UiT The Arctic University of Norway

Department of Medical Biology, Faculty of Health Science

Optimization of the use of PrepFiler *Express*™ Forensic DNA extraction kit lysis buffer discard for mRNA profiling

Kari Hamnevoll

Master's thesis in Biomedicine, MBI-3911, June 2021

Picture: Kari Hamnevoll

Abstract

mRNA profiling for body fluid or tissue type identification is able to provide contextual and circumstantial information alongside DNA profiles from biological material found at a crime scene. However, body fluid identification methods usually compromise the material destined for DNA profiling. A transition from traditional methods to RNA-profiling – especially centred on mRNA – for body fluid identification has taken place in the last two decades. To solve the challenge of material limitation, studies have successfully developed mRNA profiles of RNA extracted from lysis buffer discard, a waste product of DNA-extraction. It would be interesting to examine whether this method of RNA-extraction is possible for other combinations of DNA- and RNA-isolation kits.

The initiation of this study was to determine if the promising results from a preliminary study at the CFG in terms of RNA quantity and quality measured in the DNA-extraction lysis buffer discard and DNA eluate from body fluids extracted with the PrepFiler *Express*™ DNA Extraction system were reproducible. Obtained RNA concentrations (ng/µL) and quality measures (RIN and DV_{200}) were comparable.

The protocols of three commercially available RNA-isolation kits were optimized for the purpose of RNA-extraction from lysis buffer discard by altering kit reagent ratios. The RNA eluate of the Direct-zol™ RNA Miniprep kit gave poor results and was not further examined. The RNA eluates of the *mir*Vana™ miRNA Isolation kit and the ReliaPrep™ RNA Miniprep kit had similar levels of RNA quantity and quality. RNA concentration was measured by the Qubit® 4 Fluorometer, and level of RNA degradation was determined by the relative quantity of a short and long product of the housekeeping gene GAPDH by qPCR. A continuous problem of the ReliaPrep™ kit was contamination of genomic DNA, which l ed to the decision of the *mir*Vana™ kit as being most suited for extraction from the Prepfiler™ lysis buffer discard.

As a first step in the compilation of a RNA multiplex, a preliminary singleplex of twelve body fluid specific mRNA markers and two housekeeping genes using three different PCR programs was performed on reverse transcribed RNA extracted from five pure body fluids, and products were separated by capillary electrophoresis. An overload of the system was observed for most of the body fluid markers. Primer tests should be rerun with less RNA, e.g. RNA extracted from lysis buffer discard instead of pure body fluids, before further optimization.

Foreword

This study was completed during the spring of 2021 as a part of a master's degree in Biomedicine at the University of Tromsø (UiT) – The Arctic University of Norway, at the Forensic Genetics Centre (CFG).

My gratitude goes to the CFG at UiT – The Arctic University of Norway for choosing me as their master's candidate. The field of forensic genetics has been a matter of great interest and fascination to me for many years. The insight CFG has provided me into their methods, routines and workflows is knowledge I value greatly, and it has been a crucial addition to my education which I will carry with me into future work. I want to give a huge thanks and recognition to my supervisor Kirstin Janssen and my co-supervisor Marita Olsen, for your guidance, advice and encouragement through this whole process. Your patience as well as your orderliness kept me afloat these last few months of laboratory work and writing. The clear expectations and standards you set for me is exactly what I needed when I wasn't sure in which direction I was supposed to go.

I also want to extend my gratitude to the Genomic Support Center at UiT, for assisting me in my practical work with RNA quantification and assessment of RNA quality, and the Science Strategy Team by the Institute of Environmental Science and Research Limited, New Zealand, and the Netherlands Forensic Institute, Netherlands, who provided me with their optimized procedures and advice. A thanks goes to the Proteomics and Metabolomics Core Facility at UiT for lending me extra space in their -70°C freezer. Neighbouring laboratories also deserve a mention, for tolerating my constant running back and forth through your facilities to get my samples.

I want to thank my family, and especially my parents, for always being my supporters in everything I choose to do. My older sister who has cleared every path for me in life, has also provided me with invaluable advice and support. I also want to thank my best friend, Kaisa, who doesn't understand much because she's a dog, but has been a source of comfort and consolation anyway. Lastly, to the most annoying person I know, yet my partner in crime/ roommate, Guro. Our endless conversations about everything and nothing, and our countless hours of Netflix has been an excellent breathing space for me in these stressful times.

Kari Hamnevoll

Tromsø, June 2021

Table of Contents

List of Tables

List of Figures

Abbreviations

ACTB – Actin-beta
bp – base pairs
C_t – Cycle threshold
CYP2B7P – Cytochrome P450 Family 2 Subfamily B Member 7, Pseudogene
DNA – Deoxyribonucleic acid
    cDNA – complementary DNA
    gDNA – genomic DNA
DNase – Deoxyribonuclease
DTT – Dithiothreitol
EDNAP – The European DNA Profiling group
ESR – Institute of Environmental Science and Research
EtOH – Ethanol
FDCSP – Follicular Dendritic Cell Secreted Protein
CFG – Center of Forensic Genetics
GAPDH – Glyceraldehyde 3-phosphate dehydrogenase
GSCT – Genomics Support Center Tromsø
HBB – Hemoglobin subunit beta
KLK3 – Kallikrein Related Peptidase 3
LBD – Lysis buffer discard
MMP7/10 – Matrix metalloproteinase 7/10
MSMB – Beta-microseminoprotein
MUC4 – Mucin 4
NFI – Netherlands Forensic Institute
ng/µL – nanograms per microliter
NL60/64 – Netherlands PCR program 60/64°C annealing temperature
NSD – No Size Data
NZ – New Zealand PCR program
PCR – polymerase chain reaction
    qPCR – quantitative PCR
    RT-PCR – reverse transcriptase PCR
PRM1 – Protamine 1
rfu – relative fluorescence units
RIN – RNA integrity number
RNA – ribonucleic acid
    mRNA – messenger RNA
    miRNA – micro RNA
    ncRNA – non-coding RNA
    rRNA – ribosomal RNA
    siRNA – small interfering RNA
RNases – Ribonucleases
RT – Reverse Transcriptase
SLC4A1 – Solute carrier family 4 member 1
SA – Small Autosomal
STATH – Statherin
STR – Short tandem repeats
BL + TG – Bacterial lysis buffer + thioglycerol
TNP1 – Transition protein 1

Introduction

Forensic genetics is a branch of molecular biology in which the knowledge of human genetics is applied to legal matters and proceedings [6, 7]. In modern times, forensic genetics has taken a central role not only in criminal investigations, but also in the courtroom. The application of DNA analysis for the identification of individuals contributing to biological stains as a tool in forensic investigations was suggested as early as the 1980s and has over the course of the following forty years become an essential part in modern forensic workflows [7-9]. The identification of the cellular origin of evidential biological material can add contextual and circumstantial information about the events at a crime scene, and methods of body fluid and tissue type identification are commonly in use in forensic workflows.

To improve upon current protein-based body fluid and tissue type identification methods, the versatility of tissue specific messenger RNA and non-coding RNAs has been intensively studied in the last two decades. These studies have produced promising results towards the development of a novel, effective method. RNA profiling bears great potential as a body fluid and tissue type identification technique, in terms of circumventing the challenges with current identification methods, which are mostly presumptive in nature, and carry the potential of cross-reaction, in addition to only being available for a limited number of body fluids [10]. A restricted amount of biological test material available for the conduction of both DNA analysis and body fluid and tissue type identification is also an obstacle in forensic workflows. The decision to utilize mRNA as a tool of identification is not only based in the possibility of utilizing less material by co-extraction of DNA and RNA, but also its cell-type specific expression [5, 11-14].

The cellular basis for cell type-specific expression of RNAs and proteins

The central dogma describes the biologically fundamental concept of DNA being transcribed to single stranded messenger RNA (mRNA), which is further translated into polypeptides. Out of the 200 roughly defined cell types that exist in the human body, only a few are excepted from this process. These are cell types that do not carry DNA, such as erythrocytes, thrombocytes, and squamous cells [15].The large majority of cells carry the entire genome within their nuclei and utilize it to implement their respective tasks in the organism. The answer to how it is still

possible to distinguish between these 200 cell types and their features, despite containing the very same set of DNA is the activation and silencing of genes within the genome [16].

The human genome contains about 20 000 genes, defined as regions of DNA that are translated into polypeptides. These genes only make up about 2% of the entire genome, while the rest is referred to as “non-coding DNA”. Nevertheless, a multitude of epigenetic processes, such as DNA methylation, histone modification and regulation by small RNAs function in a synchronous sequence of events to “activate” or “silence” both coding and non-coding regions during early cell development, which in turn results in the spatial and temporal differentiation of distinct cell types [17]. As a result, each cell type will express only a subset of the genome. This unique pattern of gene expression is referred to as the cell’s transcriptome [13].

Transcription factor genes are also directly involved in the pattern of activation and silencing of cell-type specific genes, leading to an expression of unique set of genes [18]. An example of one of these cell-type distinguishing genes is the TF gene, coding for the iron-transporting protein transferrin, which is mainly produced by human liver cells. A positive test result for transferrin protein or its mRNA precursor would therefore indicate that a sample contains biological material from the liver [19].

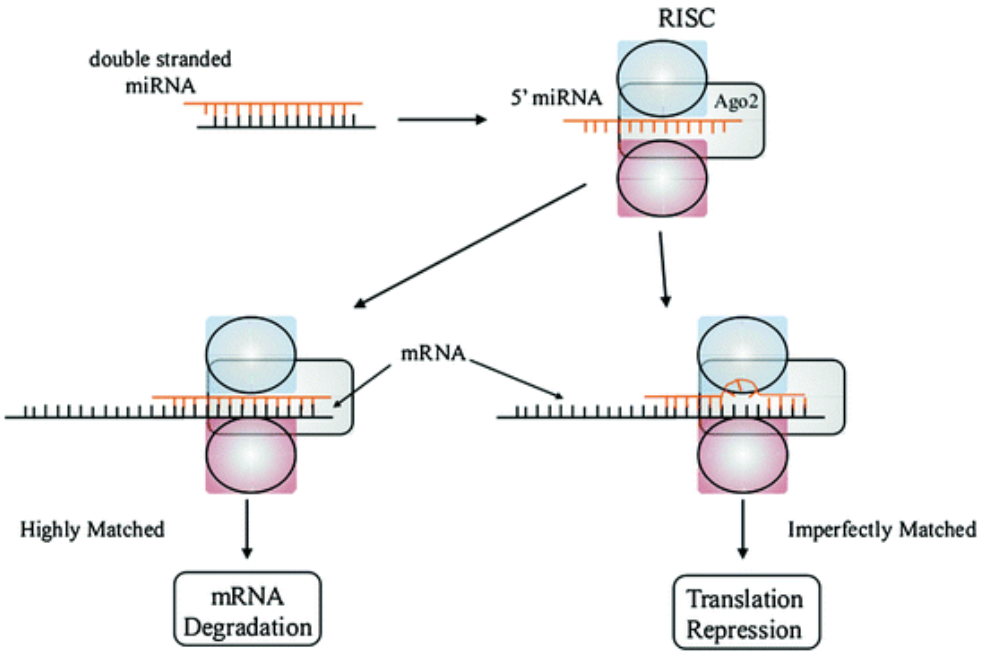


Figure 1*: miRNA is incorporated into the RISC complex and serves as a guide for base pairing with mRNA, which either leads to degradation or translation repression (Goodwin – Acute Lung Injury and Repair (2017), p.169) [3]*

Although only 2 % of the human genome is translated to proteins, about 85 % is transcribed [15]. RNA that is not translated is referred to as non-coding RNA (ncRNA). ncRNA serves many essential features within the regulation of gene expression, both on the transcriptional and post-transcriptional level. ncRNA is subdivided into classes characterized by their respective structure and function. Examples of these classes are ribosomal RNA (rRNA), small transfer RNA (tRNA), small nucleolar RNA (snoRNA), and microRNA (miRNA). microRNAs are single-stranded, short molecules, consisting of about 21-26 nucleotides, that are mainly involved in post-transcriptional gene silencing by the complementary binding of mRNA. miRNA is first incorporated into the RNA-induced silencing complex (RISC) and serves as a guide for base-pairing, which will either result in the cleavage, degradation, or translation inhibition of mRNAs (see Figure 1) [20]. miRNAs are expressed in a tissue-specific manner and are involved in the regulation of 30% of protein-coding genes [21]. Specific miRNAs, mRNAs and distinct proteins will therefore be present within respective tissue types.

It is possible to apply the knowledge of RNA tissue specificity to workflows of forensic genetics. mRNA has previously been viewed as highly unstable, because of its single- stranded nature, which renders it an accessible target for ribonucleic degradation. However, the studies

of Zubakov et al. and Setzer et al. have demonstrated that mRNA markers can have stable expression patterns after a long period of time, up to several years for certain body fluid stains [22, 23]. Nevertheless, factors such as heat and humidity will lead to a higher degree of RNA degradation. As miRNA are shorter RNA molecules, they are less prone to degradation by these environmental factors. miRNA is therefore an interesting target for further studies in the development of RNA profiling techniques [24, 25]. These techniques have been under extensive development during the last two decades and have their basis in the methods developed for DNA analysis.

The development of reliable techniques for DNA analysis in forensic work

DNA very much serves as a cookbook for the entire organism, by providing the recipes for the proteins needed to ensure homeostasis. Human DNA consists of about 5 billion base pairs divided into 46 chromosomes. These 46 chromosomes with additional mitochondrial DNA are referred to as the genome. It has now become common knowledge that the genome with its internal variations serves as a unique “nametag” for every individual. Mutation, recombination and reassortment within the genome as well as sexual reproduction and genetic drift, causes individual variations within specific loci [26].

An example of one of these variations are Short Tandem Repeats (STRs), also referred to as “microsatellite DNA". STRs consist of a series of tandemly repeated units, each between 2-7 base pairs in length. These units are usually repeated up to 100 times, and the number of repeats serves as the basis for a measurable variation between individuals. STR loci are usually found between protein-coding regions and make up about 3% of the human genome. The STRs most often used in forensic analysis are tetranucleotides, meaning four nucleotides long. The rate of mutation of these STRs produces an ample level of polymorphism, whilst remaining stable enough sequence-wise to enable easy standardization. Short tandem repeats are therefore simple and reliable to analyse, even for samples that are several decades old [27, 28].

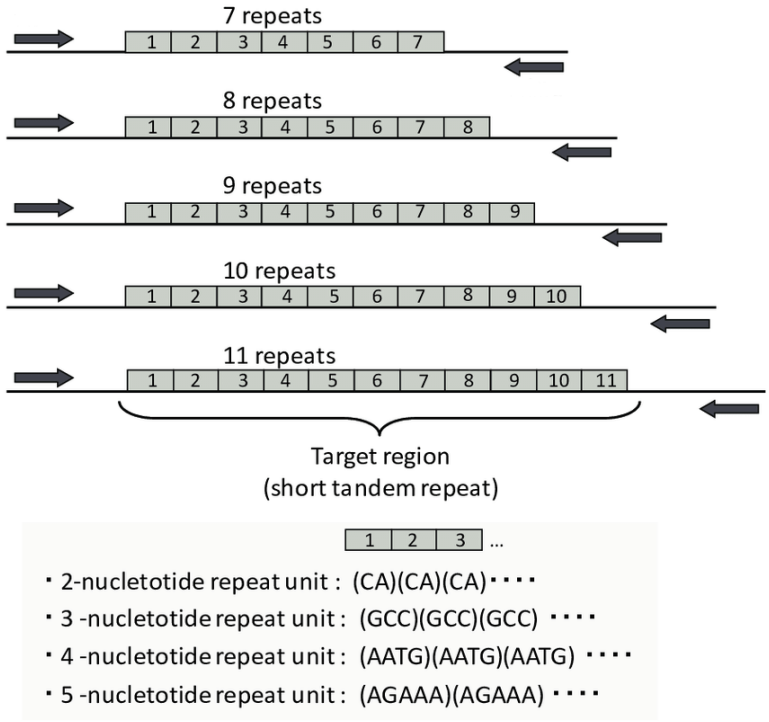


Figure 2*: Short tandem repeats consist of units repeated in a head-to-tail manner, each unit between 2-7 bp. The number of repeats is inherited and varies between individuals, and these regions are commonly found between protein-coding regions. STRs serve as a target for forensic genetic analyses (adapted from Hashiyada – DNA Biometrics (2011) [2])*

The use of variable DNA as tools in forensic work begins with the British geneticist Professor Sir Alec Jeffreys in 1984, and his discovery of so-called minisatellite DNA. Jeffreys and colleagues had taken the first, crude look at the structure of mammalian genes, and discovered that protein-coding regions were separated by apparent gaps with what Jeffreys himself referred to as "*stretches of gobbledegook*" [29]. This is what we today know as exons and introns, respectively. Jeffreys discovered that within these stretches were regions of tandemly repeated DNA sequences, in which the number of repeats in the same loci varies between individuals. These regions of repeated sequences, termed *variable number of tandem repeats*, or VNTRs, became the basis for the first DNA fingerprints, which, as the name suggests, made it possible to accurately distinguish between two individuals [9].

Jeffreys directed his focus towards the development of probes for a subset of VNTRs of 10-15 base-pairs in length, which he coined "minisatellites". Minisatellites were variable enough to provide highly specific genetic information, and probes targeting several of these regions (multi-locus probes) were utilized for the very first use of DNA typing in a forensic setting [30,

31]. In 1986, two connected cases of rape and homicide in Leicestershire, UK were solved by the use of Jeffreys' technique. First, a mass screening of blood and saliva samples from 5,500 local men towards semen samples collected from both crime scenes gave no matches. Later, it was revealed that the perpetrator Colin Pitchfork had avoided the donation of his saliva and blood by persuading a colleague to donate under his name; information which subsequently lead to Pitchfork's arrest and the matching of his DNA profile to those of the crime scene samples [27, 32].

This achievement opened the doors for a new approach in forensic investigation. However, it was not until after the development of Polymerase Chain Reaction (PCR) techniques in the early 1990s, which allowed for an easy amplification of template DNA even from miniscule amounts, that STRs were used as the standard within DNA profiling. STR marker multiplexes enable a greater discrimination between individuals, as opposed to singleplexes, as several markers are screened simultaneously. The standardization of STR marker multiplexes makes them suitable for the establishment of international DNA Databases. The European DNA Profiling Group (EDNAP) is responsible for the standardisation of STR markers in Europe, and DNA profiles have been submitted into DNA Databases since the late 1990s. DNA profiles are now widely established as a key tool in forensic work [9].

Body fluid and tissue type identification techniques

A DNA-profile alone would only make it possible to point towards individuals who may have contributed to the biological material found at a crime scene and cannot specify under which circumstances the material has been transferred. As it is possible to connect certain patterns of behaviour to a body fluid or tissue type's localization at a crime scene, a method of identification would be a useful tool during investigation [14].

An example of how a DNA-profile and tissue type identification can interplay is if a DNA profile of a victim of assault is found on the jeans of the suspect. If one cannot determine the identity of the body fluid or tissue type of the stain this DNA originated from, circumstantial information will be limited, especially if the victim and the suspect have social contact. However, if one was able to determine that the stain in question is venous blood, instead of sweat, epithelial cells or saliva, this is crucial evidence suggesting that a more heinous sequence

of events has taken place. The identity of the stain as blood provides information about the manor and circumstance of the transmission of the victim's DNA to the suspect [33].

Historically, tests for the purpose of tissue type and body fluid identification have mainly been based on chemical, immunological and protein catalytic activity. Many tests are presumptive in nature and are suitable for screening only. Few of these tests are confirmatory and are limited to only one body fluid at a time [34].

A test is classified as presumptive or confirmatory depending on its sensitivity and specificity. A test's sensitivity is its capacity to correctly identify true positives, which in forensic terms refers to its ability to positively determine the presence of a body fluid, even with miniscule amounts of catalytic agent present in the test material. A test's specificity is its capacity to correctly identify true negatives. Specificity is a forensic test's ability to accurately and reliably identify one body fluid over another [35]. Presumptive tests are based in finding the location of evidential material at a crime scene as well as evaluating its potential for screening. A confirmatory test, as its name implies, seeks to determine and confirm the identity or origin of test material. It therefore follows that presumptive tests have a high level of sensitivity and are not necessarily specific, and that confirmatory tests are highly specific, without a particular need for sensitivity.

Although proven to be very useful tools in forensic workflows, both presumptive and confirmatory tests have certain disadvantages. An example of a presumptive test, which is currently in use, but has weaknesses pertaining to its reliability is the Phadebas® Amylase-test (Phadebas). This test is based on an enzymatic reaction, in which α-amylase digests starch and forms a complex, that appears as a blue colour visible to the naked eye [36]. However, a substantial drawback of the Phadebas® Amylase-test is the fact that amylase is also expressed to some degree in other tissue types, which disrupts the test's specificity. The test cannot distinguish between salivary amylase and amylases from other tissues, such as the pancreas or vaginal secretions, in addition to losing performance efficacy over time, as the amylase enzyme is a target of degradation [22, 34, 37].

Another example of a presumptive test is luminol (3-aminophthalhydrazide), which takes advantage of a chemical reaction's bioluminescent properties. Luminol is based on a chemically oxidative reaction which is catalysed by iron in the heme-unit of haemoglobin in human blood. This reaction generates energy, which emanates as a dull, blue glow in a darkened room. A

challenge of testing for the presence of blood with luminol is that false positives in contact with strong oxidatives such as cleaning detergents, can occur. (see Figure 3) [1, 22, 38].

