1.08 | 100 |

### 3.2.8. DNA profiling of single skin flakes with mini-STRs and NGS

In order to determine if a mini-STR NGS based approach would improve DNA profiling of single skin flakes, conventional PCR combined with CE and mini-STR based NGS were tested on single skin flakes. Allele recovery of 30 single skin flakes with different types of morphology from two persons were determined using a NGS protocol amplifying all 17 forensic STR loci and compared with the allele recovery obtained with PowerPlex ESX17.

The allele recovery of skin flakes obtained with NGS were all significantly increased (P = 3.87E-06) (paired two sample t-Test) compared to the allele recovery obtained by using PowerPlex ESX17 and CE (*figure 3.2.8.1*).

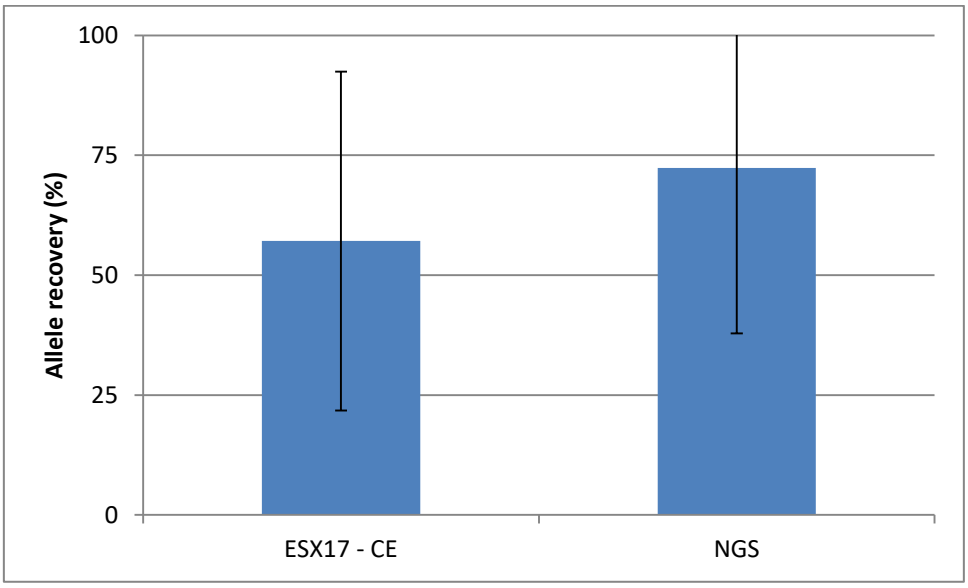


**Figure 3.2.8.1. Allele recovery of skin flakes analysed by using PowerPlex ESX17 with capillary electrophoresis (ESX17-CE) as well as by using mini-STRs with NGS (NGS). Allele recovery in % is set on the Y-axis. The bars represent the average allele recovery for each analytical method used (ESX17-CE or NGS set on the X-axis). Error bars indicate the standard deviation of allele recovery in %.**

The size of the STR marker system had a great influence on allele recovery. This can be seen in *figure 3.2.8.2* in which the capillary electrophoresis elctropherogram and the NGS-"electrophoregram", both generated from the same single skin flake, are compared with eachother. Here, D21S11 a size range of approximately 200bp - 270bp was negative in PowerPlex ESX17. However, the

same marker was positive showing two alleles in the NGS approach above. Also, the ESX17 with CE failed to recover alleles within this sample at markers D18S51, D8S1179 and FGA. Though, the alleles of the latter three markers were as well recovered by using the mini-STR NGS approach.

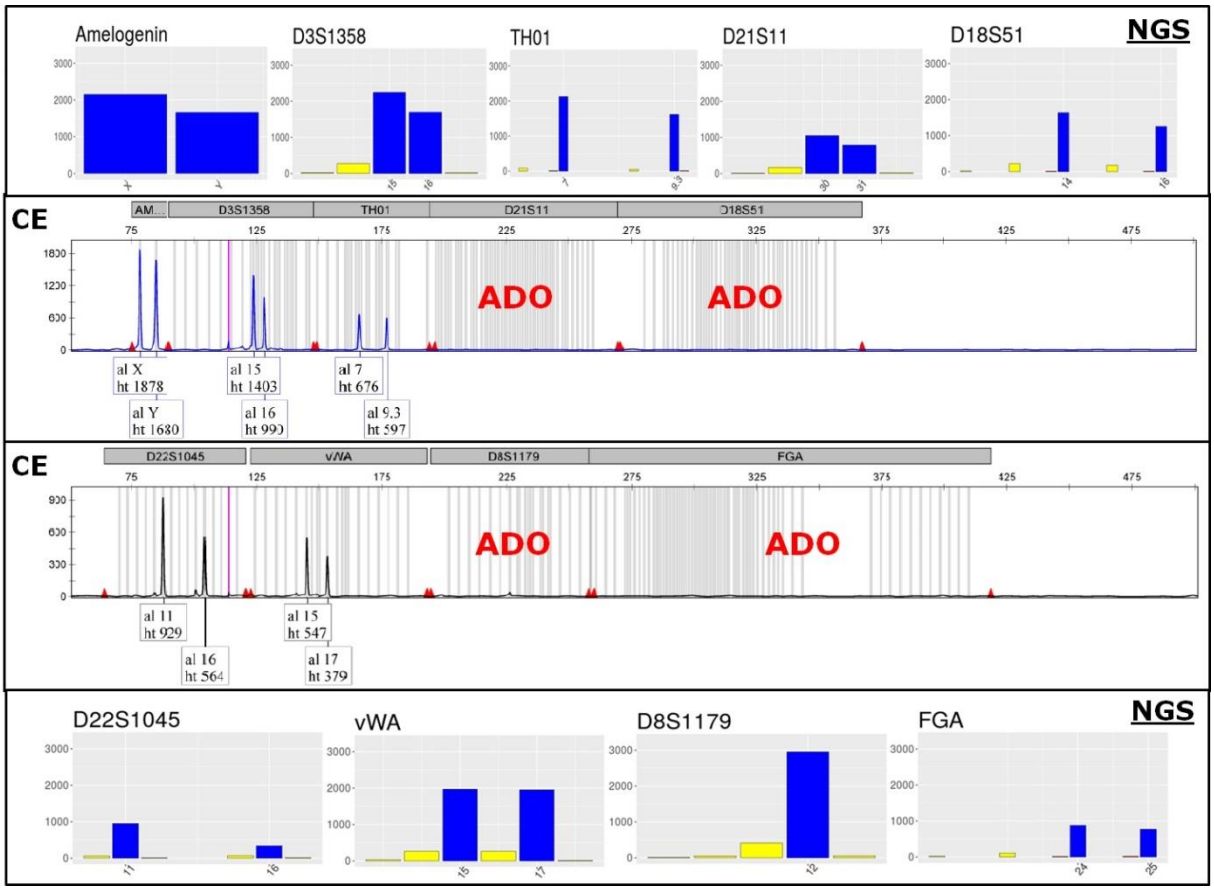


**Figure 3.2.8.2. Capillary electrophoresis electropherograms (CE) of a single skin flake compared with the NGS-"electropherogram" (NGS) of the same skin flake. ADOs (ADO) were observed at the by CE analysed markers D21S11, D18S51, D8S1179 and FGA. At the same markers analysed with NGS no ADOs were observed. In the capillary electrophoresis electropherograms the relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. On top, the X-axis represents the fragment size in base pairs (bp) with the allele corresponding STR locus name above in the grey horizontal bars. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. In the NGS-"electropherograms" the allele peak heights are shown in number of reads, vertically shown on the Y-axis. On the X-axis the allele numbers are shown. The blue bars represent the true alleles. The yellow bars represent stutter or alleles underneath the analytical threshold. Only true alleles (blue bars) were called with the corresponding allele number presented on the X-axis.**

Although, the first results on skin flakes were successful, kit optimization and post-run data analysis still require improvement. As runs were not yet optimally validated, normalisation and the exact threshold settings required could not be determined. Therefore, threshold settings were based on a previously published article with general threshold recommendations [116]. As the threshold settings were set very high, only alleles with a very high number of reads passed the threshold filter and were called a real allele. This caused many other potentially real alleles (based on matching with the reference profile) not being called. When the threshold was set lower, more alleles containing less reads were called. However, this also resulted in many ADIs. A DNA profile of a single skin flake that generated a DNA profile with 100% allele recovery, but with 3 ADIs (not seen when CE was used) is shown in appendix IV.

The results obtained from the above experiments matched the expectations. The smaller the sizes of the STR markers are, the higher the chance alleles can be recovered with degraded DNA as template. As capillary electrophoresis requires STR marker systems with different sizes, the number of small sized STR marker systems is limited. In contrast, NGS enables the use of a complete forensic set of STR marker systems which are all small in size. Therefore, analysing this set of only small STR markers was expected to increase allele recovery of samples containing degraded DNA like those from skin flakes. The findings in the present work suggest a mini-STR marker system such as the one used here combined with NGS can improve forensic DNA typing of skin flakes, thus paving the way for further research in this area.

