U.PORTO

MESTRADO EM CIÊNCIAS FORENSES

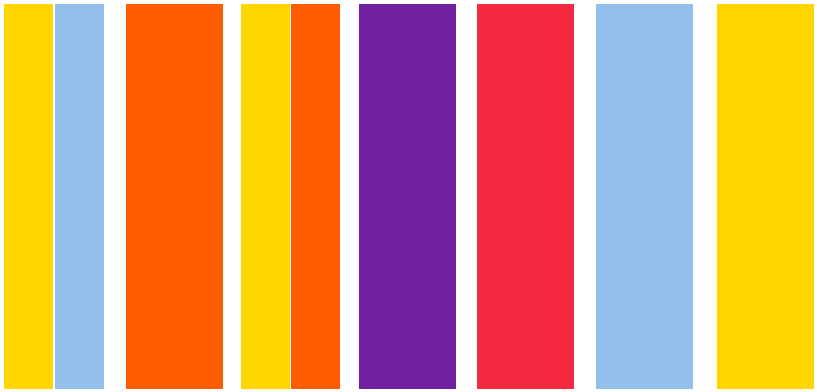
CIÊNCIAS LABORATORIAIS FORENSES

Identification of the source of biological fluids recovered from samples collected at Sexual Assaults scenarios using mRNA markers

Sara Duarte Loureiro

M

2019



Dissertation for applying for a Master's degree in Forensic Genetics, submitted to the Faculty of Medicine of the University of Porto.

The present work was developed under the scientific supervision of Laura Cainé, (INMLCF, I.P and FMUP) and Iva Gomes (i3s and IPATIMUP).

ACKNOWLEDGMENT

This dissertation would not be possible without the support of some very important people. I want to express my gratitude to all who contributed directly or indirectly to this chapter of my life.

First, I would like to thank Dr. Laura Cainé, my advisor, for accepting me as her guide, for all your help, and confidence in my work and for giving me the opportunity to do this internship at Serviço de Genética e Biologia Forense, da Delegação do Norte (SGBF-N) do Instituto Nacional de Legal Medicina Legal e Ciências Forenses, IP, (INMLCF).

To Dr. Iva Gomes, my co-advisor, a special thanks for all the teachings, all the support and availability demonstrated during these months that contributed to the evolution of my dissertation.

To INMLCF I.P. and i3s for accepting and believing in my dissertation project. A special thanks to the Forensic Genetics Laboratory and the Population Genetics Laboratory for providing me with all the means necessary to carry out this project.

To all volunteers who without them, this study would not be possible.

To my lab colleagues for your patience and for all the support.

To my true friends for all the words of encouragement and motivation during these months and for understanding all the declined invitations. Thank you all for your true friendship!

To my parents and grandfather for your love and your unconditional support. Thanks for always believing in my decisions and for accompanying me through all the stages of my life.

To all others not mentioned but no less important, a sincere thanks.

RESUMO

Em investigações forenses, nomeadamente em agressões sexuais, é necessária a deteção e determinação da natureza biológica de manchas de fluídos, uma vez que estes possuem extrema importância na resolução jurídica dos casos, nomeadamente na reconstrução das circunstâncias do crime. Muitas vezes, a identificação e análise destes vestígios constituem um obstáculo devido às limitações dos métodos convencionais comummente utilizados. Desta forma, e devido às suas potenciais vantagens, os métodos moleculares, nomeadamente a análise da expressão génica diferencial, através de RNA mensageiro (mRNA), têm sido distinguidos como técnicas emergentes. Estas técnicas têm sido cada vez mais foco de peculiar interesse por parte dos investigadores na identificação de fluidos e tecidos corporais para aplicações forenses.

Desta forma, o presente estudo tem como objetivo comprovar a eficácia do estudo de mRNA como potencial biomarcador forense analisando os padrões de expressão génica de diferentes tecidos presentes em amostras de sémen e secreções vaginais. A escolha destes fluídos prendeu-se com o facto de estes serem os mais frequentemente, colhidos em cenários de alegadas agressões sexuais. Adicionalmente, pretende-se testar a sensibilidade desta técnica através da exposição destes fluídos corporais a diferentes fatores de diluição.

Futuramente, espera-se a expansão do presente método de identificação a outros fluídos biológicos de interesse em investigação criminal e a deteção dos mesmos em misturas complexas.

Com a validação desta técnica como futura prática recorrente em laboratórios de âmbito forense Portugueses prevê-se um impacto positivo na comunidade científica. A potencial aplicação deste procedimento, sempre que necessário, juntamente com os métodos tradicionais irá permitir a superação das limitações atuais na identificação de fluídos corporais e a possível obtenção de respostas mais eficientes em cenários forenses.

Palavras-chave: Agressões Sexuais, Genética Forense, fluídos corporais, biomarcadores Forenses, mRNA profiling.

ABSTRACT

In forensic investigations, in certain circumstances, namely in cases of sexual aggressions, it may be necessary to detect and identify the biological nature of fluid stains. This is because they play an extremely important role in the resolution of legal cases, as for example, in the reconstruction of the crime circumstances. The identification and analysis of these biological traces are often very difficult due to the limitations of the conventional methods currently used. Due to its potential advantages, molecular methods, namely the analysis of differential gene expression through messenger RNA (mRNA) have been distinguished as emergent techniques. These techniques have been increasing over the past years and focus on the research interest in identifying body fluids and tissues for forensic applications.

The present study aims to prove the efficacy of mRNA as a potential forensic biomarker by analyzing the patterns of gene expression of different tissues present in semen and vaginal secretion samples. The choice of these fluids was due to the fact that they are most often collected in scenarios of alleged Sexual Assault.

In addition, it's also intended to test the sensitivity of this technique by exposing these body fluids to different dilution factors and to test the specificity of the mRNA markers chosen in this study for semen and vaginal secretion identification.

In the future, it is expected that the present identification method may be expanded to other biological fluids of interest in criminal investigation and their detection in complex mixtures.

With the validation of this mRNA-based technique as a future recurring practice in Portuguese forensic laboratories, a positive impact in the scientific community is expected. The potential application of this molecular method whenever necessary, together with the traditional methods already in practice, will allow overcoming the current limitations in body fluid identification and obtain more efficient responses in forensic scenarios.

Keywords: Sexual Assaults, Forensic Genetics, body fluids, Forensic biomarkers, mRNA profiling.

ABBREVIATIONS

APAV	Associação Portuguesa de Apoio à Vítima (in portuguese)
APS	Ammonium Persulfate
bp	base pairs
ºC	Celsius degrees
cDNA	Complementary deoxyribonucleic acid
CE	Capillary Electrophoresis
CPP	Portuguese Penal Code
CYP2A6	Cytochrome P450 Family 2 Subfamily A Member 6
CYP2B7P1	Cytochrome P450 family 2 subfamily B member 7, pseudogene
DKK4	Dickkopf-related protein 4
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Ditiotreitol
EDTA	Ethylenediamine Tetraacetic Acid
FAM83D	Family with Sequence Similarity 83 Member D
FME	Forensic Medical Examination
FUT6	Fucosyltransferase 6
g	Grams
H_2O	Water
HCL	HydroChloric Acid
HBD1	Beta-Defensin 1 Gene
Hi-Di	Highly-Deionized
INMLCF	Instituto Nacional de Medicina Legal e Ciências Forenses
IL19	Interleukin 19
IR	Infrared
KLK2	Kallikrein Related Peptidase 2 Gene
KLK3	Kallikrein Related Peptidase 3 Gene
M	Molar
min	Minutes

ml	Mililiter
MMPS	Matrix Metallopeptidase
mRNA	messenger RNA
MSLN	Mesothelin Gene
MSMB	Microseminoprotein Beta Gene
MUC4	Mucin 4 Gene
MYOZ1	Myozenin 1 Gene
NH4OAc	Ammonium acetate
NKX3-1	NK3 Homeobox 1
ng	Nanogram
nm	Nanometer
PAGE	Polyacrylamide gel electrophoresis
PAP	Prostatic Acid Phosphatase
PCR	Polymerase Chain Reaction
pg	Picogram
PMI	*Post-mortem* interval
PRM1	Protamine 1 Gene
PRM2	Protamine 2 Gene
pre-mRNA	Primary mRNA
PSA	*Prostate-Specific Antigen*
RFU	Relative Fluorescence Units
RISC	RNA-induced silencing complex
RNA	Ribonucleic Acid
RNase	Ribonuclease
RT	Reverse transcriptase
s	Seconds
SDS	Sodium Dodecyl Sulfate
SEMG1	Gene Seminogelin 1
SEMG2	Gene Seminogelin 2
SFTA2	Surfactant Associated 2
Sg	Seminogelins

Taq	*Thermus aquaticus*
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
TGM4	Transglutaminase 4 Gene
TNP1	Transition Protein 1 Gene
Tris	tris(hydroxymethyl)aminomethane
UV	Ultraviolet
V	Volts
WHO	World Health Organization
xg	Gravitational force
µl	Microliter
µM	Micromolar
%	Percentage

TABLE OF CONTENTS

FIGURE INDEX

TABLE INDEX

GRAPHICS INDEX

INTRODUCTION

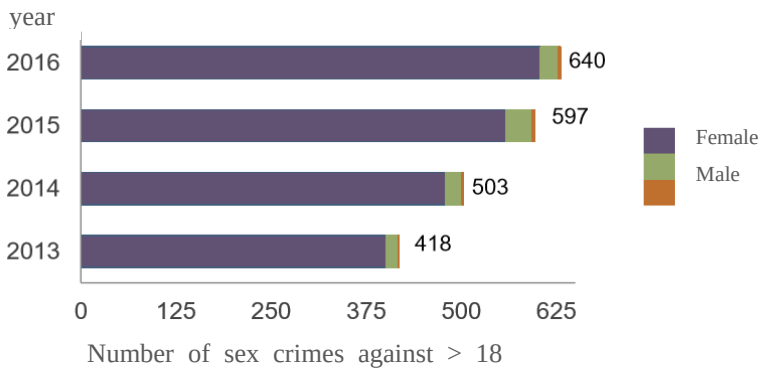
1. Sexual Assaults

The World Health Organization (WHO) presents the concept of sexual violence as "any sexual act, attempt to consummate a sexual act or unwanted sexual innuendo; or actions to market or otherwise use a person's sexuality through coercion by another person, regardless of the person's relationship with the victim, in any setting, including the home and the workplace " (WHO, 2012).

Crimes of sexual nature are considered a serious violation of human rights and are seen as a public problem that crosses all society classes. Although the social concept and its punishment have undergone a significant evolution in the last years, a resolution of these crimes still constitute a huge challenge. The concealment of cases by victims, supported by various social, cultural and psychological factors including lack of support, feelings of fear and shame, contribute to the underestimated values of world statistics (WHO, 2018).

In Portugal, the statistics provided by the Associação Portuguesa de Apoio à Vítima (APAV) regarding sexual crimes with victims aged 18 years and over, between 2013 and 2016, report a significant increase in the number of occurrences over the years (Graphic 1). The APAV's annual report of 2017 shows a total of 617 cases where approximately 85.3% of registered victims are female, and more than 80% of perpetrators are males. According to the official data, these crimes belong to the crime sector with the highest prevalence rate in our society (APAV, 2017; 2018).

Currently, to ban crimes against sexual freedom and sexual self-determination, Portuguese Law employs Chapter V of the Portuguese Penal Code (CPP), which describes in detail the various forms of sexual violence and their associated crimes.



Graphic 1. *Graphic representing the number of sexual offenses against individuals aged ≥18 years reported by APAV between 2013 and 2016. Adapted from: APAV Statistics. Sexual crimes> 18 years 2013-2016(www.apav.pt/estatisticas).*

Subtitle: DK = Don't know; DA = Don't answer

1.1 The role of biological criminalistics in cases of Sexual Assaults

The criminal investigation of Sexual Assault cases involves the connection of various areas, such as Clinical Forensics, Forensic Psychology and Forensic Genetics (genetic identification of the donor (s) of biological samples collected in investigations of biological criminalistics). The Biological Criminalistics is defined as the science that studies the biological evidence from the crime scene, and its main objective is to identify the alleged aggressor, as well as to reconstruct the circumstances related to the event (Pinheiro, 2015). As the process of producing evidence of an alleged Sexual Assault is based mostly on forensic evidence, the area of Biological Criminalistics plays a key role and is often responsible for the success of forensic investigations of these crimes (Magalhaes *et al.*, 2015; Sauer *et al.*, 2016). The forensic investigation, which includes the search for-evidence at the crime scene and the submission of the victims to a Forensic Medical Examination (FME) up to 72 hours after the occurrence (Forensic Clinic), as well as, the examination of clothing and other objects, is essential to collect the largest number of biological evidence for further analysis (Magalhaes *et al.*, 2015).

For the success of the investigation, the analysis of biological traces is crucial in order to obtain a DNA profile from a crime scene for 1) a genetic profile linked to the crime when

there is no suspect (s); 2) Confirmation of identity if there is potential (s) suspect (s) and/or 3) acquittal of an innocent suspect (s).

On the other hand, genetic analysis also allows us to infer about supplementary information in criminal investigations, which makes it possible to prove or discredit the veracity of reported statements about the alleged crime (Hanson *et al*., 2018), as well as the reconstruction of the circumstances in which the crime occurred.

1.2 Traces from Sexual Assaults

In alleged cases of Sexual Assaults many forensic evidences with interest in Biological Criminalistics are usually recovered, in particular, biological traces (stains) such as blood (peripheral or menstrual blood), semen, vaginal fluid, saliva, sweat, hair, pubic hair or other types, skin (epithelial cells) among the most frequent (Richard *et al.*, 2012). The majority of these traces often belong to both the victim (s) and/or to the aggressor (s) and are usually collected at the crime scene on various objects, clothing, or on the body of the victims during physical examination. Although the criminal importance varies according to the different biological traces, none of these traces can be ignored, since the biological evidence may be the only proof capable of directing the investigation to a successful end (Magalhaes *et al.*, 2015). The evidence collected at crime scenes not only can provide important information on the donor of the stain through DNA identification, but also the source from which the stain derived (i.e., type of body fluid/tissue) can also be identified through RNA analysis (Sijen, 2015).