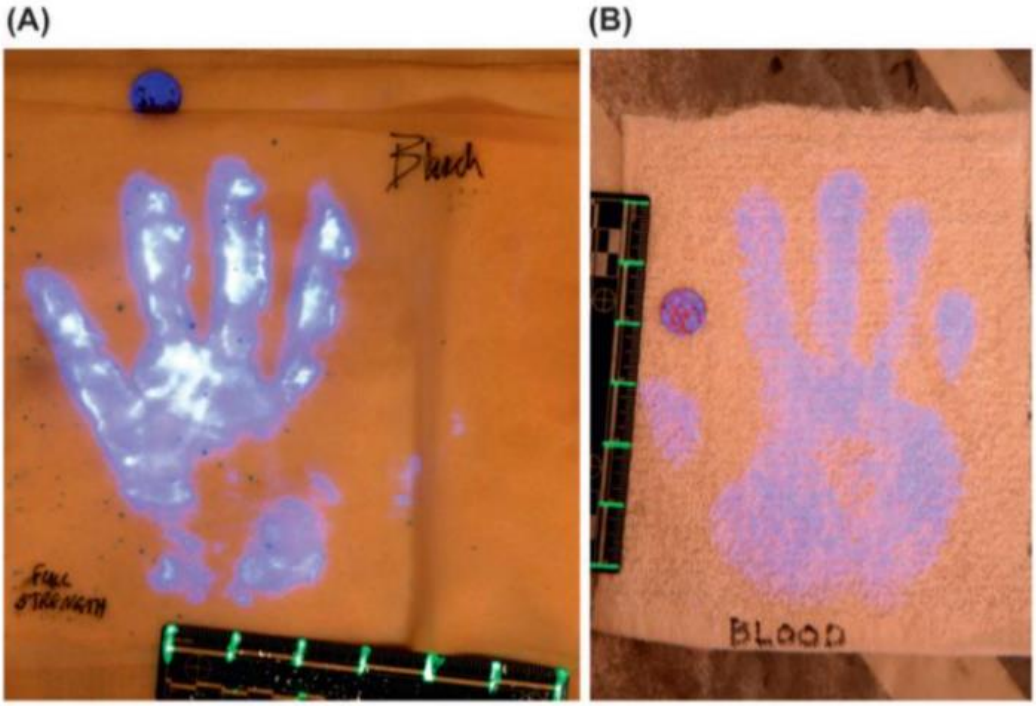


Figure 3*: A side-by-side comparison of the luminol-induced chemiluminescence of A) bleach (hydrogen peroxide) and B) blood. (From Robinson - Crime Scene Photography (2007) [1])*

Examples of confirmatory tests are immunologically based Rapid Stain Identification-tests like RSID™-blood, semen, and saliva (Independent Forensics), which detect antibody-antigen-specific reactions [34]. Although specific, these tests require relatively large amounts of test material to correctly determine its identity, which comes at the expense of the sample material needed for DNA-profiling. Another drawback of this type of confirmatory test is the limitation of testing for only one body fluid at a time. Test material would therefore be needed for several separate tests. These tests are also available a limited number of body fluids and tissue types [10]. This method of testing will therefore cost both equipment and time in addition to exploitation of total test material available.

Building on the strength and weaknesses of existing confirmatory body fluid and tissue type tests mentioned above, it is possible to set certain requirements for a new and ideal test method.

Firstly, a high level of specificity and sensitivity would be of outmost priority. This entails that the test method does not have the potential of cross-reaction and generation of false positives, whilst producing accurate true positives and true negatives despite working with minute

amounts of test material. Secondly, a test method, which does not utilize the material destined for DNA-profiling, while still deriving from the same biological stain or location, will be a significant advantage. Thirdly, the expansion of the variety of body fluids and tissues available for screening. The possibility of distinguishing between similar body fluids, for example venous blood and menstrual blood, will provide further circumstantial evidence during investigation. Finally, when developing new methods, it is obligatory to address the perpetual need to cut time, costs and equipment needed. This entails the prospect of screening for several body fluids and tissue types in the same run or setup.

This is where mRNA- and miRNA-profiling improves upon previous methods. PCR-techniques that are normally used for DNA-profiling, are applied to the detection and identification of RNAs and carry a potential for the reduction of both time and resources spent on body fluid and tissue type analyses.

The history of mRNA and miRNA-profiling

RNA was first mentioned in the forensic literature by Oemichen et al. (1984). This study involves RNA- and DNA-synthesis in post-mortem tissues and showed that RNA-synthesis can provide information about the cells' vitality and functionality [39]. In 1994, Phang et al. successfully amplified reverse transcribed mRNA transcripts from post-mortem tissues by PCR [40]. Real-time quantitative PCR methods, introduced by Heid et al., made it possible to amplify, and simultaneously measure PCR products [41]. Real-time PCR methods enabled the screening for tissue-specific expression patterns, as in the studies of Bauer et al. [42, 43]. They identified matrix metalloproteinase mRNA as a possible marker for menstrual blood in forensic blood stains.

By the utilization of markers for mRNA and miRNA, several studies have shown that it is possible to accurately and efficiently determine the identity of a body fluid or tissue type. Studies such as those by Hanson et al, Fleming et al., and Haas et al. have examined several markers to choose those with highest specificity and reliability for use in body fluid and tissue type identification [44-46]. The development of multiplex assays of mRNA markers have allowed for the simultaneous analysis of several body fluids and tissue types from the same samples. In an extensive cooperation between EDNAP labs across Europe, the marker multiplexes have proven to be highly specific during blind tests. These multiplexes have also

shown to be reliable for use in a forensic workflow, as they give consistent results even with different RNA-extraction methods.

Markers had previously been selected from the monitoring of gene expression in various tissues by micro-arrays of complementary DNAs or sequence specific long oligonucleotides in a chip format (see Figure 4). The most promising gene candidates of these comprehensive expression arrays have been further investigated by checking against existing gene databases. Finally, these candidates have been confirmed through quantitative PCR [4, 24].

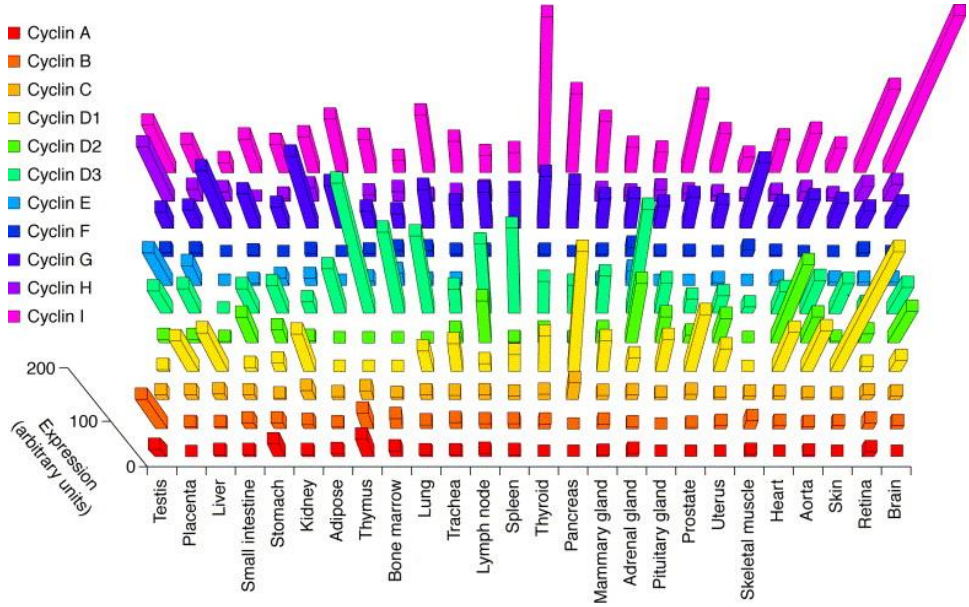


Figure 4: *The result of a microarray of the expression level of genes that encode cyclins, monitored in different tissues. A variation of cyclin-type expression between tissue types is seen as a difference in expression units. For this study, expression levels were normalized to the standard levels of housekeeping genes. (from Gerhold – DNA chips (1999) [4])*

Whole transcriptome sequencing (RNA-Seq) through Massively parallel sequencing (MPS) techniques, is another approach for the selection of novel body fluid markers, as both known and unknown mRNA base sequences are screened [47]. Sequenced reads are aligned to a reference genome with known genes [48]. Whether a mRNA marker is suitable for incorporation into a multiplex, is dependent on its specificity and level of expression. Cross-reaction with non-target body fluids will decrease the marker's reliability. For example, in a case of assault, a cross-reaction of a marker for saliva to vaginal fluid could completely alter the perception and overall outcome. One must also take into consideration that expression levels vary greatly among individuals, due to various aspects such as age, gender, physiology, and

environmental factors. A high level of expression of one marker could potentially influence other markers in a multiplex and disrupt the results [47].

The typical RNA-profiling workflow begins with the extraction of RNA followed by DNase-treatment for the removal of genomic DNA. The RNA sample is then reverse transcribed to complementary DNA, which is amplified through endpoint PCR using marker-specific primers. The amplified sample can then be detected and analysed with capillary electrophoresis, or by quantitative PCR [49].

Newer MPS methods developed in the last decade are also applicable for RNA profiling workflows. This advancement has enabled the development of MPS assays specifically designed for body fluid identification. One such mRNA sequencing assay was developed by Hanson et al., consisting of 33 tissue-specific mRNA markers [49-51].

MPS technologies have several advantages, including the ability to target a larger number of markers in one assay than PCR and capillary electrophoresis methods. This limits the use of available material even further. As MPS methods also utilize shorter amplicons, they will provide reliable results even for degraded or low-level samples [52-54].

The development of RNA and DNA co-isolation methods for the purpose of RNA profiling

The studies of Alvarez et al., and Bowden et al. have focused on developing a method of RNA isolation from the lysis buffer discard, a waste-product of DNA extraction, to enable mRNA profiling without compromising the sample material needed for DNA profiling [5, 55]. In particular, Bowden et al. aimed to detect the presence of RNA purified from the flow through or “waste” product of DNA-extraction, using an RNA-isolation kit. They were able to obtain mRNA profiles from all samples (see Figure 5).

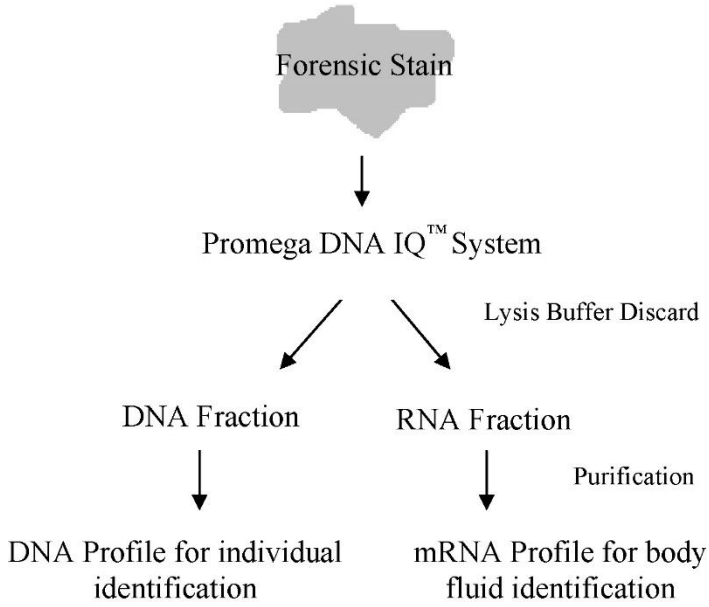


Figure 5*: The workflow utilized in the studies by Bowden et al. (2009). Promega DNA IQ™ method with purification using the Zymo Research Mini RNA Isolation Kit™ II. (from Bowden et al. - A method for DNA and RNA co-extraction for use on forensic samples using the Promega DNA IQTM system* (2009) [5])

The finding of Bowden et al. is a promising step towards an incorporation of a co-isolation method of RNA and DNA, as well as a development of both DNA-profiles and mRNA body fluid/tissue type profiles from the same biological stain. This implies that the biological material used for DNA profiling will not be compromised by the body fluid/tissue type identification.

More recent studies have successfully utilized the QIAamp DNA mini Kit (QIAGEN) and the *mir*Vana™ miRNA Isolation Kit (Applied Biosystems™, Ambion®) for the purpose of DNA/RNA co-isolation and multiplex RNA profiling [11, 12]. These studies were also able to identify forensically relevant body fluids and tissue types with high sensitivity and specificity. Thus, the DNA lysate discard from any DNA extraction kit frequently used in forensic genetic laboratories may have a potential and should be tested for RNA extraction and profiling.

Preliminary analysis at the Centre of Forensic Genetics, UiT, using the PrepFiler *Express*™ Forensic DNA Extraction Kit, showed that RNA was present in both DNA eluate and DNA lysis buffer discard, and is promising for downstream analysis [56].

These findings are the basis for this master's project.

Aims of study

The Center of Forensic Genetics (CFG) seeks to implement a method for body fluid and tissue type identification into their routine workflow, utilizing the lysis buffer discard of the currently used PrepFiler *Express*™ DNA extraction kit for mRNA profiling. The aims of this study were as follows:

- Determine if preliminary findings showing that the PrepFiler *Express*™ DNA lysis buffer discard as well as DNA-eluate may contain RNA that could potentially be used for body fluid identification are reproducible
- Optimize the protocol for RNA-extraction from the PrepFiler *Express*™ DNA lysis buffer discard and compare three commercially available RNA-isolation kits with emphasis on RNA quantity and quality as well as workflow integrity.
- Optimize PCR amplification conditions for a set of published mRNA body fluid markers

Material and methods

Materials

Venous blood, saliva, vaginal fluid, and menstrual blood samples were collected from a volunteer at the Center of Forensic Genetics (CFG). Venous blood was retrieved through venepuncture, and a saliva sample was collected in a test tube and aliquoted. Vaginal fluid and menstrual blood samples were collected on sterile cotton swabs. Semen was donated from the Fertility Outpatient Clinic (In vitro fertilization, IVF) at The University Hospital of North Norway (UNN). All donors were informed that they were contributing to a research project and have given their written consent.

Body fluids chosen for this project were based on their relevancy in forensic work.

Samples were stored at -70°C a soon as possible after collection for deactivation of RNases and reduction of degradation.

1. The potential of extracting RNA from different steps when using the PrepFiler *Express*™ DNA extraction kit

As earlier mentioned, it is advantageous in terms of conserving test material destined for DNA analysis if RNA can be extracted from DNA-extraction waste product for the purpose of downstream RNA analysis. The PrepFiler *Express*™ Forensic DNA Extraction Kit, is a DNA-extraction method integrated into routine workflows at the CFG. We therefore want to examine if RNA is present in DNA-extraction steps from this kit, and whether this RNA may be qualitatively suited for downstream analysis.

The starting point for this study were preliminary results from Gårdvik (2020) that showed that RNA was present in the DNA eluate and lysis buffer discard when extracting DNA from blood, semen and saliva with the PrepFiler *Express*™ Forensic DNA Extraction Kit [56]. Two sets of triplets of the three body fluids had undergone RNA quantification by two methods of RNA concentration measurement, and quality analysis by measurement of DV_{200} and RIN. RNA isolated from pure body fluid with the *mir*Vana™ miRNA Isolation Kit (Invitrogen), was used as a comparison to the samples generated by the PrepFiler™ DNA extraction, both in terms of RNA quantity and quality. These methods were repeated in this study to see if the results were reproducible.

1.1 Sample preparation

Two sets of triplets of 40 µL from each of the three body fluids were prepared from the same stock as used by Gårdvik (2020). Samples had been kept for approximately six months in a -70°C freezer. One set of triplets was pipetted onto PrepFiler™ LySep columns with collection tubes attached (Thermo Fisher Scientific), for a routine DNA extraction. The other set of body fluid triplets was pipetted into DNA LoBind Tubes (Eppendorf) for use in RNA extraction.

1.2 DNA extraction

All DNA extractions were carried out with the PrepFiler *Express*™ Forensic DNA Extraction Kit and the semi-automated AutoMate *Express*™ Forensic DNA Extraction System (Thermo Fisher Scientific) according to manufacturer's protocol [57, 58]. Lysis solution was added to each LySep column containing up to 40 µL body fluid, which were tightly closed and placed on a thermomixer at 70°C and 750 revolutions per minute (rpm) for a total of 40 minutes. LySep columns were centrifuged at 10 200 rpm for 2 minutes to collect the lysate in the attached collection tubes.

The sample tubes with lysate were then loaded onto the rack of the AutoMate *Express*™ Forensic DNA Extraction System instrument, together with pipetting tips, collection tubes and reagent cassettes for the automated extraction step. An elution volume of 50 µL was chosen for all samples.

Succeeding the automated extraction was the collection and storage of elution tubes containing DNA eluate, as well as the transfer of lysis buffer discard (hereafter referred to as LBD) from the reagent cassette to respective LoBind-tubes. LBD in the cassette has an approximate volume of 700 µL. All samples were stored at -70°C, until further quantification and quality analyses were performed.

1.2.1 Principles of the method

The PrepFiler *Express*™ Forensic DNA Extraction kit contains a lysis solution of lysis buffer and dithiothreitol (DTT) - a strong reducing agent which stabilizes enzymes and reduces disulphide bonds - which is added to LySep columns containing biological test material such as cotton swab tips, body fluid or other tissue samples [59, 60].

Lysis and denaturation of cell components for the dissociation of free DNA is performed in LySep columns on a thermomixer. Following the lysis step is a centrifugation of the LySep columns with attached collection tubes for the separation of lysate from any solid components.

As there are no solid components in the pure body fluids used in the DNA-extraction, this step has no function in the workflows of this study.

As a result of the centrifugation, the lysate is collected in the collection tubes, and the LySep columns containing solid parts can be discarded. Collection tubes with lysate can then be loaded onto the AutoMate *Express*™ Forensic DNA Extraction System.

The AutoMate *Express*™ System incorporates a magnetic syringe unit for the separation of DNA from other biological components. The reagent cassette contains washing buffers and magnetic beads that are mixed with the samples by the instrument. As DNA is a negatively charged molecule, it will be fixated onto the magnetic beads that are held along the sides of the pipette tips, as other cell components, waste products and impurities are washed away during the extraction process. The lysis buffer discard will be disposed within a well of the cassette, whilst DNA is eluted into the elution tubes [58].

1.3 RNA extraction

The other set of triplets from blood, saliva and semen underwent a routine RNA extraction with the *mir*Vana™ miRNA Isolation Kit (Invitrogen) according to the manufacturer's protocol for total RNA [61].

The RNA was eluted with nuclease-free water that was heated to 95°C, for a total volume of 70 µL. All RNA eluate samples were stored at -70˚C, until further quantification and quality analyses were performed.

1.3.1 Principles of the method

The *mir*Vana™ miRNA Isolation Kit combines both methods of RNA purification by organic extraction and solid-phase extraction. The samples are lysed by the addition of a lysis solution for the inactivation of RNases and stabilisation of RNA. Homogenate additive further ensures the homogenisation of the lysed sample. The organic extraction step consists of the addition of Acid-Phenol:Chloroform, followed by a centrifugation step to separate the organic from the aqueous phase. Because of its negative charge, RNA will have an affinity towards a hydrophilic environment and separate into the aqueous phase, whilst other cellular components are fixed within the organic phase. The aqueous phase is pipetted after centrifugation and transferred into a new LoBind tube, whilst the remaining organic phase can be discarded.

The solid-phase step of extraction involves increasing the RNA molecules' affinity for solid support by the addition of absolute ethanol (>99.8 % EtOH). The appropriate volume of EtOH is adjusted to the total volume of aqueous upper layer retrieved during phase separation, in the ratio 1:1.25 (aqueous layer:EtOH). Then, the mixture is transferred to collection tubes with filter cartridges, 700 µL at a time, and centrifuged for the fixation of RNA to the glass fibre filters. The filters with fixed RNA then undergo three wash steps for the removal of waste products and impurities before elution of RNA with nuclease free water.

1.4 RNA quantification

The RNA in the samples was quantified using both the Qubit® RNA HS assay on the Qubit® 4 Fluorometer (Invitrogen) according to manufacturer's protocol, and the Agilent RNA 6000 Pico Kit on the Agilent 2100 Bioanalyzer system (Agilent) [62]. The latter analyses were carried out by the Genomics Support Center Tromsø (GSCT).

A High Sensitivity RNA measurement of nanograms per µL was conducted with the Qubit® 4 fluorometer with 10 µL sample input, and all samples were measured three times for the calculation of a mean value. RNA quantity was measured in picograms per µL by the Agilent 2100 Bioanalyzer system (Agilent), and results were converted into nanograms per µL for further comparisons.

1.4.1 Principles of the method

The Qubit® 4 Fluorometer is a benchtop fluorometer that conducts highly sensitive measures of fluorescence-based quantity and quality assays of biological molecules. The Qubit® RNA HS assay uses a fluorescent dye reagent as a tag for RNA molecules, where a stronger signal of fluorescence indicates a higher concentration of RNA in the sample. The fluorescent dye reagent is diluted with a buffer provided in the assay kit to form a working solution and up to 20µL of sample can be added to working solution for a total of 200µL assay volume. Standards for calibration of the Qubit® 4 Fluorometer, also provided in the kit, are prepared and measured for calibration before measuring samples [62].

The Agilent 2100 Bioanalyzer system (Agilent) automated instrument serves as a tool for measurement and analysis in biochemical workflows, based on electrophoretic separation of biomolecules. With the Agilent RNA 6000 Pico Kit, the Bioanalyzer system separates nucleic acid fragments by their respective sizes into gel-filled interconnected microchannels on a microfluidic chip. The chip contains fluorescent dyes that will bind to nucleic acids and

produce a measurable laser-induced fluorescent signal, which will be depicted in an electropherogram in a size/signal strength-based manner [63, 64].