A further advantage of NGS is that it provides a higher throughput of samples and the opportunity to analyse further sets of markers, such as isoalleles, X-and Y-STRs and SNPs [88,117]. Isoalleles can be of advantage when dealing with mixtures. STR alleles with the same length from two contributors would overlap when analysed with CE and therefore the DNA profiles of the contributors cannot be completely separated from each other. This can be solved when reference DNA profiles from both contributors is available. However, this gets problematic when

only one reference profile is available plus the electropherogram is difficult to interpret, due to allelic imbalances and ADO when dealing with degraded DNA. As with NGS not only the length of the STR allele can be determined, but the DNA sequence itself as well, SNPs within the STR amplicon can be detected. These SNPs can differ between persons. Thus, more "versions" of the same alleles (based on length) exist based on their different SNPs. These other "versions" are called isoalleles. For example, at STR locus TH01 person A and person B are both homozygous containing allele 8. Though, at position 93bp, in the DNA sequence of allele 8, person A contains a G in contrast to person B which contains an A. Therefore, overlapping alleles can be separated by using NGS based on these isoalleles. This can make mixtures with any overlapping alleles easier to analyse. Thus, mixed DNA on tape liftings could be analysed by determining the isoalleles using NGS without the need of isolating single skin flakes. However, when from one allele no two isoalleles are determined, the allele cannot be separated. Also when no reference profile from both contributors is available and the number of reads from the alleles are equally high, then even with two determined isoalleles it is still difficult to establish two DNA profiles from the mixed profile. Thus single source DNA by isolating single skin flakes would be still the best. Also in cases of DNA samples with a high number of contributors each contributing in different proportions can make DNA profiling problematic. This may be due to the fact that the contributors of the DNA evidence which are not related to the crime, could also be present in abundance and therefore can mask the crime related contributor(s). This could result into not detecting the crime related contributor(s). Other SNPs which are included in the Signature Prep kit from Illumina are identity informative SNPs (iiSNPs), phenotypic informative SNPs (piSNPs) and ancestry-informative SNPs (aiSNPs). The iiSNPs are determined by small amplicons (most are <125bp). This makes them ideal for degraded DNA in such a level that STRs cannot be determined. As skin flakes can include highly degraded DNA, the iiSNPs could therefore be used to profile single skin flakes. However, criminal DNA databases are currently only based on STRs, which makes a search with a SNP based DNA profile impossible. With the piSNPs hair and eye colour can be estimated of the person the DNA originated from. Furthermore,

biographical ancestry can be estimated using the aiSNPs. Altogether, using the Illumina NGS platform combined with the Signature Prep kit, much more than just autosomal STRs can be determined from a DNA sample at the same time [117]. Though, the recommended amount of template DNA using the Signature Prep kit is 1ng intact DNA. Previous sensitivity studies have however shown successfully using the Signature Prep kit with less than 1ng. However, when <100pg was used the loci calling rates dropped substantially. Furthermore, the efficiency will also be dependent on the quality of the DNA. This generates the questions whether it is possible to determine both genotypic information as well as phenotypic information from single skin flakes. Therefore, more research is required to determine the performance of the Signature Prep kit using DNA from single skin flakes.

### 3.2.9. Conclusion

Skin flakes can be hard to distinguish from other particles on tape liftings. Though, now, skin flakes can be stained with Orange G making them distinguishable from other particles, for example particles from hair cosmetic products. Furthermore, a DNA extraction method was developed which was compatible with PCR with no DNA purification steps required. As a result keeping DNA sample loss to a minimum. Also difference in DNA quantity was seen between different persons. Results also showed that allele recovery of single skin flakes was not only dependent on the DNA quantity, but also on the DNA integrity. Additionally, analysis with NGS using mini-STRs significantly improved allele recovery in comparison with DNA profiling of single skin flakes by using a multiplex STR PCR kit and CE.

# Chapter 4

# Overall conclusion and future work

# Chapter 4. Overall conclusion and future work

## 4.1. Sperm

By using micromanipulation and WGA as described in this thesis not only enables amplification of autosomal STRs from single spermatozoa but can be used for multiple STR analyses like Y- and X-chromosomal STRs at the same time. By using micromanipulation of single spermatozoa in combination with WGA, mixtures can be circumvented and DNA profiles can be used for criminal DNA database searches, no matter if a single or multiple known or unknown male contributed to the sample.
As all experiments were performed under controlled conditions, the method described here, should therefore be tested on casework samples first to guarantee success in casework. Casework samples could be split in half. One half will go through the conventional process, the other half could then be tested with the single spermatozoa method. Though this could be difficult, as the number of sexual assault cases including multiple male offenders are low compared to sexual assault cases including a single offender.

Furthermore, the single spermatozoa DNA typing method should be tested on some additional mock swabs with more diluted semen to determine the applicability on cases involving swabs with limited amounts of semen.
To prevent DNA degradation as much as possible, semen samples and prepared mock intimate swabs should not be frozen like the ones used in this thesis. Instead, dilutions of semen should be prepared from fresh semen and mock intimate swabs should be stored in a dry and cool place.

The research described in this thesis was conducted in a forensic genetic context. However, the techniques described here could also be applied in other fields of research such as whole genome sequencing of single cells. Single cell analysis is already much conducted in bacterial research and in cancer research where single tumor cells are analysed for determining specific mutations [25,72,118]. Also

single spermatozoa are sequenced to obtain more knowledge on meiotic recombination and infertility in animals [119,120]. For all these single cell DNA sequencing analyses, a pre-amplification of the single genomes is necessary to generate enough template DNA for a (multiplex) PCR required for (whole) genome sequencing. Thus, a whole genome amplification used as described in this thesis could be used to facilitate these other fields of research in which single cells are analysed.

### 4.2. Skin flakes

Skin flakes can be hard to distinguish from other particles on tape liftings. Though, now, skin flakes can be highlighted by Orange G which discriminates them more from other particles for example from particles from hair cosmetic products.
A DNA extraction was designed without the need of any DNA purification steps and enabled for a direct-PCR of the crude lysate. This enabled to keep loss of DNA sample to a minimum.

However, a few aspects should be considered and might require more research in the future of which one aspect is the staining method of choice. Tape used for tape liftings can be made from many different materials. Since staining using Orange G was only performed on one type of tape liftings, the staining should be tested on tape liftings made from other material as well. Tape liftings are also used to isolate fibres for fibre analysis. Since Orange G stains wool [121], the staining should be tested on fibres from animal, plant and synthetic origin as well. Should Orange G stain a certain type of fibre then it should be determined whether it is possible to remove the Orange G without affecting fibre analysis by e.g. a washing step.

As the method of DNA profiling of single skin flakes contains no DNA purification step, the lysate only needs to be neutralized and transferred into the PCR master mix. This was all conducted by manual pipetting. Though, the number of single

skin flakes for case work could be quite high and could lead to a high number of DNA samples. Therefore, the pipetting steps after the lysis should be tested by automated pipetting platforms to integrate the DNA profiling of single skin flakes in a higher throughput environment.

The single skin flake DNA profiling method as described in this thesis could further be tested on other high keratin containing single source sample material such as single telogenic hairs. Similar to skin flakes, the DNA in hairs is also degraded. An alkaline based DNA extraction used in a previous study was successful and better than a grinding and organic based method [122]. However, the method used a DNA purification step instead of adding the crude lysate directly in the PCR. Hairs include melanin which is known to inhibit PCR [123]. For this reason, a direct PCR of the hair's lysate might not function. Though, more research is necessary to determine the performance of the direct PCR approach combined with the alkaline skin flake lysis buffer on single telegenic hairs.

Also more research is required to be able to determine the effect of REPLI-g Single Cell WGA on single skin flakes without being inhibited by the crude lysate. Here, a clean-up is probably necessary. By using low-DNA-binding tubes and tips for purification of the DNA from the crude lysate, the loss of DNA could be reduced.

The in-house mini-STR-NGS based DNA profiling method showed significantly increased allele recovery of single skin flakes compared to the conventional method (PCR and CE). As the NGS based kit from Illumina, the ForenSeq Singature Prep kit, also uses STRs it would be interesting to compare the ForenSeq Signature Prep kit with the in-house NGS method based to determine any differences in allele recovery of autosomal STRs of single skin flakes between these two NGS based methods. Until now, only autosomal STRs were analysed during NGS experiments described in this thesis. The ForenSeq Signature Kit from Illumina enables next to STRs also the analyses of other loci such as X/Y-

STRs, iiSNPs, and phenotypic/ancestry SNPs. Though, more research is required to determine the performance of the Signature Prep kit based on coverage of the additional loci from single skin flake DNA.

# References


[1] J.M. Butler, Genetics and genomics of core short tandem repeat loci used in human identity testing, J. Forensic Sci. 51 (2006) 253–265. https://doi.org/10.1111/j.1556-4029.2006.00046.x.

[2] I. Sibille, C. Duverneuil, G. Lorin de la Grandmaison, K. Guerrouache, F. Teissière, M. Durigon, P. De Mazancourt, Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa, Forensic Sci. Int. 125 (2002) 212–216. https://doi.org/10.1016/S0379-0738(01)00650-8.

[3] J.M. Butler, Chapter 11 - Statistical Interpretation: Evaluating the Strength of Forensic DNA Evidence, in: J.M. Butler (Ed.), Fundam. Forensic DNA Typing, Academic Press, San Diego, 2010: pp. 229–258. https://doi.org/https://doi.org/10.1016/B978-0-12-374999-4.00011-4.

[4] M.L. Gaines, P.W. Wojtkiewicz, J. a Valentine, C.L. Brown, Reduced volume PCR amplification reactions using the AmpFlSTR Profiler Plus kit., J. Forensic Sci. 47 (2002) 1224–37. http://www.ncbi.nlm.nih.gov/pubmed/12455643.

[5] A. Piovesan, M.C. Pelleri, F. Antonaros, P. Strippoli, M. Caracausi, L. Vitale, On the length, weight and GC content of the human genome, BMC Res. Notes. 12 (2019) 1–7. https://doi.org/10.1186/s13104-019-4137-z.