1.2.1 Semen

The presence of semen stains in forensic investigations is the most valued and credible trait when presented in court as evidence in the context of sexual crimes (Fujimoto *et al.*, 2017; Takamura *et al*., 2017; Westring *et al*., 2014).

In a healthy individual, the semen consists of a suspension of spermatozoa in the seminal plasma that is formed by fluids from the seminal vesicles, prostate and bulbourethral glands (Cotton and Fisher, 2015; Sijen, 2015). In a Sexual Assault context, this biological

fluid is typically thick and opaque (usually whitish) and can be found in stains, on the body surface or in tissues, mixed with other fluids or can be, for example, obtained directly from condoms.

On the surface of a body, semen stains acquire the appearance of a thin layer on the skin. On other surfaces, the morphology and the color of these stains varies according to the type of the material (more or less absorbent) and the type of the ejaculation (direct or indirect) (Pinheiro, 2008). However, since this biological trace is not always macroscopically detected, the use of IR or UV forensic lights is necessary to identify its location (Magalhaes *et al.*, 2015).
Preliminary screening tests such as the Prostatic Acid Phosphatase (PAP), Prostate-Specific Antigen test (PSA) and Seminogelin tests (Sg I and Sg II) may be performed in the presence of semen stains. The PAP focuses on detecting the enzymatic activity of an enzyme released by the prostate gland that is present in high concentrations in seminal fluid and at lower concentrations in other body fluids. The PSA test consists of detecting a glycoprotein secreted by prostate epithelial cells in the seminal fluid but also detected at low concentrations in other biological fluids. The detection test for Sg I and Sg II is based on the search for one of the most important proteins in the formation of the seminal clot (Cotton and Fisher, 2015; Romero-Montoya *et al.*, 2011; Suttipasit and Wongwittayapanich, 2018; Virkler and Lednev, 2009).

However, these orientation tests are only presumptive tests since these semen constituents are not unique to this body fluid and can be detected at low concentrations in other fluids (Vennemann *et al.*, 2014).

The presence of semen can be confirmed using a 'test of certainty', which consists of the microscopic visualization of spermatozoa (Cotton and Fisher, 2015; Romero-Montoya *et al.*, 2011; Virkler and Lednev, 2009). However, microscopic detection of spermatozoa in cases of azoospermia (absence of spermatozoa) or oligospermia (reduced numbers of spermatozoa) may be difficult. The same can happen when the sample is impregnated in very absorbent materials with a strong adhesion of the spermatozoa to this type of tissue, and in cases of dehydration caused by the lack of necessary precautions for the correct storage of the sample (Pinheiro, 2008).

1.2.2 Vaginal Secretions

Vaginal secretions are another example of an extremely important body fluid that can be found in scenarios of alleged sexual crimes (Hanson and Ballantyne, 2013). In the context of Sexual Assaults, this trace is usually detected in the form of stains, impregnated in clothes or other fabrics recovered from the crime scene. Fluorescence emitted by vaginal fluid samples when exposed to alternating light sources is the most usual method for detecting these stains since they are not always visible macroscopically (Sikirzhytskaya *et al*., 2012).

This typically colorless or light colored fluid is characterized by a mixture of mucus secreted by glands of the vaginal mucosa or cervix, genital tract proteins, glycogenated epithelial cells, immunoglobulins, among other constituents (Sikirzhytskaya *et al.*, 2012). According to the literature, the constitution of this fluid is significantly influenced by the stage of the menstrual cycle. For this reason, there is great difficulty in accepting specific orientation tests for the identification of vaginal fluid, such as the periodic-schiff acid staining technique (Bexon and Williams, 2015). However, the Lugol iodine test is suggested as a presumptive method capable of identifying glycogenated cells of the vaginal epithelium (Harbison and Fleming, 2016). Additionally, male cells from the urinary tract may also react in the presence of Lugol iodine leading to false-positive results. Other methods, such as enzymatic and immunohistochemical tests, are described for the purpose of identifying the nature of this fluid (Virkler and Lednev, 2009). At present, no reliable confirmatory test capable of proving the presence of vaginal fluid is known.

2. Messenger RNA

2.1 Definition of messenger RNA

Messenger RNA (mRNA) is defined as a single-stranded ribonucleic acid sequence (RNA), which is originated from the transcription of a deoxyribonucleic acid template strand (DNA). Although these molecules are considered the least abundant type of RNA,

they present an enormous sequence and size heterogeneity because they are constituted by a very variable number of nucleotides complementary to the DNA nucleotides (Pelley, 2007). The role played by mRNAs is crucial since they participate directly in the regulation of protein synthesis.

In the past years, several authors have demonstrated their growing interest in the study of mRNA, highlighting the relevance of these molecules in the expression of human genes. These fundamental endogenous regulators involved in the gene expression program are able to inactivate and activate the genes and thereby influence the protein synthesis. (Hanson *et al.*, 2018). Despite the instability of these molecules, the specificity of mRNA expression, typical of each cell, is an advantage that contributes to the high emphasis attributed to the study of these molecules.

Recent studies suggest that, to date, the number of validated specific mRNAs is still reduced (Sijen, 2015). However, the increase of this number shows a trend of expansion in research focused on the study of mRNAs expression as potential biomarkers in the field of Biomedicine (associated with the diseases detection, e.g., Del Pino *et al.*, 2015; Goffinet *et al.*, 2005; Sorbye *et al.*, 2016; Tian *et al.*, 2016) and in the area of Forensic Science e.g., Afolabia *et al.*, 2017; Albani and Fleming, 2018; Blackman *et al.*, 2018; Hanson and Ballantyne, 2013; Richard *et al.*, 2012; Vennemann and Koppelkamm, 2010.

2.2 Biogenesis of Messenger RNA

First, the mRNAs synthesis and processing occurs in the nucleus of cells and then migrate to the cytoplasm where they participate in protein synthesis (Pelley, 2007). Initially, transcription of genes occurs whose DNA sequences are used as template and are transcribed by RNA polymerase II originating primary mRNAs (pre-mRNAs). These molecules are defined as immature mRNAs composed of coding sequences (exons) and non-coding sequences (introns). The combination of transcription factors during the process, the state of chromatin and the different promoter or silencing DNA regions are mainly responsible for the variation of gene expression in mRNA transcription. (Sijen, 2015). Still in the nucleus, pre-mRNAs undergo a maturation process in order to become functional mRNA. During processing of these molecules, occurs the removal of introns, the addition of cap (7-methylguanosine-RNA) at the 5 'end and the addition of the poly

(A) tail at the 3' end (variable number adenine sequence) (Aregger and Cowling, 2017). After these modifications, only a minority percentage of the transcribed mRNA is able to migrate to the cytoplasm and initiate the amino acids translation process for protein formation. (Sijen, 2015).

During gene expression, this whole process is controlled by regulatory systems responsible for maintaining the quality of mRNA (Sijen, 2015).

2.3 Messenger RNA as a forensic marker

The mRNAs have become a field of interest in Biomedical research and Forensic Genetics (Richard *et al.,* 2012). In recent years, there has been an increase in the number of published works for the purpose of validating mRNAs as biomarkers, as well as, development and homogenization of the techniques used for application of this method in casework. In some cases, it is difficult to identify the biological source (type of fluid or body tissue) of a DNA profile obtained from a crime scene. Thus, recent studies have suggested the analysis of mRNA expression as a confirmatory test to identify the nature and source of biological fluids to aid in the reconstruction of the criminal investigation line (Afolabia *et al.*, 2017; Albani and Fleming, 2018; Blackman *et al.*, 2018).

Specific mRNAs of each cell type have been found in abundance in biological fluids such as peripheral blood, menstrual blood, vaginal secretions, and seminal fluid, saliva, urine, sweat, sperm and other organs such as skin, brain, lung, liver, kidney and others (Sijen, 2015). Despite the limitations of these methods, several studies suggest these molecules as potential forensic biomarkers by proving their efficacy in body fluid and tissue identification. These considerations are based on the mRNA specificity for target fluids and on the consistent abundance of mRNA between different donors of the same fluid (Hanson *et al.*, 2018).

According to the scientific literature, an ideal biomarker is characterized by being specific, stable, presenting a long half-life, remaining unalterable by external factors and be able to be analyzed by non-invasive, fast, accurate, reproducible and economical methods as much as possible (Etheridge *et al.,* 2011). Despite some of the limitations of these principles when using RNA markers, studies consider and demonstrate the efficacy of mRNAs as potential forensic markers in aiding criminal investigations (Vennemann

and Koppelkamm, 2010; Sijen, 2015). Some authors also describe other important applications of mRNA markers in Forensic Sciences, such as, .determination of the cause of death; the post-mortem interval (PMI) by quantifying mRNA expression; ante-mortem or post-mortem burns; the time of a dermal lesion; the age of a stain, among others are some of the additional applications (Sijen, 2015).

A vast number of studies using mRNA markers for body fluid identification in forensic casework have already identified promising genes with tissue-specific expression that can be used for this aim (Gomes *et al.*,2013; Hanson *et al.,* 2018; Haas *et al.*, 2013 and 2014; Ingold *et al.,* 2018). As the current work focuses specifically on the identification of the source of samples that are usually recovered from Sexual Assaults scenarios, a list of mRNA markers for semen and vaginal secretion identification has been complied (Table 1).

Table 1. *Resume of the different types of mRNA markers (genes) reported for the identification of semen and vaginal secretions in Forensic Genetics.*

Body Fluid	**Gene**	**References**
Semen	PRM1	a, b, c, d, e, f, h, i, l, o
	PRM2	b, d, e, f, h, i, n, o
	TNP1	b
	TGM4	a, b, d, f, h, i, n, o
	SEMG1	c, d, e, f, h, i, l, o
	SEMG2	d, h, i
	KLK2	b
	KLK3	d, f, h, i, o
	MSMB	a, m
	NKX3-1	a,
Vaginal secretions	CYP2B7P1	a, b, c, d, g, h, i,
	DKK4	d, g, h, i,
	FAM83D	d, h, i,
	CYP2A6	d, h, i,
	MUC4	b, c, j, l, n, o, p
	HBD1	b, j, l, o, p
	MMPS	b
	MYOZ1	g
	SFTA2	a, g
	FUT6	g
	IL19	g
	MSLN	m

Subtitle: a) (Afolabia et al., 2017), b) (Albani and Fleming, 2018), c) (van den Berge et al., 2014), d) (Dorum et al., 2018), e) (Haas et al., 2009), f) (Haas et al., 2013), g) (Hanson and Ballantyne, 2013), h*) (Hanson et al., 2018), i) (Ingold et al., 2018), j) (Jakubowska et al., 2013), l) (Lindenbergh et al., 2012), m) (Park et al., 2013), n) (Richard et al., 2012), o) (Roeder and Haas, 2013), p) (Cossu et al., 2009)*

Data available in the literature regarding body fluid and tissue identification in forensics were extensively analyzed. This resulted in the selection of two "semen" markers, PRM1 (Protamine 1) and SEMG1 (Semenogelin 1), and in two "vaginal secretion" markers, MUC4 (Mucin 4) and MYOZ1 (Myozenin 1) for the present project. These genes revealed to be among the most robust and consistent for semen and vaginal secretion identification in forensic investigations and are therefore described in detail below.

2.3.1 mRNAs in semen

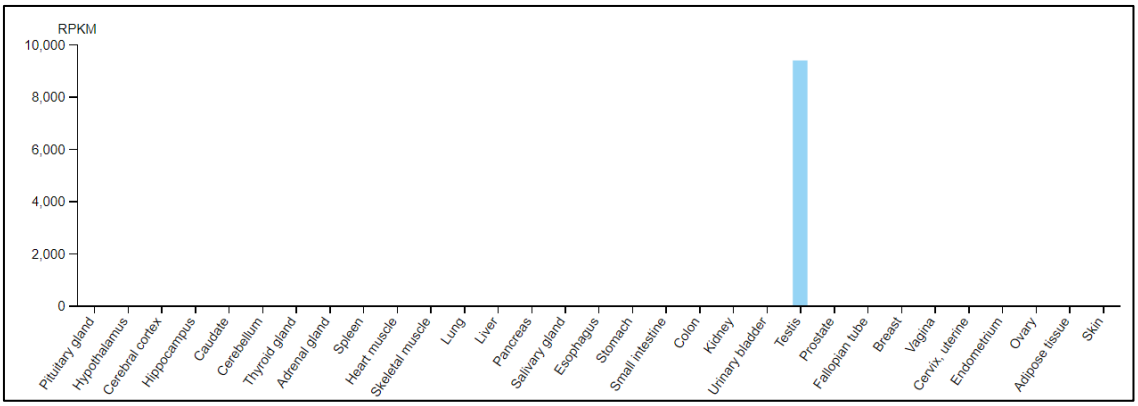
Research studies, have demonstrated that mRNAs may be useful biomarkers for the diagnosis of pathologies such as neoplasms and others strictly related to male infertility (e.g., Dere *et al*., 2016; Tian *et al.*, 2016). On the other hand, in the field of Forensic Genetics, advances made for the validation of mRNAs markers for the identification of semen stains in casework are also notorious (e.g., Afolabia *et al*., 2017; Albani and Fleming, 2018). According to recent investigations (Albani and Fleming, 2018; Haas *et al.*, 2009) specific mRNAs were detected in both sperm cell (sperm) and seminal fluid, which are the constituents of semen. This gives an enormous advantage in cases of azoospermia or oligospermia since this allows the confirmation of the presence of semen in stains deposited by males suffering from these conditions. According to several studies (e.g., Afolabia *et al*., 2017; Haas *et al*., 2013; Park *et al*., 2013; Richard *et al.*, 2012), to date, PRM1 and PRM2 are described as specific mRNAs of the sperm cell, whereas TGM4, PSA, KLK3, SEMG1, MSMB are conserved and abundant in seminal plasma. Both groups have been recognized as possible markers for the identification of semen in forensic investigations. Nevertheless, it should be noted that due to the lack of reproducibility of the conclusions based on this molecular technique, not all of these mRNAs are confirmed by all studies. Therefore, when selecting markers from the literature this must be taken into consideration and a careful selection of the most promising markers should be done to avoid the selection of poor markers.