1.5 RNA quality analysis – RIN and DV_{200}

The 2100 Bioanalyzer system was also used to measure RNA quality, namely the RNA Integrity Number (RIN) and DV_{200}. These values will give an indication to the degree of degradation of the RNA in the samples.

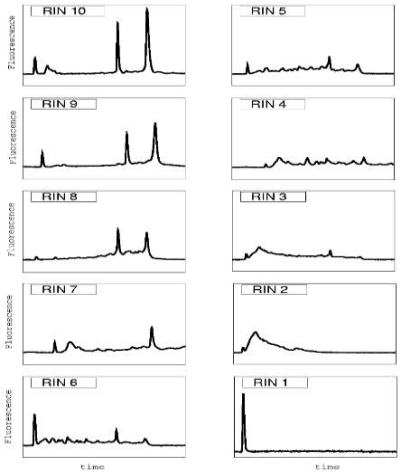
1.5.1 Principles of the method

Biological material in trace evidence from crime scenes is very likely to be exposed to several degradative factors, such as chemicals, oxidation, UV-radiation, and of course innate enzymes (RNases and DNases). Strongly degraded RNA may not be fit for the use in further analysis, and we therefore performed two measurements of degradation.

Agilent developed the RIN value algorithm for a standardized measurement of RNA integrity. Using machine learning methods, the algorithm analyses the features of the measurements and compares them to already established expert-assigned categories. In this way, as the method adapts and learns with increasing empirics, the measurement of RNA integrity becomes user independent. RNA integrity is measured on a scale from one to ten, with one indicating a total RNA degradation, and ten a totally intact sample (See Figure 6A). A sample RIN value between seven and ten is considered ideal for the purpose of downstream analysis. If a sample has a RIN value between four and six, there is no guarantee that further analyses will give satisfactory results. Finally, a sample with a RIN value between one and three indicates high levels of degradation, and is not suited for downstream analysis ([63, 65].

DV_{200} is a metric developed by Illumina, originally for usage in measurement of RNA degradation in samples originating from formalin-fixed paraffin-embedded tissue. Its purpose is to measure fragment size distribution by indicating the proportion of total RNA in a sample with a length of minimum 200 nucleotides. A length below this threshold would be regarded as partially or totally degraded RNA. An overview of DV_{200} categories can be seen in Figure 6B [66, 67].

A) B)



DV_{200} proportion	Quality
>70%	High
50-70%	Medium
30-50%	Low
<30%	Too degraded

Figure 6*: A) RIN quality categories 1-10 represented in electropherograms of fluorescence/time. The RNA shifts- from shorter fragments 1(degraded) to longer fragments 10 (intact) (From Schroeder - The RIN: an RNA integrity number for assigning integrity values to RNA measurements [62]) B) DV_{200} quality categories in relation to DV_{200} proportion* [67].

2 Comparison of three commercially available RNA-isolation kits for extraction from PrepFiler™ LBD

One of the aims of this project was to optimize the protocol of commercially available RNA-extraction kits for the extraction of RNA from DNA-extraction lysis buffer discard. This was to determine which kit was the most effective for this purpose. The kits compared in this study were the *mir*Vana™ miRNA Isolation Kit, the Direct-zol™ RNA Miniprep kit (Zymo Research Corporation), and the ReliaPrep™ RNA Miniprep System (Promega Corporation), which were chosen based on their successful use in previous studies [5, 48]. The comparison of the kits was based on RNA yield and level of degradation in the final RNA-eluates, as well as the overall workflow efficiency when working with the LBD from the PrepFiler *Express*™ Forensic DNA Extraction Kit.

2.1 Preparation of sample stock and means of measurement

40 µL blood was used in a standard DNA-extraction with the PrepFiler *Express*™ Forensic DNA Extraction kit as described in paragraph 1.2, with the maximal number of samples tubes (13 tubes) on the AutoMate *Express*™ Forensic DNA Extraction System. A sample stock was

prepared by collecting the LBD from all 13 reagent cassettes in a 50 mL centrifuge tube which was thoroughly vortexed to achieve a homogenous solution. This would enable a direct comparison of RNA-isolation kit yield and quality in downstream analysis. The stock was pipetted into DNA LoBind tubes in 500 µL aliquots and stored at -70°C until further use.

The Qubit® RNA HS assay on the Qubit® 4 Fluorometer was used to measure the RNA concentration and to determine the efficacy of the optimization of all three kits (see paragraph 1.4).

2.2 Optimization of the *mir*Vana™ miRNA Isolation Kit

The *mir*Vana™ miRNA Isolation Kit is developed for miRNA extraction, with an additional protocol for total RNA extraction from tissues and pure body fluids and yields RNA even from small volumes. The manufacturer's protocol is not adapted for extraction from DNA-extraction LBD. Therefore, the protocol needs to be optimized for this purpose. As forensic samples are expected to yield low concentrations of RNA, the aim was to utilize the highest possible input volume of LBD for RNA extraction. Volumes of water and Lysis/Binding Buffer of the *mir*Vana™ kit had in previous attempts been added to the LBD to increase its polarity, as it was suspected that a large majority of solutes and substances ended in the organic phase, leading to a very small aqueous phase. In addition, the volume of LBD combined with the RNA extraction kit reagents, needed to be within the input limit of the 2 mL DNA LoBind tubes. The volume of three components in the workflow of the *mir*Vana™ Kit may be varied for a higher RNA yield: the LBD-input, the Lysis/Binding Buffer and phenol:chloroform.

All extractions were performed according to the procedure described in paragraph 1.3. Aliquots of LBD were thawed on ice, thoroughly vortexed, and appropriate volumes were pipetted into new DNA LoBind tubes. The volume ratios of LBD, the Lysis/Binding Buffer and phenol:chloroform were varied according to Table 1. All extractions were performed in triplets. The most promising ratios were re-tested for more accurate results (highlighted in Table 1). The reason for re-testing the 1:1:2 ratio is based on the manufacturer's recommendation of a 1:1 ratio between lysed sample and phenol:chloroform [61].

Table 1*: Ratios and volumes of LBD : lysis/binding : phenol-chloroform (µL) used during the optimization of the mirVana™ miRNA Isolation kit. Ratios marked with * were re-tested.*

LBD : lysis/binding buffer : phenol-chloroform	
Ratio	**(µL)**
(1:1:2)	450:450:900*
(1:1:0.8)	500:500:400
(1:1:1.2)	500:500:600
(1:1:1.6)	500:500:800*
(1:1:2)	500:500:1000
(1:0.82:1.45)	550:450:800
(1:0.67:1.33)	600:400:800
(1:0.67:1.67)	600:400:1000
(1:0.43:1.43)	700:300:1000

As the extraction process with the *mir*Vana™ Kit involves a high number of transfers, in addition to the use of toxic phenol:chloroform [68], it is not only a laborious and time-consuming procedure, but also poses a health and safety risk for the laboratory operant. It is therefore interesting to explore safer and less time-consuming options.

The studies of Bowden et al. and Han Lin et al. have successfully extracted mRNA from DNA-extraction waste products with two other kits, the Zymo Research Mini RNA Isolation KitTM II (Ngaio Diagnostics) and the ReliaPrep™ RNA Miniprep System (Promega Corporation) [5, 48]. The former kit is no longer in production, and therefore the Direct-zol™ RNA Miniprep kit was chosen as a substitute.

2.3 Direct-zol™ RNA Miniprep kit optimization

As with the *mir*Vana™ Kit, the Direct-zol™ RNA Miniprep kit's protocol is not optimized for the extraction of RNA from the LBD of the DNA-extraction process. For optimization, the approach in paragraph 2.2 was used as a guideline. The sequential steps of RNA-isolation with the Direct-zol™ RNA Miniprep kit was performed according to the manufacturers protocol, but the volume ratio of LBD, TRIzol Reagent, and EtOH was varied according to Table 2 [69]

Table 2*: Ratios and volumes of LBD : TRIzol Reagent : EtOH (µL)used during the optimization of the Direct-zol™ RNA Miniprep kit.*

LBD : TRIzol Reagent : EtOH	
Ratio	**(µL)**
1:1.88:1.88	400:750:750
1:3:3	100:300:300
1:3:3	200:600:600
1:3:4	200:600:800
1:3.5:4.5	200:700:900
1:0.5:3.5:4.5	200:(100uL H_2O):700:900

The Direct-zol™ RNA Miniprep kit (Zymo Research Corporation) utilizes TRIzol® reagent (Total RNA Isolation Reagent) for the isolation of RNA. TRIzol reagent is an acid-guanadine-phenol mix in monophase that eliminates the need for phase-separation by phenol:chloroform or alcohol precipitation [70]. It also limits the amount of transfers and wash-steps needed or RNA-isolation. The kit is supplied with an in-procedure DNase treatment which can be performed directly on the Zymo-Spin™ IICR Column [69].

The manufacturer's protocol involves a lysis step with TRIzol reagent, before adding EtOH, which is followed by a centrifugation through a Zymo-Spin™ IICR Column in a collection Tube, and lastly three wash steps before elution in nuclease free water.

2.4 ReliaPrep™ RNA Miniprep System optimization

The sequential steps of RNA-isolation with the ReliaPrep™ RNA Miniprep System (Promega Corporation) was performed according to the manufacturers protocol, with the LBD, BL + TG buffer and isopropanol volumes added at different ratios, according to Table 3 [71].

Table 3*: Ratios and volumes of LBD : BL + TG buffer: isopropanol (µL)used during the optimization of the ReliaPrep™ RNA Miniprep System*

LBD : BL + TG buffer: isopropanol	
Ratio	**(µL)**
1:4:1.7	100:400:170
1:0.67:1.77	300:200:170
1:0.25:0.43	400:100:170
1:0:0.34	500:0:170

The ReliaPrep™ RNA Miniprep System is an RNA-isolation kit which does not utilize phenol:chloroform extractions or ethanol precipitations, making it a safer RNA-isolation kit option. It combines a buffer containing guanidine thiocyanate and 1-thioglycerol (BL + TG buffer) for the lysis of cell components and inactivation of endogenous RNases with an addition of isopropanol for an increased polarity of the solution. The lysed sample is filtered through a ReliaPrep™ Minicolumn with a collection tube attached, followed by a DNase treatment (incubation of 15 minutes at room temperature) included in the kit and additional three wash steps before elution in nuclease free water [71].

RNA-isolation with the ReliaPrep™ was performed in accordance with the in-house protocol of the Institute of Environmental Science and Research Limited (ESR), New Zealand, which

was kindly provided by Dr Sally Ann Harbison, and is slightly different from the manufacturer's protocol (see Appendix 3). This includes eliminating the lysis step with the BL + TG buffer, as the LBD is already lysed, as well as a longer incubation time for the DNase treatment (40 minutes at 37°C).

The in-house protocol from ESR, is developed to befit the RNA-isolation from LBD. This protocol was utilized in all the following RNA-isolation with the ReliaPrep™ Kit and reflects the 500:0:170 volume ratio in Table 3.

2.5 DNase-treatment

To further analyse the RNA, all genomic DNA (gDNA) must be eliminated from the sample. If a substantial amount of gDNA is present during quantitative PCR (qPCR), this may be amplifiable and show up as a signal alongside any results deriving from cDNA from RNA, leading to unreliable results.

On-column DNase treatment is included in both the Direct-zol™ RNA Miniprep kit and the Promega™ ReliaPrep™ RNA Miniprep System, and these samples are treated during the RNA-extraction process.

The *mir*Vana™ Kit does not include a DNase treatment, and all samples extracted were treated with the TURBO™ DNase (2 U/µL) (Invitrogen), according to the manufacturer's protocol for a standard DNAase-treatment [72].

2.6 Reverse transcriptase reaction

For the use in real time PCR (qPCR), extracted RNA from all three RNA-isolation kits was reversely transcribed to complementary DNA (cDNA). cDNA was prepared from DNase-treated samples using the reverse transcriptase SuperScript™ IV VILO™ Master Mix (Invitrogen) according to the manufacturer's protocol, using 16 µL of RNA in the reaction [73]. Two reverse transcription parallels were prepared from each RNA sample: a RT+ parallel with reverse transcriptase enzyme and an RT- parallel with no reverse transcriptase enzyme. RT-parallels were run for detection of any contaminating gDNA. An RNA control standard was prepared using the Applied Biosystems™ Total RNA Control (Human) for RT reaction with an input of 2 µL.

2.7 Real time PCR (qPCR) of the housekeeping gene GAPDH

qPCR of a housekeeping gene was performed for the comparison of relative RNA quantity in samples from the three RNA-isolation kits. The housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH), was chosen for qPCR as endogenous control, as its expression levels have been shown to be rather stable under differing experimental conditions [74]. As GAPDH is expressed in all cell types, qPCR results are comparable in terms of mRNA quality between the three RNA-isolation kits.

qPCR was performed for GAPDH, using the primers and probe from Juusola & Ballentine (2007), TaqMan™ Fast Advanced Master Mix (Applied Biosystems™), and cDNA template in the volume ratios shown in Table 4 [75]. PCR samples were prepared in a MicroAmp™ Optical 96-Well Reaction Plate (Applied Biosystems™), which was covered with MicroAmp™ Optical Adhesive film (Applied Biosystems™). Reactions were run on the Applied Biosystems™ 7500 Real-Time PCR System using the HID Real-Time PCR Analysis Software v1.2, according to the run method described in the manufacturer's protocol [76].

A random selection of samples from all three RNA-isolation kits was chosen for GAPDH qPCR. A standard curve with five concentrations was made using a 1:10 dilution series from the RNA control standard cDNA.

Table 4: *Reagents and their respective volumes, as well as primer and probe (with NED dye) concentrations and sequences used for qPCR of GAPDH.*

Reagents		**Volume per well**
TaqMan™ Fast Advanced Master Mix		10 µL
	Sequence (5'-3')	
20µM Forward primer GAPDH	AAT GGA AAT CCC ATC ACC ATC TT	0.9 µL
20µM Reverse primer GAPDH	GCG GGG TGA TCA AAA ACC	0.9 µL
Probe GAPDH	NED-CAG GAG CGA GAT CC	0.5 µL
cDNA template		2 µL
Nuclease free water		5.7 µL
Volume per well		**20 µL**

2.7.1 Principles of the method

Polymerase chain reaction (PCR) amplifies DNA through the annealing of sequence-specific primers to a template strand, followed by a synthesis of a complimentary strand by DNA polymerase. Real time PCR, or quantitative PCR (qPCR), combines both amplification and detection (quantitation) of DNA present in a sample. This is achieved by the addition of a sequence-specific probe, which will emit fluorescence as it is cleaved by DNA polymerase activity during the extension step of the PCR. The amount of DNA present will therefore be correlated to the fluorescence intensity, which will be detected at the point in time when it supersedes the background fluorescence, referred to as the cycle threshold (C_t) [41, 77]. The higher amount of DNA is present in a sample, the earlier this threshold is reached and consequently the lower the C_t-value. A standard of known quantity is used to develop a standard curve to which C_t-values of samples are compared to calculate relative quantity.

3 A comparison of ReliaPrep™ kit and *mir*Vana™ kit for blood, saliva and semen

Based on the performance of the three RNA extraction kits in paragraph 2, a decision was made to only continue with the *mir*Vana™ miRNA Isolation Kit and the ReliaPrep™ RNA Miniprep System in further analyses. A comparison of the quantitative and qualitative performance of these two RNA-isolation kits on lysis buffer discard from 4 µL and 40 µL of blood, saliva, and semen was performed with emphasis on RNA yield and level of degradation, as well as gDNA contamination, and overall workflow efficiency. An overview of the workflow is shown in Figure 7.

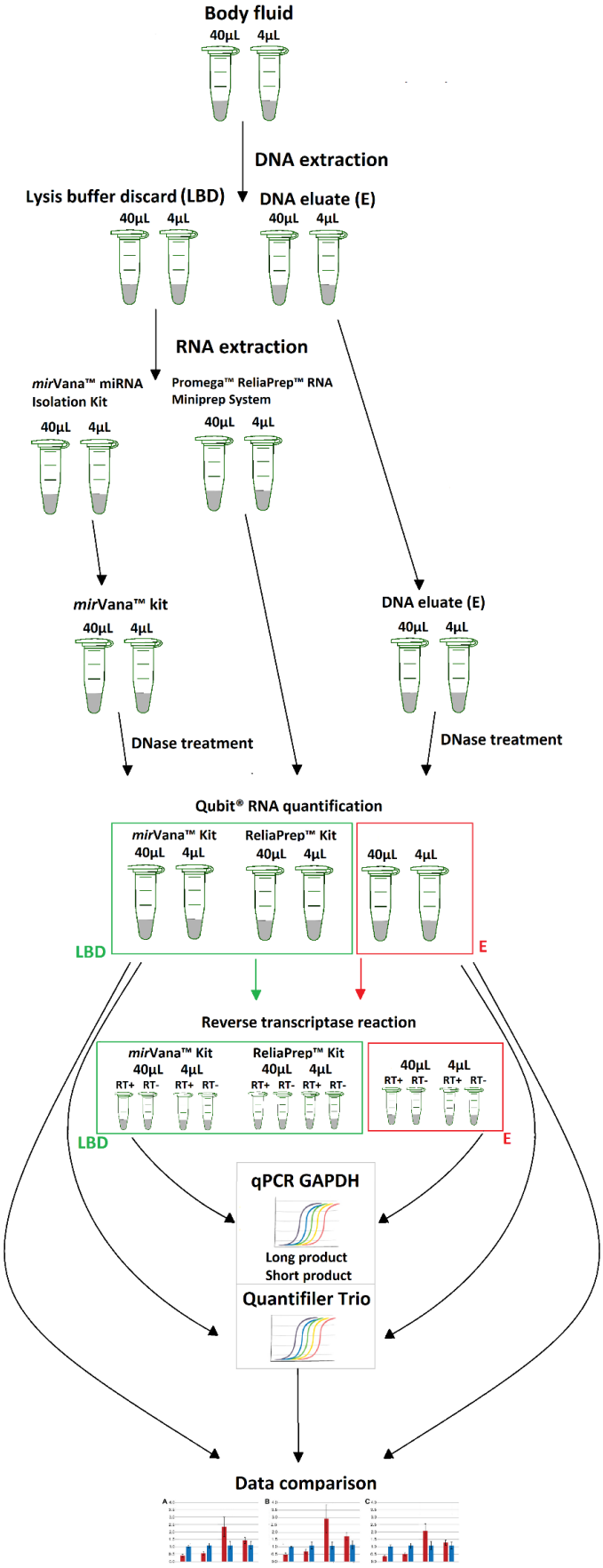


Figure 7*: A flowchart of the workflow for paragraph 3. The figure displays sample preparation (DNA- and RNA-extraction, DNase-treatment, reverse transcriptase reaction), and qPCR using GAPDH short and long product, as well as with the Quantifiler™ Trio DNA Quantification Kit.*

3.1 DNA extraction and lysis buffer discard preparation

Volumes of 4 µL and 40 µL of each body fluid were pipetted onto nine LySep columns with collection tubes attached, and DNA extractions were performed as described in paragraph 1.2. Triplets of cassettes of LBD were reserved for the *mir*Vana™ Kit extractions and ReliaPrep™ Kit extractions, respectively. Tubes of DNA eluate were combined accordingly and labelled to the corresponding cassette LBD. This was done to directly compare values obtained from both LBD and eluate.

3.2 RNA extraction and sample preparation

An RNA extraction with the *mir*Vana™ Kit as described in paragraph 1.2 was performed from 500 µL LBD of each body fluid. 500 µL Lysis/binding buffer, and 800 µL of phenol:chloroform was used during the extraction process, as this is the optimal volume ratio from the optimization experiments (paragraph 2). RNA extraction eluates, as well as DNA extraction eluates were DNase-treated as described in paragraph 2.5. This included all negative extraction controls. An RNA extraction with the ReliaPrep™ Kit was performed from 500 µL LBD of each body fluid as described in paragraph 2.4, using 170 µL isopropanol.

All DNase-treated RNA-eluates were quantified for RNA concentration with the Qubit® 4 Fluorometer. RNA eluates from all body fluids and both kits were used to prepare RT+ (samples with reverse transcriptase enzyme) and RT- (samples with no reverse transcriptase enzyme) samples in a reverse transcriptase reaction as described in paragraph 2.6.

3.3 Assessment of RNA degradation by qPCR

To reveal the level of degradation of the RNA in the samples, primers for a longer GAPDH product (114 bp) were used. The results for the longer product would serve as comparison to those of the shorter product (58 bp), as degradation typically leads to more fragmented polynucleotides.

RT+ and RT- for all samples (both RNA extractions and DNA extraction eluate) were run for the short GAPDH product as described in Table 4 and paragraph 2.7. A qPCR setup for the longer GAPDH product was prepared from RT+ of all samples according to the volume ratios in Table 4, with the same reverse primer and probe, but a different forward primer (5'-**CCA CCC ATG GCA AAT TCC)**. This longer GAPDH product forward primer was designed with Primer Express™ Software v3.0.1 (Applied Biosystems), and cross-checked with primer-BLAST (NCBI), and OligoEvaluator™ (Sigma-Aldrich®). Short and long GAPDH products were run

according to the run method described in the manufacturer's protocol [76]. Relative quantities from qPCR results were used to calculate a short/long GAPDH product ratio number (Formula 1). A ratio number < and closer to 1 will imply intact GAPDH products, as opposed to a ratio number >1, which will suggest more fragmented (degraded) long GAPDH products. This will act as a general representation of RNA degradation levels in the samples.