[6] K.S. Grisedale, A. van Daal, Comparison of STR profiling from low template DNA extracts with and without the consensus profiling method, Investig. Genet. 3 (2012) 1. https://doi.org/10.1186/2041-2223-3-14.

[7] G. Axler-DiPerte, S. Orans, a. Singh, T. Caragine, M. Prinz, Z.M. Budimlja, Comparison and optimization of DNA recovery from sperm vs. epithelial cells using laser capture microdissection technology and an immunofluorescent staining system, Forensic Sci. Int. Genet. Suppl. Ser. 3 (2011) e224–e225. https://doi.org/10.1016/j.fsigss.2011.08.112.

[8] D. Taylor, J.A. Bright, J. Buckleton, The interpretation of single source and mixed DNA profiles, Forensic Sci. Int. Genet. 7 (2013) 516–528. https://doi.org/10.1016/j.fsigen.2013.05.011.

[9] T.R. Moretti, R.S. Just, S.C. Kehl, L.E. Willis, J.S. Buckleton, J.A. Bright, D.A. Taylor, A.J. Onorato, Internal validation of STRmix™ for the interpretation of single source and mixed DNA profiles, Forensic Sci. Int. Genet. 29 (2017) 126–144. https://doi.org/10.1016/j.fsigen.2017.04.004.

[10] J. Pereira, R. Neves, S. Forat, W. Huckenbeck, K. Olek, MtDNA typing of single-sperm cells isolated by micromanipulation, Forensic Sci. Int. Genet. 6 (2012) 228–235. https://doi.org/10.1016/j.fsigen.2011.05.005.

[11] K.M. Horsman, S.L.R. Barker, J.P. Ferrance, K. a. Forrest, K. a. Koen, J.P.

Landers, Separation of sperm and epithelial cells in a microfabricated device: Potential application to forensic analysis of sexual assault evidence, Anal. Chem. 77 (2005) 742–749. https://doi.org/10.1021/ac0486239.

[12] J.M. Butler, DNA Extraction Methods Collection/Storage, Adv. Top. Forensic DNA Typing Methodol. (2012) 29–47. https://doi.org/10.1016/B978-0-12-374513-2.00002-6.

[13] H.G. Wong, J.S. Mihalovich, G. Sensabaugh, Automating the Differential Digestion Method in the Analysis of Sexual Assault Cases using Selective Degredation, (2013) 1–45.

[14] K. Yoshida, K. Sekiguchi, N. Mizuno, K. Kasai, I. Sakai, H. Sato, S. Seta, The modified method of two-step differential extraction of sperm and vaginal epithelial cell DNA from vaginal fluid mixed with semen, Forensic Sci. Int. 72 (1995) 25–33. https://doi.org/10.1016/0379-0738(94)01668-U.

[15] I. Quinones, B. Daniel, Forensic Science International : Genetics Cell free DNA as a component of forensic evidence recovered from touched surfaces, Forensic Sci. Int. Genet. 6 (2012) 26–30. https://doi.org/10.1016/j.fsigen.2011.01.004.

[16] M.D. Timken, S.B. Klein, M.R. Buoncristiani, Improving the efficacy of the standard DNA differential extraction method for sexual assault evidence, Forensic Sci. Int. Genet. 34 (2018) 170–177. https://doi.org/10.1016/j.fsigen.2018.02.014.

[17] I. Findlay, A. Taylor, P. Quirke, DNA fingerprinting from single cells Archaeopteryx-like skull in Enantiornithine bird, Nature. (1997) 555–556.

[18] R.A.H. van Oorschot, K.N. Ballantyne, R.J. Mitchell, Forensic trace DNA: A review, Investig. Genet. 1 (2010) 14. https://doi.org/10.1186/2041-2223-1-14.

[19] J.P. Allery, N. Telmon, R. Mieusset, a Blanc, D. Rougé, Cytological detection of spermatozoa: comparison of three staining methods., J. Forensic Sci. 46 (2001) 349–351.

[20] K. Anslinger, B. Mack, B. Bayer, Fluorescent labeling and isolation of male cells, Int. Congr. Ser. 1288 (2006) 613–615. https://doi.org/10.1016/j.ics.2005.09.016.

[21] S. Yilmaz, M.F. Haroon, B.A. Rabkin, G.W. Tyson, P. Hugenholtz, Fixation-free fluorescence in situ hybridization for targeted enrichment of microbial populations, ISME J. 4 (2010) 1352–1356. https://doi.org/10.1038/ismej.2010.73.

[22] C. McAlister, The use of fluorescence in situ hybridisation and laser microdissection to identify and isolate male cells in an azoospermic sexual assault case, Forensic Sci. Int. Genet. 5 (2011) 69–73.

https://doi.org/10.1016/j.fsigen.2010.04.008.

[23] K.W.P. Miller, J. Old, B.R. Fischer, B. Schweers, S. Stipinaite, K. Reich, Developmental validation of the SPERM HY-LITER TM Kit for the identification of human spermatozoa in forensic samples, J. Forensic Sci. 56 (2011) 853–865. https://doi.org/10.1111/j.1556-4029.2011.01796.x.

[24] Y.S. Ping, X.L.S. Chan, S.K. Goh, C.K.C. Syn, Optimization of spermatozoa detection using immunofluorescent staining and laser micro-dissection, Forensic Sci. Int. (2015). https://doi.org/10.1016/j.forsciint.2015.07.010.

[25] P.C. Blainey, The future is now: Single-cell genomics of bacteria and archaea, FEMS Microbiol. Rev. 37 (2013) 407–427. https://doi.org/10.1111/1574-6976.12015.

[26] T.C. Shih, K.H. Chu, C.H. Liu, A programmable biochip for the applications of trapping and adaptive multisorting using dielectrophoresis array, J. Microelectromechanical Syst. 16 (2007) 816–825. https://doi.org/10.1109/JMEMS.2007.897558.

[27] K.M. Horsman, J.M. Bienvenue, K.R. Blasier, J.P. Landers, Forensic DNA analysis on microfluidic devices: A review, J. Forensic Sci. 52 (2007) 784–799. https://doi.org/10.1111/j.1556-4029.2007.00468.x.

[28] C.X. Li, J.P. Han, W.Y. Ren, A.Q. Ji, X.L. Xu, L. Hu, DNA profiling of spermatozoa by laser capture microdissection and low volume-PCR, PLoS One. 6 (2011) 1–7. https://doi.org/10.1371/journal.pone.0022316.

[29] T. Miyazaki, M. Hara, A. Ichiki, Y. Yamamoto, A. Takada, An efficient novel method for analyzing STR loci from a single sperm captured by laser microdissection, 1 (2008) 437–438. https://doi.org/10.1016/j.fsigss.2007.10.074.

[30] T. Geng, R. Novak, R. Mathies, Single-Cell Forensic Short Tandem Repeat Typing within Microfluidic Droplets, Anal. Chem. (2013). http://pubs.acs.org/doi/abs/10.1021/ac403137h.

[31] J.M. Bienvenue, N. Duncalf, D. Marchiarullo, J.P. Ferrance, J.P. Landers, Microchip-based cell lysis and DNA extraction from sperm cells for application to forensic analysis, J. Forensic Sci. 51 (2006) 266–273. https://doi.org/10.1111/j.1556-4029.2006.00054.x.

[32] B. Polzer, G. Medoro, S. Pasch, F. Fontana, L. Zorzino, A. Pestka, U. Andergassen, F. Meier-Stiegen, Z.T. Czyz, B. Alberter, S. Treitschke, T. Schamberger, M. Sergio, G. Bregola, A. Doffini, S. Gianni, A. Calanca, G. Signorini, C. Bolognesi, A. Hartmann, P.A. Fasching, M.T. Sandri, B. Rack, T. Fehm, G. Giorgini, N. Manaresi, C.A. Klein, Molecular profiling of single circulating tumor cells with diagnostic intention., EMBO Mol. Med. 6 (2014) 1371–86. https://doi.org/10.15252/emmm.201404033.

[33] K. Farash, E.K. Hanson, J. Ballantyne, Enhanced Genetic Analysis of Single

Human Bioparticles ... Materials, J. Vis. Exp. (2015) 1–2. https://doi.org/10.3791/52612.

[34] B. Herber, K. Herold, DNA typing of human dandruff, J. Forensic Sci. 43 (1998) 648—656. http://europepmc.org/abstract/MED/9608703.

[35] M. Lorente, C. Entrala, J. a Lorente, J.C. Alvarez, E. Villanueva, B. Budowle, Dandruff as a potential source of DNA in forensic casework., J. Forensic Sci. 43 (1998) 901–902. https://doi.org/10.1016/S1353-1131(99)90191-8.

[36] L. Winkler, DNA typing of shed epithelial cells, (2011).

[37] H. Schneider, T. Sommerer, S. Rand, P. Wiegand, Hot flakes in cold cases, Int. J. Legal Med. 125 (2011) 543–548. https://doi.org/10.1007/s00414-011-0548-7.

[38] S.I. Kim, H.J. Kim, H.J. Lee, K. Lee, D. Hong, H. Lim, K. Cho, N. Jung, Y.W. Yi, Application of a non-hazardous vital dye for cell counting with automated cell counters, Anal. Biochem. 492 (2016) 8–12. https://doi.org/10.1016/j.ab.2015.09.010.

[39] T.M. Elmorsi, Equilibrium Isotherms and Kinetic Studies of Removal of Methylene Blue Dye by Adsorption onto Miswak Leaves as a Natural Adsorbent, J. Environ. Prot. (Irvine,. Calif). 02 (2011) 817–827. https://doi.org/10.4236/jep.2011.26093.