PRM1

PRM1 (Protamine 1) is a Protein Coding gene. According to GeneCards and BioGPS data these proteins have a predominant expression in Testis seminiferous tubules, testis Leydig cells, testis interstitial cells, testis Germ cells (Graphic 2). They play an important role in the process of formation of spermatozoa since the Protamines substitute for histones in the chromatin of sperm during the haploid phase of spermatogenesis. They compact sperm DNA into a highly condensed, stable and inactive complex. According to medical studies, diseases associated with PRM1 include Male Infertility and Pelvic Varices (Garcia-Peiro *et al.*, 2011; Oliva, 2006). Other main biological functions as well as the genomic location and the size are listed in the Table 2.

In forensics, PRM1 has been described as a specific marker for semen detection. The presence of this gene has been proven in fresh ejaculated semen samples as well as in dried stains (Fleming and Harbison, 2010). However, its efficiency is only verified in semen samples from fertile individuals, since this marker is only expressed in the presence of sperm cells (Roeder and Haas, 2013). Thus, as reported by Lindenbergh *et al.*, 2012, PRM1 expression varies according to the amount of spermatozoa produced.

Given the high specificity (Graphic 2) verified by the absence of cross reactions with other cell types, and its high sensitivity (Juusola and Ballantyne, 2005), many authors include PRM1 in multiplex to identify simultaneously different body fluids (e.g. Haas *et al.*, 2009; Haas *et al.*, 2013; Juusola and Ballantyne, 2005; Lindenbergh *et al.*, 2012).



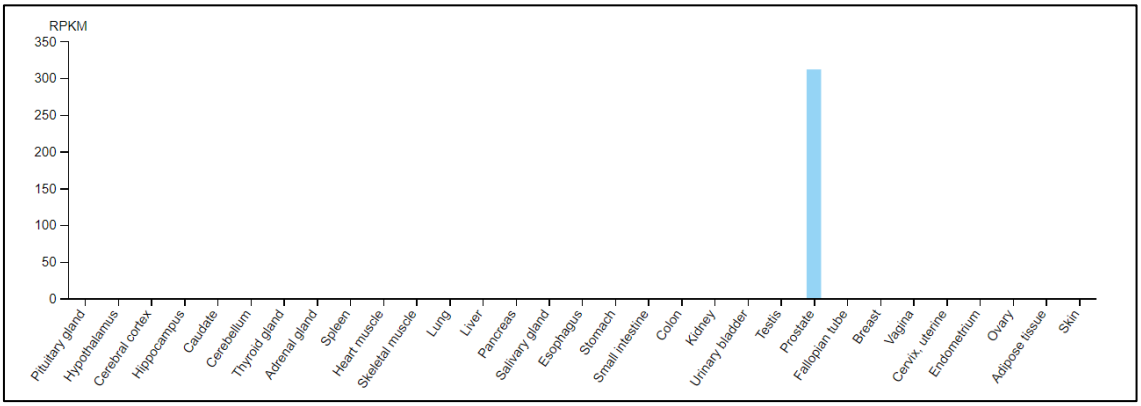
Graphic 2. *Dataset of PRM1 expression in different tissues generated by the Genotype-Tissue Expression database (GTEx).*

SEMG1

SEMG1 (Semenogelin 1) is a Protein Coding gene. The encoded secreted protein is predominant in seminal plasm and is involved in the formation of a gel matrix that encases ejaculated spermatozoa. Studies in molecular medicine state that diseases associated with SEMG1 include Male Infertility (Heidary *et al.*, 2019; Zhou *et al.*, 2019). Other important characteristics such as genomic location and biological functions of this gene are listed in the Table 2.

In forensics, SEMG1 has been described as another semen specific marker. However, unlike PRM1, SEMG1 is capable of detecting semen samples from not only fertile but also infertile individuals, as this marker is specific for seminal fluid. Thus, some authors have demonstrated the need to include SEMG1 for the identification of body fluids in order to ensure the identification of semen from azoospermic and vasectomized men in forensic cases (Haas *et al*., 2009; Haas *et al.*, 2013; Lindenbergh *et al.*, 2012).

Due to the specificity of this marker (Graphic 3), many authors have incorporated SEMG1 in multiplexes to identify different body fluids in forensic cases (e.g. Haas *et al.*, 2013; Lindenbergh *et al.*, 2012*;* ; Song *et al*., 2015).



Graphic 3. *Dataset of SEMG1 expression in different tissues generated by the Genotype-Tissue Expression database (GTEx).*

2.3.2 mRNAs in vaginal secretions

In vaginal secretions as well as in semen, expressed mRNA levels are strictly associated with the diagnosis of infections and some diseases such as tumors in the uterine cervix and have been commonly designated as potential biomarkers in Molecular Medicine (e.g., Del Pino *et al.*, 2015; Goffinet *et al.*, 2005; Sorbye *et al.*, 2016).

Likewise, in a forensic context, and due to the lack of adequate and specific vaginal fluid identification methodologies, studies have been published in order to recognize some of these RNA molecules as forensic biomarkers capable of identifying the nature of this body fluid (e.g., Haas *et al.*, 2014; Jakubowska *et al.*, 2013). As shown in Table 1, recent studies propose HBD1, MUC4, MSLN, CYP2B7P1, MYOZ1, among others, as possible specific mRNAs for the detection of vaginal secretions.

Although the specificity of some mRNA is demonstrated by many assays, some authors (e.g. Cossu *et al*., 2009; Park *et al.*, 2013) demonstrate the high heterogeneity of gene expression of vaginal secretions. Studies admit that the variation in the expression of the specific biomarkers is strictly associated with the oscillations that occurred during the menstrual cycle (Evans *et al.*, 2018; Park *et al.*, 2013; Richard *et al.*, 2012).

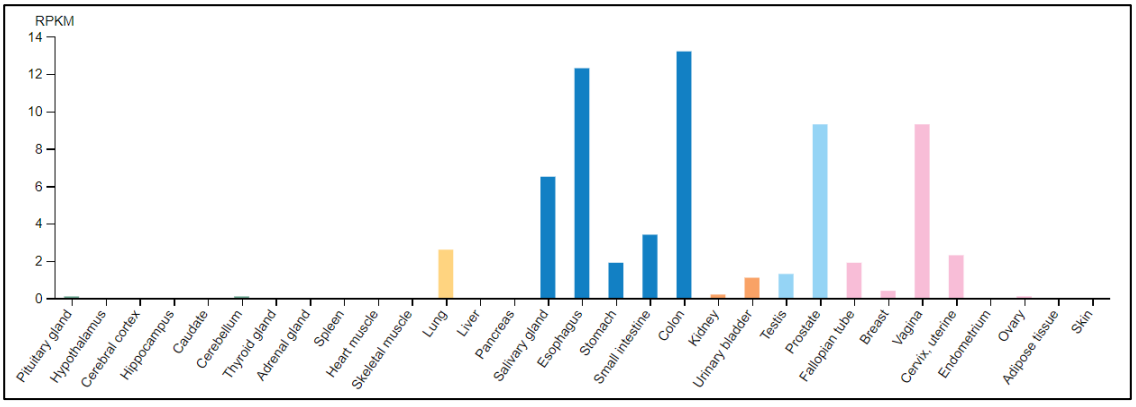
MUC4

MUC4 (Mucin 4) is a Protein Coding gene. According to GeneCards and BioGPS data, the encoded proteins called mucins are the major constituents of mucus. In this way they have a high expression in the viscous secretion that covers epithelial surfaces of the trachea, colon, and cervix. These highly glycosylated proteins play important roles in the protection of the epithelial cells and they are involved in regeneration and differentiation processes. Some medicine molecular studies report that Diseases associated with MUC4 include Filamentary Keratitis, Biliary Tract Cancer, Infertility or Endometrial pathologies (Dharmaraj *et al.*, 2014; Matull *et al.*, 2008; Tanioka *et al.* 2009). Other important characteristics such as genomic location and biological functions of this gene are listed in the Table 2.

Despite the significant MUC4 gene expression in many body tissues (Graphic 4), in Forensics it has been described as a specific marker for vaginal fluid, since many studies have confirmed the high expression in the presence of the fluid which internally coats the endothelium endocervical (Jakubowska *et al.*, 2013; Juusola and Ballantyne, 2005). However, it is reported that marker expression suffers some fluctuations as it decreases as progesterone levels increase (Roeder and Haas, 2013).

The stability and sensitivity of MUC4 has been proven by some authors (Jakubowska *et al.*, 2013; Juusola and Ballantyne, 2005; Setzer *et al.*, 2008). However, in recent years, some controversy has arisen about its specificity. Although some authors have stated the specificity of MUC4 for detection of vaginal secretions, the inclusion of this marker in some studies has revealed cross reactivity in saliva stains and buccal epithelial cells, not being conclusive for the full distinction between oral fluid and vaginal fluid (Cossu *et al.*, 2009; Roeder and Haas, 2013). These authors believe that this is due to the biochemical similarity of the two fluids (Lindenbergh *et al.*, 2012; Richard *et al.*, 2012; Song *et al.*, 2015). Another study reveals sporadic detection of MUC4 in semen samples (Jakubowska *et al.*, 2013).

Thus, due to the expression of this marker, although weaker in the remaining fluids, Jakubowska *et al.*, 2013 reported that some questions arose about the usefulness in forensic cases. However, other studies demonstrate the effectiveness of MUC4 when incorporated into mRNA multiplex systems associated with a careful interpretation of the results (Haas *et al.*, 2009; Haas *et al.* 2014; Lindenbergh *et al.*, 2012)



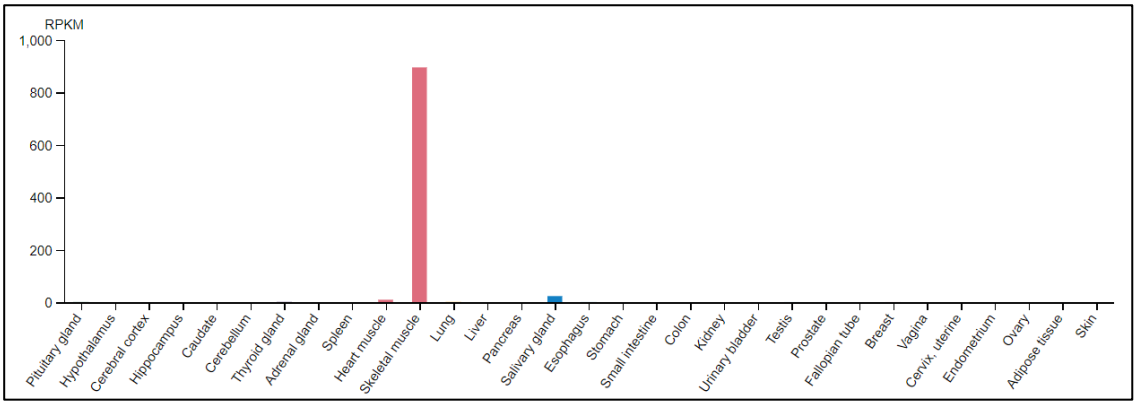
Graphic 4. *Dataset of MUC4 expression in different tissues generated by the Genotype-Tissue Expression database (GTEx).*

MYOZ1

MYOZ1 (Myozenin 1) is a Protein Coding gene. The protein encoded is found in large quantities in the skeletal muscle. According to GeneCards and BioGPS data, members of the myozenin family play an important role in modulation of calcineurin signaling to the sarcomere of cardiac and skeletal muscle. This protein is very important in myofibrillogenesis and according to MalaGene data, alterations in myoz1 expression are associated with some diseases such as Dilated Cardiomyopathy, Atrial Fibrillation among others. Other main biological functions as well as the genomic location and the size are listed in the Table 2.

Information of GeneCards databases reveals significant RNA expression of this Gene in skeletal muscle (Graphic 5). However, in forensic studies, MYOZ1 has been described as a good marker for identifying vaginal epithelial cells or their secretions. Although an exact function of MYOZ1 in vaginal secretions is unknown and some occurrences of cross reactivity with saliva have been reported by some studies, its specificity for vaginal fluid identification has been demonstrated in several multiplex mRNA profiling assays (Hanson and Ballantyne, 2013; Hass *et al.*, 2014).

Although in these studies MYOZ1 is not as strongly expressed compared to other specific vaginal markers, stability and sensitivity tests have shown promising results for use in forensic caseworks (Hanson and Ballantyne, 2013; Hass *et al.*, 2014).



Graphic 5. *Dataset of MYOZ1 expression in different tissues generated by the Genotype-Tissue Expression database (GTEx).*

Table 2. *Characteristics of the biomarkers selected in this study for the identification of the body fluids semen and vaginal secretions Adapted table from information obtained from BioGPS and GeneCards databases.*

Gene	Genomic Location	Main Biological Functions
PRM1 (Protamine1)	chr16: 11280836 - 11281350	DNA packing Multicellular organism development Spermatogenesis Cell differentiation Chromossome condensation
SEMG1 (Seminogelina 1)	chr20: 45206997 - 45209773	Insemination Antimicrobial humoral response Antibacterial humoral response Killing of cells of other organism Celular protein metabolic process Sperm capacitation Coagolation Protein heteroologomerizaton Negative regulation of calcium ion import Positive regulation of flagellated sperm motility
MUC4 (Mucin 4)	chr3: 195746765 - 195812277	Extracellular matrix constituent and lubricant activity Stimulatory C-type lectin receptor signaling pathway Cell matrix adhesion Regulation of signaling receptor activity O-glycan processing Maintenance of gastrointestinal epithelium
MYOZ1 (Myozenin 1)	chr10: 73631612 - 73641757	Sketetal muscle tissue development Myofibril assembly Negative regulation of skeletal muscle tissue regeneration Skeletal muscle fiber adaptation Sarcomere organization Negative regulation of calcinerin N-FAT signaling cascade

2.4 mRNA stability

Forensic samples from Sexual Assaults, as in most forensic cases, are typically found in low quantities and often under adverse conditions (e.g., Richard *et al.*, 2012; Sirker *et al.*, 2016). An advanced state of degradation interferes with the quality of the biological evidence and can significantly compromise laboratory investigation because it does not allow the use of current techniques for DNA or RNA analysis or the obtaining of complete results. Studies based on the recovery and stability of mRNA from biological material are very important because they have demonstrated the effects of the exposure of different biological fluids stains to various abiotic factors such as temperature, humidity/dehydration, UV and luminescent radiation, and even rain during different periods of time (e.g., Setzer *et al.*, 2008; Sirker *et al.*, 2016). These authors state that environmental conditions are still a major challenge, since they can alter the amount of total RNA recovered as well as the stability of mRNA expression, given the higher susceptibility of these molecules to degradation by physical factors.