Formula 1 $$\frac{Relative\ quantity\ \boldsymbol{short\ GAPDH\ product}}{Relative\ quantity\ \boldsymbol{long\ GAPDH\ product}} = ratio\ of\ GAPDH\ product\ intactness$$

3.4 Detection of genomic DNA contamination with the Quantifiler™ Trio kit

As mentioned in paragraph 2.6, samples without no reverse transcriptase was used to examine for any gDNA contamination by a detectable reaction on qPCR. To further evaluate the efficacy of DNase-treatments utilized in this project, DNase-treated RNA extraction eluates, as well as the DNase-treated DNA extraction eluates were run with the highly sensitive Quantifiler™ Trio DNA Quantification Kit according to the manufacturer's protocol, for detection of potential gDNA contamination [78]. As the aim is to determine any gDNA contamination, and not the nature of the contaminating DNA, only the results for the small autosomal targets of the Quantifiler™ Trio kit, which is used as a measure of quantity, were used.

3.4.1 Principles of the method

The Quantifiler™ Trio DNA Quantification Kit amplifies multiple copies of small and large autosomal loci, as well as several loci on the Y-chromosome for quantitation of male DNA present. These loci have been chosen as they have conserved primer- and probe-binding sites within the genome, and the results are therefore not highly influenced by variability between individuals. The kit is highly sensitive and will detect tiny amounts of amplifiable DNA [78].

4 A singleplex of twelve body fluid markers for optimization of PCR and capillary electrophoresis

Twelve body fluid markers for four types of body fluids and two markers for housekeeping genes (see Table 5) utilized in previous studies were chosen for this study. These markers were selected for their proven specificity and sensitivity, and primers were run on four dilutions of body fluids under three different PCR programs. As PCR and capillary electrophoresis of body fluid markers had not previously been performed at CFG, there was a need to optimize, incorporate, and establish this method into current workflows. A singleplex of one marker run on the corresponding body fluid would be the first approach in establishing the optimal PCR run method, primer concentrations and settings of capillary electrophoresis assay settings.

4.1 Sample preparation

RNA-extraction from pure body fluid was performed using the *mir*Vana™ Isolation Kit as described in paragraph 1.3 on 40 µL of the body fluids venous blood, saliva and semen, and one, whole cotton swab was used per sample of vaginal fluid and menstrual. DNase-treatment and RT-reaction was conducted with all samples as described in paragraph 2.5 and 2.6. cDNA was diluted in nuclease free water in a dilution series of 1:10:100:1000. The purpose of this dilution step was to enable an estimate of maximum material input level in downstream capillary electrophoresis.

4.2 Body fluid marker PCR

The primers of the markers listed in Table 5 were mixed with their corresponding body fluid cDNA, 2x QIAGEN Multiplex PCR Master Mix, and RNase-Free Water in volumes according to Table 6 [11, 79]

Table 5: *Overview of markers used in the singleplex PCR reactions, with corresponding tissue types, primer concentrations, primer sequences, amplicon sizes and references to the studies from which these were collected [11, 79].*

Marker name	Tissue	[primer] µM	Forward primer (5'-3') Reverse rimer (5'-3')	Expected fragment length	Dye	Ref
18S-rRNA	Housekeeping gene	2.0	**PET**-CTC AAC ACG GGA AAC CTC AC CGC TCC ACC AAC TAA GAA CG	110 bp	PET™	[11]
ACTB	Housekeeping gene	2.0	**PET**-TGA CCC AGA TCA TGT TTG AG TTT CGT ACA GGG ATA GCA CAG[1]	78 bp	PET™	[11]
HBB	Venous blood	2.0	**FAM**-GCA CGT GGA TCC TGA GAA C ATG GGC CAG CAC ACA GAC	61 bp	FAM™	[11]
SLC4A1	Venous blood	2.0	**NED**-AAC TGG ACA CTC AGG ACC AC TGGA TGT CTG GGT CTT CAT ATT CCT[2]	103 bp	NED™	[79]
STATH	Saliva	2.0	**FAM**-TTT GCC TTC ATC TTG GCT CT CCC ATA ACC GAA TCT TCC AA	93 bp	FAM™	[11]
FDCSP	Saliva	2.0	**NED**-CTC TCA AGA CCA GGA ACG AGA A GGG CAG ATT CAG GTA TTG GAA TAG[3]	170 bp	NED™	[79]
PRM1	Spermatozoa	2.0	**NED**-AGA CAA AGA AGT CGC AGA C TAC ATC GCG GTC TGT ACC	91 bp	NED™	[11]
TNP1	Spermatozoa	2.0	TTT GAT GAC GCC AAT CGC AAT TAC C **FAM**-TTCCT TCT GCT GTT CTT GTT GCT G[1]	107 bp	FAM™	[79]
MSMB	Seminal fluid	2.0	CTT TGC CAC CTT CGT GAC TTT ATG **PET**-TACA GTT GTC AGT CTG CCA CT[4]	143 bp	PET™	[79]
KLK3	Seminal fluid	2.0	**PET**-GAC GTG GAT TGG TGC TGC ACC CTT CTC GCA CTC CCA GCC TC	64 bp	PET™	[11]
CYP2B7P	Vaginal fluid	2.0	**VIC**-AGT CTA CCA GGG ATA TGG CAT G CTA TCA GAC ACT GAG CCT CGT CC	146 bp	VIC®	[11]
MUC4	Vaginal fluid	2.0	**FAM**-CTG CTA CAA TCA AGG CCA AAG GGA AGT TCT AGG TTG AC	141 bp	FAM™	[11]
MMP7	Menstrual blood	2.0	**VIC**-GAA CAG GCT CAG GAC TAT CTC TTA ACA TTC CAG TTA TAG GTA GGC C[1]	127 bp	VIC®	[11]
MMP10	Menstrual blood	2.0	**NED**-TA CCC ACT CTA CAA CTC ATT CAC AGA G TTGGT TCC TCA GTA GAG GCA GG[2]	112 bp	NED™	[79]

Comments:
1. Underlined nucleotide was added as a 5' tail for better spacing in multiplex
2. Underlined nucleotide was added as a 5' tail for better spacing in multiplex. HEX™ dye swapped for NED™.
3. HEX™ dye swapped for NED™.
4. Underlined nucleotide was added as a 5' tail for better spacing in multiplex. FAM™ dye swapped for PET™.

Table 6*: Reagents and respective volumes, as well as primer concentrations used in PCR for body fluid markers.*

Reagents	**Volume per reaction**
2x QIAGEN Multiplex PCR Master Mix	12.5 µL
2µM Forward primer	2.5 µL
2µM Reverse primer	2.5 µL
cDNA template	1 µL
Nuclease free water	6.5 µL
Volume per reaction	**25 µL**

Markers with all dilutions of body fluids were run on the Veriti Thermo Cycler (Thermo Fisher Scientific) with two PCR run methods from Albani et al. and Lindenbergh et al. [12, 47], called NZ and NL60 here, and an in-house PCR program from the Netherlands Forensic Institute, which was kindly provided by Professor Titia Sijen, called NZ64 here (see Table 7).

Table 7*: PCR run methods used for body fluid markers*

	NZ			**NL60**			**NL64**		
	°C	Time		°C	Time		°C	Time	
Stage 1	95°C	15 min		95°C	15 min		95°C	15 min	
Stage 2	94°C	30 sec		94°C	20 sec		94°C	20 sec	
	60°C	3 min	35 cycles	60°C	30 sec	33 cycles	64°C	30 sec	33 cycles
	72°C	1 min		72°C	40 sec		72°C	40 sec	
Stage 3	72°C	10 min		60°C	45 min		60°C	45 min	
	4°C	∞		4°C	∞		4°C	∞	

4.3 Capillary electrophoresis

PCR products from cDNA were added to a mixture of Hi-Di™ Formamide and GeneScan™ 600 LIZ® Size Standard v2.0 (Applied Biosystems™), in volumes according to Table 8, on a MicroAmp™ Optical 96-Well Reaction Plate. With plate septa attached, plates were spun briefly, then denatured at 95°C for 3 minutes, and snap-cooled on a cooling block for at least 3 minutes before being placed on the 3500xL Genetic Analyzer (Applied Biosystems™) [80].

Amplified PCR products were separated and detected by the 3500xL Genetic Analyzer with Data Collection software v3.1, using a 36 cm capillary array, performance optimized polymer 4 (POP-4™ Applied Biosystems™) as separation matrix, and default instrument parameters for fragment analysis. PCR samples were injected at 1.2 kV for 24 seconds.

Expected fragment lengths as shown in Table 5 are only theoretical. Observed fragment lengths may be different as they can be affected by molecular, environmental, and systemic factors. Therefore, a deviation of ±6 bp from the theoretical fragment length was interpreted as "expected base pair area". Profiles were analysed with GeneMapper ID-X v1.5 (Applied Biosystems™) in order to find fragments of expected bp length for the respective body fluids.

Table 8*: Reagents and respective volumes used per sample in capillary electrophoresis on the 3500xL Genetic Analyzer*

Reagents	**Volume**
Hi-Di Formamide	9.5 µL
GeneScan™ 600 LIZ® Size Standard v2.0	0.5 µL
PCR product	1.0 µL
Total per sample	**11.0 µL**

4.3.1 Principles of the method

Capillary electrophoresis is a DNA separation technique that outranks previous slab gel electrophoresis by allowing a much shorter analysis time under higher voltages, in addition to eliminating the need for further analysis as quantitative data is electronically available as a part of the run software [27, 64]. Glass capillaries, about 50µm in diameter, filled with a flexible polymer serving as the sieving matrix, enables the use of only a miniscule amount of sample. The capillaries are connected in either end to two electrodes, respectively placed in anode and cathode buffers, which are coupled to a high-voltage power supply. A laser shines into a detection window on the capillary and excites fluorescent dyes attached to DNA molecules as they pass by. The polymer gel functions as an obstacle leading longer molecules to migrate slower than shorter molecules.

Samples include an internal size standard of DNA fragments of known sizes, which serves as a basis for computational estimates of sample fragment sizes. The estimation of sample fragment sizes is achieved by plotting relative fluorescence intensity from the sample, to the timespan from sample injection to sample detection. Fluorescence intensity is measured in relative fluorescence units (rfu) [81].

This enables the estimation of RT-PCR mRNA fragment lengths by comparing the internal size standard fragments to the sample data. The raw data can be visualized in the GeneMapper ID-X (Applied Biosystems™) software as a plot of DNA size in base pairs to rfu.

Guidelines for working with RNA

In contrast to the highly stable, double helical structure of native DNA, single stranded RNA has a much higher free energy [82]. In addition to the base thymine being replaced by the base uracil, the sugars of RNA have an OH-group, which is a site of high reactivity. This will not only provide RNA with the ability to form intramolecular structures through hydrogen bonding, but also leaves the RNA molecule as a chemically labile target. It is therefore of great importance to establish a routine for working with RNA to avoid molecular alteration and degradation as much as possible [83].

RNases are highly active enzymes that degrade RNA molecules into smaller components and are produced by every organism ever studied. It therefore follows that RNases are found everywhere in the environment around us. These enzymes act as a part of the human immune system as they have cytotoxic effects and are found extracellularly (exoribonucleases) and intracellularly (endoribonucleases) in several cell types, as for example skin cells [84].

To reduce contact with any RNase contaminated surfaces, it is recommended to dedicate a working station to workflows involving the handling of RNA. This includes a fume hood for the eradication of potential RNases produced by fungi spores in the air, as well as pipettes and other required equipment [85]. Single use plastics such as pipette tips and nuclease free tubes are also essential. A frequent change of gloves and avoiding touching contaminated surfaces decreases the risk of transferring RNases originating from skin.

To reduce the risk of reaction with external factors such as RNases and other chemical factors produced by microorganisms, a cleaning of all surfaces of the workstation with absolute ethanol should be performed before and after every procedure [86]. This also applies to all working equipment, such as pipettes and tube racks. Various manufacturers recommend the use of different products developed for this purpose, such as RNaseZap™ (Invitrogen), but the experience at the Center of Forensic Genetics is that a strict cleaning procedure, a tidy workstation, and an overall orderly work routine is adequate to minimize contamination by RNases.

Factors such as high temperatures and alkaline pH-levels can cause strand cleavage of RNA. Therefore, it is recommended to maintain a neutral pH to reduce reactivity, and to keep a working temperature of 0-4°C. Keeping fresh ice at the workstation to store RNA samples

before and during various procedures, is an effective mean of maintaining a low temperature [83].

Most RNases are deactivated at -70°C, and all samples should therefore be preferably stored at this temperature. It should also be added here that minimizing the amount of freezing-thawing cycles of samples and reagents by aliquoting into working volumes will restrict any major degradation.

Results

1. Quantity and quality of RNA in PrepFiler *Express*™ Forensic DNA Extraction Kit lysis buffer discard and eluate

RNA quantity

Both measurements from the Qubit® Fluorometer and the Agilent 2100 Bioanalyzer were used to evaluate whether RNA could be found at different stages in the DNA extraction process and if so, where it had highest quantity and best quality. The PrepFiler™ DNA eluate and LBD (lysis buffer discard) were the two DNA-extraction components measured. The concentration measurements of RNA-eluate samples extracted with the *mir*Vana™ miRNA Isolation Kit were used as a comparison. The sample setup was the same as used by Gårdvik(2020) , and the same sample stock of body fluids was used to allow for direct comparison [56].

Table 9 displays all RNA concentration values (mean ± SD) for the Qubit® Fluorometer and the Bioanalyzer measured in this study, as well as the values obtained by Gårdvik (2020) for comparison. See Appendix 1 for raw data of this study.

As seen in Table 9, the RNA concentration measures were similar when using the Qubit® Fluorometer and the Bioanalyzer, with the exception of the concentrations for semen. The Qubit® Fluorometer RNA concentration for blood were slightly higher, yet similar to the RNA concentrations measured by the Bioanalyzer. The RNA concentrations measured for saliva were highly similar in both methods. For semen, the Qubit® Fluorometer RNA concentrations were three times higher for both the DNA eluate and for the RNA eluate than the RNA concentrations measured by the Bioanalyzer.

With the exception of saliva, RNA concentrations were higher in the RNA eluate than in the DNA eluate, independent of quantification method (Table 9). For saliva, the RNA concentration was slightly higher in the DNA eluate than in the RNA eluate. The LBD from the DNA extraction gave the lowest RNA concentration with both methods across all three body fluids. LBD for blood measured by the Qubit® Fluorometer was the only LBD with an RNA-concentration above detection limits (0.67 ± 0.02 ng/µL). The rest of the LBD samples had lower RNA-concentrations than the detection threshold (25 pg/µL).

Gårdvik found similar patterns with concentrations being the lowest in the LBD from the DNA extraction for all body fluids and both methods of measurement (Table 9). For saliva, the RNA concentration in the DNA eluate was also higher than in the RNA eluate. Thus, the RNA-concentration of the samples was reproducible between studies.

Table 9*: Mean RNA concentrations and standard deviations in ng/µl for blood, saliva and semen from RNA eluate from the mirVana™ miRNA Isolation kit, and DNA eluate and lysis buffer discard (LBD) from the PrepFiler Express™ Forensic DNA Extraction Kit. The table contains results from both the Qubit® fluorometer and the Bioanalyzer of this study and Gårdvik (2020). See Appendix 1 for raw data.*

** The Qubit® Fluorometer would mark samples that fell below the detection threshold at 25 pg/µL as "Too low. Out of range". These samples are marked as "0" in this study for comparison* purposes.

Body fluid		**RNA quantity with Qubit® Mean ± Standard deviation (ng/µL)**		**RNA quantity with Bioanalyzer Mean ± Standard deviation (ng/µL)**	
		This study	Gårdvik, 2020	This study	Gårdvik, 2020
Blood	RNA eluate	7.57 ± 0.56	7.61 ± 0.96	4.17 ± 0.95	1.87 ± 0.11
	DNA eluate	3.18 ± 0.90	5.03 ± 1.12	2.43 ± 0.70	1.62 ± 0.48
	LBD	0.67 ± 0.02	1.54 ± 0.06	0	0.03 ± 0.03
Saliva	RNA eluate	4.15 ± 0.81	5.38 ± 1.73	3.45 ± 1.36	2.53 ± 1.09
	DNA eluate	5.12 ± 0.29	5.73 ± 1.76	6.28 ± 2.72	3.29 ± 0.66
	LBD	0*	0.89 ± 0.09	0	0.04 ± 0.01
Semen	RNA eluate	5.62 ± 0.17	6.38 ± 0.32	1.75 ± 0.02	1.99 ± 0.15
	DNA eluate	1.07 ± 0.29	3.69 ± 0.21	0.29 ± 0.21	1.62 ± 0.56
	LBD	0	0.90 ± 0.04	0	0.18 ± 0.28

RNA quality by DV_{200} and RIN

The quality parameters RIN and DV_{200} were used to determine the level of RNA degradation in the *mir*Vana™ RNA eluate, the PrepFiler™ DNA eluate, and the PrepFiler™ LBD. RIN results are divided into the three quality categories described in methods paragraph 1.5: RIN 7-10 is "High quality", RIN 4-6 is "Low quality", and RIN <4 is "Too degraded" (see Table 10). DV_{200} results are divided into four categories according to Figure 6B in paragraph 1.5, where

DV_{200} >70% is “High quality”, DV_{200} 50-70% is “Medium quality”, DV_{200} is “Low quality”, and DV_{200} <30% is “Too degraded”. (see Table 11).

Table 10: *Heatmap for RIN from this study and Gårdvik (2020). High quality (RIN 7-10)* ----, *low quality (RIN 4-6)* ----, *and too degraded (RIN <4)* ----. *Samples for which no RIN values could be calculated are marked with ----. See Appendix 2 for raw data.*

	RIN					
	RNA eluate		DNA eluate		LBD	
	This study	Gårdvik (2020)	This study	Gårdvik (2020)	This study	Gårdvik (2020)
Blood						
Saliva						
Semen						

Table 11: *Heatmap for DV_{200} from this study and Gårdvik (2020). High quality* ----, *medium quality* ----, *low quality* ---- *and too degraded RNA* ----. *Samples for which no DV_{200} values could be calculated are marked with ----. See Appendix 2 for raw data.*

	DV_{200}					
	RNA eluate		DNA eluate		LBD	
	This study	Gårdvik (2020)	This study	Gårdvik (2020)	This study	Gårdvik (2020)
Blood						
Saliva						
Semen						

The highest RIN overall can be found for the RNA eluate. Blood performed the best where all triplets had RIN values in the “low-quality” range. Semen and saliva performed similar, with low RIN values equivalent to “too degraded”. For the DNA eluate, blood had RIN values for all samples, in the “too degraded” range. Only two samples of the triplets had readable “too degraded” RIN values for semen, and no RIN values were obtained for saliva. The LBD gave no RIN values.

The RIN results of Gårdvik (2020) show a similar pattern for the RNA eluate, with blood having RIN values equivalent to “high quality”. The DNA eluate only yielded a single, “low quality” RIN value, which was for blood. Finally, the LBD had no calculated RIN values.

Calculated DV_{200} for all samples can be seen in Table 11. For this study, the highest and least variable DV_{200} values were found for the DNA eluate. All blood triplets had DV_{200} in the “high

quality" range, with semen coming in second with DV_{200} in the "high quality" and "medium quality" range, and saliva coming in third, with DV_{200} ranging from "high quality" to "too degraded". The RNA eluate had DV_{200} equivalent to the "medium quality" to "high quality" categories for blood, and saliva had DV_{200} calculated to be in the "medium quality" to "low quality" range. "Low quality" DV_{200} was calculated for semen, but values were consistent among the triplets. DV_{200} could not be calculated for the LBD.

The DV_{200} results of Gårdvik (2020) were much more variable between triplets for the DNA eluate for all three body fluids, with semen giving the most consistent calculated DV_{200} results in the "medium quality" to "low quality" categories. The RNA eluate had the highest, and least varied calculated DV_{200} for blood with values equivalent to "medium quality", followed by semen with "medium quality" to "low quality", and last saliva with calculated values in the "too degraded" category. Only one DV_{200} was calculated for the LBD for semen in the results of Gårdvik.

RNA quantity and quality - implications for further work

Overall, the similarities in the results of this study to those of Gårdvik (2020) demonstrate their reproducibility. Despite the low RNA-concentrations and lack of calculated RIN and DV_{200}, it was interesting to further investigate the LBD. Only a small volume of LBD directly from the large volume available in reagent cassette was used during the above-mentioned measurements which might account for the low concentrations and low-quality results. Preliminary tests of purifying RNA from LBD gave promising RNA concentration results. Isolating RNA from a larger volume of LBD by the means of an RNA-isolation kit, to concentrate available material, may be a better option for RNA quality assessment.

2. Optimalization with three RNA-isolation kits

The protocols of the three RNA isolation kits, namely the *mir*Vana™ miRNA Isolation Kit, the Direct-zol™ RNA Miniprep kit, and the ReliaPrep™ RNA Miniprep System, were optimized for the extraction of RNA from the DNA extraction lysis buffer discard.

The purpose of the optimization was to investigate which RNA-isolation kit was most suited for the extraction of RNA from the PrepFiler™ LBD, both in terms of RNA concentration and quality. A common denominator for the optimization of all three kits was the adjustment of reagent ratios. The reagents that needed to be adjusted are involved in lysis, nucleic acid

purification and fixation. As the LBD had previously yielded low RNA concentrations, the goal was to use the highest possible volume of LBD as input for RNA isolation.