[40] R.O.Y.Y.A. Nayaka, Optical and electrical properties of organic dye sensitized Cr – ZnO and Ni – CdO nanoparticles, SN Appl. Sci. 2 (2020) 1–14. https://doi.org/10.1007/s42452-020-2257-7.

[41] K. Farash, E.K. Hanson, J. Ballantyne, Single source DNA profile recovery from single cells isolated from skin and fabric from touch DNA mixtures in mock physical assaults, Sci. Justice. 58 (2018) 191–199. https://doi.org/10.1016/j.scijus.2017.12.006.

[42] J. Liesche, M. Marek, T. Günther-Pomorski, Cell wall staining with Trypan blue enables quantitative analysis of morphological changes in yeast cells, Front. Microbiol. 6 (2015) 1–8. https://doi.org/10.3389/fmicb.2015.00107.

[43] M.K. Mondal, S. Singh, M. Umareddy, B. Dasgupta, Removal of Orange G from aqueous solution by hematite: Isotherm and mass transfer studies, Korean J. Chem. Eng. 27 (2010) 1811–1815. https://doi.org/10.1007/s11814-010-0301-9.

[44] N. Chantziantoniou, A.D. Donnelly, M. Mukherjee, M.E. Boon, R.M. Austin, Inception and development of the papanicolaou stain method, Acta Cytol. 61 (2017) 266–280. https://doi.org/10.1159/000457827.

[45] F.A. Tanious, J.M. Veal, H. Buczak, L.S. Ratmeyer, W.D. Wilson, DAPI (4', 6-Diamidino-2-Phenylindóle) Binds Differently to DNA and RNA: Minor-Groove Binding at AT Sites and Intercalation at AU Sites, Biochemistry. 31

(1992) 3103–3112. https://doi.org/10.1021/bi00127a010.

[46] S. Szabo, K. Jaeger, H. Fischer, E. Tschachler, W. Parson, L. Eckhart, In situ labeling of DNA reveals interindividual variation in nuclear DNA breakdown in hair and may be useful to predict success of forensic genotyping of hair., Int. J. Legal Med. 126 (2012) 63–70. https://doi.org/10.1007/s00414-011-0566-5.

[47] C. Proff, M. a. Rothschild, P.M. Schneider, Low volume PCR (LV-PCR) for STR typing of forensic casework samples, Int. Congr. Ser. 1288 (2006) 645–647. https://doi.org/10.1016/j.ics.2005.10.012.

[48] R. Daniel, A. Poy, N. Pedersen, S. Thorpe, R.A.H. van Oorschot, Preliminary trials of low volume AmpFlSTR® Profiler Plus® amplification using AmpliGrid (AG480F) slides, Forensic Sci. Int. Genet. Suppl. Ser. 2 (2009) 117–118. https://doi.org/10.1016/j.fsigss.2009.08.062.

[49] D. McNevin, J. Edson, J. Robertson, J.J. Austin, Reduced reaction volumes and increased Taq DNA polymerase concentration improve STR profiling outcomes from a real-world low template DNA source: telogen hairs, Forensic Sci. Med. Pathol. 11 (2015) 326–338. https://doi.org/10.1007/s12024-015-9679-3.

[50] U. Schmidt, S. Lutz-Bonengel, H.J. Weisser, T. Sänger, S. Pollak, U. Schön, T. Zacher, W. Mann, Low-volume amplification on chemically structured chips using the PowerPlex16 DNA amplification kit, Int. J. Legal Med. 120 (2006) 42–48. https://doi.org/10.1007/s00414-005-0041-2.

[51] C. Li, B. Qi, A. Ji, X. Xu, L. Hu, The combination of single cell micromanipulation with LV-PCR system and its application in forensic science, Forensic Sci. Int. Genet. Suppl. Ser. 2 (2009) 516–517. https://doi.org/10.1016/j.fsigss.2009.08.016.

[52] Advalytix, AmpliGrid, (n.d.). http://www.gene-quantification.de/advalytix/index.html (accessed 3 February 2016).

[53] Kyongshin, AmpliGrid, (n.d.). http://www.kyongshin.com/pro/pro_view.asp?first_code=14&second_code=14-02&three_code=14-02&idx=144 (accessed 2 March 2016).

[54] F. Diehl, M. Li, Y. He, K.W. Kinzler, B. Vogelstein, D. Dressman, T. Geng, R. Novak, R. a Mathies, J.G. Wetmur, M. Kumar, L. Zhang, C. Palomeque, S. Wallenstein, J. Chen, M.J. Blow, T. Zhang, T. Woyke, C.F. Spelle, a Krivoshapkin, D. Yang, a Derevianko, E.M. Rubin, Single-Cell Forensic Short Tandem Repeat Typing within Micro fl uidic Droplets, Genome Res. 33 (2014) 551–559. https://doi.org/10.1038/nmeth898.

[55] A.M. Sidore, F. Lan, S.W. Lim, A.R. Abate, Enhanced sequencing coverage with digital droplet multiple displacement amplification, Nucleic Acids Res. (2015) gkv1493. https://doi.org/10.1093/nar/gkv1493.

[56] T. Geng, R. a. Mathies, Minimizing inhibition of PCR-STR typing using digital agarose droplet microfluidics, Forensic Sci. Int. Genet. 14 (2015) 203–209. https://doi.org/10.1016/j.fsigen.2014.10.007.

[57] P.C. Blainey, S.R. Quake, Digital MDA for enumeration of total nucleic acid contamination, Nucleic Acids Res. 39 (2011) 1–9. https://doi.org/10.1093/nar/gkq1074.

[58] Y. Zhu, Y.X. Zhang, L.F. Cai, Q. Fang, Sequential operation droplet array: An automated microfluidic platform for picoliter-scale liquid handling, analysis, and screening, Anal. Chem. 85 (2013) 6723–6731. https://doi.org/10.1021/ac4006414.

[59] T. Geng, R. Novak, R. a. Mathies, Single-cell forensic short tandem repeat typing within microfluidic droplets, Anal. Chem. 86 (2014) 703–712. https://doi.org/10.1021/ac403137h.

[60] K.R. Pandit, P.E. Rueger, R. V Calabrese, S.R. Raghavan, I.M. White, Assessment of surfactants for efficient droplet PCR in mineral oil using the pendant drop technique., Colloids Surf. B. Biointerfaces. 126 (2015) 489–95. https://doi.org/10.1016/j.colsurfb.2015.01.001.

[61] X. Ma, W. Xu, C. Chen, Z. Lu, J. Li, A microfabrication-free nanoliter droplet array for nucleic acid detection combined with isothermal amplification, Analyst. 140 (2015) 4370–4373. https://doi.org/10.1039/C5AN00573F.

[62] C. Zong, S. Lu, A.R. Chapman, X.S. Xie, Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell, Sci. . 338 (2012) 1622–1626. https://doi.org/10.1126/science.1229164.

[63] Y. Fu, C. Li, S. Lu, W. Zhou, F. Tang, X.S. Xie, Y. Huang, Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification, 112 (2015) 2–7. https://doi.org/10.1073/pnas.1513988112.

[64] C.F. a de Bourcy, I. De Vlaminck, J.N. Kanbar, J. Wang, C. Gawad, S.R. Quake, A quantitative comparison of single-cell whole genome amplification methods., PLoS One. 9 (2014) e105585. https://doi.org/10.1371/journal.pone.0105585.

[65] S. Lu, C. Zong, W. Fan, M. Yang, J. Li, A.R. Chapman, P. Zhu, X. Hu, L. Xu, L. Yan, F. Bai, J. Qiao, F. Tang, R. Li, X.S. Xie, Probing Meiotic Recombination and Aneuploidy of Single Sperm Cells by Whole Genome Sequencing, Science. 338 (2012) 1627–1630. https://doi.org/10.1126/science.1229112.

[66] A. Ambers, M. Turnbough, R. Benjamin, H. Gill-King, J. King, A. Sajantila, B. Budowle, Modified DOP-PCR for Improved STR Typing of Degraded DNA from Human Skeletal Remains and Bloodstains, Leg. Med. 18 (2015) 7–12. https://doi.org/10.1016/j.legalmed.2015.10.013.

[67] E.K. Hanson, J. Ballantyne, Whole genome amplification strategy for

forensic genetic analysis using single or few cell equivalents of genomic DNA, Anal. Biochem. 346 (2005) 246–257. https://doi.org/10.1016/j.ab.2005.08.017.

[68] C. Spits, C. Le Caignec, M. De Rycke, L. Van Haute, A. Van Steirteghem, I. Liebaers, K. Sermon, Whole-genome multiple displacement amplification from single cells., Nat. Protoc. 1 (2006) 1965–1970. https://doi.org/10.1038/nprot.2006.326.

[69] K.N. Ballantyne, R. a H. van Oorschot, I. Muharam, A. van Daal, R. John Mitchell, Decreasing amplification bias associated with multiple displacement amplification and short tandem repeat genotyping, Anal. Biochem. 368 (2007) 222–229. https://doi.org/10.1016/j.ab.2007.05.017.

[70] T. a. Guilliam, S.K. Jozwiakowski, a. Ehlinger, R.P. Barnes, S.G. Rudd, L.J. Bailey, J.M. Skehel, K. a. Eckert, W.J. Chazin, a. J. Doherty, Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins, Nucleic Acids Res. 43 (2015) 1056–1068. https://doi.org/10.1093/nar/gku1321.

[71] G. Wang, E. Maher, C. Brennan, L. Chin, C. Leo, M. Kaur, P. Zhu, M. Rook, J.L. Wolfe, G.M. Makrigiorgos, DNA amplification method tolerant to sample degradation11, Genome Res. 14 (2004) 2357–2366. https://doi.org/10.1101/gr.2813404.analyses.