Based on the literature, the amount of total RNA recovered from biological samples decreases significantly when exposed to environmental conditions especially when not protected from rain because it causes a substantial loss of biological material (Setzer *et al.*, 2008). However, the observed effects, when exposed to UV light, vary depending on the biological material. According to the findings of Setzer *et al.*, 2008 UV light can even promote the increase of total RNA in semen samples since it contributes to the rupture of the sperm membrane which results in the release of abundant amount of RNA. In contrast, the opposite effect is described in samples of vaginal secretions when irradiated by UV light. A decrease in mRNA expression of samples exposed to uncontrolled environmental conditions is noticeable. The changes in differential expression vary, not only depending on the biological fluid and the surrounding environmental conditions, but also on the sensitivity, regarding the detection of differential expression of these markers (Setzer *et al.*, 2008).

Thus, although the mRNA class is characterized by its instability due to the presence of ribonucleases (Sirker *et al.*, 2016), several authors value this molecule as being stable enough for forensic use (e.g, An *et al.*, 2012; Setzer *et al.*, 2008; Vennemann and Koppelkamm, 2010). This statement is corroborated by studies demonstrating the ability

to identify the biological nature of very old and/or degraded samples by mRNA analyses (Hara *et al.*, 2016; Kohlmeier and Schneider, 2012; Setzer *et al.*, 2008; Sirker *et al.*, 2016).

Therefore, it is crucial to have robust and stable enough markers to overcome the limitations of RNA quantity and quality either due to degradation environmental factors or simply due to low quantity of biological material recovered at a crime scene.

2.5. Body fluid mixtures in Sexual Assaults

Mixtures of different biological fluids are typically found in forensic caseworks (Song *et al.*, 2015). The analysis of these mixtures is very important not only to identify the contributors of this sample (different cell types can come from a single donor or multiple donors) but also to identify the biological source of the various body fluids (Sijen, 2015). However, these complex mixtures remain a major challenge as it is not possible to associate the biological source of the fluid with its genetic profile except for gender-specific body fluids under certain circumstances (Harteveld *et al.*, 2013; Sijen, 2015).

In alleged cases of sexual assault, the identification of fluids in mixed samples may be crucial to produce evidence. In these cases, positive identification of semen/vaginal fluid stain is a commonly found sample and it is the main evidence of sexual contact (Hanson *et al.*, 2013; Song *et al.*, 2015).

According to the literature, in last years some authors have developed molecular methods namely multiplexes based on mRNAs, which have shown promising results in identifying and distinguishing different biological fluids (Haas *et al.*, 2013; Juusola and Ballantyne, 2005; Lindenbergh *et al.*, 2012; van den Berge *et al.*, 2014).

3. The use of mRNA in biological criminalistics

3.1 Advantages of using mRNA

Conventional methods of identification for biological fluids based on chemical, enzymatic/catalytic techniques, immunological tests, among others, are characterized by

the ease of handling and the ability to rapidly present results. However, these methods reveal several disadvantages (Fujimoto *et al.*, 2018). Low specificity, lack of sensitivity, inability to identify and distinguish fluids, incompatibility with individual genetic identification, destruction and increased sample consumption among others, are some of the problems in the use of these methods, often requiring meticulous analysis and additional tests (Kulstein *et al.*, 2018; Sauer *et al.*, 2016). Additionally, these conventional identification methods are based on the composition of biological fluids. However, these components are not always unique to a single fluid and may lead to false positive results and consequently misinterpretation (Song *et al*., 2015).

Recent studies have demonstrated a great evolution in order to overcome the presented limitations, allowing a growing interest to investigations using molecular methods based on the differential expression of mRNA among different tissues. Its high fluid identification spectrum, the possibility of detecting different fluids in the same reaction (multiplex) and the ability to co-extract DNA and mRNA from the same biological trait (consumption of less sample without loss of DNA) are strong advantages of this method (Ingold *et al.*, 2018; Richard *et al.*, 2012). Additionally, its high specificity (Albani and Fleming, 2018; Harbison and Fleming, 2016), sufficient stability for forensic use (Harbison and Fleming, 2016; Vennemann and Koppelkamm, 2010), the use of non-invasive protocols, the possibility of automating extraction techniques (Richard *et al*., 2012) and the use of already existent platforms in most forensic laboratories are also key features of this molecular method.

On the other hand, the difficult and time-consuming process of creating robust and reliable multiplexes, the high prices of the required technology (commercial kits and laboratory techniques), the detailed study required, and the operator's practice in the analysis of results are disadvantages of the use of method a mRNA body fluid identification method.

However, the previously described pros corroborate the researchers' interest in the study of these potential forensic biomarkers compared to conventional methods (An *et al.*, 2012; Ingold *et al.*, 2018; Kulstein *et al.*, 2018; Richard *et al.*, 2012). Nevertheless, despite the fact that analysis of mRNA expression fills the gap of traditional techniques, some authors predict that the total replacement of conventional by molecular methods

will not be feasible. The use of these emerging methods, besides requiring complex and costly analysis, is not justified by the availability of confirmatory tests to identify the nature of fluids (e.g., microscopic visualization of spermatozoa in the seminal fluid identification). Therefore, the combination and availability of both methods in forensic laboratory settings is proposed for the identification of the biological source of evidence in criminal casework (Kulstein *et al.*, 2018).

3.2 Comparison with other molecular methods for biological fluids identification

In addition to the study of mRNA expression, other molecular methods have been developed such as the study of differential expression of Micro RNA (miRNA), and the study of DNA methylation patterns in different biological fluids (An *et al.*, 2012; Sijen, 2015; Sirker *et al.*, 2016).

miRNAs are also considered good forensic biomarkers since they confer a unique gene expression of each biological sample. These well-preserved and highly stable molecules are small, which promotes greater resistance to degradation by environmental factors- an advantage of miRNAs in relation to mRNAs highly desirable in Forensic Genetics. However, the analysis of biological samples through miRNA still presents some inconsistencies in the analysis of different biological samples when analyzed by different techniques in addition to the high co-expression (lower specificity then mRNAs) observed in miRNAs. Thus, the methodology based on the study of miRNA expression requires rigorous procedural validation, miRNA quantification methods and accurate models for data analysis (Sirker *et al.*, 2017).

The study of DNA methylation patterns is another method used in the forensic identification of biological fluids (Jung *et al.*, 2016; Sijen 2015). This recent technique quantifies the proportion of methylated and unmethylated cytosines (methylation levels) creating unique patterns of each biological sample. Although, this method is not a susceptible to degradation due to the use of DNA instead of RNA, studies prove that these markers undergo changes over time, being influenced by endogenous and exogenous factors. The high consumption of DNA sample and the insufficiency of validated markers for some biological fluids such as saliva are still other disadvantages to consider (Gomaa

et al., 2017; Ingold *et al.*, 2018; Park *et al.*, 2014). Although this technique is promising, more studies are needed to prove and validate DNA methylation markers to aid forensic investigations.

Finally, the advantages of mRNAs as molecular biomarkers for biological fluids identification are notorious in comparison to conventional methods, to miRNA analyses and to DNA methylation studies. In order to increase sensitivity in the identification of low amounts of samples, some studies argue that the combination of several molecular techniques would be the most useful and efficient approach (Sijen, 2015).

AIM

The identification of human body fluids and tissues in forensic crime scene investigation has become a popular and increasing topic of research in the last years. mRNA markers have obtained a distinct place in this field of research and are the most reliable, stable and best studied markers in forensic body fluid identification.

The present dissertation has, as main objective, the identification and differentiation of semen and vaginal secretions samples using mRNA markers. With this main aim traced, objectives were proposed and different points of study were investigated:

- To develop and optimize a tetraplex system consisting of four mRNA markers (PRM1, SEMG1, MYOZ1, MUC4) selected from the literature.

- To compare two methods for DNA/RNA co-extraction by testing the applicability and robustness of the developed multiplex.

- To prove the efficacy of the selected mRNAs as forensic biomarkers to identify and distinguish semen (PRM1 and SEMG1) and vaginal secretions (MYOZ1 and MUC4) from samples common recovered from sexual crimes.

- To test the sensitivity of the selected markers and technique used in low quantity samples.

- To test the performance of the tetraplex system in mixed sample settings of semen and vaginal secretions.

- To infer on the specificity of the semen and vaginal mRNA markers.

MATERIALS AND METHODS

The practical component of the present work focuses on the simultaneous extraction of nucleic acids of DNA and RNA for the possible simultaneous identification of the source of a donor, as well as the source of the biological material. All common and necessary procedures to avoid contamination of samples as well as degradation of RNA during extraction of nucleic acids were applied, which are of high importance in forensic settings. Also, samples were processed in a laboratory hotte when toxic chemicals were in use. The whole material and method process employed in this work is represented in the following scheme (Figure 1):

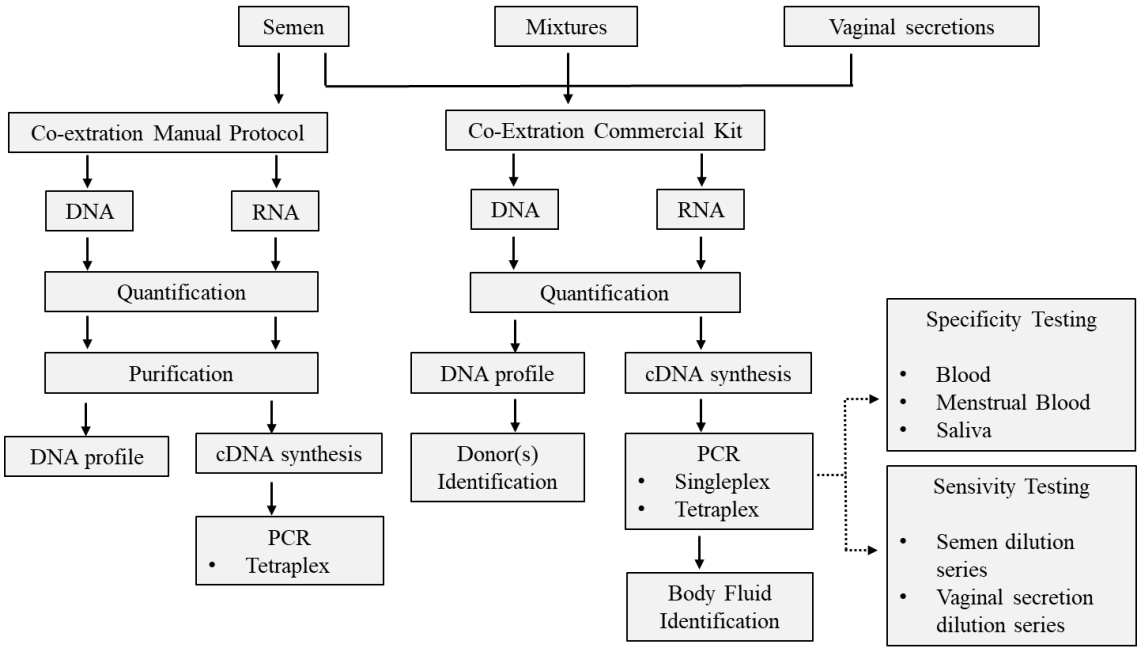


Figure 1. *Summary of the sequence of methods included in this study.*

The protocols used in all the experimental procedures and specific steps are described in detail further below.

1. Collection of Biological Samples

The present study included the collection of semen and vaginal secretion samples from (apparently) healthy individuals aged between 20 and 60 years. A 'Participant Information' document was given to all volunteered donors describing the purpose of the study as well as all information about the sampling procedure (Attachment 1 and 2). Subsequently, informed written consent was obtained from each participant according to the recommendations of the Helsinki Declaration and the anonymity of all samples collected was guaranteed (Attachment 3).

Freshly liquid ejaculated semen samples, in sterile and sealed plastic tubes, were provided by 14 male volunteers and 3 sterile vaginal swabs per women were required from 10 female volunteers. In addition, all women involved in this study completed a questionnaire about hormonal cycles as well as the intake of birth control medication.

For additional testing (specificity testing), RNA of other types of body fluids were also extracted: a fresh saliva swab, a menstrual blood swab and a peripheral blood sample were also collected from different donors. For collection of blood, a lancet was used, and a drop of sample was placed on a FTA paper.

2. Coding and storage of biological samples

The samples collected were stored guaranteeing all the conditions to avoid degradation and contamination. In this way, 200µl of semen as working aliquots were prepared and stored frozen at -20°C. Vaginal secretions samples were dried at room temperature for a few hours and then frozen at -20°C, until further analysis.

To guarantee the confidentiality of the participants all samples were coded according to Table 3. The semen samples were marked with the code Sx.1/2/3 (x is a number from 1 to 14 randomly assigned to the male volunteer and 1/2/3 is the replica number per individual) and the vaginal secretion samples were represented by VSx.1/2/3 (x is a number from 1 to 10 randomly assigned to female volunteers and 1/2/3 is the replica number per individual).