Table 12*: Ratios, reagent volumes used in RNA-extraction with the mirVana™ kit and mean ± SD of RNA concentration measurements ng/µL performed with the of Qubit® Fluorometer. LBD (lysis buffer discard)*

LBD : lysis/binding buffer : phenol:chloroform		
Ratio	**(µL)**	**Mean RNA ng/µL ± SD**
(1:1:2)	450:450:900	2.81 ± 0.05
(1:1:0.8)	500:500:400	0.51 ± 0.25
(1:1:1.2)	500:500:600	1.83 ± 0.05
(1:1:1.6)	500:500:800	3.27 ± 0.06
(1:1:2)	500:500:1000	1.71 ± 0.13
(1:0.82:1.45)	550:450:800	3.18 ± 0.05
(1:1:1.3)	600:400:800	2.96 ± 0.19
(1:0.67:1.67)	600:400:1000	0.85 ± 0.06
(1:0.43:1.43)	700:300:1000	0.68 ± 0.20

For the *mir*Vana™ kit the ratio between LBD, lysis/binding buffer and phenol-chloroform was adjusted, and each ratio was tested in triplets. Table 12 shows the mean and SD of the RNA concentration in ng/µL measured by the Qubit® Fluorometer for each of the investigated ratios.

Increasing the volume of LBD did not yield higher RNA-concentrations, but rather displayed similar or lower values than those ratios with an LBD volume of 500 µL. Keeping LBD and lysis/binding buffer volumes constant whilst increasing the volume of phenol:chloroform, showed the highest concentrations around the 1:1:1.6 ratio, including 800 µL of phenol:chloroform. A further increase in the volume of phenol:chloroform lead to lower RNA concentrations.

Decreasing both LBD and lysis/binding buffer volumes, and increasing phenol:chloroform volume to a ratio of 1:1:2 gave similar RNA concentrations as the 1:1:1.6 ratio. As the manufacturer recommends a 1:1:2 ratio, it was chosen as one of two ratios that were re-investigated [61]. Figure 8 shows the mean of all obtained concentration values for the two ratios. RNA concentration obtained with the 1:1:1.6 ratio was slightly higher compared to the 1:1:2 ratio, 3.24 ng/µL versus 2.73 ng/µL, respectively. The 1:1:1.6 ratio was therefore considered as optimal when isolating RNA with the *mir*Vana™ kit

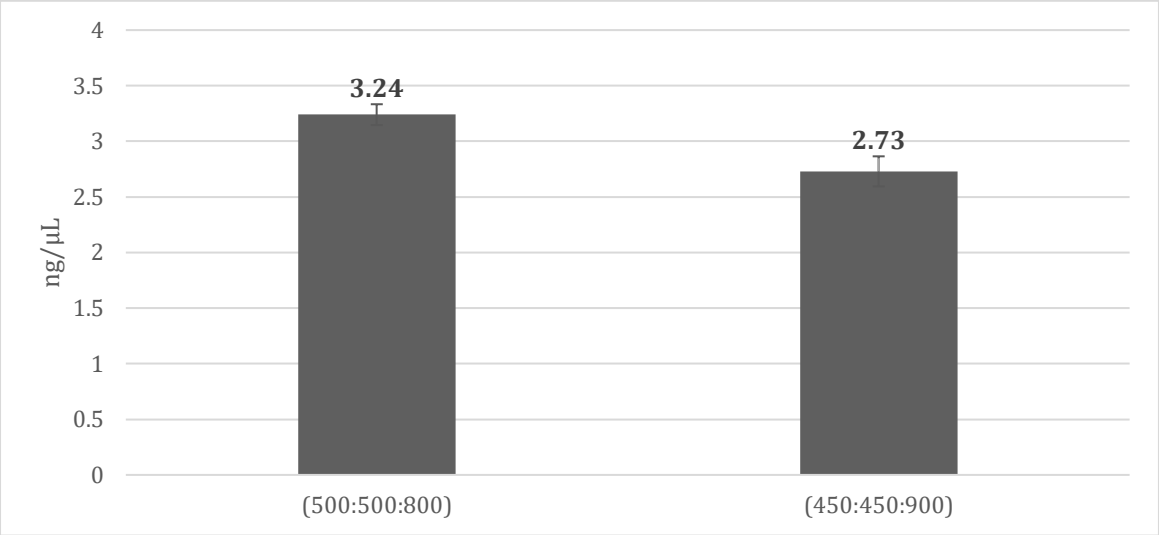


Figure 8*: Mean RNA concentrations in ng/µL of the two re-tested ratios of the mirVana™ kit of 1:1:1.6 (500:500:800 µL) and 1:1:2 (450:450:900 µL) when extracting RNA from LBD.*

The optimization of the Direct-zol™ RNA Miniprep kit, and the ReliaPrep™ RNA Miniprep System was not as comprehensive as for the *mir*Vana™ kit, as no reproducible RNA concentrations could be measured for any of the examined reagent ratios.

None of the extraction samples from the Direct-zol™ kit gave readable concentrations on the Qubit® Fluorometer, and consistently came up as “Too low, sample out of range”. Several ratios of LBD, TRIzol Reagent, and EtOH were tested for this kit, with no observable difference in outcome.

The investigated ratios of LBD, BL + TG buffer, and isopropanol for the ReliaPrep™ kit isolated too little RNA to consistently exceed the 25 pg/µL detection limit of the Qubit® Fluorometer. Only one or two of the triplets measured produced RNA concentration results. The highest concentration measured was 0.67 ng/µL, for the 300:200:170 ratio. However, one of the triplets of this ratio had an RNA-concentration below the detection limit, leaving this result inconsistent.

Although no reliable concentration measurements could be obtained for the DirectZol™ and ReliaPrep™ kits with the Qubit® Fluorometer, the presence of RNA in the extracted LBD samples was further investigated, amplifying the housekeeping gene GAPDH by qPCR. Table 13 shows the reagent volume ratios and the mean C_t values from a random selection of reverse transcribed RNA eluate samples extracted from LBD from the three RNA-isolation kits. Both samples with (RT+) and without reverse transcriptase (RT-) are shown.

Table 13*: Random selection of reverse transcribed RNA eluate samples from LBD from all three kits with reagent volume ratios and sample names (RT+ or RT-). Mean C_t values of triplets were calculated from C_t values obtained from qPCR on the housekeeping gene GAPDH.*

Sample	**Mean C_t**
***mir*Vana™ kit**	
500 µL LBD : 500 µL lysis/binding buffer : 800 µL phenol:chloroform	
RT +	27.42
RT -	Undetermined
Direct-Zol™ kit	
200 µL LBD : 600 µL TRI reagent : 600 µL EtOH	
RT +	33.56
RT -	Undetermined
100 µL LBD : 300 µL TRI reagent : 300 µL EtOH	
RT +	34.32
RT -	Undetermined
ReliaPrep™ kit	
500 µL LBD : 0 µL TG+BL buffer : 170 µL isopropanol	
RT +	27.50
RT -	37.41
400 µL LBD : 100 µL BL + TG buffer : 170 µL isopropanol	
RT +	26.88
RT -	38.03

The triplets with reverse transcriptase enzyme (RT+) from *mir*Vana™ kit extractions consistently showed up with a mean cycle threshold value of 27.42. The parallels with no enzyme added (RT-) were all marked as "Undetermined".

The DirectZol™ kit triplets had higher C_t values than the *mir*Vana™ kit, suggesting that the RNA (cDNA) content is relatively lower than for the samples processed with the *mir*Vana™ kit (Table 13). RT- parallels for the DirectZol™ kit, were also marked as "Undetermined" for both ratios tested.

The ReliaPrep™ kit triplets had similar mean C_t values as the *mir*Vana™ kit, 27.50 and 26.88, respectively, suggesting a similar RNA content. C_t values of the RT- parallels for the ReliaPrep™ kit were in the range of 37-38.

The *mir*Vana™ and ReliaPrep™ kit produced the lowest mean C_t values, equating to higher RNA content in the samples. However, the ReliaPrep™ kit RT- samples produced high but not negligible C_t values, suggesting that gDNA is still present in the samples after the on-column DNase treatment. The DirectZol™ kit samples gave inconsistent or high C_t values for the RT+ parallels, indicating a low content of intact RNA. These results lead to the decision of continuing the comparative work with the *mir*Vana™ and ReliaPrep™ kit only.

3. Comparison of RNA quantity and quality of PrepFiler™ LBD extraction with *mir*Vana™ and ReliaPrep™ kit from blood, saliva and semen

The workflow seen in Figure 7 was the basis for an accurate comparison of the *mir*Vana™ and ReliaPrep™ kit in the isolation of RNA from LBD (lysis buffer discard), using two volumes of the body fluids blood, saliva, and semen. The generous 40µL volume of body fluid was used not only to determine the RNA-isolation capability of the two kits, but also to enable further investigation of RNA-quality of the isolated RNA. The smaller 4µL volume was used as it is closer to the setting of a realistic forensic investigation. This comparison included RNA-quantitation by Qubit® Fluorometer and amplifiable RNA of RT+ samples of GAPDH short product, RNA degradation determination using the ratio between the relative quantities of short and long GAPDH product, assessment gDNA contamination and DNase treatment efficiency with the Quantifiler™ Trio DNA Quantification kit and C_t of RT- samples of GAPDH short product, as well as overall workflow efficiency.

RNA quantity in DNAse-treated samples

All DNase-treated RNA eluates and DNA eluates were quantified using the Qubit® Fluorometer for RNA-concentration. Table 14 shows mean RNA-concentrations measured for all samples. No samples of 4µL of body fluid had RNA-concentrations above the detection limit of 25 pg/µL and showed as "Too Low". These samples are set as "0" in Table 14, for comparison purposes. With the exception of the 40µL LBD *mir*Vana™, all PrepFiler™ eluates yielded higher concentrations than the RNA-eluates of both RNA-isolation kits. The highest RNA-concentrations were measured for saliva, with semen coming in second, and lastly blood.

The RNA-concentration of the LBD ReliaPrep™ were lower than those of the *mir*Vana™ eluates, for saliva and semen. No RNA-concentrations were above detection limit for RNA-isolation kit 40µL blood samples and could not be compared.

For qPCR for short GAPDH of RT+ samples, the mean C_t values for 40µL samples were lower than C_t values for 4µL samples of all body fluids. This was seen for LBD samples of both isolation kits and the PrepFiler™ eluates, and points towards the higher volume of body fluid initially used in the DNA-extraction. The mean C_t of the RT+ samples showed similar trends between the *mir*Vana™ kit and the ReliaPrep™ kit for both volumes of all three body fluids. Generally, there was a difference of 1-3 cycles between the two kits, where the *mir*Vana™ kit had the lowest C_t, with the exception of 4µL and 40µL saliva. This indicates similar relative quantities of RNA in LBD *mir*Vana™ and LBD ReliaPrep™. The C_t of RT+ PrepFiler™ eluates, was generally lower than the C_t of the corresponding LBD RNA-isolation kit RT+ sample. This implies higher relative quantities of RNA in the PrepFiler™ eluates than in the LBD RNA-isolation kit samples.

Table 14: *Mean RNA-concentrations of DNAase-treated samples, measured by the Qubit® Fluorometer 4, and mean C_t of RT+ samples for qPCR of short GAPDH. Shown are mirVana™ kit and ReliaPrep™ kit samples extracted from PrepFiler LDB (lysis buffer discard), from 4µL and 40µL blood, saliva, and semen. Both RNA-isolation kits have corresponding PrepFiler™ kit DNA eluates deriving from the same reagent cassette. Samples that read as "Too low" on the Qubit® Fluorometer, are set as "0" for comparison purposes.*

Body fluid	Body fluid volume	Kit	Mean [RNA] (ng/µL)	Mean C_t of RT+ GAPDH short
Blood				
	4 uL			
		LBD *mir*Vana™	0	28.58
		Eluate PrepFiler™	0	25.75
		LBD ReliaPrep™	0	29.52
		Eluate PrepFiler™	0	25.07
	40 uL			
		LBD *mir*Vana™	0	23.29
		Eluate PrepFiler™	2.15	27.18
		LBD ReliaPrep™	0	24.03
		Eluate PrepFiler™	2.00	28.05
Saliva				
	4 uL			
		LBD *mir*Vana™	0	35.33
		Eluate PrepFiler™	0	29.29
		LBD ReliaPrep™	0	32.58
		Eluate PrepFiler™	0	30.21
	40 uL			
		LBD *mir*Vana™	7.40	29.24
		Eluate PrepFiler™	10.60	27.00
		LBD ReliaPrep™	4.27	28.59
		Eluate PrepFiler™	8.97	24.04
Semen				
	4 uL			
		LBD *mir*Vana™	0	26.92
		Eluate PrepFiler™	0	23.31
		LBD ReliaPrep™	0	27.97
		Eluate PrepFiler™	0	24.35
	40uL			
		LBD *mir*Vana™	6.78	22.79
		Eluate PrepFiler™	3.84	22.32
		LBD ReliaPrep™	2.10	23.31
		Eluate PrepFiler™	7.24	18.48

gDNA contamination and DNAse-treatment efficiency

ReliaPrep™ kit RT- samples have previously shown to have detectable reactions during qPCR of the short GAPDH product, indicating the presence of gDNA and an inefficient on-column DNase-treatment. To determine gDNA contamination further, DNase-treated RNA- and DNA extraction samples were analysed with qPCR using the small autosomal target (SA) in the Quantifiler™ Trio DNA Quantification Kit. All RT- samples made for the RNA-isolation kit comparison were also run on qPCR for the short GAPDH product for assessment of gDNA contamination and DNAse-treatment efficiency.

Table 15: *An assessment of DNase treatment efficiency with the Quantifiler™ Trio DNA Quantification Kit and qPCR of RT- samples fort short GAPDH. Shown are mean C_t values for the Small Autosomal human target (SA), and mean C_t values of RT- samples for short GAPDH, of mirVana™ kit and ReliaPrep™ kit samples extracted from PrepFiler LDB (lysis buffer discard), from 4µL and 40µL blood, saliva, and semen. Both RNA-isolation kits have corresponding PrepFiler™ kit DNA eluates deriving from the same reagent cassette. Ct-values marked as "Undetermined" were set as Ct = 40 for the purpose of calculating a mean value.*

Body fluid	Body fluid volume	Kit	Mean Ct SA	Mean Ct of RT-GAPDH short
Blood				
	4 uL			
		LBD *mir*Vana™	40.00	40.00
		Eluate PrepFiler™	38.76	33.37
		LBD ReliaPrep™	31.74	32.91
		Eluate PrepFiler™	38.00	37.64
	40 uL			
		LBD *mir*Vana™	38.93	30.37
		Eluate PrepFiler™	38.21	39.09
		LBD ReliaPrep™	32.98	34.37
		Eluate PrepFiler™	37.20	38.25
Saliva				
	4 uL			
		LBD *mir*Vana™	40.00	40.00
		Eluate PrepFiler™	38.80	35.38
		LBD ReliaPrep™	31.90	32.65
		Eluate PrepFiler™	37.33	39.01
	40 uL			
		LBD *mir*Vana™	40.00	30.15
		Eluate PrepFiler™	34.99	35.41
		LBD ReliaPrep™	34.02	34.37
		Eluate PrepFiler™	37.86	37.69
Semen				
	4 uL			
		LBD *mir*Vana™	40.00	40.00
		Eluate PrepFiler™	36.35	40.00
		LBD ReliaPrep™	31.33	34.09
		Eluate PrepFiler™	35.12	40.00
	40uL			
		LBD *mir*Vana™	36.34	39.00
		Eluate PrepFiler™	36.22	29.25
		LBD ReliaPrep™	30.24	32.44
		Eluate PrepFiler™	36.60	27.19

As seen in Table 15, LBD *mir*Vana™ kit samples were "Undetermined" for the small autosomal (SA) target, with the exception of 40µL blood and semen. Corresponding DNase-treated PrepFiler™ DNA eluates had high C_t values in the range of 36-39, indicating a small amount of gDNA present in these samples. In several cases, either one or two of the samples in a triplet were "Undetermined". This was especially true for LBD *mir*Vana™ samples and corresponding PrepFiler™ DNA eluate samples. These samples had DNA amounts below the threshold value and will have a Ct of 40 or above. For the purpose of calculating mean values of triplets in Table 15, samples marked as "Undetermined" were given the value of 40.

LBD ReliaPrep™ kit samples were in the C_t range of 30-34 for both volumes in all three body fluids. C_t values for the 40µL samples were 1-2 cycles lower than for the 4µL samples, indicating a two to four-fold higher content of gDNA. Out of all LBD ReliaPrep™ samples, only one was marked as "Undetermined", unlike the LBD *mir*Vana™ samples, for which a majority was "Undetermined".

LBD *mir*Vana™ kit RT- samples for the short GAPDH qPCR were "Undetermined" for 4µL of all three body fluids. For *mir*Vana™, 40µL body fluid RT- samples had C_t values in the range of 30-39.

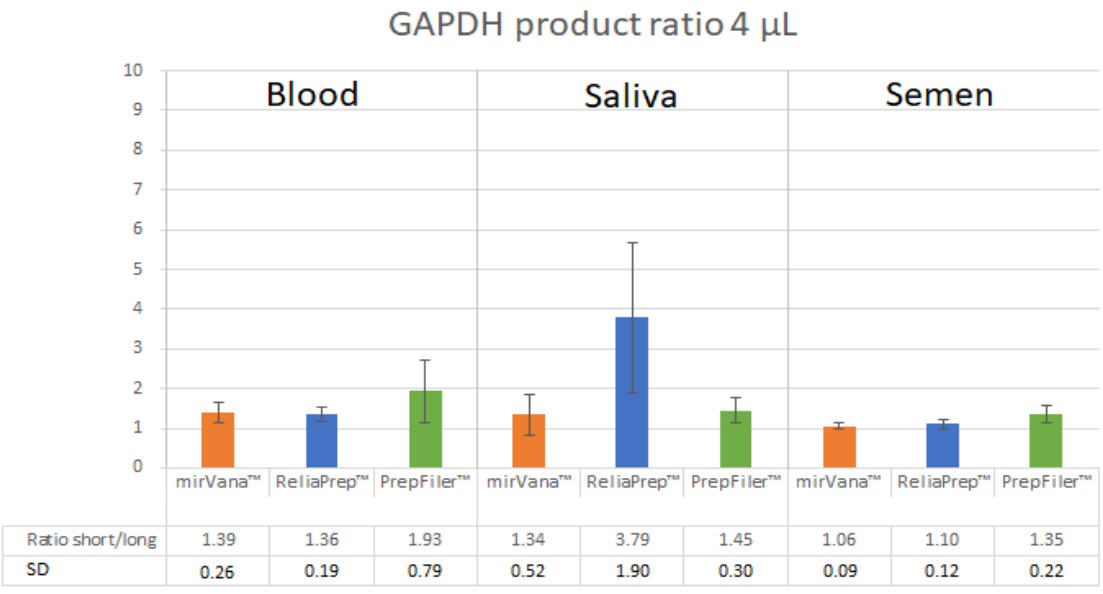
LBD ReliaPrep™ kit RT-samples were in the C_t range of 32-34, where mean C_t of blood and saliva was higher for 40µL than 4µL. Corresponding PrepFiler™ DNA eluates for both RNA-isolation kits had C_t in the range of 27-40. Overall, ReliaPrep™ kit RT-samples had lower C_t values across all three body fluids and volumes than the *mir*Vana™ kit RT- samples.

When comparing the two qPCR runs for the two RNA-isolation kits, DNase-treatment with TURBO™ DNase kit was more efficient for 4µL of body fluid, as higher C_t values were seen. However, PrepFiler™ DNA eluate triplets treated with the TURBO™ DNase kit had C_t values in the range of 34-38 for SA targets and C_t values as low as 27 and up to 40, independent of body fluid and volume, indicating that a considerable amount of gDNA was still present in many samples, and that the DNase-treatment was not sufficient.

RNA degradation

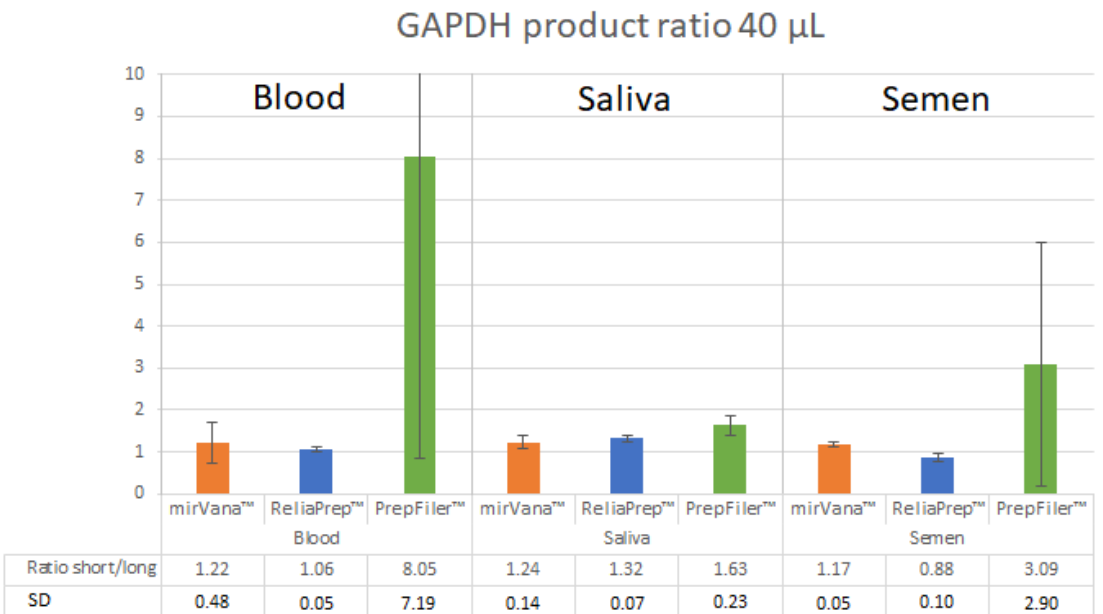
For an overview of RNA degradation, a ratio of the relative quantity of the shorter GAPDH (58 bp) product over the relative quantity of the longer GAPDH (114 bp) product was calculated. A short/long GAPDH product ratio of ≤ 1 and around 1 indicates no degradation, a rising short/long GAPDH product ratio above 1 indicates increasing levels of degradation. Figure 9 shows the short/long GAPDH product ratios calculated from the mean relative quantities for the RNA-eluate (*mir*Vana™ and ReliaPrep™) and DNA-eluate (PrepFiler™) triplets.