[72] J. Li, L. Harris, H. Mamon, M.H. Kulke, W.-H. Liu, P. Zhu, G. Mike Makrigiorgos, Whole genome amplification of plasma-circulating DNA enables expanded screening for allelic imbalance in plasma., J. Mol. Diagn. 8 (2006) 22–30. https://doi.org/10.2353/jmoldx.2006.050074.

[73] H. Sun, J. Chen, B. Ni, X. Yang, Y. Wu, Recent advances and current issues in single-cell sequencing of tumors, Cancer Lett. 365 (2015) 1–10. https://doi.org/10.1016/j.canlet.2015.04.022.

[74] R.S. Lasken, Single-cell sequencing in its prime, Nat Biotech. 31 (2013) 211–212. http://dx.doi.org/10.1038/nbt.2523.

[75] N.R. Treff, J. Su, X. Tao, L.E. Northrop, R.T. Scott, Single-cell whole-genome amplification technique impacts the accuracy of SNP microarray-based genotyping and copy number analyses, Mol. Hum. Reprod. 17 (2011) 335–343. https://doi.org/10.1093/molehr/gaq103.

[76] Á.J. Picher, B. Budeus, O. Wafzig, C. Krüger, S. García-Gómez, M.I. Martínez-Jiménez, A. Díaz-Talavera, D. Weber, L. Blanco, A. Schneider, TruePrime is a novel method for whole-genome amplification from single cells based on TthPrimPol, Nat. Commun. 7 (2016). https://doi.org/10.1038/ncomms13296.

[77] A.S. Van Der Plaetsen, L. Deleye, S. Cornelis, L. Tilleman, F. Van Nieuwerburgh, D. Deforce, STR profiling and copy number variation analysis on single, preserved cells using current Whole Genome Amplification

methods, Sci. Rep. 7 (2017) 1–9. https://doi.org/10.1038/s41598-017-17525-5.

[78] S.T. Motley, J.M. Picuri, C.D. Crowder, J.J. Minich, S. a Hofstadler, M.W. Eshoo, Improved multiple displacement amplification (iMDA) and ultraclean reagents., BMC Genomics. 15 (2014) 443. https://doi.org/10.1186/1471-2164-15-443.

[79] Y. Marcy, T. Ishoey, R.S. Lasken, T.B. Stockwell, B.P. Walenz, A.L. Halpern, K.Y. Beeson, S.M.D. Goldberg, S.R. Quake, Nanoliter reactors improve multiple displacement amplification of genomes from single cells, PLoS Genet. 3 (2007) 1702–1708. https://doi.org/10.1371/journal.pgen.0030155.

[80] G. Kumar, E. Garnova, M. Reagin, A. Vidali, Improved multiple displacement amplification with phi29 DNA polymerase for genotyping of single human cells., Biotechniques. 44 (2008) 879–90. https://doi.org/10.2144/000112755.

[81] K.N. Ballantyne, R. a H. van Oorschot, R.J. Mitchell, Comparison of two whole genome amplification methods for STR genotyping of LCN and degraded DNA samples, Forensic Sci. Int. 166 (2007) 35–41. https://doi.org/10.1016/j.forsciint.2006.03.022.

[82] Z. Jiang, X. Zhang, R. Deka, L. Jin, Genome amplification of single sperm using multiple displacement amplification, Nucleic Acids Res. 33 (2005) 1–9. https://doi.org/10.1093/nar/gni089.

[83] C. Spits, C. Le Caignec, M. De Rycke, L. Van Haute, Optimization and evaluation of single-cell whole genome multiple displacement amplification, Hum. Mutat. 27 (2006) 796–802. https://doi.org/10.1002/humu.

[84] E.H. Kim, H.Y. Lee, I.S. Yang, S.E. Jung, W.I. Yang, K.J. Shin, Massively parallel sequencing of 17 commonly used forensic autosomal STRs and amelogenin with small amplicons, Forensic Sci. Int. Genet. 22 (2016) 1–7. https://doi.org/10.1016/j.fsigen.2016.01.001.

[85] Illumina, Targeted Next-Generation Sequencing for Forensic Genomics, (2015) 12. https://www.illumina.com/content/dam/illumina-marketing/documents/products/appspotlights/app_spotlight_forensics.pdf (accessed 30 October 2019).

[86] W.J. Ansorge, Next-generation DNA sequencing techniques, N. Biotechnol. 25 (2009) 195–203. https://doi.org/10.1016/j.nbt.2008.12.009.

[87] J.M. Rothberg, W. Hinz, T.M. Rearick, J. Schultz, W. Mileski, M. Davey, J.H. Leamon, K. Johnson, M.J. Milgrew, M. Edwards, J. Hoon, J.F. Simons, D. Marran, J.W. Myers, J.F. Davidson, A. Branting, J.R. Nobile, B.P. Puc, D. Light, T.A. Clark, M. Huber, J.T. Branciforte, I.B. Stoner, S.E. Cawley, M. Lyons, Y. Fu, N. Homer, M. Sedova, X. Miao, B. Reed, J. Sabina, E. Feierstein, M. Schorn, M. Alanjary, E. Dimalanta, D. Dressman, R.

Kasinskas, T. Sokolsky, J.A. Fidanza, E. Namsaraev, K.J. McKernan, A. Williams, G.T. Roth, J. Bustillo, An integrated semiconductor device enabling non-optical genome sequencing, Nature. 475 (2011) 348–352. https://doi.org/10.1038/nature10242.

[88] L.I. Moreno, M.B. Galusha, R. Just, A closer look at Verogen's Forenseq™ DNA Signature Prep kit autosomal and Y-STR data for streamlined analysis of routine reference samples, Electrophoresis. 39 (2018) 2685–2693. https://doi.org/10.1002/elps.201800087.

[89] I. Sinkiewicz, A. Śliwińska, H. Staroszczyk, I. Kołodziejska, Alternative Methods of Preparation of Soluble Keratin from Chicken Feathers, Waste and Biomass Valorization. 8 (2017) 1043–1048. https://doi.org/10.1007/s12649-016-9678-y.

[90] K.L. Opel, E.L. Fleishaker, J.A. Nicklas, E. Buel, B.R. McCord, Evaluation and quantification of nuclear DNA from human telogen hairs., J. Forensic Sci. 53 (2008) 853–7. https://doi.org/10.1111/j.1556-4029.2008.00777.x.

[91] F.B. Dean, S. Hosono, L. Fang, X. Wu, A.F. Faruqi, P. Bray-Ward, Z. Sun, Q. Zong, Y. Du, J. Du, M. Driscoll, W. Song, S.F. Kingsmore, M. Egholm, R.S. Lasken, Comprehensive human genome amplification using multiple displacement amplification, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 5261–5266. https://doi.org/10.1073/pnas.082089499.

[92] S. Bigdeli, R.O. Dettloff, C.W. Frank, R.W. Davis, L.D. Crosby, A Simple Method for Encapsulating Single Cells in Alginate Microspheres Allows for Direct PCR and Whole Genome Amplification, PLoS One. 10 (2015) e0117738. https://doi.org/10.1371/journal.pone.0117738.

[93] M.L. Leung, Y. Wang, J. Waters, N.E. Navin, SNES: single nucleus exome sequencing, Genome Biol. 16 (2015). https://doi.org/10.1186/s13059-015-0616-2.

[94] R. Xue, R. Li, F. Bai, Single cell sequencing: technique, application, and future development, Sci. Bull. 60 (2014) 33–42. https://doi.org/10.1007/s11434-014-0634-6.

[95] N. Li, L. Wang, H. Wang, M. Ma, X. Wang, Y. Li, W. Zhang, J. Zhang, D.S. Cram, Y. Yao, The Performance of Whole Genome Amplification Methods and Next-Generation Sequencing for Pre-Implantation Genetic Diagnosis of Chromosomal Abnormalities, J. Genet. Genomics. (2015). https://doi.org/10.1016/j.jgg.2015.03.001.

[96] B.A. Tereba, L. Flanagan, P. Mandrekar, R. Olson, DIFFEREX ™ SYSTEM A New , Rapid Method to Separate Sperm and Epithelial Cells, Profiles DNA. (2004) 8–10. www.promega.de (accessed 31 October 2019).

[97] G. Tarone, G. Galetto, M. Prat, P.M. Comoglio, Cell surface molecules and fibronectin-mediated cell adhesion: Effect of proteolytic digestion of membrane proteins, J. Cell Biol. 94 (1982) 179–186.

https://doi.org/10.1083/jcb.94.1.179.

[98] M. Enciso, G. Pieczenik, J. Cohen, D. Wells, Development of a novel synthetic oligopeptide for the detection of DNA damage in human spermatozoa, Hum. Reprod. 27 (2012) 2254–2266. https://doi.org/10.1093/humrep/des201.

[99] M. Davison, E. Hall, R. Zare, D. Bhaya, Challenges of metagenomics and single-cell genomics approaches for exploring cyanobacterial diversity, Photosynth. Res. (2014). https://doi.org/10.1007/s11120-014-0066-9.

[100] B. Brinkmann, M. Klintschar, F. Neuhuber, J. Hühne, B. Rolf, Mutation rate in human microsatellites: Influence of the structure and length of the tandem repeat, Am. J. Hum. Genet. 62 (1998) 1408–1415. https://doi.org/10.1086/301869.

[101] J.L. Weber, C. Wong, Mutation of human short tandem repeats, Hum. Mol. Genet. 2 (1993) 1123–1128. https://doi.org/10.1093/hmg/2.8.1123.