Table 3. *Coding of semen, vaginal secretion, peripheral blood, menstrual blood and saliva samples.*

Semen Samples			Vaginal Secrection Samples			Saliva Samples	Peripheral Blood Samples	Menstrual Blood Samples
S1.1	S1.2	S1.3	VS1.1	VS1.2	VS1.3	SAL1	B1	MB1
S2.1	S2.2	S2.3	VS2.1	VS2.2	VS2.3			
S3.1	S3.2	S3.3	VS3.1	VS3.2	VS3.3			
S4.1	S4.2	S4.3	VS4.1	VS4.2	VS4.3			
S5.1	S5.2	S5.3	VS5.1	VS5.2	VS5.3			
S6.1	S6.2	S6.3	VS6.1	VS6.2	VS6.3			
S7.1	S7.2	S7.3	VS7.1	VS7.2	VS7.3			
S8.1	S8.2	S8.3	VS8.1	VS8.2	VS8.3			
S9.1	S9.2	S9.3	VS9.1	VS9.2	VS9.3			
S10.1	S10.2	S10.3	VS10.1	VS10.2	VS10.3			
S11.1	S11.2	S11.3						
S12.1	S12.2	S12.3						
S13.1	S13.2	S13.3						
S14.1	S14.2	S14.3						

3. Preparation of body fluid mixtures

In order to mimic the analysis of a swab commonly collected in real cases of sexual aggression (i.e., a mixture of female and male component), mixed samples were produced with different ratios of semen and vaginal secretions from the same donors according to Table 4. Thus, 1 male sample and 1 female sample (VS3 and S7) were randomly selected and used in the preparation of the mixtures. For comparison purposes, samples from the same donors with equal quantities of biological fluid (swabs for the vaginal secretions samples and volume for semen samples) were tested and used as reference (Table 4).

Table 4*. Mixtures with different ratios of semen and vaginal secretions as well as the respective samples with the reference volumes (Ref) and swab quantities prepared for the body fluid mixture testing.*

Mixture	VS samples	+	S samples
Mix 1	1 swab	+	10µl
Mix 2	1 swab	+	5µl
Mix 3	1/2 swab	+	10µl
Mix 4	1/2 swab	+	5µl
Ref 1	---		10µl
Ref 2	---		5µl
Ref 3	1 swab		---
Ref 4	1/2 swab		---

4. Co-extraction of DNA and RNA

In this work, two methods for DNA and RNA co-extraction, were tested and compared. The commercial kit ExtractME RNA & DNA Kit (Extractme, BLIRT S.A., Poland) was used to simultaneously extract both types of nucleic acids from semen and vaginal secretions samples, according to manufacturer instructions. Additionally, a quick and simple manual based protocol, adapted from Shojaie N. *et al*. [Protocol Exchange (2014), doi:10.1038/protex.2014.036], was also tested for co-extraction of DNA and RNA only for semen samples. Some modifications were performed in both methods to adapt to the specific type of samples and materials used in this work.

4.1 DNA/ RNA Co-extraction with the commercial kit

4.1.1 Semen and Vaginal secretions samples

The commercial kit ExtractME RNA & DNA (Extractme, BLIRT S.A., Poland) was used for the simultaneous co-extraction of DNA and RNA for all samples described above (section 1). According manufacture's instructions, this commercial kit utilizes spin columns with membranes which efficiently and selectively bind nucleic acids at high concentration of chaotropic salts. The kit protocol is divided into two phases: 1º) the Isolation protocol that includes the disintegrate of intercellular bonds and fragmentize the membranes and proteins allowing the DNA selectively bounds to DNA Purification Column and RNA bounds to RNA Purification Column; 2º) Total Purification protocol including Total RNA purification and Genomic DNA purification. These procedures are intended to effectively remove impurities and enzyme inhibitors. They also allow that purified RNA and DNA extracts to be eluted with low ionic strength buffers for be used in further applications.

The Isolation protocol included the following steps:

1. The biological samples (200µl of semen or 1 vaginal swab) were placed in a 2ml tube. 600µl of Lysis Buffer containing guanidine thiocyanate and B-mercaptoethanol was added and mixed well with a vortex for 60s.

2. The samples were centrifuged for 120s at 15 000xg.
3. The supernatant was transferred into DNA Purification Column (Red column in the Figure 2) placed in a collection tube and centrifuged for 30s at 12 000xg. The DNA column was saved for further DNA purification and the filtrate was used in step 4.
4. 600µl of 70% ethanol was added to the filtrate from step 3 and mixed well by pipetting.
5. The obtained mixture was transferred into RNA Purification Column (Purple column in the Figure 2) centrifuged for 30s at 12 000xg. Then, the RNA column was saved for further RNA purification and the filtrate was discarded.

In an initial assay, four semen samples were co-extracted following the instructions of the commercial kit described above. Subsequently, in order to improve the efficacy of the kit, during the first step of the isolation protocol, 20µl of DTT and 600µl of Lysis Buffer was added to all remaining 200µl semen samples, followed by an incubation at 56°C for 30 minutes (Figure 2). The remaining steps were the same previously described.

All vaginal samples were also co-extracted following the instructions described above and an additional incubation at 56°C for 30 minutes was also added to the protocol after the first step.

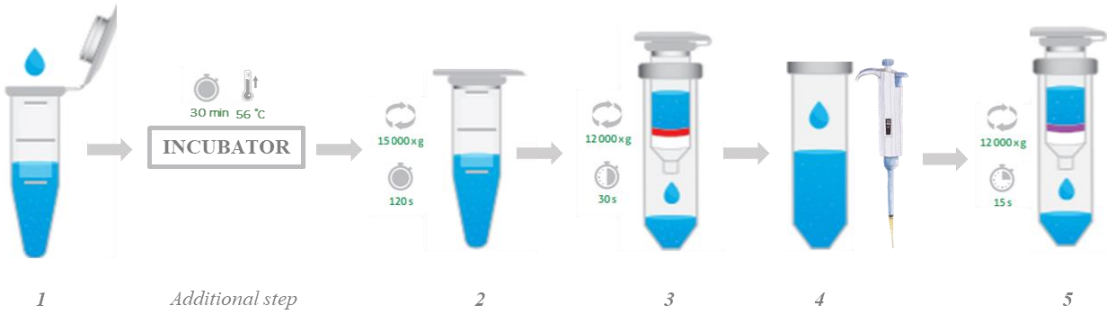


Figure 2. *Schematic Diagram of the Isolation protocol. Adapted from manufacturer's instructions of the commercial kit ExtractME RNA & DNA (Extractme, BLIRT S.A., Poland).*

After the isolation protocol, the purification protocol was performed for both DNA and RNA extracts following the manufacturer's instructions. Thus, the total RNA purification protocol included the following steps:

1. 700µl of RNA Wash Buffer 1 was added into RNA Purification column and then centrifuged for 15s at 12000xg. The filtrate obtained was discarded.
2. 500µl of Wash Buffer 2 (containing 100% ethanol as recommended by manufacturer) was added into the same RNA Purification column and then centrifuged for 15s at 12000xg. The filtrate obtained was discarded. This step was repeated one more time.
3. The RNA Purification column was centrifuged again for 90s at 15000xg. The filtrate obtained was discarded and the purification minicolumn was carefully transferred to a sterile Eppendorf tube.
4. As a final step, purified RNA was eluted in 50µl of RNA Elution Buffer by centrifugation for 60s at 12000xg.

Additionally, at the end of the experimental procedure, a second 20µl RNA elution was performed for both semen and secretion vaginal samples.

To Genomic DNA purification the following steps were followed:

1. 700µl of DNA Wash Buffer 1 were added into DNA Purification column and then centrifuged for 15s at 12000xg. The filtrate obtained was discarded.
2. 500µl of Wash Buffer 2 (containing 100% ethanol as recommended by manufacturer) were added into the same DNA Purification column and then centrifuged for 15s at 12000xg. The filtrate obtained was discarded. This step was repeated one more time.
3. The DNA Purification column was centrifuged again for 90s at 15000xg. The filtrate obtained was discarded and the purification minicolumn was carefully transferred to a sterile Eppendorf tube.
4. As a final step, purified DNA was eluted in 50µl of DNA Elution Buffer by centrifugation for 60s at 12000xg.

At the end of the co-extraction procedure purified DNA and RNA extracts were preserved at -20°C for further analysis.

4.1.2 Mix samples

The commercial kit ExtractME RNA & DNA (Extractme, BLIRT S.A., Poland) was also used for the DNA and RNA co-extraction for mixed samples previously prepared as described in section 3.

The mixed samples were coextracted following the protocol described above. However, in the first assay, during the first step of the isolation protocol, 20µl of DTT and 600µl of Lysis Buffer was also added to all mixed samples, followed by an incubation at 56°C for 30 minutes. The remaining steps of the isolation and purification protocol were the same previously described.

In order to improve the effectiveness of the procedure, a second assay using the mixed samples from the same donors was performed. Thus, after the first step of the isolation protocol (Add 600µl of Lysis Buffer and 20µl of DTT to mixed samples and an incubation at 56ºC for 30 minutes) a spin shot in the centrifuge using homemade spin baskets was included. It was intended to ensure full filtrate recovery for use in the following steps.

The remaining steps of the isolation and purification protocol were the same previously described.

All the extracted DNA and RNA extracts from the mixed samples were also preserved at -20°C for further analysis.

4.2 DNA/ RNA Co-extraction with the manual protocol

A second method using a manual protocol were also used for the simultaneous co-extraction of DNA and RNA only for semen samples.

For each co-extraction reaction, the following protocol was used:

1. To 200µl semen aliquots, 20µl of DTT, 25µl of proteinase K and 300µl of extraction buffer (4.4ml 0.5M Tris; 4.4 ml Glycerol; 2.2ml 20% SDS; 0.5ml 2-Mercaptoethanol) were added and mixed well with a vortex.
2. Mixtures were incubated at 56°C for 30 minutes.

3. 800µl of Isopropanol were added to the lysate and mixed well by manual inversion. Then the mixture was left for 5 minutes and it was possible to observe a white coil DNA.
4. The DNA coil observed was transferred into an Eppendorf tube using a sterile spatula. The aqueous phase was kept as the RNA extract sample.
5. 800µl of 100% ethanol (stored at -20°C) were added to both DNA and RNA extracts and centrifuged at 10000xg for 5 minutes.
6. The supernatant was removed. 800µl of 70% ethanol (stored at -20°C) were added to both DNA and RNA and then centrifuged at 10000xg for 5 minutes.
7. The supernatant was removed again. To ensure complete evaporation of ethanol traces, the DNA and RNA extracts were incubated at 65°C for 30 minutes.
8. As a final step, the pellet was re-suspended in 70µl of TE Buffer by incubation at 65°C for 30 minutes.

The DNA and RNA extracts of semen samples were preserved at -20°C for further analysis.

5. DNA and RNA quantification

Quantifications of all DNA and RNA extracts obtained from both extraction procedures were estimated using the spectrophotometer Nanodrop 1000 (Thermo Fisher Scientific, USA). As a measure of quality control (QC) check of the DNA and RNA extracts purity ratios (260/230 and 260/280) were also obtained.

Additionally, DNA extracts from the semen samples extracted with the manual protocol followed by purification were also quantified using 'Quantifiler Trio DNA Quantification Kit' (Thermo Fisher Scientific®). All steps were performed following the manufacturer's instructions.

6. Purification of DNA and RNA extracts

DNA and RNA samples extracted with the manual co-extraction protocol were subjected to a purification process to remove the presence of potential inhibitors that are carried over from the extraction procedure. The following protocol was used:

1. To 50µl of RNA or DNA extracts, 25µl of NH4OAc and 125µl of absolute ethanol (stored at -20°C) were added and mixed well with a vortex.
2. Samples were then incubated at -20°C for 1 hour and then centrifuged at 12000xg at room temperature for 20 minutes.
3. The supernatant from each sample was removed, and the pellet was washed with 70% ethanol and centrifuged at 12000xg at room temperature for 5 minutes. This step was repeated one more time.
4. To ensure complete evaporation of ethanol traces, the samples were left to dry at room temperature for 24 hours.
5. As a final step, the pellet was re-suspended in 30µl of TE Buffer and samples were further stored at -20ºC.

7. Reverse Transcription reaction: cDNA synthesis

To allow the amplification of the body fluid specific mRNA markers, complementary DNA (cDNA) must be generated to serve as template in the end-point PCR reaction.

The RT+ (Reverse transcription of the RNA sample) was performed for each RNA sample of semen and vaginal secretion obtained in this study. cDNA was synthesized using the commercial 'SensiFAST cDNA Synthesis kit' (BIRLT S.A., Poland) in a final volume of 10µl. The reagents of RT reaction were added according to Table 5A.

For some RNA samples (S6/ S7/ VS1/ VS8), a RT- (No RNA control of the RT reaction) without reverse transcriptase enzyme, was also performed to ensure the absence of genomic DNA in the extract. The reagents were added according to Table 5B.

The thermal cycle conditions for the cDNA synthesis reaction were 25ºC for 10 minutes (primer annealing), 42ºC for 15 minutes (reverse transcription), 85ºC for 5 minutes

(inactivation). Then, the final cDNA reaction products were stored at -20ºC until further downstream applications.

Table 5. A. *Reagents and respective volumes for one sample for cDNA (RT+) reaction;* ***B.*** *Reagents and respective volume for one sample for cDNA with no Transcriptase (RT-) reaction.*

A.

Reagents	Volume per sample
5x TransAmp Buffer	2 µl
Dnase/Rnase free-water	2.5µl
Reverse transcriptase	0.5µl
	TOTAL: 5µl Reaction MIX + 5µl RNA

B.