A)



	Blood			Saliva			Semen		
	mirVana™	ReliaPrep™	PrepFiler™	mirVana™	ReliaPrep™	PrepFiler™	mirVana™	ReliaPrep™	PrepFiler™
Ratio short/long	1.39	1.36	1.93	1.34	3.79	1.45	1.06	1.10	1.35
SD	0.26	0.19	0.79	0.52	1.90	0.30	0.09	0.12	0.22

B)



	Blood			Saliva			Semen		
	mirVana™	ReliaPrep™	PrepFiler™	mirVana™	ReliaPrep™	PrepFiler™	mirVana™	ReliaPrep™	PrepFiler™
Ratio short/long	1.22	1.06	8.05	1.24	1.32	1.63	1.17	0.88	3.09
SD	0.48	0.05	7.19	0.14	0.07	0.23	0.05	0.10	2.90

Figure 9: *Short/long GAPDH product ratios and SD for assessment of RNA degradation levels of A) 4µL LBD extraction samples and PrepFiler™ eluate and B) 40µL LBD extraction samples and PrepFiler™ eluate, for blood, saliva, and semen. Short/long GAPDH product ratios are calculated from mean relative RNA quantity values obtained from C_t values referenced towards the standard curve.*

For 4µL and 40µL of body fluid, the short/long GAPDH ratios of RNA-extractions from LBD are close to 1, indicating a low level of RNA degradation. This is true for both the *mir*Vana™ and the ReliaPrep™ kit extraction products, with the exception of ReliaPrep™ 4µL saliva. However, the standard deviation of *mir*Vana™ and ReliaPrep™ for 4 µL saliva slightly overlap, and a difference between the two kits cannot be determined.

The PrepFiler™ eluate triplets had higher short/long GAPDH product ratios for both 4µL and 40µL body fluid. In addition, these triplets were highly variable for 40µL body fluid. This indicates that the RNA in PrepFiler™ samples are much more likely to be degraded.

Overall, the *mir*Vana™ kit samples had lower C_t values and higher RNA quantities for all body fluids and both volumes while giving similar short/long GAPDH product ratios to the ReliaPrep™ kit. This indicates that the overall RNA content is higher for the *mir*Vana™ extractions, and that the level of degradation is similar for both kits, speaking for an equal quality of the extraction products.

4. Optimization of PCR conditions for twelve selected body fluid markers

Twelve body fluid markers and two housekeeping genes were chosen for analysis with PCR and capillary electrophoresis. These were markers for venous blood, saliva, semen, vaginal secretion, and menstrual blood. The future goal is to incorporate these markers into a multiplex. Before this can be undertaken, a determination of possible primer concentrations adjustments, and most ideal body fluid PCR program for the downstream capillary electrophoresis, needed to be done. Four dilutions of reverse transcribed *mir*Vana™ RNA eluate that were extracted from 40 µL of pure body fluid (venous blood, saliva, and semen) and cotton swabs (vaginal fluid and menstrual blood) were used.

Table 16 gives an overview of the criteria used for the categorization of samples for the heatmap in Table 17. Observed fragment lengths may be different from the theoretical lengths in Table 5, as they can be affected by molecular, environmental, and systemic factors. Therefore, a deviation of ±6 bp from the theoretical fragment length was interpreted as “expected base pair area”. Table 17 is a heatmap of capillary electrophoresis results of all body fluid markers and housekeeping genes on four dilutions of their specific body fluid, run under three different PCR programs.

For examples of electropherograms that fit the criteria of Table 16, see Appendix 4.

Table 16: *Colour code and description of criteria used to interpret electrophoresis results of body fluid markers and housekeeping genes in Table 17.*

Colour	Criteria description
	Overload. One or more high peaks in expected base pair area that give very high pull-ups in several of the other colour channels. Samples that give no results because of total overload of the system (No Size Data)
	A clear, high peak in expected base pair area, but peak of interest is <4x higher than pull-ups and other extra peaks.
	A clear, high peak in expected base pair area. Peak of interest is 4x-7x higher than other extra peaks.
	A clear peak in expected base pair area. Peak of interest is >7x higher than other extra peaks.
White	No peaks in expected base pair area or peaks <700rfu.

Table 17: H*eatmap of capillary electrophoresis results for body fluid markers and housekeeping genes. Four dilutions of each body fluid were used. Each body fluid primer pair was run on all dilutions of the respective body fluid with three PCR programs (NZ, NL60, and NL64).*

	Dilution	18S-rRNA		ACTB		HBB			SLC41A		
		NZ	NL60	NZ	NL60	NZ	NL60	NL64	NZ	NL60	NL64
Blood	1										
	0.1										
	0.01										
	0.001										
						STATH			FDCSP		
						NZ	NL60	NL64	NZ	NL60	NL64
Saliva	1										
	0.1										
	0.01										
	0.001										
						PRM1			TNP1		
						NZ	NL60	NL64	NZ	NL60	NL64
Spermatozoa	1										
	0.1										
	0.01										
	0.001										
						MSMB			KLK3		
						NZ	NL60	NL64	NZ	NL60	NL64
Seminal fluid	1										
	0.1										
	0.01										
	0.001										
						CYP2B7P			MUC4		
						NZ	NL60	NL64	NZ	NL60	NL64
Vaginal secretion	1										
	0.1										
	0.01										
	0.001										
						MMP7			MMP10		
		NZ	NL60	NZ	NL60	NZ	NL60	NL64	NZ	NL60	NL64
Menstrual blood	1										
	0.1										
	0.01										
	0.001										

A common trend for all body fluids was the apparent overload of material (cDNA), as the capillary electrophoresis system would either mark samples with "No size Data" or there were major pull-ups in several colour channels in the electropherogram (see Table 17). Peaks of up to 33 000 rfu were found for several samples, exceeding the saturation limit of the camera in the instrument.

The NZ program samples for the house keeping gene 18s-rRNA were all marked with "No size data" (NSD). With the NL60 program, electropherograms for 18s-rRNA had pull-ups in several colour channels. With the NZ program electropherograms for ACTB had a clear peak in the expected bp area for the 0.001 dilution samples for all body fluids except venous blood and menstrual blood. With the NL60 program, electropherograms for ACTB had a clear peak in the expected bp area at the 0.001 dilution samples for all body fluids.

With the exception of the spermatozoa markers PRM1 and TNP1, electropherograms from more diluted samples showed that markers generally had clear peaks in the expected areas. For the NZ program, the clearest peaks could be seen for the 0.001 dilution samples. However, this program would also produce the highest amount of overload, and the markers PRM1, TNP1, MSMB and MMP7 would not give any clear results.

With the NL60 and NL64 programs, electropherograms had clear peaks in more diluted samples, and markers MUC4, STATH and FDCSP had small and unspecific peaks in the 0.001 dilution samples. This indicates that vaginal fluid and saliva have less amplifiable material for these markers, and that the current primer concentrations and body fluid dilutions are suited for the two NL programs. The NZ program seems to give an overload of material for these markers with the tested primer concentrations.

In terms of body fluid performance, semen (spermatozoa and seminal fluid) electropherograms had overload or were marked with "NSD" to a much higher degree than for other body fluids. This was also the case for menstrual blood, and points towards an extremely high amount of cDNA for body fluid specific markers in these samples.

In contrast, saliva and vaginal fluid had clear results for more diluted samples for all three programs, in addition to unreadable results for the highest dilutions of saliva for the NL64 program.

Discussion

The studies of Bowden (2011) established a co-extraction method of DNA and RNA from body fluids by the combination of the Promega DNA IQ™ system and the Zymo Research Mini RNA Isolation Kit™, and were able to successfully perform downstream RNA profiling of the RNA from the LBD of the DNA extraction [5]. This method would enable the performance of accurate body fluid identification while removing the need to utilize the material destined for DNA profiling. The project by Gårdvik (2020) aimed to combine the PrepFiler *Express*™ DNA Extraction kit and AutoMate *Express*™ Forensic DNA Extraction System used at CFG with the *mir*Vana™ miRNA Isolation kit for the same purpose [56].

Determination of RNA quantity and quality in PrepFiler *Express*™ DNA Extraction kit components

The first aim of this thesis was to determine whether the findings of the project by Gårdvik (2020) were reproducible. This was to examine if RNA was present in LBD and DNA eluate of the PrepFiler *Express*™ DNA Extraction method, and further to determine the quantity and quality of this RNA using two methods of quantitation (the Qubit® Fluorometer and the Agilent 2100 Bioanalyzer) and two measures of quality (DV_{200} and RIN). Simultaneously, RNA extracted from pure body fluid with the *mir*Vana™ miRNA Isolation Kit, would also be quantitatively and qualitatively measured to serve as a comparison. The body fluids examined were blood, saliva and semen from the same stock as Gårdvik.

The RNA concentrations found by both the Qubit® Fluorometer and the Agilent 2100 Bioanalyzer for all three body fluids were comparable to those found in the project of Gårdvik, both in terms of mean values and standard deviation trends. For both projects, the RNA concentrations are lowest in the PrepFiler™ LBD and the highest in the *mir*Vana™ eluate. The latter is expected, as the *mir*Vana™ kit is aimed at RNA-isolation, whereas the purpose of the PrepFiler™ kit is DNA-extraction.

The highest calculated DV_{200} was for the PrepFiler™ eluate and RIN measurements had the highest values for the *mir*Vana™ eluate. For DV_{200}, this suggests that the PrepFiler™ eluate has a higher proportion of fragments with a size of 200 nucleotides or longer. For RIN, the longer and most intact fragments were found for the *mir*Vana™ eluate.

The PrepFiler™ LBD produced few to no quality values, both in this project and the project of Gårdvik. As RNA-concentrations were low for LBD, there might have been too little material

upon which to calculate any quality. The LBD contains not only cell components and debris, but also high levels of solutes derived from the wash buffers and reagents of the DNA extraction process. It is therefore likely that measurements of both quantity and quality of the LBD were disturbed by the various contaminants. It is not certain exactly what RNA material the Qubit® Fluorometer is measuring, nor what criterion said material must fulfil to appear as a valid concentration measurement. The matter is the same for the Agilent 2100 Bioanalyzer, which in the studies of Caruana and McInnes was proven to give inaccurate RNA-concentration readings caused by gDNA contamination [87]. As both the PrepFiler™ eluate and LBD contains a mixture of nucleic acids and cell components, it is possible that RNA concentration readings are inaccurate. In addition, RIN algorithm computation cannot account for anomalies, such as gDNA contamination, which potentially explains why few values were obtained from the DNA-rich PrepFiler™ eluate [63]. It can therefore not be concluded which type of measurement is the most accurate and reliable.

In this study and the study of Gårdvik (2020), there was a difference between body fluids, both in terms of quantity and quality. Blood had both the highest RNA-concentration and the highest calculated DV_{200} and RIN. Saliva had higher RNA-concentrations, but lower RIN-values than semen. Semen and saliva had similar trends of calculated DV_{200}. This implies that not all body fluids are equally suited for the workflows and downstream analysis of this study, as they initially contain variable amounts of material. The variation between available material in body fluids is further discussed in the optimization of PCR conditions of body fluid markers.

Overall, the quality and RNA concentration measurements gave similar patterns to that of Gårdvik (2020), which demonstrates their reproducibility. Based on these results, a decision was made to move forward with testing several methods to extract RNA from the PrepFiler™ LBD. The RNA-concentrations and quality results found for the PrepFiler™ DNA-eluate led to choosing it for further analysis in the comparison of the *mir*Vana™ kit and ReliaPrep™ kit discussed below.

Optimization of three RNA-isolation kits for determination of workflow and RNA-isolation efficacy

The protocol of three commercially available RNA-isolation kits were optimized for the extraction of RNA from the PrepFiler™ LBD (lysis buffer discard); the *mir*Vana™ miRNA Isolation kit, the Direct-zol™ RNA Miniprep kit and the ReliaPrep™ RNA Miniprep System.

The chosen approach for all three kits was varying the ratios of kit reagents, as the *mir*Vana™ kit in preliminary studies had proven to yield higher RNA-concentrations by this method.

*mir*Vana™ miRNA Isolation kit

It was shown that a ratio of 1:1:1.6 for LBD, Lysis/Binding buffer and phenol:chloroform (500:µL 500µL:800µL) gave the clearest phase separation and also the highest yield of RNA on the Qubit® Fluorometer for the *mir*Vana™ kit. This ratio was also the most advantageous in terms of handling, as the total volume did not override the capacity of the 2 mL Lo-Bind tubes. Ratios such as 1:0.67:1.67 and 1:1:2 were challenging to handle as they proved difficult to vortex and easily spilled on gloves and tube lids. This is not acceptable regarding health and safety when working with dangerous chemicals such as phenol:chloroform [68]. An option for avoiding this challenge would be to use 5 mL Lo-Bind tubes. However, the spin columns have a maximal capacity of 700 µL, which would demand an increased number of transfers, in turn increasing the potential for error and contamination in a forensic setting. A high number of transfers would also prolong extraction duration. As it is most beneficial to keep the handling time of RNA in room temperature as short as possible, using increased volumes might not be a better solution.

Direct-zol™ RNA Miniprep kit

The volumes of the LBD, TRIzol™ Reagent and EtOH (µL) were chosen as the variable reagents of the Direct-zol™ RNA Miniprep kit, due to their function in the RNA-isolation process. Two of the components that make up TRIzol™ Reagent are guanidine isothiocyanate and phenol, which functionally correlates to the Lysis/Binding Buffer and phenol:chloroform of the *mir*Vana™ kit [70]. As the increase in volume of these components successfully elevated the RNA yield of the *mir*Vana™ kit, it was interesting to test whether the same was true for the Direct-zol™ kit. However, every sample from this kit was “too low” on the Qubit® Fluorometer, which implies that the RNA-concentration does not exceed the detection threshold of 25 pg/µL. As previously mentioned, the PrepFiler™ LBD contains high levels of solutes deriving from the DNA-extraction process, which might not be compatible with the reagents and methods of all RNA-isolation kits. The result of this incompatibility may be a very low, or no yield of RNA. A few samples were kept for further analysis by qPCR, as this method may be more sensitive in detecting RNA.

ReliaPrep™ RNA Miniprep System

The ReliaPrep™ RNA Miniprep System was recommended as an alternative RNA-isolation kit by Dr Harbison, Institute of Environmental Science and Research Limited (ESR), New Zealand. The volume ratios of the LBD, BL + TG buffer and isopropanol were varied during the optimization of this kit. Whilst optimizing this method, the RNA eluate would not produce any consistent results on the Qubit® Fluorometer. Triplets of equal ratios would vary greatly in RNA yield; from comparable values of the *mir*Vana™ kit, to turning up as "Too low". As results were inconclusive, it was difficult to make a decision on any further adjustments. An in-house procedure supplied by the ESR was tested, in which the step of adding BL + TG buffer was removed, and the incubation time of the DNase-treatment was elongated (40 minutes). This method yielded similar RNA concentrations as those acquired with the previously tested ratios but was more consistent in readable results on the Qubit® Fluorometer. As the BL + TG buffer is designed to lyse cells and homogenize a lysed sample, it was strictly not necessary to add it to the already lysed PrepFiler™ LBD. A longer incubation time would also ensure a more effective DNase-treatment. However, as discussed later, the on-column DNase-treatment included in the ReliaPrep™ kit was not as effective as post-extraction DNase-treatment.

qPCR of the housekeeping gene GAPDH on RNA-isolation kit samples for kit comparison

A qPCR using the primers and probe of the housekeeping gene GAPDH on a random selection of RNA eluates extracted from LBD with the three isolation kits were further analysed for the quantity and quality of RNA. This analysis gave a notion of overall RNA (cDNA) content in the RT+ sample, in addition to potential gDNA contamination in all samples by using the RT- sample as a reference. As reverse transcriptase enzyme had not been added to the RT- samples, any detectable reaction by qPCR will indicate gDNA contamination and an ineffective DNase treatment. The RT- samples of *mir*Vana™ and DirectZol™ samples were all marked as "Undetermined". This indicates that the TURBO™ DNase kit used for *mir*Vana™, and the on-column DNase treatment of the DirectZol™ kit was effective in removing contaminating DNA.

All RT- for ReliaPrep™ samples gave a detectable reaction on qPCR, suggesting that the on-column DNase treatment of this kit was not sufficient to remove all contaminating DNA. The samples with the 400 µL LBD : 100 µL TG+BL buffer : 170 µL isopropanol samples did not receive the 40 minute long DNase incubation time, as the 500 µL LBD : 0 µL TG+BL buffer :

170 µL isopropanol samples, which may have contributed to a lower degree of gDNA removal. DNase-treatment inefficiency is not an unknown phenomenon of the ReliaPrep™ kit, and studies have tackled this challenge by treating contaminated samples with an additional DNase treatment (TURBO™ DNase kit), with the Quantifiler™ Trio kit to confirm the removal of gDNA [47]. The strategy of an additional DNase-treatment was tested in this project, which resulted in a considerable increase in C_t-values for RT- samples, but also for RT+ samples on short GAPDH qPCR. This implies that the additional DNase treatment does remove gDNA but might lead to loss of RNA content in the sample as well, as the DNase-treatment involves several transfers. As previously mentioned, RNA is sensitive to higher temperatures, and as an additional DNase-treatment requires an incubation at 37°C, this might lead to sample degradation. The additional DNase treatment was therefore not used further in this project.

The low RNA content found in the DirectZol™ samples lead to the decision of moving forward with only the *mir*Vana™ kit and ReliaPrep™ kit, as these proved to be most promising in the extraction of RNA from the PrepFiler™ LBD.

Comparison of PrepFiler™ LBD from blood, saliva and semen extracted with *mir*Vana™ and ReliaPrep™ kit

A more extensive set of analyses was performed for blood, saliva and semen with the *mir*Vana™ and ReliaPrep™ kit to further determine which of the kits was best suited for the extraction of RNA from PrepFiler™ LBD, both in terms of RNA quantity and quality for downstream analyses, as well as workflow efficiency (see Figure 7).

This workflow attempted as identical conditions as possible for both kits, to enable the direct comparison of results. As the PrepFiler™ eluate previously had been found to contain RNA of promising concentration and quality in preliminary analyses, this was also included in the workflow for comparison with the RNA extracts from the PrepFiler™ LBD. Reagent cassettes were therefore marked according to the corresponding DNA eluate, to enable the comparison of values obtained from the various types of measurements.

DNA extractions were performed from both 40µL and 4µL of body fluid. 40µL was the volume utilized in all previous analyses of PrepFiler™ LBD in this project and is a high amount in forensic terms. The quantity of 4µL was therefore also used for assessment of the same methods with less material to work with. 4µL is still a generous amount, but closer to a realistic forensic setting.

RNA quantity

The RNA-concentration of DNase-treated samples was measured with the Qubit® Fluorometer. The 4µL body fluid samples had no RNA-concentrations above the detection limit, neither for the RNA-isolation kits, nor the PrepFiler™ eluates. This is probably due to too low amounts of material in these samples for this method of measurement. As previously seen, a challenge when extracting RNA from LBD of 40µL of body fluid, is that the low amount of extractable material might be very diluted in the large volume of the reagent cassette. With a smaller volume of body fluid used in the DNA-extraction, the total amount in the reagent cassette is even lower.

Higher RNA-concentrations were measured for the LBD *mir*Vana™ kit samples than the LBD ReliaPrep™ kit samples. However, quantitation with short GAPDH qPCR, showed more similarity between relative quantities of *mir*Vana™ kit samples and ReliaPrep™ samples, for both volumes of all three body fluids.

The RNA-concentrations measured for the PrepFiler™ eluate, is equivalent to the previous results in paragraph 1, and RT+ samples of PrepFiler™ eluates generally had higher relative quantities than corresponding LBD RNA-isolation kit samples in short GAPDH qPCR. PrepFiler™ eluate samples have a high DNA content for the purpose of DNA analysis, and it is suspected that this might affect the RNA-concentration measurement to some degree. Because of the high DNA content, one DNase-treatment might not be sufficient for total removal. This was the reason for the two methods of gDNA contamination assessment performed for all LBD *mir*Vana™, LBD ReliaPrep™ and PrepFiler™ eluate samples.