[102] M.G. García, L. Gusmão, C.I. Catanesi, G.A. Penacino, N. Pinto, Mutation rate of 12 X-STRs from investigator Argus X-12 kit in Argentine population, Forensic Sci. Int. Genet. Suppl. Ser. (2017). https://doi.org/10.1016/j.fsigss.2017.09.219.

[103] Y. Nishikawa, M. Hosokawa, T. Maruyama, K. Yamagishi, T. Mori, H. Takeyama, Monodisperse Picoliter Droplets for Low-Bias and Contamination-Free Reactions in Single-Cell Whole Genome Amplification, PLoS One. 10 (2015) e0138733. https://doi.org/10.1371/journal.pone.0138733.

[104] J.E. Allard, The collection of data from findings in cases of sexual assault and the significance of spermatozoa on vaginal, anal and oral swabs, Sci. Justice - J. Forensic Sci. Soc. 37 (1997) 99–108. https://doi.org/10.1016/S1355-0306(97)72154-3.

[105] D. Durairajanayagam, Lifestyle causes of male infertility, Arab J. Urol. 16 (2018) 10–20. https://doi.org/10.1016/j.aju.2017.12.004.

[106] D.W. Gjertson, C.H. Brenner, M.P. Baur, A. Carracedo, F. Guidet, J.A. Luque, R. Lessig, W.R. Mayr, V.L. Pascali, M. Prinz, P.M. Schneider, N. Morling, ISFG: Recommendations on biostatistics in paternity testing, Forensic Sci. Int. Genet. 1 (2007) 223–231. https://doi.org/10.1016/j.fsigen.2007.06.006.

[107] L. Feng, C. Xu, X. Zeng, H. Zhang, F. Yang, W. Li, Z. Tu, C. Li, L. Hu, Y-chromosomal haplotyping of single sperm cells isolated from semen mixtures – a successful identification of three perpetrators in a multi-suspect sexual assault case, Croat. Med. J. 55 (2014) 537–541. https://doi.org/10.3325/cmj.2014.55.537.

[108] L. Goddard, David R and Michaelis, A study on keratin, J. Biol. Chem. 106

(1934) 605–614.

[109] P. Mokrejs, O. Krejci, P. Svoboda, Producing keratin hydrolysates from sheep wool, Orient. J. Chem. 27 (2011) 1303–1309.

[110] S.W. Youn, J.I. Na, S.Y. Choi, C.H. Huh, K.C. Park, Regional and seasonal variations in facial sebum secretions: A proposal for the definition of combination skin type, Ski. Res. Technol. 11 (2005) 189–195. https://doi.org/10.1111/j.1600-0846.2005.00119.x.

[111] B.M. Kemp, M. Winters, C. Monroe, J.L. Barta, How much DNA is lost? Measuring DNA loss of short-tandem-repeat length fragments targeted by the powerplex 16® system using the Qiagen Minelute purification kit, Hum. Biol. 86 (2015) 313–329. https://doi.org/10.13110/humanbiology.86.4.0313.

[112] M. Goray, J.R. Mitchell, R.A.H. Van Oorschot, Evaluation of multiple transfer of DNA using mock case scenarios, Leg. Med. 14 (2012) 40–46. https://doi.org/10.1016/j.legalmed.2011.09.006.

[113] R.A.H. van Oorschot, B. Szkuta, G.E. Meakin, B. Kokshoorn, M. Goray, DNA transfer in forensic science: A review, Forensic Sci. Int. Genet. 38 (2019) 140–166. https://doi.org/10.1016/j.fsigen.2018.10.014.

[114] B. Herber, K. Herold, DNA Typing of Human Dandruff, J. Forensic Sci. 43 (1998) 648–656.

[115] F.Q. Gavazaj, I.I. Mikerezi, V.H. Morina, F.A. Cakaj, E.B. Maloku, B.B. Gavazaj, D.S. Kastrati, B.A. Duriqi-Maloku, Optimization of DNA concentration to amplify short tandem repeats of human genomic DNA, Bosn. J. Basic Med. Sci. 12 (2012) 236–239. https://doi.org/10.17305/bjbms.2012.2444.

[116] S. Ganschow, J. Silvery, J. Kalinowski, C. Tiemann, toaSTR: A web application for forensic STR genotyping by massively parallel sequencing, Forensic Sci. Int. Genet. 37 (2018) 21–28. https://doi.org/10.1016/j.fsigen.2018.07.006.

[117] A.C. Jäger, M.L. Alvarez, C.P. Davis, E. Guzmán, Y. Han, L. Way, P. Walichiewicz, D. Silva, N. Pham, G. Caves, J. Bruand, F. Schlesinger, S.J.K. Pond, J. Varlaro, K.M. Stephens, C.L. Holt, Developmental validation of the MiSeq FGx Forensic Genomics System for Targeted Next Generation Sequencing in Forensic DNA Casework and Database Laboratories, Forensic Sci. Int. Genet. 28 (2017) 52–70. https://doi.org/10.1016/j.fsigen.2017.01.011.

[118] F. Salvianti, G. Rotunno, F. Galardi, F. De Luca, M. Pestrin, A.M. Vannucchi, A. Di Leo, M. Pazzagli, P. Pinzani, Feasibility of a workflow for the molecular characterization of single cells by next generation sequencing, Biomol. Detect. Quantif. 5 (2015) 23–29. https://doi.org/10.1016/j.bdq.2015.07.002.

[119] W. Vegetti, Correlation between semen parameters and sperm aneuploidy rates investigated by fluorescence in-situ hybridization in infertile men, Hum. Reprod. 15 (2000) 351–365. https://doi.org/10.1093/humrep/15.2.351.

[120] A.G. Hinch, G. Zhang, P.W. Becker, D. Moralli, R. Hinch, B. Davies, R. Bowden, P. Donnelly, Factors influencing meiotic recombination revealed by whole-genome sequencing of single sperm, Science (80-. ). 363 (2019). https://doi.org/10.1126/science.aau8861.

[121] A.R. Khataee, M.N. Pons, O. Zahraa, Photocatalytic degradation of three azo dyes using immobilized TiO2 nanoparticles on glass plates activated by UV light irradiation: Influence of dye molecular structure, J. Hazard. Mater. 168 (2009) 451–457. https://doi.org/10.1016/j.jhazmat.2009.02.052.

[122] E.A. Graffy, D.R. Foran, A Simplified Method for Mitochondrial DNA Extraction from Head Hair Shafts, J. Forensic Sci. 50 (2005) 1–4. https://doi.org/10.1520/jfs2005126.

[123] L. Eckhart, J. Bach, J. Ban, E. Tschachler, Melanin binds reversibly to thermostable DNA polymerase and inhibits its activity, Biochem. Biophys. Res. Commun. 271 (2000) 726–730. https://doi.org/10.1006/bbrc.2000.2716.