Reagents	Volume per sample
5x TransAmp Buffer	2µl
Dnase/Rnase free-water	3µl
Reverse transcriptase	---
	TOTAL: 5µl Reaction MIX + 5µl RNA

8. mRNA markers and primers

In this study, two vaginal secretions markers (MUC4 and MYOZ1), one seminal plasma marker (SEMG1) and one sperm marker (PRM1) were selected for body fluid identification. Sequence primers for each mRNA marker were adopted from the literature (Table 6). The forward primer of each primer set was labeled with 6-FAM™ dye and ordered in a lyophilized state (Sigma-Aldrich). Sequences, concentrations and the expected amplicon sizes are listed in Table 6.

Table 6. *mRNA markers, the respective primers sequences, expected amplicon sizes and 5´ label for the body fluid identification of semen and vaginal secretions.*

Body fluid	Gene		Primers	Lable 5'	Size (bp)	Reference
Semen	PRM1	Protamine 1	f: GCC AGG TAC AGA TGC TGT CGC AG r: TTA GTG TCT TCT ACA TCT CGG TCT	6-FAM –	153	*a*
	SEMG1	Semenogelin 1	f: TCG GTA ACC ATG TGA AAG GA r: GTG TCA TCC ATG GAC CAA GA	6-FAM –	120	*b*
Vaginal Secretion	MYOZ1	Myozenin 1	f: GGG TTG GTG AGA CAG GAT CA r: TCC CAT GGG GAA ATA TAG GT	6-FAM –	81	*c*
	MUC4	Mucin 4	f: GGA CCA CAT TTT ATC AGG AA r: TAG AGA AAC AGG GCA TAG GA	6-FAM –	235	*a*

References: a) (Juusola and Ballantyne, 2005), b) (Haas et al., 2009), c) (Hanson and Ballantyne, 2013)

9. Polimerase Chain Reaction (*PCR*)

9.1 Singleplex development

In an initial step, and to ensure the functioning of the selected primers, the semen and vaginal secretions specific-primers were tested in singleplex PCR reactions using RNA extracts from two samples of each body fluid that showed good concentration and purity values (S6, S7, VS1, VS8).

Each PCR reaction was amplified in a final volume of 8µl and for each marker all reagents were added according to Table 7. Each primer was used in a final concentration of 0.25µM. For each PCR reaction, a negative control was added without the cDNA template.

The thermocycling PCR reaction conditions used were: denaturing step at 95ºC for 15 minutes; 35 cycles of 94ºC for 20s, 55ºC for 30s, 72ºC for 40s and a final extension at 72ºC for 15 minutes.

Table 7. *Reagents and respective volume for one singleplex.*

Reagents	Volume per sample
Water nuclease-free	2.2µl
Primer Foward	0.4µl
Primer Reverse	0.4µl
Qiagen Amplification Kit	4µl
TOTAL: 8µl (7µl Reaction MIX + 1µl cDNA)	

9.2 Multiplex development and optimization

A tetraplex system including the four mRNA markers (PRM1, SEMG1, MUC4 and MYOZ1) was developed. 1µl of cDNA reaction product was amplified in 8µl of PCR reaction final volume. Negative control without cDNA product was also included. For each PCR reaction, all reagents were added according to Table 8.

As optimization procedures, two different Primer Mixes (Primer Mix 1 and Primer Mix 2), were tested following the primer concentrations listed in Table 9.

The thermocycling PCR reaction conditions used in the tetraplex PCR development were: initial denaturation step at 95ºC for 15 minutes; 35 cycles of 94ºC for 30s, 57ºC for 90s, 72ºC for 40s and final extension at 72ºC for 30 minutes.

Table 8. *Reagents and respective volumes for one multiplex PCR reaction.*

Reagents	**Volume per sample**
Água nuclease-free	2µl
Primer Mix	1µl
Qiagen Amplification Kit	4µl
	TOTAL: 8µl (7µl Reaction MIX + 1µl cDNA)

Table 9. A. *Primer Mix concentrations (µM)* ***B.*** *Primer Mix concentrations (µM) in PCR reaction.*

A.

	mRNA markers concentrations							
	PRM1 F	PRM1 R	SEMG1 F	SEMG1 R	MUC4 F	MUC4 R	MYOZ1 F	MYOZ1 R
Primer Mix 1	2.7	2.7	2.5	2.7	2.1	2.7	6.7	6.7
Primer Mix 2	1.7	1.7	1.6	1.7	1.3	1.7	12.5	12.5

B.

	mRNA markers concentrations in PCR							
	PRM1 F	PRM1 R	SEMG1 F	SEMG1 R	MUC4 F	MUC4 R	MYOZ1 F	MYOZ1 R
Primer Mix 1	0.3	0.3	0.31	0.3	0.26	0.3	0.8	0.8
Primer Mix 2	0.2	0.2	0.19	0.2	0.16	0.2	1.6	1.6

10. Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gel electrophoresis is a technique based on the separation of fragments according to their molecular size. When electric current is applied, negatively charged fragments migrate through the polyacrylamide gel towards the positive pole. Since the mobility of molecules is inversely proportional to its size, smaller fragments migrate more rapidly through the gel than the larger ones. Then, to visualize the bands of fragments, the gel can be stained used colored dyes.

This technique was initially used in order to test the amplification of the PCR products from the singleplex and multiplex developed by visualizing the amplified fragments in the gel.

For the preparation of a 10 x 4.5cm gel, a solution was prepared with the following reagents: 2ml of Acrylamide stock (25ml of 1.5M Tris-HCL Buffer pH8.8; 20ml of Acrylamide:Bisacrylamide 19:1; 7ml of Glycerol and 43ml of H2O), 114µl APS 2.5% and 7µl of TEMED (Thermo Fisher Scientific®).

For the electrophoresis: 1.3µl of amplified cDNA from the study samples and respective negative controls were loaded into the wells of the polyacrylamide gel. A 100bp Ladder (Zymo Research®) was also added and used as a standard reference containing DNA fragments of known lengths. The samples were run through the gel by action of the electric current, at 120V for about 1hour and 30 minutes using a Multiphor™ II electrophoresis system.

To visualize the bands, the silver staining technique was used. This consists of immersing the gel in the following solutions: 1) ethanol (10%) for 10 minutes; 2) Nitric acid (1%) for 5 minutes; 3) 2 washes with distilled water for 10s each; 4) silver nitrate (0.2M) for 20 minutes and in the absence of light; 5) 2 washes with distilled water for 10s each; 6) a solution containing 100ml of distilled water, 3g of carbonate and 1ml of formaldehyde until the appearance of the bands; 7) acetic acid (10%) for 2 minutes; 8) and finally rinsing again with distilled water. Due to the toxicity of some reagents, this entire process was carried out inside a hood fume.

11. Capillary Electrophoresis (CE)

Capillary electrophoresis is a very sensitive technique based on the separation of the DNA amplicons according to their molecular size and allow the detection of these fragments marked with a fluorescent labelled marker. The PCR products migrate through a capillary filled with a polymer by the action of the electronic current. Due to the different lengths of DNA fragments, they migrate at different speeds and consequently have different detection times. These previously fluorescent labeled fragments are detected as they pass through the laser located at the end of the column.

To accurately prove the efficacy of singleplex and multiplex development, the PCR products were analyzed by capillary electrophoresis technique using a Genetic Analyzer 3500 automated sequencer (Applied Biosystems®).

Post PCR preparation of each sample for electrophoresis was performed in a 12µl final volume as following: 11µl of a mix solution containing 0.4µl of GeneScan 500 LIZ dye Size Standard (Applied Biosystems) and 9.6µl of HI-DI Formamide (Applied Biosystems) added to 1µl of PCR product. The following steps (loading of samples onto the CE and running instrument) were performed according to the manufacturer's protocol. Electropherograms were analyzed using GeneMapper ID-X 1.4 Software (Applied Biosystems).

12. Sensivity Testing

The sensitivity of the final developed tetraplex system was tested using a dilution series approach of both biological fluids (semen and vaginal secretions). The detection limits of these biological materials were determined by the lowest amount of sample (concentration) in which it was possible to detect specific mRNA markers for each body fluid. For this, five quantifications of one RNA vaginal secretion sample (RNA-VS9) and one RNA semen sample (RNA-S14) were obtained by the Nanodrop 1000 Spectrophotometer (Thermo Fisher Scientific). The final means of quantification of each sample were then divided by a factor of 2 (as a conservative approach) since other internal validations (data not shown) have shown that this spectrophotometer results are usually overestimated.

Dilutions were performed using the cDNA samples. Due to the high efficiency (near 100%) of the SensiFAST cDNA Synthesis kit (BIRLT S.A., Poland), as most other kits, it is assumed that 1ng of RNA is reverse transcribed into 1ng of cDNA. Thus, the two selected cDNA samples, cDNA-VS9 and cDNA-S14, were used for the vaginal and semen sensitivity testing, respectively. The dilutions were performed in concentrations from 10 ng/µl to 10pg/µl according to Table 10.

The tetraplex system was used to test all diluted samples and by maintaining the PCR conditions previously described. The PCR products (mRNA markers) were also analyzed using GeneMapper ID-X 1.4 Software.

Table 10. *Assay for sensitivity study: cDNA concentrations used for the dilution series.*

Dilution	Dilution Fator	Concentration (ng/µl)
D1	1:2	10
D2	1:2	5
D3	1:2	2.5
D4	1:2	1.25
D5	1:2	0.625
D6	1:2	0.313
D7	1:2	0.156
D8	1:4	0.039
D9	1:4	0.010
D10	1:4	0.002

13. Specificity Testing

To test the specificity of the four markers selected in the detection of semen and vaginal secretions, analyses were performed on different and forensically relevant biological fluids. Therefore, one saliva sample, one peripheral blood sample and one menstrual blood sample were included and submitted to the same previously described procedures.

14. Genotyping

DNA profiles were obtained using the 'GlobalFiler PCR Amplification kit' (Thermo Fisher Scientific®). This kit allows the amplification of 22 autosomal STRs (D3S1358, vWA, D16S539, CSF1PO, TPOX, D8S1179, D21S11, D18S51, DYS391, D2S441, D19S433, TH01, FGA, D22S1045, D5S818, D13S317, D7S820, SE33, D10S1248, D1S1656, D12S391, D2S1338) and amelogenin gene for gender determination.

The entire process was carried out on a previously decontaminated UV (Aura PCR) camera. The amplification protocol in use for casework at the Serviço de Genética e Biologia Forenses, da Delegação do Norte (SGBF-N) do Instituto Nacional de Legal Medicina Legal e Ciências Forenses, IP, (INMLCF) was followed:

Each DNA sample was amplified in 13.5µl of PCR reaction final volume. For each MicroAmp® Reaction tube (ThermoFisher - Applied Biosystems®), 1µl of DNA sample was added to 3.75µl of Master Mix, 1.25µl of Primer Set and 7.5µl of low TE buffer. A negative control (without DNA sample) and a positive control were also performed.

All tubes were sealed with MicroAmp® Caps (ThermoFisher - Applied Biosystems®) and subsequently, they were placed in a GeneAmp® PCR System 9700 thermocycler for amplification according to the manufacturer's protocol.

Post PCR preparation of each sample for electrophoresis was performed in a 14µl final volume. For each well of MicroAmp® Optical 96-Well Reaction Plate (ThermoFisher - Applied Biosystems®), 13µl of a mix solution containing 0.5µl of GeneScan™ 600 LIZ® Size Standard v2.0 (Applied Biosystems) and 12.5µl of HI-DI™ Formamide (Applied Biosystems) were added to 1µl of PCR product. In addition, as reference 1µl of Allelic Ladder included in the kit was added to 13µl of a mix solution previously described.

The plates were denatured at 95°C in a GeneAmp® PCR System 9700 thermocycler for 4 minutes and subsequently subjected to capillary electrophoresis using a Genetic Analyzer 3500 automated sequencer. The following steps were performed according to the manufacturer's protocol.

Electropherograms were analyzed using GeneMapper ID-X version 1.4 to determine the genetic profiles of all extracted DNA samples. The DNA profiles were analyzed using a detection threshold of 150 RFU (relative fluorescent units).

RESULTS AND DISCUSSION

The first step of this study consisted in the co-extraction of DNA and RNA using two different methodologies. This aim was set to compare the efficiencies of a column-based extraction kit (ExtractMe, BLIRT S.A., Poland) and a simple, low cost manual based procedure adapted from (Shojaie *et al.*, 2014) for forensic samples. The objective was to test if the latter procedure would perform at the same level as a commercially developed and optimized kit. This would allow the implementation of a straight forward and low cost DNA/RNA co-extraction protocol in any forensic laboratory with interest in body fluid identification.

Extractions of nucleic acid were divided into two parts: part 1, the semen samples (n=14 x2) were extracted with two different procedures, the manufactured kit ExtractMe (n=14) (BLIRT S.A., Poland) and the manual protocol (n=14) (adapted (Shojaie *et al.*, 2014). Part 2, in addition to the semen samples, the vaginal secretion swabs (n=10) were also extracted with the DNA/RNA kit (ExtractMe). Therefore, for semen samples a total of 28 RNA and 14 DNA extracts were processed, followed by a total of 10 DNA and 10 RNA extracts resulted from the vaginal secretion swabs.

1. Analysis of DNA and RNA quantifications

Results of DNA and RNA extractions using different procedures were analyzed regarding nucleic acid concentrations and purity ratios assessment using a spectrophotometer, the Nanodrop.

The 260/280 is the main measure of nucleic acid purity that is obtained when using a spectrophotometer. A ratio of ~1.8 is generally accepted as “pure” for DNA and of ~2.0 for “pure” RNA. Low ratios of 260/280 are indicative of the presence of proteins or contaminants that absorb strongly at or near 280nm.

The ratio of absorbance at 260nm and 230nm is a secondary measure of nucleic acid purity and the ideal values are between 1.8-2.2 but should be at least greater than 1.5. Typically, lower values for 260/230 reveal the presence of inhibitors that absorb at 230nm that may interfere with downstream applications such as PCR.

The values obtained by the spectrophotometer readings are presented in Attachment 4 - Supplementary table 1 and results are discussed below individually for each part.