Several of the samples that did not exceed the detection limit of the Qubit® Fluorometer, had detectable reactions on short GAPDH qPCR. This was also seen for the qPCR of random RNA-isolation kit samples described above. The Qubit® Fluorometer and GAPDH qPCR have different targets of measurement and cannot be directly compared. However, using the Qubit® Fluorometer on samples extracted from LBD to assess the efficiency of the RNA-isolation kit may not be the most reliable method of quantitation. If any further protocol optimization of RNA-isolation kits is undertaken, the Qubit® Fluorometer is not recommended for quantitation, based on the findings of this study.

gDNA contamination

gDNA contamination could lead to erroneous qPCR results, as this gDNA can be amplifiable with the chosen primers. This is especially true for housekeeping genes, as they are widely expressed. The small autosomal target in the highly sensitive Quantifiler™ Trio DNA Quantification Kit was employed for detection of gDNA contamination of DNase-treated PrepFiler™ DNA-eluate and LBD RNA-eluates before reverse transcriptase reaction. The results for the RT- samples in the short GAPDH qPCR were also used for this purpose. As 16µL of RNA eluate was used in the reverse transcriptase reaction, the RT-samples were slightly more diluted, yet comparable to the DNase-treated samples used with the Quantifiler™ Trio kit, as no reverse transcriptase was added to the RT- samples. qPCR for short GAPDH is less sensitive than Quantifiler™ Trio kit, as only one target was detected and quantified.

*mir*Vana™ samples were all treated with the TURBO™ DNase kit and a majority were marked as "Undetermined" in the Quantfiler™ Trio qPCR and short GAPDH qPCR. Any results where one triplet gave a detectable reaction might be due to external contamination or a sporadic signal obtained during the qPCR. According to the manufacturer's protocol, a C_t >38 might be obtained, even for negative control or samples with minute amounts of DNA [78]. For semen samples, which are known to contain high amounts of DNA due to a high number of cells, a single DNase treatment might not be sufficient for total removal of contaminating DNA. The case is the same for all PrepFiler™ eluate samples, which contain high levels of DNA.

Lower C_t values were obtained for ReliaPrep™ samples across all body fluids and volumes, than for *mir*Vana™ and PrepFiler™ eluate samples for both the Quantfiler™ Trio qPCR and short GAPDH qPCR. The on-column DNase-treatment of the ReliaPrep™ seems to be less effective than TURBO™ DNase (2 U/µL) treatment, despite the prolonged incubation time as recommended by ESR.

DNase-treatment with the TURBO™ DNase (2 U/µL) is the most effective for removal of DNA for the purpose of treating RNA-eluate samples from LBD.

RNA degradation using two GAPDH products

The relative quantity of longer GAPDH product was used in comparison to that of the shorter GAPDH product in qPCR. By comparing the quantities of the two products, this would give an indication of the intactness of the cDNA in the samples, as degraded RNA is more fragmented in nature. As the two GAPDH products are equally efficient under the same qPCR conditions, their results are comparable. The relative quantities obtained from qPCR were used to calculate ratios between the shorter and longer product.

Overall, the short/long GAPDH product ratios are similar between the mirVana™ and ReliaPrep™. This indicates that the level of degradation is similar for both kits and speaks for an equal quality of the extraction products. However, for all ReliaPrep™ kit results, it is important to take the possible gDNA contamination from an incomplete DNase-treatment into account.

The relative RNA quantities for the PrepFiler™ eluate samples were much more variable and resulted in higher short/long GAPDH product ratios with large SD. In light of the findings of gDNA contamination in these samples, a substantial portion of the GAPDH products detected could be derived from gDNA from an inefficient DNase-treatment. In addition, GAPDH also exists in the genome as a processed pseudogene, which could contribute to gDNA contamination [75, 88]. Additional DNase-treatments or a more robust DNase-treatment may be employed on the PrepFiler™ DNA eluate for more efficient removal of DNA, as both the quantity and quality of RNA measured for these samples still make them an interesting target for further analysis. As there is no need for an additional RNA-extraction, using the PrepFiler™ DNA eluate for RNA-analysis is advantageous in terms of saving costs and time.

Although the ReliaPrep™ kit involves less transfers and hazardous reagents, the apparent gDNA contamination might influence downstream analysis results. As the *mir*Vana™ has given the most consistent results overall, both in RNA-concentration measured by the Qubit® Fluorometer, the lack of contamination, and relative quantities of GAPDH in qPCR, it seems to be the most sufficient for the purpose of extracting RNA from LBD.

Optimization of PCR conditions for twelve selected body fluid markers

The singleplex performed with body fluid markers and housekeeping genes selected from previous studies on venous blood, saliva, semen, vaginal fluid and menstrual blood, demonstrated that the cDNA concentration was too high for several body fluids, even for the most diluted samples. Peaks of up to 34 000 rfu with pull-ups in other colour channels and additional high peaks in other bp areas were seen for both body fluid markers and housekeeping genes. Exceptions were vaginal fluid and saliva body fluid markers, which had clear peaks in expected bp areas at the 0.01 dilution samples on both the NL60 and NL64 PCR program. However, a trend can still be seen for the other body fluids, where there seemed to be less overload of the system with more diluted samples.

The shorter annealing time for the NL-programs, as compared to the NZ-program, contribute to a less PCR products in the samples with current primer concentrations, which in turn shows up as a lesser overloaded system in capillary electrophoresis.

Several markers, including TNP1, PRM1 and HBB were expected to give high peaks in the electropherograms, as they are known to be abundantly expressed. The HBB marker has previously shown to be a highly stable marker and has demonstrated to produce a clear mRNA profile in aged samples and samples subjected to harsh environmental factors such as heat and humidity. However, this extreme abundance could also influence the sensitivity of other mRNA markers, and HBB has therefore been removed from multiplexes in some studies [46, 89-92].

Further project plan and future perspectives

The work performed in this study is the starting point of establishing a method for DNA and RNA co-isolation and mRNA profiling into the routine workflows at CFG.

The protocols of both the *mir*Vana™ and ReliaPrep™ kits were optimized for RNA-isolation from PrepFiler™ LBD. A comparison of the level of degradation of the housekeeping gene GAPDH by qPCR, demonstrated that the kits isolate RNA of similar quality. In terms of overall performance in this study, the *mir*Vana™ kit is recommended for further use.

The next step would therefore be isolate RNA from LBD, amplify body fluid markers and separate them by capillary electrophoresis to determine whether this material is suited for body

fluid identification. The RNA-eluates from LBD contain less usable material and must therefore be tested before making any further adjustments or determining which PCR program is best suited.

Further work would include adjusting primer concentrations for the incorporation into a multiplex. The input of material to body fluid PCR also needs to be adjusted according to initial capillary electrophoresis results. The input of vaginal fluid and saliva seems to be adequate for current primer concentrations. However, the input of semen, menstrual blood and venous blood must be downregulated as they overload the system, even in the most diluted samples. A purification of RNA products could also be performed, to examine whether it has a potential improvement on capillary electrophoresis results.

Furthermore, a decision must be made as to whether all markers are to be incorporated into a larger multiplex, or alternatively only putting a smaller assortment together in a case-orientated fashion. Additional markers for other body fluids and tissue types may be tested. Several primers are already available at CFG and will be tested.

The method for assessing marker expression used in this study is capillary electrophoresis. It would be interesting to also evaluate marker expression by massively parallel sequencing methods, in order to determine which method is most suited for the purpose of body fluid identification.

Before establishing this mRNA profiling method into the workflows at CFG, it needs to be thoroughly tested and validated. This involves assessing marker specificity by testing samples of other body fluids, in addition to samples of other species. Sensitivity and reproducibility in more complex samples must also be examined by the use of several body fluids of high and low volumes, both pure and mixed, as well as old and new samples, and finally case samples. Donor variation (i.e. age, sex, point in time of menstrual cycle, fertility/non-fertility, use of different contraceptives, and time since last meal) must be assessed in terms of sensitivity and reproducibility.

Conclusion

This Master's study aimed to optimize a method of extracting RNA from DNA-extraction waste product (lysis buffer discard) of the PrepFiler *Express* ™ Forensic DNA Extraction Kit, as well as optimizing PCR amplification conditions for a set of published mRNA body fluid markers. This study showed reproducible results to preliminary findings indicating that the PrepFiler *Express*™ DNA lysis buffer discard as well as DNA-eluate contain RNA that could potentially be used for body fluid identification. The optimization of the protocols of three RNA-isolation kits for the purification of RNA from lysis buffer discard, and the further analyses for comparison of RNA quantity and quality, showed that the *mir*Vana™ miRNA Isolation Kit is best suited for this purpose. An adjustment of reagent ratios for the peak RNA-concentration was achieved for the *mir*Vana™ miRNA Isolation Kit and the ReliaPrep™ RNA Miniprep System, which also showed similar relative quantities and degradation levels of RNA in a qPCR of two products of the housekeeping gene GAPDH. A reoccurring challenge was an apparent inefficiency of the on-column DNAse-treatment of the ReliaPrep™ kit. To which degree the contamination gDNA affects GAPDH qPCR results is not determined. The Direct-zol™ RNA Miniprep kit was not further analysed, as it had no samples exceeding the detection limit of the Qubit® Fluorometer, and low relative quantities in samples run on GAPDH qPCR. An overload of the system was observed for most of the twelve body fluid markers under all three PCR conditions in the capillary electrophoresis. Primer tests should be rerun with less RNA, e.g. RNA extracted from lysis buffer discard instead of pure body fluids, before further optimization.

References

[1] E. M. Robinson, *Crime scene photography*. Academic Press, 2016.

[2] M. Hashiyada, "DNA biometrics," in *Biometrics*: IntechOpen, 2011.

[3] A. J. Goodwin, *Acute Lung Injury and Repair: Scientific Fundamentals and Methods*. Humana Press, 2016.

[4] D. Gerhold, T. Rushmore, and C. T. Caskey, "DNA chips: promising toys have become powerful tools," *Trends in biochemical sciences,* vol. 24, no. 5, pp. 168-173, 1999.

[5] A. Bowden, R. Fleming, and S. Harbison, "A method for DNA and RNA co-extraction for use on forensic samples using the Promega DNA IQ™ system," *Forensic Science International: Genetics,* vol. 5, no. 1, pp. 64-68, 2011.

[6] M. Arenas *et al.,* "Forensic genetics and genomics: Much more than just a human affair," *PLOS Genetics,* vol. 13, no. 9, p. e1006960, 2017, doi: 10.1371/journal.pgen.1006960.

[7] P. M. Schneider, "Scientific standards for studies in forensic genetics," *Forensic science international,* vol. 165, no. 2-3, pp. 238-243, 2007.

[8] M. A. Jobling and P. Gill, "Encoded evidence: DNA in forensic analysis," *Nature Reviews Genetics,* vol. 5, no. 10, pp. 739-751, 2004.

[9] P. D. Martin, H. Schmitter, and P. M. Schneider, "A brief history of the formation of DNA databases in forensic science within Europe," *Forensic science international,* vol. 119, no. 2, pp. 225-231, 2001.

[10] S. Blackman *et al.,* "Developmental validation of the ParaDNA® Body Fluid ID System—A rapid multiplex mRNA-profiling system for the forensic identification of body fluids," *Forensic Science International: Genetics,* vol. 37, pp. 151-161, 2018.

[11] M. van den Berge, B. Bhoelai, J. Harteveld, A. Matai, and T. Sijen, "Advancing forensic RNA typing: on non-target secretions, a nasal mucosa marker, a differential co-extraction protocol and the sensitivity of DNA and RNA profiling," *Forensic Science International: Genetics,* vol. 20, pp. 119-129, 2016.

[12] A. Lindenbergh *et al.,* "A multiplex (m) RNA-profiling system for the forensic identification of body fluids and contact traces," *Forensic Science International: Genetics,* vol. 6, no. 5, pp. 565-577, 2012.

[13] J. Juusola and J. Ballantyne, "Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification," *Forensic science international,* vol. 135, no. 2, pp. 85-96, 2003.

[14] J. Juusola and J. Ballantyne, "Multiplex mRNA profiling for the identification of body fluids," *Forensic science international,* vol. 152, no. 1, pp. 1-12, 2005.

[15] T. Sijen, "Molecular approaches for forensic cell type identification: on mRNA, miRNA, DNA methylation and microbial markers," *Forensic Science International: Genetics,* vol. 18, pp. 21-32, 2015.

[16] C. Trapnell, "Defining cell types and states with single-cell genomics," *Genome research,* vol. 25, no. 10, pp. 1491-1498, 2015.

[17] J. Walter and A. Hümpel, "Introduction to epigenetics," in *Epigenetics*: Springer, 2017, pp. 11-29.

[18] D. Arendt *et al.*, "The origin and evolution of cell types," *Nature Reviews Genetics,* vol. 17, no. 12, pp. 744-757, 2016.

[19] UniProtKB - P02787 (TRFE_HUMAN) [Online] Available: https://www.uniprot.org/uniprot/P02787

[20] L.-A. MacFarlane and P. R Murphy, "MicroRNA: biogenesis, function and role in cancer," *Current genomics,* vol. 11, no. 7, pp. 537-561, 2010.

[21] M. Vennemann and A. Koppelkamm, "mRNA profiling in forensic genetics I: Possibilities and limitations," *Forensic science international,* vol. 203, no. 1-3, pp. 71-75, 2010.

[22] D. Zubakov, E. Hanekamp, M. Kokshoorn, W. van IJcken, and M. Kayser, "Stable RNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples," *International journal of legal medicine,* vol. 122, no. 2, pp. 135-142, 2008.

[23] M. Setzer, J. Juusola, and J. Ballantyne, "Recovery and stability of RNA in vaginal swabs and blood, semen, and saliva stains," *Journal of forensic sciences,* vol. 53, no. 2, pp. 296-305, 2008.

[24] D. Zubakov, A. W. Boersma, Y. Choi, P. F. van Kuijk, E. A. Wiemer, and M. Kayser, "MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation," *International journal of legal medicine,* vol. 124, no. 3, pp. 217-226, 2010.

[25] E. K. Hanson, H. Lubenow, and J. Ballantyne, "Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs," *Analytical biochemistry,* vol. 387, no. 2, pp. 303-314, 2009.

[26] W. J. Ewens, *Population genetics*. Springer Science & Business Media, 2013.

[27] J. M. Butler, *Fundamentals of forensic DNA typing*. Academic press, 2009.

[28] M. Lynch, S. A. Cole, R. McNally, and K. Jordan, *Truth machine: The contentious history of DNA fingerprinting*. University of Chicago Press, 2010.

[29] "The history of genetic fingerprinting." https://le.ac.uk/dna-fingerprinting/history (accessed.

[30] A. J. Jeffreys, V. Wilson, R. Neumann, and J. Keyte, "Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells," *Nucleic Acids Research,* vol. 16, no. 23, pp. 10953-10971, 1988.

[31] A. J. Jeffreys, V. Wilson, and S. L. Thein, "Hypervariable 'minisatellite'regions in human DNA," *Nature,* vol. 314, no. 6006, pp. 67-73, 1985.

[32] W. contributors. "Colin Pitchfork." Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Colin_Pitchfork&oldid=1017090416 (accessed 14.05.2021.

[33] G. A. Williams, "Body fluid identification: A case for more research and innovation," *Forensic Science International: Reports,* vol. 2, p. 100096, 2020.

[34] J. H. An, K.-J. Shin, W. I. Yang, and H. Y. Lee, "Body fluid identification in forensics," 2012.

[35] D. G. Altman and J. M. Bland, "Diagnostic tests. 1: Sensitivity and specificity," *BMJ: British Medical Journal,* vol. 308, no. 6943, p. 1552, 1994.

[36] Phadebas. "Phadebas in Forensic Biology." https://www.phadebas.com/areas-of-use/forensic-biology/ (accessed.

[37] H. D. Janowitz and D. A. Dreiling, "The plasma amylase: source, regulation and diagnostic significance," *The American journal of medicine,* vol. 27, no. 6, pp. 924-935, 1959.

[38] E. H. White, O. Zafiriou, H. H. Kagi, and J. H. Hill, "Chemilunimescence of luminol: The chemcial reaction," *Journal of the American Chemical Society,* vol. 86, no. 5, pp. 940-941, 1964.

[39] M. Oehmichen and K. Zilles, "Postmortem DNA and RNA synthesis. Preliminary studies in human cadavers," *Zeitschrift fur Rechtsmedizin. Journal of legal medicine,* vol. 91, no. 4, pp. 287-294, 1984.

[40] T.-W. Phang, C. Shi, J. Chia, and C. Ong, "Amplification of cDNA via RT-PCR using RNA extracted from postmortem tissues," *Journal of Forensic Science,* vol. 39, no. 5, pp. 1275-1279, 1994.

[41] C. A. Heid, J. Stevens, K. J. Livak, and P. M. Williams, "Real time quantitative PCR," *Genome research,* vol. 6, no. 10, pp. 986-994, 1996.

[42] M. Bauer and D. Patzelt, "Evaluation of mRNA markers for the identification of menstrual blood," *Journal of Forensic Science,* vol. 47, no. 6, pp. 1278-1282, 2002.

[43] M. Bauer, A. Kraus, and D. Patzelt, "Detection of epithelial cells in dried blood stains by reverse transcriptase-polymerase chain reaction," *Journal of Forensic Science,* vol. 44, no. 6, pp. 1232-1236, 1999.

[44] E. K. Hanson and J. Ballantyne, "Rapid and inexpensive body fluid identification by RNA profiling-based multiplex High Resolution Melt (HRM) analysis," *F1000Research,* vol. 2, 2013.

[45] R. I. Fleming and S. Harbison, "The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids," (in eng), *Forensic Sci Int Genet,* vol. 4, no. 4, pp. 244-56, Jul 2010, doi: 10.1016/j.fsigen.2009.10.006.

[46] C. Haas *et al.*, "mRNA profiling for the identification of blood—results of a collaborative EDNAP exercise," *Forensic Science International: Genetics,* vol. 5, no. 1, pp. 21-26, 2011.

[47] P. P. Albani and R. Fleming, "Novel messenger RNAs for body fluid identification," *Science & justice,* vol. 58, no. 2, pp. 145-152, 2018.

[48] M.-H. Lin, D. F. Jones, and R. Fleming, "Transcriptomic analysis of degraded forensic body fluids," *Forensic Science International: Genetics,* vol. 17, pp. 35-42, 2015.

[49] A. P. Salzmann *et al.*, "mRNA profiling of mock casework samples: Results of a FoRNAP collaborative exercise," *Forensic Science International: Genetics,* vol. 50, p. 102409, 2021.

[50] E. Hanson, S. Ingold, C. Haas, and J. Ballantyne, "Messenger RNA biomarker signatures for forensic body fluid identification revealed by targeted RNA sequencing," *Forensic Science International: Genetics,* vol. 34, pp. 206-221, 2018.

[51] S. Ingold *et al.*, "Body fluid identification and assignment to donors using a targeted mRNA massively parallel sequencing approach–results of a second EUROFORGEN/EDNAP collaborative exercise," *Forensic Science International: Genetics,* vol. 45, p. 102208, 2020.

[52] A. C. Jäger *et al.*, "Developmental validation of the MiSeq FGx forensic genomics system for targeted next generation sequencing in forensic DNA casework and database laboratories," *Forensic Science International: Genetics*, vol. 28, pp. 52-70, 2017.

[53] Q. Zhang *et al.*, "Evaluation of the performance of Illumina's ForenSeq™ system on serially degraded samples," *Electrophoresis*, vol. 39, no. 21, pp. 2674-2684, 2018.

[54] P. R. Haddrill, "Developments in forensic DNA analysis," *Emerging Topics in Life Sciences*, 2021.

[55] M. Alvarez, J. Juusola, and J. Ballantyne, "An mRNA and DNA co-isolation method for forensic casework samples," *Analytical biochemistry*, vol. 335, no. 2, pp. 289-298, 2004.

[56] S. S. Gårdvik, "Development of a method for co-extraction of RNA and DNA from biological traces for use in forensic analysis," Bachelor Bachelor thesis, Det Helsevitenskapelige fakultet - Institutt for medisinsk biologi, UiT - The Arctic University of Norway, 2020.

[57] "PrepFiler Express™ and PrepFiler Express™ Forensic DNA Extraction Kits." [Online]. Available: https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets/LSG/manuals/4443104_PrepFilrExPFExBTA_QR.pdf

[58] "AutoMate Express™ Instrument User Guide." [Online]. Available: https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets/LSG/manuals/4441982_AutoMateExp_UG.pdf

[59] S. Fjelstrup *et al.*, "The effects of dithiothreitol on DNA," *Sensors*, vol. 17, no. 6, p. 1201, 2017.

[60] H. McKiernan and P. Danielson, "Molecular diagnostic applications in forensic science," in *Molecular diagnostics*: Elsevier, 2017, pp. 371-394.

[61] "mirVana™ miRNA Isolation Kit." [Online]. Available: https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2Fcms_055423.pdf&title=bWlyVmFuYSZ0cmFkZTsgbWlSTkEgSXNvbGF0aW9uIEtpdCAoRW5nbGlzaCAp

[62] "Qubit® RNA HS Assay Kits."

[63] A. Schroeder *et al.*, "The RIN: an RNA integrity number for assigning integrity values to RNA measurements," *BMC molecular biology*, vol. 7, no. 1, pp. 1-14, 2006.

[64] O. Mueller *et al.*, "A microfluidic system for high-speed reproducible DNA sizing and quantitation," *ELECTROPHORESIS: An International Journal*, vol. 21, no. 1, pp. 128-134, 2000.

[65] "Applications for DNA, RNA, and Protein Analysis," *Agilent 2100 Bioanalyzer System*.

[66] T. Matsubara *et al.*, "DV200 index for assessing RNA integrity in next-generation sequencing," *BioMed research international*, vol. 2020, 2020.

[67] E. Graf, "Simplified DV200 Evaluation with the Agilent 2100 Bioanalyzer System." [Online]. Available: https://www.agilent.com/cs/library/technicaloverviews/public/5991-8287EN.pdf

[68] "Acid-Phenol: Chloroform."