# Appendix I – X-STR electropherogram from spermatozoa 9 of person A

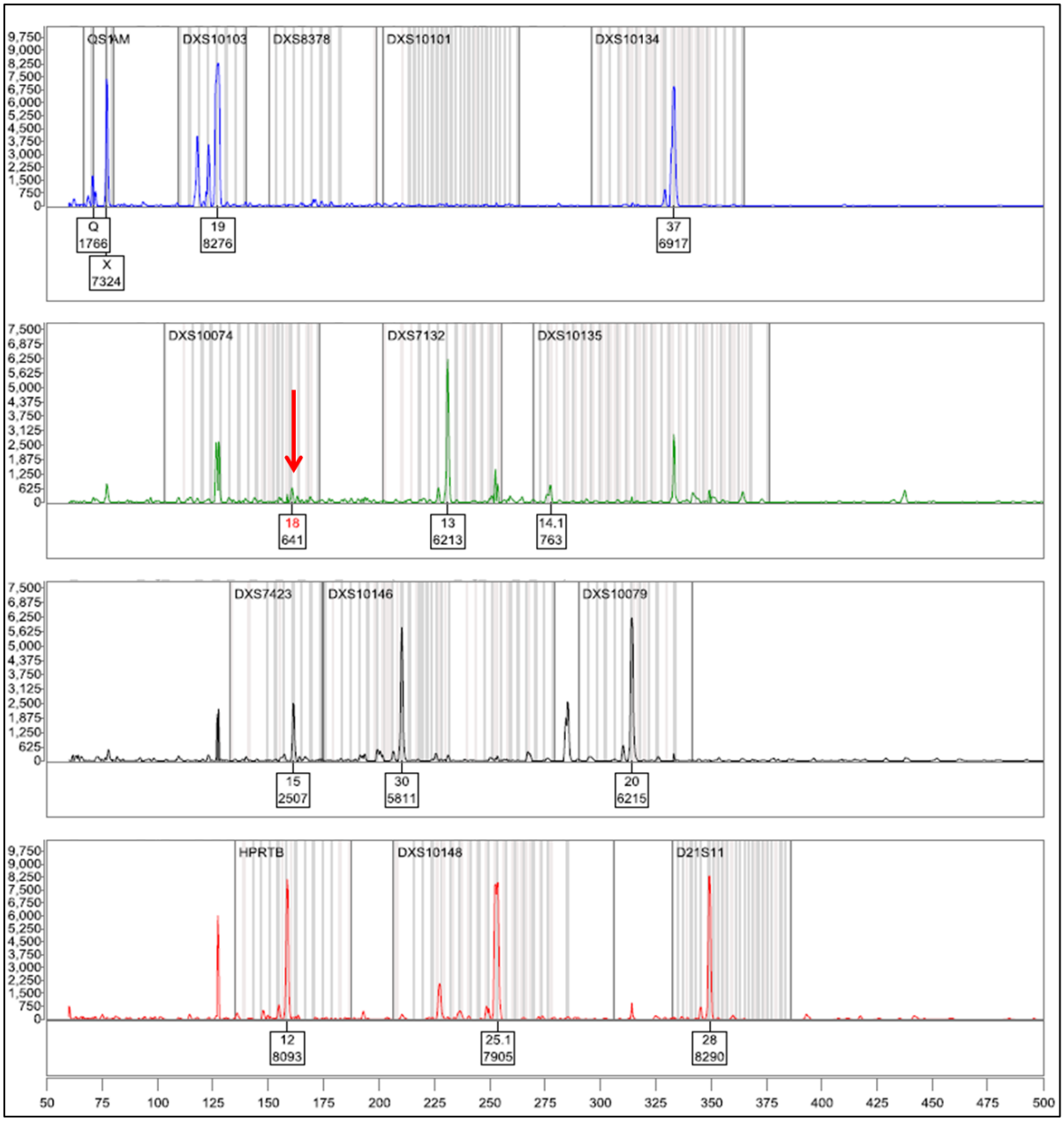


**Figure I.I. Electropherogram X-chromosomal STR profile from spermatozoon 9 from person A and analysed with WGA and Argus X-12 QS. The relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. Below the red channel (with red allele peaks), the X-axis represents the fragment sizes for all detection channels in base pairs (bp). The STR locus name is shown on top in each detection channel above the allele peaks. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. False alleles were not called and therefore no label box is present underneath a false allele peak. The red arrow at locus DXS10074 indicates allele 18 that was not in concordance with allele 19 of the reference profile of person A.**

# Appendix II – X-STR electropherogram from spermatozoa 11 of person A

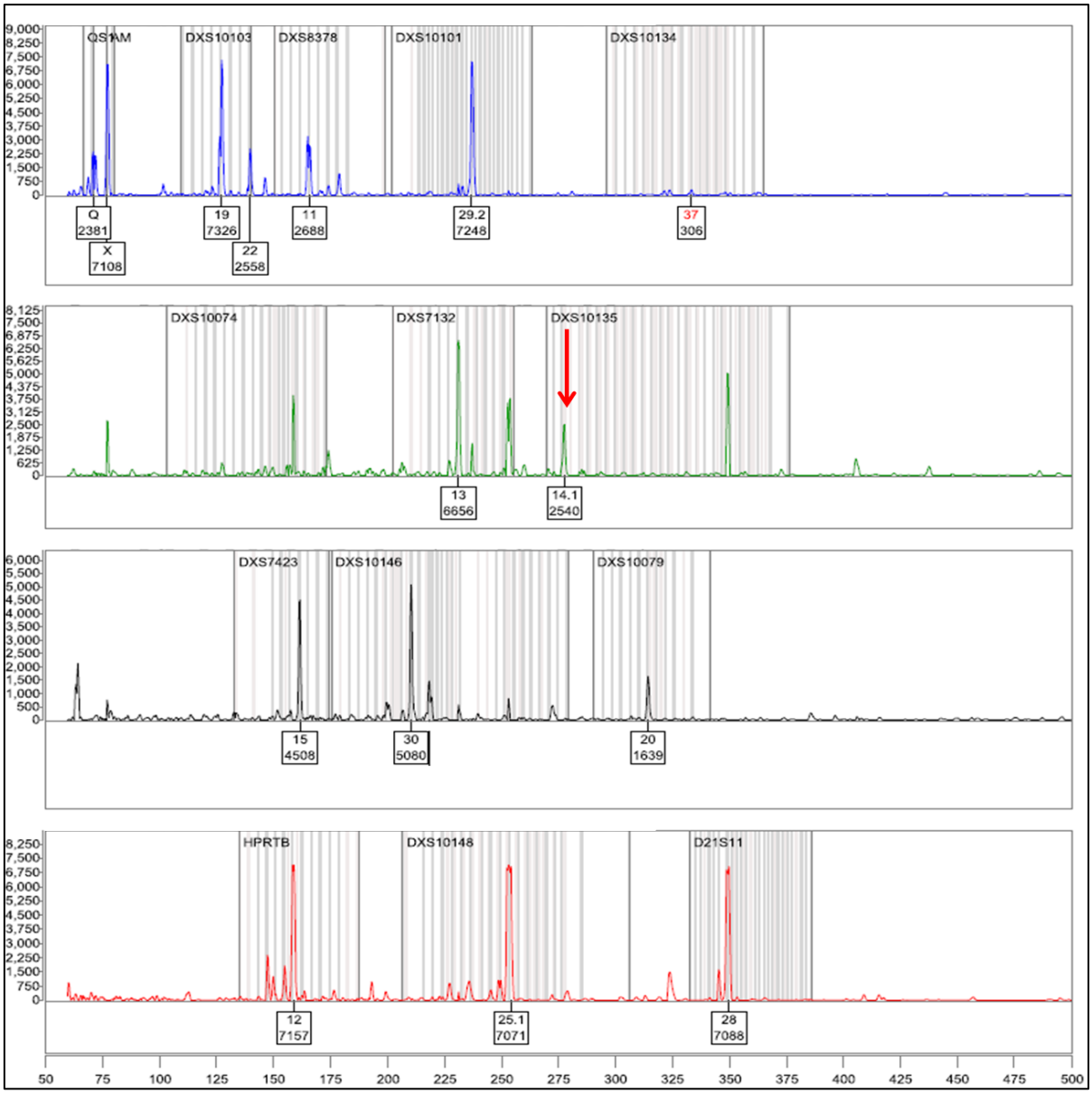


**Figure II.I. Electropherogram X-chromosomal STR profile from spermatozoon 9 from person A and analysed with WGA and Argus X-12 QS. The relative fluorescent units (RFU) is set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. Below the red channel (with red allele peaks), the X-axis represents the fragment sizes for all detection channels in base pairs (bp). The STR locus name is shown on top in each detection channel above the allele peaks. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected. False alleles were not called and therefore no label box is present underneath a false allele peak. The red arrow at locus DXS10135 indicates allele 14.1 that was not in concordance with allele 29 of the reference profile of person A.**

# Appendix III – Autosomal STR electropherogram from single skin flake of person D

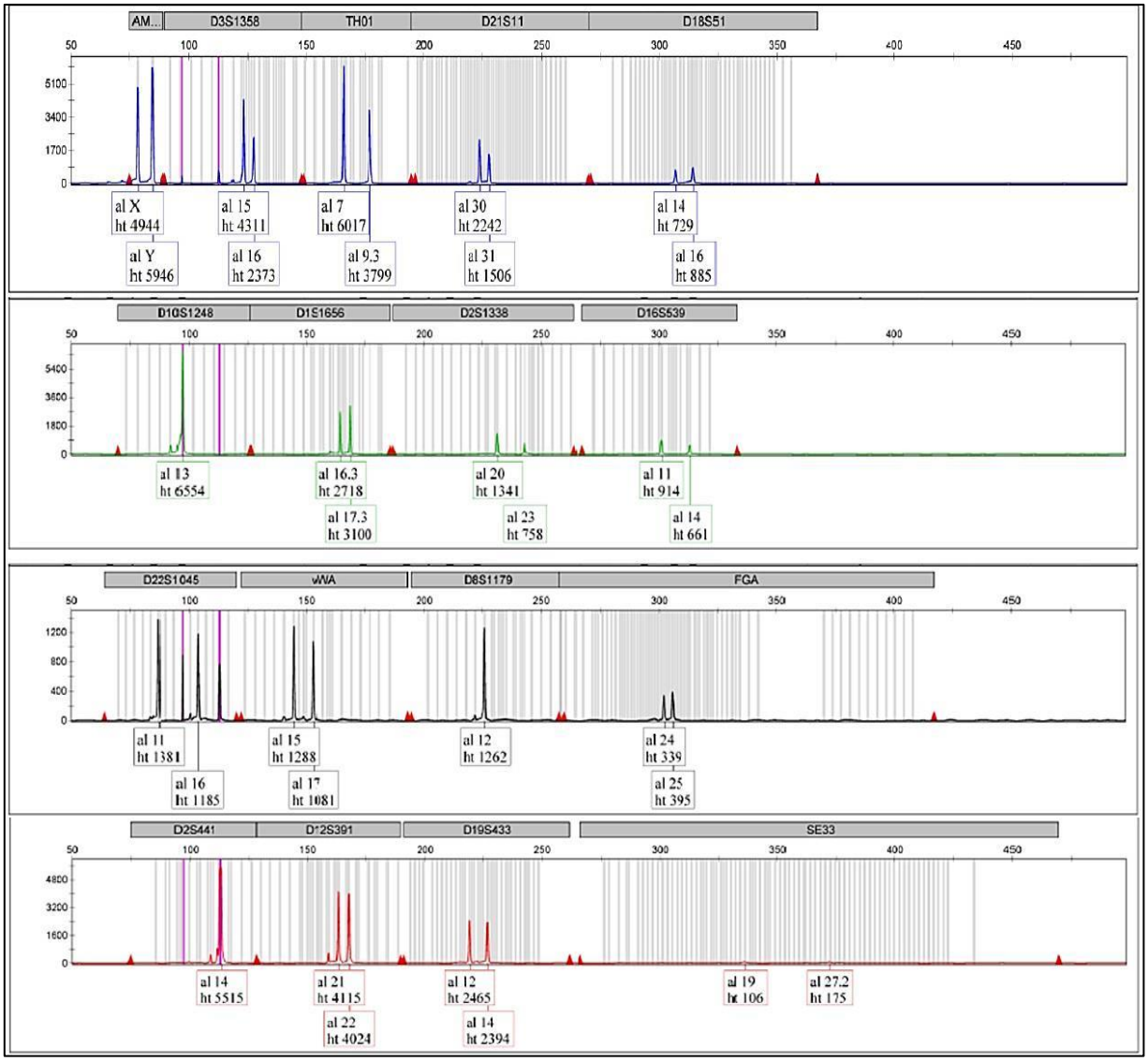


**Figure III.I. Electropherogram with a full autosomal STR profile (100% allele revovery) from a single skin flake from person D. A single skin flake of a pillow where person D had laid down on was secured by tape lifting stained with Orange G and the DNA was extracted in order to perform a direct-PCR (PowerPlex ESX17) on the lysate. Peaks represent the amplicons of different alleles amplified using multiplex PCR and detected by capillary electrophoresis. The different allele amplicons were detected and shown using 4 colour channels (blue, green, black and red). The relative fluorescent units (RFU) are set on the Y-axis and the allele numbers with the corresponding RFU is labelled in a box underneath each allele peak. On top, the X-axis represents the fragment size in base pairs (bp) with the allele corresponding STR locus name above in the grey horizontal bars. The vertical grey lines are allele bins which are located at each fragment size at which a potential allele could be detected.**