1.1 DNA/RNA co-extraction kit ExtractMe

1.1.1 Semen samples

Initially, to test the efficacy of the kit's protocol four samples (samples S1 to S4) were extracted and evaluated. In this first assessment, DNA extracts showed an average concentration of 1.8ng/µl, ranging from 1.3 to 2.6ng/µl. Two out of the four samples showed purity values measured by the 260/280 ratio, greater than 1.8 (Table 11). Regarding RNA extracts, the concentration values varied between 4.0 and 15ng/µl, and the ratio 260/280 presented values above 2.4. For these four preliminary tested samples, the 260/230 ratio values in both DNA and RNA extracts are far below the intended values (<0.45) showing the presence of contaminants. In a second assay, the remaining semen samples (n=10) were extracted with some protocol modifications to improve quantification values.

In addition to the inclusion of DTT and a 56°C incubation step for 30 minutes during lysis, in the co-extraction procedure (samples S5 to S14), a significant increase in the quantity and quality of extracted nucleic acids was observed. Thus, as summarized in Table 11, after the optimization of this method, the average concentrations of DNA and RNA showed values about 26 and 8 times higher, respectively, compared to the average concentrations of the first four samples. Additionally, the DNA and RNA quality assessed by 260/280 was improved, showing values close to the recommended values (Graphic 3a and 3b). The majority of extracted semen samples presented a satisfactory DNA quality assessment values of 260/280 (~1.8) (Attachment 4 - Supplementary table 1). Among the 14 samples extracted only two samples presented low ratios (~1.0). The sample applies for the RNA quantification results: most of the samples (10 out of 14) presented values for accepted "pure" RNA (~2.0). It is believed that, the improvement in quantification results of both DNA and RNA is due to the addition of DTT as this reagent is able to break down the protein disulfide bonds in the acrosome. The destruction of this structure, which covers the sperm head facilitates the release of nucleic acids located inside. The

additional incubation process is an added advantage as it facilitates lysis buffer and DTT performance.

The 230/260 ratios measured on both purified nucleic acids are very low. Only in 4 out of 10 DNA extracts and in 2 out of 10 RNA extracts, the values are above 1.5 (minimum ideal value). This may be due to remaining of buffers present in the eluate that absorb at or near 230nm. Increasing the spin speed in the pre-elution step to remove all alcohol traces from the wash solutions that may interfere with enzymatic reactions would be a possible solution.

As described in manufacturer's protocol of the ExtractMe kit (BLIRT S.A., Poland), metabolically active biological materials contain more RNA than DNA. Therefore, as expected, for the same semen sample, extractions using the commercial kit generally revealed higher values of RNA concentrations than DNA (Attachment 4 - Supplementary Table 1).

1.1.2 Vaginal secretions samples

The DNA and RNA concentrations from vaginal swabs estimated by the spectrophotometer were much lower than the quantities obtained in semen samples (Table 11). These results are not unexpected as the quantity of biological material present in a swab (vaginal secretions) is predicted to be much lower compared to the 200µl of used semen. It is recognized that these values could be improved by including spin baskets in the extraction procedure, ensuring the full recovery of all swab material and therefore maximizing DNA and RNA yields.

The low recovery of DNA and RNA from vaginal secretion samples can be also explained by the high protection of nucleic acids by the large number of bacteria present in the vaginal mucus.

The purity of DNA extracts measured by the 260/280 ratio showed purity assessments of the vaginal secretion samples between 1.1 and 2.0. Most values (9 out of 10 samples) were ≥1.4 (Attachment 4 - Supplementary Table 1). RNA extracts except for one sample showed acceptable values for pure RNA with a purity assessment average of 1.9 (Table 11).

However, as with semen samples extracted with the kit, it is believed that for the same reason as above, the observed 230/260 ratio values were much lower than the intended values, possibly due to carry over of reagents that absorb at 230nm during the extraction procedure.

As known, the vaginal fluid contains a high number of proteins. Thus, the inclusion of proteinase k would be recommended during the process of nucleic acids co-extraction. This enzyme would promote the digestion of membrane proteins facilitating the recovery of DNA and RNA. Additionally, it would be also useful in the nuclease inactivation ensuring the quality of the extracted DNA and RNA.

1.2 DNA/RNA co-extraction with manual kit

1.2.1 Semen

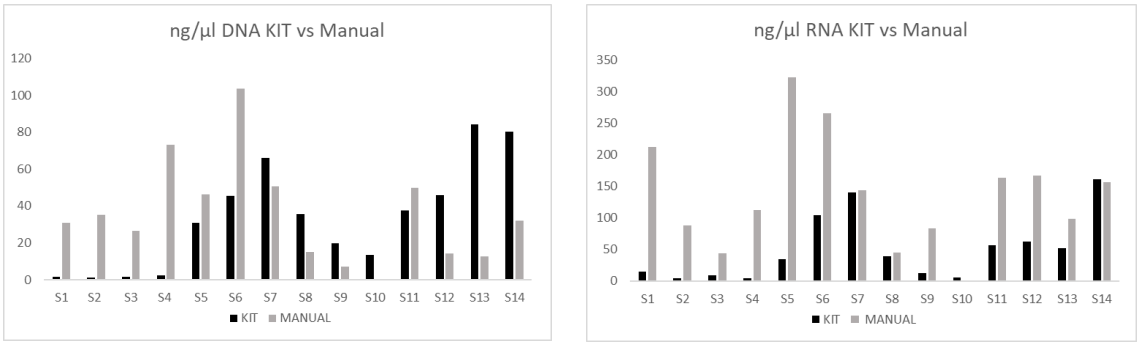
The 14 semen samples from the same donors extracted with the ExtractMe kit were also subjected to the manual DNA and RNA co-extraction method (Shojaie *et al.,* 2014) (Attachment 4 - Supplementary Table 1). However, one sample (S10M) was excluded due to experimental issues because the white coil DNA was not observed (Material and Methods, section 4.2).

Overall, manual method quantifications estimated by Nanodrop spectrophotometer (Table 11) report that only 3 out of 9 DNA samples (Graphic 6A) and in 8 out of 9 RNA samples show higher concentrations (Graphic 6B) comparing with the values of the optimized commercial kit (sample S5 to S14).

Despite the satisfactory nucleic acid concentration values, the purity of DNA and RNA extracts proved to be lower than those obtained by the commercial kit method (Attachment 4 - Supplementary Table 1). None of the RNA extracts and only 2 out of 13 DNA extracts of semen samples revealed acceptable values for purity ratios 260/280 (Graphics 7A and 7B). All values measured by the 260/230 ratio were shown to be less than 1.5, except for one RNA sample. As the manual method is considered to be a rougher methodology due to the lack of the commercial optimization it was expected that the

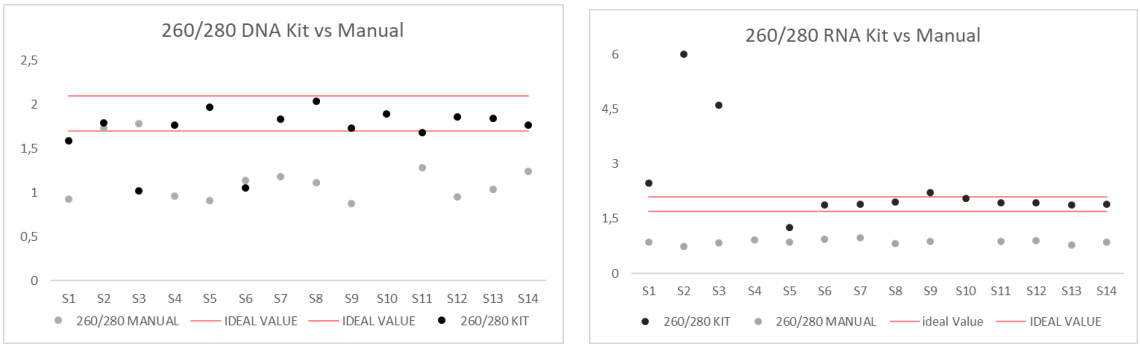
purity values presented for both extracted nucleic acids would be below the limit described in the literature.

A. ***B.***



Graphic 6*. Comparison of* ***A.*** *DNA and* ***B.*** *RNA concentration (ng/µl) in samples extracted with the commercial kit (black bars) and the manual method (grey bars). The S10 sample was not extracted using the manual method.*

A. ***B.***



Graphic 7. *Comparison of* ***A.*** *DNA and* ***B.*** *RNA purity ratios (260/280) in samples extracted with the commercial kit ExtractMe (black points) and the manual method (grey points). The red lines represent the range of ideal values described in the literature. The S10 sample was not extracted using the manual method (missing data).*

Table 11. *Summarized quantification results for* ***A.*** *DNA and* ***B.*** *RNA co-extracted from semen and vaginal secretions samples.*

A.

Sample	Extraction method	*n*	Average of DNA concentration (ng/µL)	Range of DNA concentration (ng/µL)	Average of DNA 260/280	Range of DNA 260/280	Average of DNA 260/230	Range of DNA 260/230
Vaginal Secretions	Commercial Kit	10	13.2	7.4 - 19.3	1.5	1.1 - 2.0	0.5	0.1 - 1.6
Semen	Commercial Kit	14	1.8 *	1.3 - 2.6*	1.5*	1.0 - 1.8*	0.2*	0.0 - 0.4*
			46.0**	19.9 - 84.1**	1.8**	1.1 - 2.0**	1.5**	0.1 - 2.4**
	Manual	13	52.8	7.1 - 213.4	1.2	0.9 - 1.8	0.4	0.2 - 1.3

*1º Assay

** optimized Assay

B.

Sample	Extraction method	*n*	Average of RNA concentration (ng/µL)	Range of RNA concentration (ng/µL)	Average of RNA 260/280	Range of RNA 260/280	Average of RNA 260/230	Range of RNA 260/230
Vaginal Secretions	Commercial Kit	10	15.9	2.0 - 33.8	1.9	1.4 - 2.2	0.2	0.0 - 0.6
Semen	Commercial Kit	14	8.2*	4.0 - 15.0*	4.5*	2.5 - 6.2*	0.1*	0.1 - 0.2*
			67.1**	6.0 - 161.0**	1.9**	1.3 - 2.2**	1.0**	0.1 - 2.4**
	Manual	13	146.5	44.1 - 323.1	0.9	0.7 - 1.0	0.6	0.2 - 2.0

*1º Assay

** optimized Assay

1.2.2 Purified semen samples

As the 260/280 ratios are quite low in the manual adapted protocol (Shojaie *et al.,* 2014), an additional purification step after of both extracted nucleic acids can be performed in order to improve DNA and RNA purity values. Therefore, four DNA and RNA samples extracted by manual method and submitted to the purification process were quantified with a spectrophotometer. The results shown in Table 12 demonstrate a decrease in RNA concentration (average of RNA concentration = 29.65ng/µl) compared to results from the same samples prior to the purification process.

Nucleic acids (both DNA and RNA) absorb light at a wavelength of 260nm. If DNA is still present in the RNA extract, then the Nanodrop spectrophotometer results will be overestimated. The presence of inhibitors that absorb light at the same wavelength will also interfere with this result. Thus, it is presumed that after a purification process and consequently elimination of all extra constituents of the samples concerned, the concentration value will decrease.

Additionally, to test the reliability of the estimated values by the spectrophotometer, the DNA extracts from the purified semen samples were also quantified with the 'Quantifiler Trio DNA Quantification Kit' (Thermo Fisher Scientific®). The DNA concentrations obtained with this kit were even lower than the values obtained with the Nanodrop spectrophotometer (Table 13). As expected, these results confirm the overestimation of the spectrophotometer values, proving the lack of precision of this instrument.

The increase of the purity ratios 260/280 and 260/230 of both DNA and RNA samples reveals, as expected, an improvement in the quality of extracted nucleic acids (Table 12).

Even so, none of these samples show the values within the ideal parameters suggesting the permanence of inhibitors.

Thus, even after inclusion of the purification procedure in this manual methodology, and despite favorable DNA and RNA concentrations, the lack of nucleic acid purity is a major disadvantage as it may compromise further studies.

These conclusions may prove the lack of robustness of this method. It was already expected, since this methodology requires human intervention in every steps, decreasing the accuracy of the entire process and increasing the risk of contamination.

The comparison of results and overall performance between the extraction kit ExtractMe (BLIRT S.A., Poland) and the adapted manual protocol (Shojaie *et al.*, 2014) have resulted in a conference proceeding paper as well as a poster presented at the International Society of Forensic Genetics meeting (ISFG, 2019, Prague). Therefore, more details can be found in Attachment 8 on the results and conclusions of this study.

Table 12. *Summarized quantification results by spectrophotometer for* ***A.*** *DNA and* ***B.*** *RNA co-extracted from semen samples with the manual protocol (M), on the left, and with the manual protocol followed by a purification (MP), on the right.*

A.

	DNA M			DNA MP		
	ng/µl	260/280	260/230	ng/µl	260/280	260/230
S4	73.3	1.0	0.2	28.8	1.7	0.4
S5	46.4	0.9	0.2	72.1	2.0	1.0
S6	103.7	1.1	0.3	163.1	1.8	1.5
S7	50.6	1.2	0.3	48.9	1.5	1.9

B.

	RNA M			RNA MP		
	ng/µl	260/280	260/230	ng/µl	260/280	260/230
S4	112.3	0.9	0.4	12.3	1.5	0.5
S5	323.1	0.9	2.0	5.9	1.4	*n.d.*
S6	265.5	0.9	1.2	37.4	1.7	0.8
S7	143.7	1.0	0.5	63.0	1.6	0.6

n.d. – not determined

***Table 13**. Summarized quantification results by spectrophotometer and 'Quantifiler Trio DNA Quantification Kit' for DNA extracted from semen samples with the manual protocol followed by a purification (MP).*

	ng/µl DNA MP	
	Nanodrop spectophotometer	Quantifiler Trio DNA Quantification Kit
S4	28.8	4.2
S5	72.1	26.7
S6	163.1	62.6
S7	48.9	17.6

2. Analysis of Singleplex system

To develop a small size multiplex capable of identifying the presence of semen and vaginal secretions, four mRNAs were selected from the literature (as described in Materials and Methods, section 8). In order to study the effectiveness of the chosen mRNA markers, for the desired biological fluids, the semen and vaginal secretion specific primers were tested in singleplex. Two different techniques were used to analyze the results.