[69] "Direct-zol™ DNA/RNA Miniprep." [Online]. Available: https://files.zymoresearch.com/protocols/r2080t_r2080_r2081_direct-zol_dna_rna_miniprep.pdf

[70] invitrogen. *TRIzol™ Reagent*, Revision B.0 ed., ThermoFisher Scientific, 15596026 15596018, 2020. [Online]. Available: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/trizol_reagent.pdf.

[71] "ReliaPrep™ RNA Cell Miniprep System." [Online]. Available: file:///C:/Users/kari_/Downloads/ReliaPrep%20RNA%20Cell%20Miniprep%20System%20TM370%20(2).pdf

[72] "TURBO DNA-free™ Kit." [Online]. Available: https://assets.fishersci.com/TFS-Assets/LSG/manuals/1907M_turbodnafree_UG.pdf

[73] "SuperScript™ IV VILO™ Master Mix." [Online]. Available: https://www.thermofisher.com/order/catalog/product/11754050#/11754050

[74] K. Gorzelniak, J. Janke, S. Engeli, and A. M. Sharma, "Validation of endogenous controls for gene expression studies in human adipocytes and preadipocytes," (in eng), *Horm Metab Res*, vol. 33, no. 10, pp. 625-7, Oct 2001, doi: 10.1055/s-2001-17911.

[75] J. Juusola and J. Ballantyne, "mRNA profiling for body fluid identification by multiplex quantitative RT-PCR," (in eng), *J Forensic Sci*, vol. 52, no. 6, pp. 1252-62, Nov 2007, doi: 10.1111/j.1556-4029.2007.00550.x.

[76] "TaqMan® Fast Advanced Master Mix USER GUIDE." [Online]. Available: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/cms_084554.pdf

[77] M. L. Wong and J. F. Medrano, "Real-time PCR for mRNA quantitation," (in eng), *Biotechniques*, vol. 39, no. 1, pp. 75-85, Jul 2005, doi: 10.2144/05391rv01.

[78] "Quantifiler™ HP and Trio DNA Quantification Kits USER GUIDE." [Online]. Available: https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets/LSG/manuals/4485354.pdf

[79] P. P. Albani and R. Fleming, "Developmental validation of an enhanced mRNA-based multiplex system for body fluid and cell type identification," (in eng), *Sci Justice*, vol. 59, no. 3, pp. 217-227, May 2019, doi: 10.1016/j.scijus.2019.01.001.

[80] "DNA Fragment Analysis by Capillary Electrophoresis." [Online]. Available: https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets/LSG/manuals/4474504.pdf

[81] J. M. Butler, *Advanced topics in forensic DNA typing: methodology*. Academic press, 2011.

[82] A. Wada and A. Suyama, "Local stability of DNA and RNA secondary structure and its relation to biological functions," *Progress in biophysics and molecular biology*, vol. 47, no. 2, pp. 113-157, 1986.

[83] H. Nielsen, "Working with RNA," in *RNA*: Springer, 2011, pp. 15-28.

[84] W. contributors, "Ribonuclease," in *Wikipedia, The Free Encyclopedia*, ed: Wikipedia, The Free Encyclopedia, 2020.

[85] "Working with RNA: the basics." [Online]. Available: http://assets.thermofisher.com/TFS-Assets/BID/Technical-Notes/rnasezap-rnase-decontamination-solution-tech-note.pdf

[86] "Precautions for Handling RNA." LifeScience. https://lifescience.roche.com/en_no/articles/precautions-for-handling-of-rna.html (accessed 2021).

[87] G. Caruana and R. Mcinnes, "Assessing genomic DNA contamination of total RNA isolated from kidney cells obtained by Laser Capture Microdissection using the Agilent RNA 6000 Pico assay Application," 2004.

[88] Z. Zhang, P. M. Harrison, Y. Liu, and M. Gerstein, "Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human

genome," (in eng), *Genome Res,* vol. 13, no. 12, pp. 2541-58, Dec 2003, doi: 10.1101/gr.1429003.

[89] G. Kulstein, U. Schacker, and P. Wiegand, "Old meets new: Comparative examination of conventional and innovative RNA-based methods for body fluid identification of laundered seminal fluid stains after modular extraction of DNA and RNA," *Forensic Science International: Genetics,* vol. 36, pp. 130-140, 2018/09/01/ 2018, doi: https://doi.org/10.1016/j.fsigen.2018.06.017.

[90] F. Kohlmeier and P. M. Schneider, "Successful mRNA profiling of 23 years old blood stains," (in eng), *Forensic Sci Int Genet,* vol. 6, no. 2, pp. 274-6, Mar 2012, doi: 10.1016/j.fsigen.2011.04.007.

[91] C. Mayes, R. Houston, S. Seashols-Williams, B. LaRue, and S. Hughes-Stamm, "The stability and persistence of blood and semen mRNA and miRNA targets for body fluid identification in environmentally challenged and laundered samples," *Legal Medicine,* vol. 38, pp. 45-50, 2019/05/01/ 2019, doi: https://doi.org/10.1016/j.legalmed.2019.03.007.

[92] Z. Liu *et al.*, "A method of identifying the blood contributor in mixture stains through detecting blood-specific mRNA polymorphism," *ELECTROPHORESIS,* vol. 41, no. 15, pp. 1364-1373, 2020, doi: https://doi.org/10.1002/elps.202000053.

Appendix 1: RNA quantification results

Appendix table 1: *The measured RNA concentrations of all samples by the Qubit® 4 Fluorometer and the Agilent 2100 BioAnalyzer.*

	Qubit® 4 Fluorometer (ng/µL)								
	RNA eluate			DNA eluate			LBD		
Blood	8.23	7.60	6.87	2.05	3.23	4.26	0.65	0.67	0.69
Semen	5.41	5.81	5.68	1.09	1.40	0.70	0.00	0.00	0.00
Saliva	4.30	5.05	3.09	3.93	4.55	4.55	0.00	0.00	0.00
	Agilent 2100 BioAnalyzer (pg/µL)								
	RNA eluate			DNA eluate			LBD		
Blood	5,479	3,236	3,784	1,650	2,290	3,357	0	56	29
Semen	1,713	1,767	1,758	237	565	65	23	30	5
Saliva	4,005	4,772	1,581	10,129	4,282	4,432	6	19	3

Appendix 2: Calculated values for RIN and DV_{200}

Appendix table 2: *Calculated values for RIN and DV_{200} for all samples*

	RIN								
	RNA eluate			DNA eluate			LBD		
Blood	6.5	6.6	6.7	2.7	2.9	2.8	N/A	N/A	N/A
Semen	2.4	2.5	2.5	2.7	2.3	N/A	N/A	N/A	N/A
Saliva	2.5	2.6	2.4	N/A	N/A	N/A	N/A	N/A	N/A
	DV_{200}								
	RNA eluate			DNA eluate			LBD		
Blood	65 %	66 %	82 %	76 %	83 %	83 %	N/A	N/A	N/A
Semen	50 %	45 %	43 %	17 %	64 %	74 %	N/A	N/A	N/A
Saliva	27 %	25 %	20 %	52 %	69 %	82 %	N/A	N/A	N/A

Appendix 3: In-house procedure of the Institute of Environmental Science and Research Limited (ESR), New Zealand

ReliaPrep™ RNA Purification

Date:_____ Analyst:_____ Batch no(s):_____

PCR Workstation:_____

Note: Ensure ethanol has been added to RNA Wash Solution and Column Wash Solution

		Samples	Controls
1.	If lysis is in a 0.5mL nunc tube, transfer lysis to a new labelled 1.5mL tube. Verify:		
2.	Add the appropriate amount of 100% isopropanol (see table) to each tube containing lysis buffer and mix by vortexing for 5 seconds		
3.	Transfer all lysate to a minicolumn in a collection tube. Retain the lysate tube. Verify:		
4.	Centrifuge at 12,000 rpm for 45 seconds and discard flow through using pipette		
5.	Add 500µL of RNA Wash Solution to the minicolumn, centrifuge at 12,000 rpm for 45 seconds and again at 13,000 rpm for 2 minutes. Discard flow through.		
6.	Air dry the minicolumns for 2 minutes		
7.	Add 30µL of prepared DNase I Incubation Solution to the minicolumn. *Use specific pipette*		
8.	Incubate at 37°C for 40 minutes		
9.	Add 200µL of Column Wash Solution to the minicolumn and centrifuge at 12,000 rpm for 30 seconds		
10.	Add 500µL of RNA Wash Solution to the minicolumn and centrifuge at 12,000 rpm for 45 seconds		
11.	Place the minicolumn into a new collection tube and discard the old collection tube containing the flow through		
12.	Add 300µL RNA Wash Solution to the minicolumn, centrifuge at 13,000 rpm for 2 minutes and then again for 1 minute		
13.	Air dry the minicolumns for 5 minutes		
14.	Transfer the minicolumn from the collection tube to a new labelled elution tube Verify:		
15.	Add 35µL of nuclease free water to the minicolumn and incubate at room temperature for 1 minute		
16.	Centrifuge at 12,000 rpm for 1 minute		
17.	Add a further 35µL of nuclease free water to the minicolumn and incubate at room temperature for 1 minute		
18.	Centrifuge at 12,000 rpm for 1 minute for a total volume of 70 µL		
19.	Discard the minicolumn and store the elution tube at -20°C until required		

DNase I Incubation Solution:

Number of samples + 1: _____ (A)

Yellow Core Buffer:	24µL x A =_____ (B)
$MnCl_2$, 0.009M	3µL x A = _____ (C)
DNase I *(use specific pipette)*:	3µL x A = _____ (D)
Total:	_____

Combine B, C and D in an appropriate container and mix well

Isopropanol:

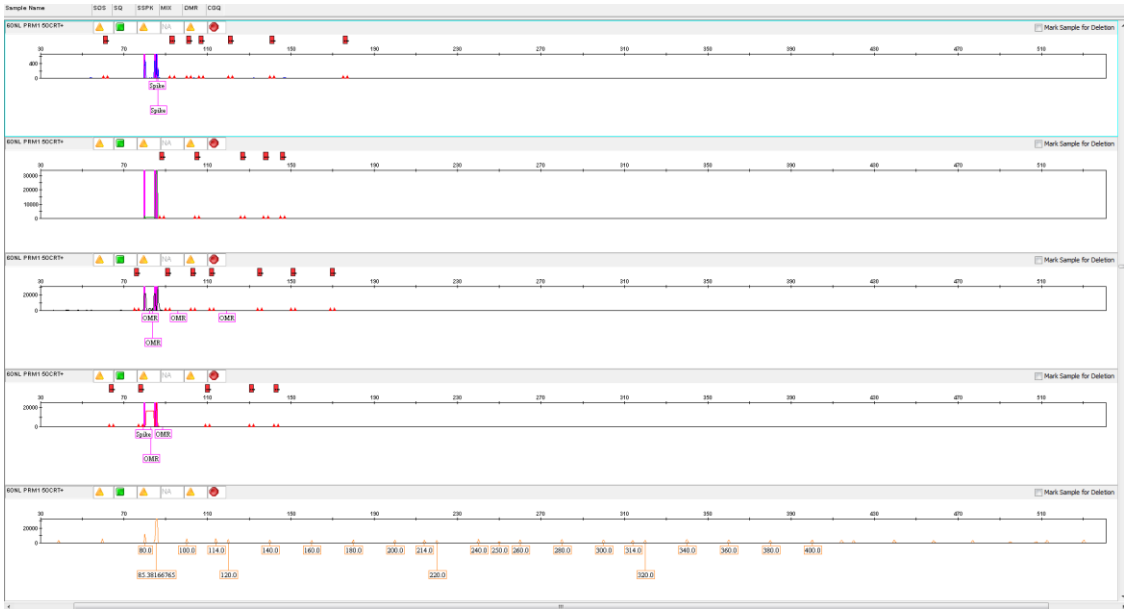
Lysis buffer used	100% isopropanol	No. of samples	Volume required
100µL	34µL		
250µL	85µL		
350µL	120µL		
475µL	161µL		
500µL	170µL		
Total Volume			

Reagent	**Lot Number and date received**	**Expiry**	**QC pass**
ReliaPrep™ kit			
Isopropanol			
Purification Positive Control(s)			

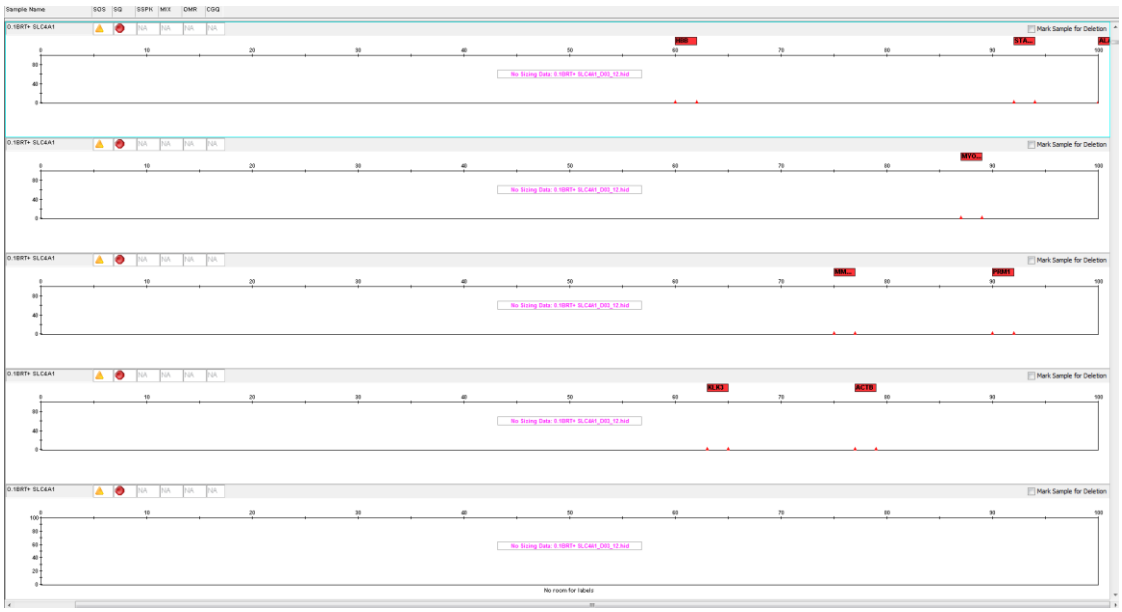
Appendix 4: Heatmap criteria electropherograms

	1. Overload. One or more high peaks in expected base pair area that give very high pull-ups in several of the other colour channels. 2. Samples that give no results because of total overload of system (No Size Data)

1.

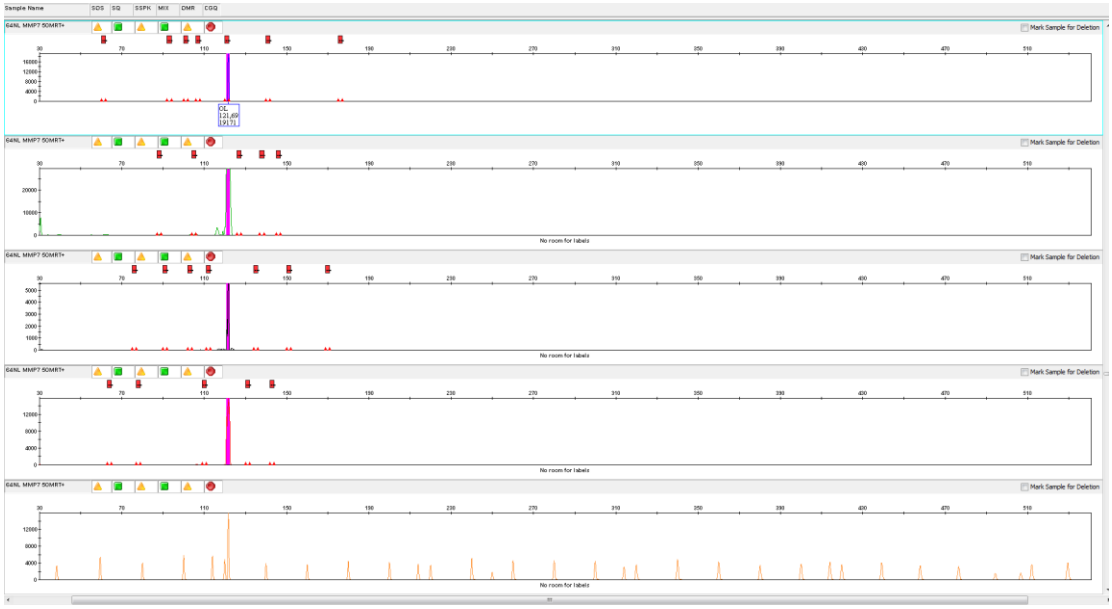


2.



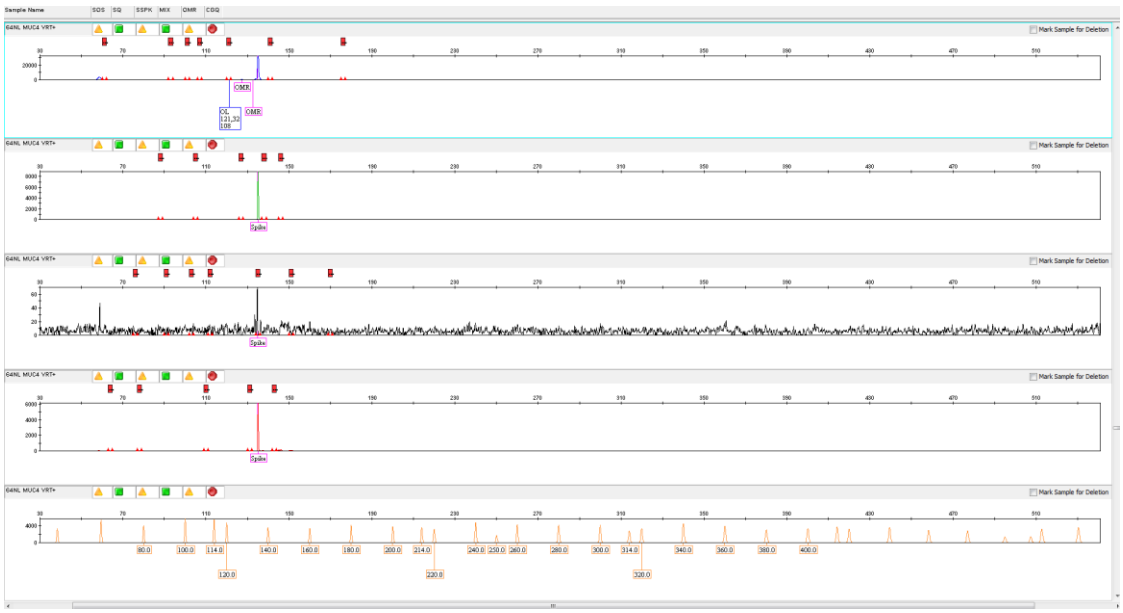
	3. A clear, high peak in expected base pair area, but peak of interest is <4x higher than pull-ups and other, extra peaks.

3.



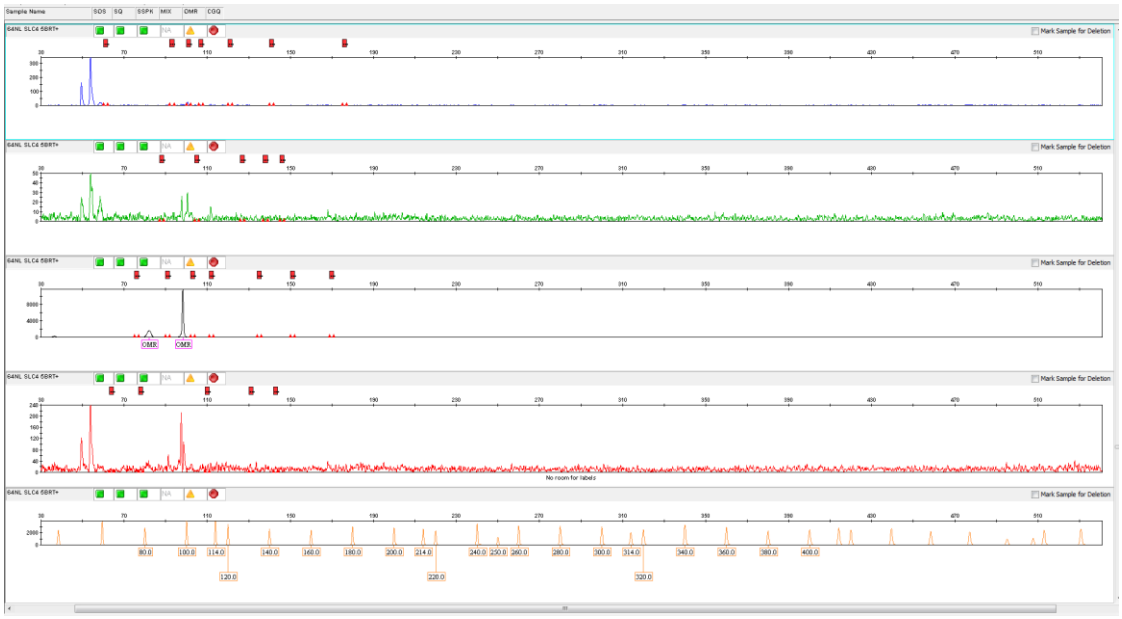
	4. A clear, high peak in expected base pair area. Peak of interest is 4x-7x higher than other extra peaks in other areas. Pull-ups are >8x lower than the main peak.

4.



	5. A clear peak in expected base pair area. Peak of interest is >x higher than other extra peaks.

5.



White	6. No peaks in expected base pair area or peaks of <700rfu

6.