# Appendix IV – Autosomal STR output from single skin flake of person A generated by NGS

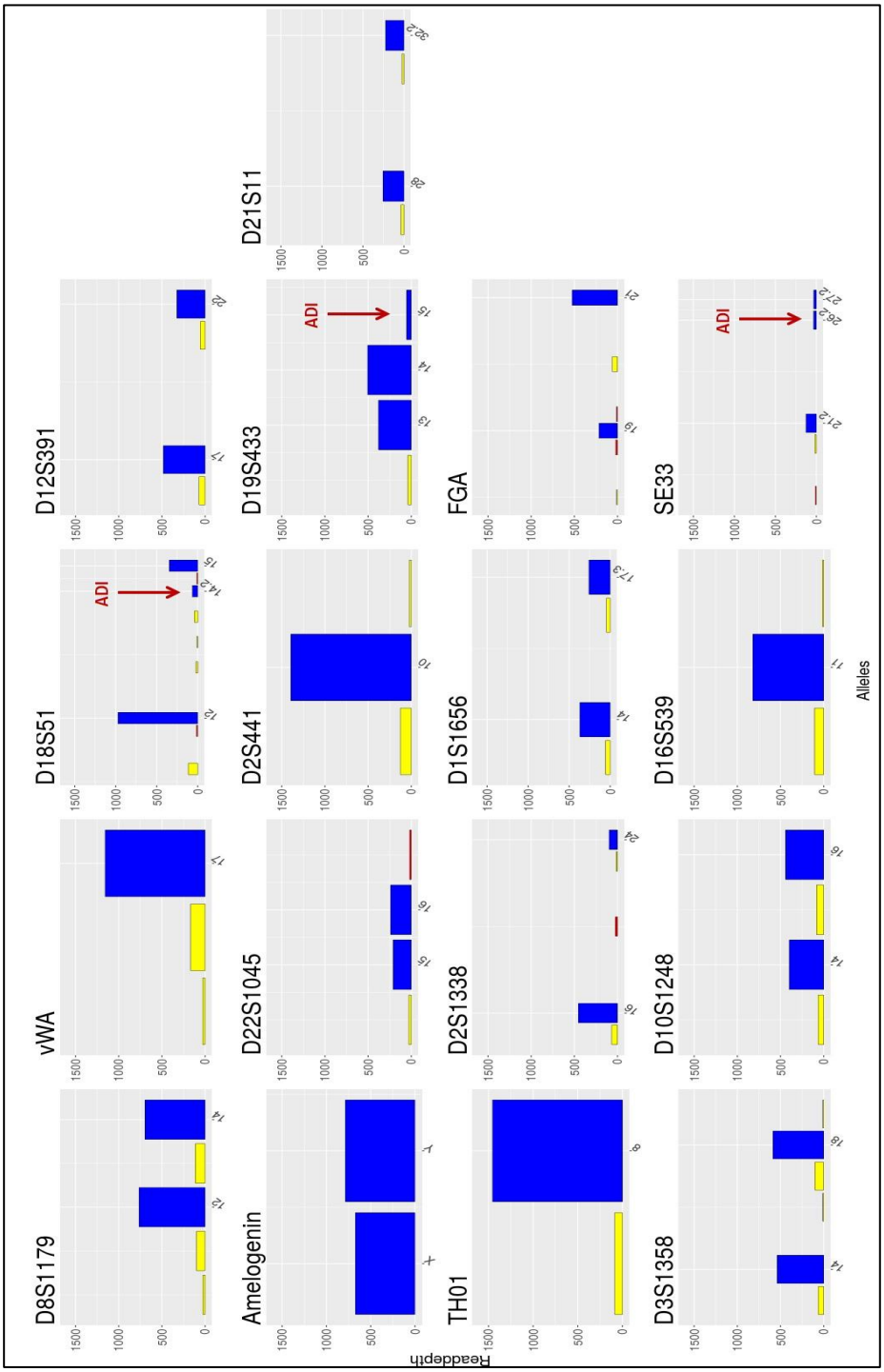


**Figure IV.I. NGS output of a full DNA profile (100% allele recovery) from a single skin flake from person A. The isolated skin flake was lysed and the lysate was directly used in the in-house developed mini-STR PCR for NGS. The NGS was run on the Illumina MiSeq FGx device (Illumina, San Diego, USA). STR loci and Amelogenin are presented on top of each diagram. The number of reads is shown on the Y-axis of the diagrams. The allele numbers are shown on the X-axis of the diagrams. Only alleles (blue bars) that passed the calling filters were called. Stutters are represented as yellow bars. Three ADIs (red arrow) were seen, namely at locus D18S51, D19S433 and SE33.**

# Appendix V – Copy of ethical approval

SCHOOL OF PHARMACY & LIFE SCIENCES
Robert Gordon University
Sir Ian Wood Building
Garthdee Road
Aberdeen
AB10 7GJ
United Kingdom
Tel: 01224 262500/2800
www.rgu.ac.uk

12 May 2017

Dear Glenn,

Re.: WGA on single sperm cells and skin flakes as a new strategy to overcome mixed profiles in forensic DNA analysis.

The School Research Ethics Committee has assessed your application and the overall decision is that there are no ethical issues with your project. Additionally, the Convener of the committee was comfortable with the changes to the project.

I can now confirm that you are able to proceed with your research and any further ethics applications.

Should there be any amendments to this project during the research we would advise you to consult with the convener of the ethics committee as to whether a further ethical review would be required.

We wish you success with your project.

Regards

Convener of the School Ethics Review Panel

# Appendix VI – Donor consent form (sperm)

## Donor Consent Form

Project: DNA profiling on single sperm cells

You have volunteered to participate in a research project on behalf of the Bonn-Rhein-Sieg University of Applied Sciences (H-BRS) in Rheinbach, Germany in cooperation with Robert Gordon University in Aberdeen, UK and the Scottish Police Authority – Forensic Services, UK.

For male volunteers (non-vasectomized):
We ask for you to provide a buccal mouth swab and a sample of your sperm.

For female volunteers:
We ask for you to provide a buccal mouth swab and two mouth swabs using medical swabs. Please obtain the sample with the buccal mouth swab at a separate time of the day or on a separate day to ensure a sufficient collection of cells.

Your samples and DNA profile will be used for research purposes within this project only and not disclosed to any other parties under any circumstances. Your samples and DNA profile will be saved in our internal DNA database and only be used for validating new DNA profiling techniques and not for forensic and/or clinical diagnostic purposes. Your name and complete profile will not be published or shared with other parties. Samples will not contain your name and will be kept anonymously. If anything other than this is to be required then written permission must be obtained from you. If desired, we will remove your profile from our internal DNA database after a period of 1 year from the completion of the research project. (Consent can be withdrawn at any time prior to this and, once a written request to remove the profile has been received, the profile will be removed from the internal database.)

Please complete **this box**:

I **give my consent** to my biological material and DNA profile being held for research purposes within H-BRS and understand the purpose of me providing these samples.

Name:_____

Gender: ☐ male (buccal and sperm)
☐ female (buccal and two medical mouth swabs)

Donor ID (on tubes/universals):

Signature: _____ Date:_____

# Appendix VII – Donor consent form (skin flakes)

## Donor Consent Form

Project: DNA profiling on single skin flakes

You have volunteered to participate in a research project on behalf of the Bonn-Rhein-Sieg University of Applied Sciences (H-BRS) in Rheinbach, Germany in cooperation with Robert Gordon University in Aberdeen, UK and the Scottish Police Authority – Forensic Services, UK.

We ask for you to provide a buccal mouth swab, skin flakes from your head and/or to lie down on several pillows as part of an experiment.

Your samples and DNA profile will be used for research purposes within this project only and not disclosed to any other parties under any circumstances. Your samples and DNA profile will be saved in our internal DNA database and only be used for validating new DNA profiling techniques and not for forensic and/or clinical diagnostic purposes. Your name and complete profile will not be published or shared with other parties. Samples will not contain your name and will be kept anonymously. If anything other than this is to be required then written permission must be obtained from you. If desired, we will remove your profile from our internal DNA database after a period of 1 year from the completion of the research project. (Consent can be withdrawn at any time prior to this and, once a written request to remove the profile has been received, the profile will be removed from the internal database.)

Please complete **this box**:

I **give my consent** to my biological material and DNA profile being held for research purposes within H-BRS and understand the purpose of me providing these samples.

Name:_____

Donor ID (on tubes/universals):

Signature: _____ Date:_____