In the first instance, PAGE results were analyzed. In the gel shown in Figure 3A the expression of the four markers is visible in the following order: PRM1, SEMG1, MUC4, MYOZ1. These mRNAs markers were tested using RNA extracts from two samples of each fluid: PMR1 and SEMG1 in S6 and S7 samples; MUC4 and MYOZ1 in VS1 and VS8 samples. For each marker the respective negative control is also observed.

The RT reactions of the singleplex system showed the expression of the expected bands (red circles), as well as some nonspecific bands marked with black circles in Figure 3A. In order to improve the results, two more tests varying annealing temperature to 60ºC and 65ºC, were performed. However, extra bands of PRM1 marker remained present (Figures 3A and 3B).

Controls with non-RT reactions (RT-) of the singleplex system (Figure 3B), which ideally expected the absence of any expression, also revealed the existence of nonspecific bands. Thus, based on previously published studies (Haas *et al.*, 2013; Roeder and Haas, 2013) it is assumed that extra bands are an expression resulting from genomic contamination.

Since the expected size of the tetraplex is between 81bp and 235bp, it is believed that bands outside these limits will not interfere with the results. However, the extra band marked with an arrow in Figure 3A appears to be the same size as the MUC4 bands. If this happened, the results of MUC4 marker would be invalid.

To accurately visualize the band sizes, the results of capillary electrophoresis (CE) were analyzed using the GeneMapper ID-X version 1.4 Software. Based on the expected (described in literature (Haas *et al.*, 2009; Hanson and Ballantyne, 2013; Juusola and Ballantyne, 2005) and observed size values of the four mRNA markers summarized in Table 14, no interference is expected between the peak of MUC4 and the PMR1 extraneous peak as they exhibit significantly different sizes. According to the results of the electropherogram, it is found that the PRM1 extra peak at 240.29bp is also outside the tetraplex size limits.

According to the literature, once the contaminating gDNA acts as a second competitor in RNA-PCR assays and in order to avoid false positive results, the samples should have been subjected to a DNase I treatment (Haas *et al*., 2013). This enzyme removes the DNA that act as a RNA contaminant and preserve the respective mRNAs. This essential treatment was not carried out due to the lack of financial support necessary to acquire the reagents for this purpose.

Electropherograms of the singleplex system (Figure 4) confirmed the efficacy of the selected biomarkers to identify and distinguish these specific fluids since the specificity is attributed when the expression of a marker is verified in a single body fluid or when its expression is significantly high in a target body fluid compared to others non-target (Haas *et al*., 2013).

All markers showed high expression except for MYOZ1, which revealed a weaker expression than the other markers (MUC4, SEMG1, PRM1). This was not unexpected since similar data are reported in the literature. This weak expression of MYOZ1 cannot be explained as its exact function in vaginal fluid is not known. Despite the promising results of the specificity of MYOZ1 for vaginal fluid is described by some authors, the gene databases do not report the high expression of this biomarker in this body fluid (Hanson and Ballantyne, 2013). For the remaining markers, since PRM1 is strongly expressed in sperm and SEMG1 is responsible for encoding the major seminal fluid protein, the high expression of these markers in semen samples was expected. The high

expression of MUC4 in vaginal samples is also easily explained as this gene encodes the abundant constituent in cervical mucus.

Table 14. *Expected and observed size markers.*

Markers	Expected Size (bp)	Observed Size (bp)
MYOZ1	81	81.05
SEMG1	120	118.44
PRM1	153	149.59 240.29*
MUC4	235	232.21

*Contamination with gDNA - extra band

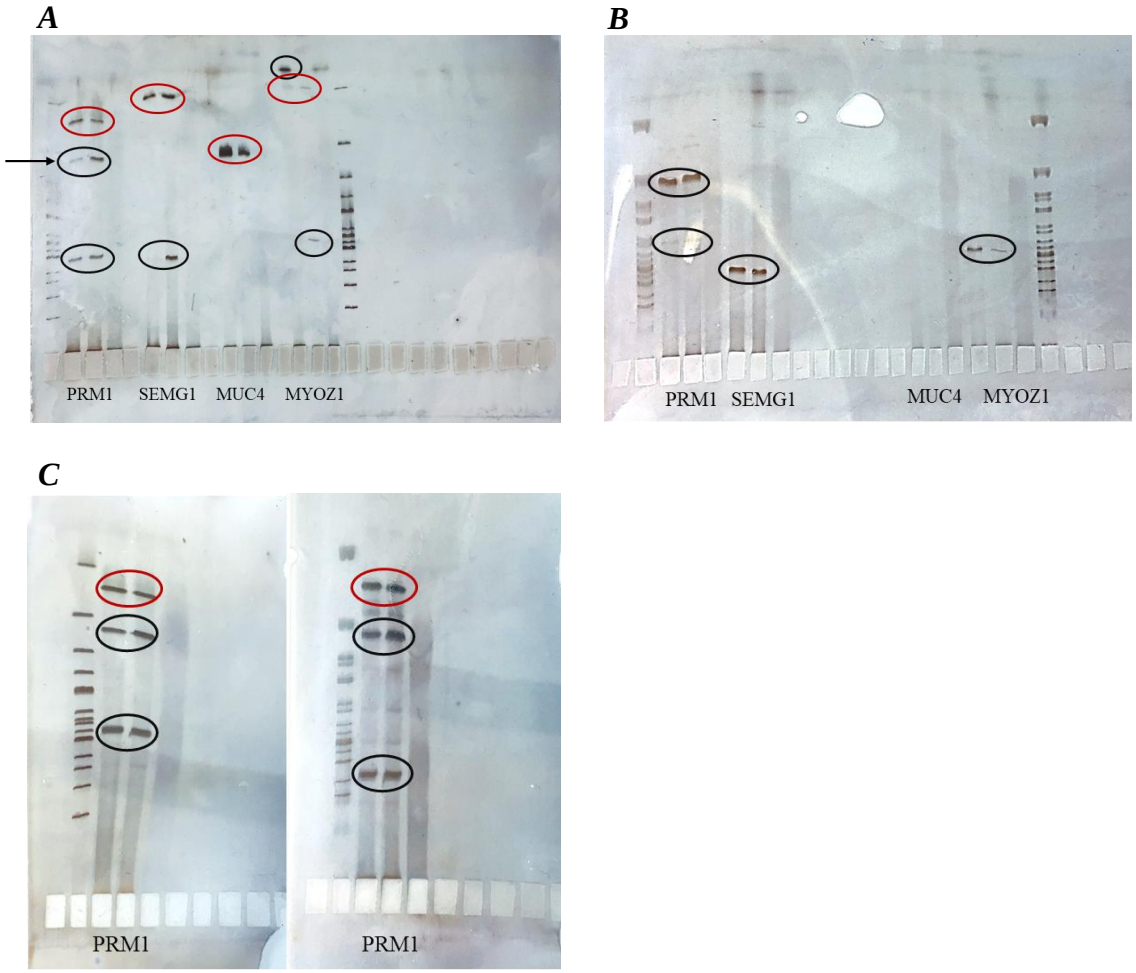


Figure 3. *Results of singleplex assay in agarose gel electrophoresis format.*

A*. RT + reactions. Each mRNA marker (PRM1, SEMG1, MUC4, MYOZ1) was tested on two samples and the respective Negative control was run in parallel.* ***B.*** *For each mRNA markers (PRM1, SEMG1, MUC4, MYOZ1), controls without RT (RT-) were also performed;* ***C.*** *PRM1 marker was tested in singleplex with two different annealing temperatures (left figure with T= 60ºC, right figure with T=65ºC). The expected mRNA bands are indicated with red circles and the non-specific bands with black circles. The PRM1 extra-band with similar size to MUC4 is marked with an arrow.*

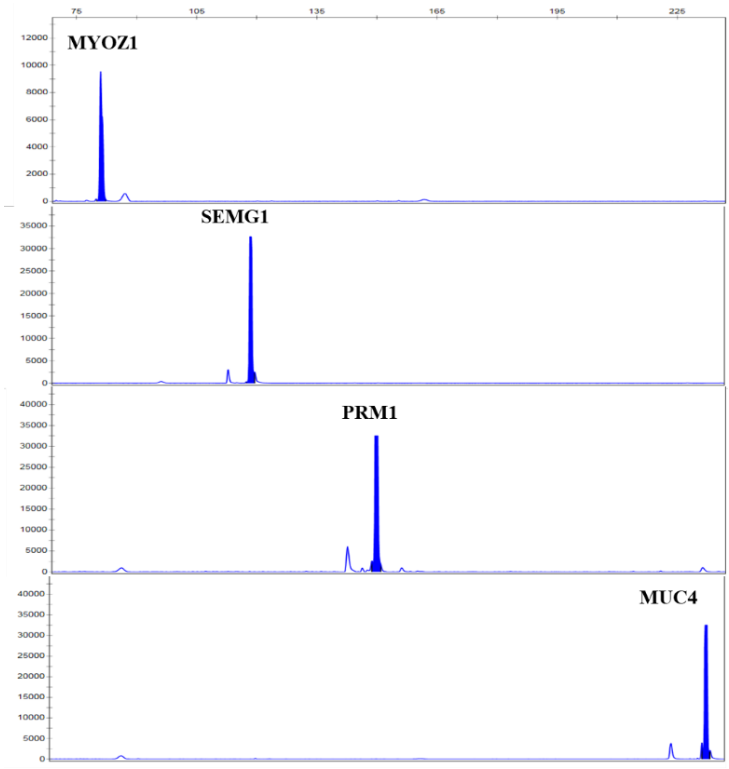


Figure 4. *Results of singleplex assay in capillary electrophoresis format. Representative electropherograms of each mRNA marker expression in their specific body fluids (MYOZ1 and MUC4 in vaginal secretions; SEMG1 and PRM1 in semen).*

3. Analysis of Tetraplex system

3.1 Analysis of Tetraplex system in Semen and Vaginal secretion samples

After demonstrating the specificity of the selected markers, the tetraplex was developed for simultaneous analysis of semen and vaginal secretion samples. Two semen-specific and two vaginal secretion-specific mRNA markers were analyzed in one reaction. Initially, the multiplex with the four markers was tested on samples previously used in the singleplex system: two semen samples (S6, S7) and two vaginal secretion samples (VS1, VS8). The PCR products were run on a polyacrylamide gel and the results show the expression of the specific markers of each fluid (in semen samples: PMR1 and SEMG1; in vaginal secretion samples: MUC4 and MYOZ1). Additionally, an expression of a set of nonspecific bands with size greater than 500bp was also observed. The extra bands of the PMR1 marker at 240bp, observed in the semen samples singleplex, were

also detected in the tetraplex analysis (Figure 5). As previously mentioned, due to the absence of DNase I treatment it is assumed that these bands are derived from genomic contaminations. However, no interferences with the results are expected since they are larger than the expected limit of the tetraplex.

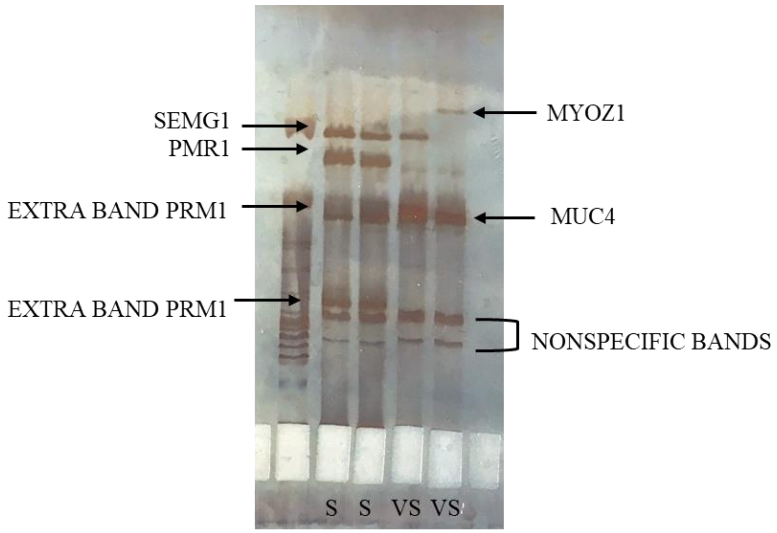


Figure 5. *Results of tetraplex assay in polyacrylamide gel electrophoresis format. The expected sizes of the tetraplex markers (MYOZ1, SEMG1, PRM1, MUC4) as well as extra bands and some nonspecific bands are indicated with an arrow.*

In order to visualize PCR products more accurately, the results were analyzed by capillary electrophoresis. The bands of polyacrylamide gel electrophoresis were confirmed by electropherograms which, as intended, in a range between 75bp and 235bp only demonstrated the peak expression of markers in the respective fluids (e.g., Figures 6A and 6B).

In these four samples, all markers, except MYOZ1, were successfully amplified (Table 15). Due to undetectability of this marker in the VS1 sample and in order to improve its poor expression detected in the VS8 sample, a new primer mix (Primer Mix 2) with some adjustments in the concentrations of the different markers (described in Materials and Methods, section 9.2.) was included in this study.

To optimize the multiplex assay, the remaining semen and vaginal secretions samples were amplified with Primer Mix 2 and the respective electropherograms were analyzed. The results demonstrated the expression of all markers in the respective samples (Table 15) which confirm the ability of the multiplex to identify the desired fluids.

Although, MYOZ1 was successfully amplified in the remaining vaginal samples, in 7 out of 8, the expression of this specific marker was poor compared to the remaining markers. These evidences coincide with the results demonstrated in the singleplex analysis and they are corroborated by other previously published studies (Hanson and Ballantyne, 2013; Hass *et al.*, 2014). As noted, the reason for poor expression is thought to be intrinsically related to unknown gene function in vaginal samples.

In addition to the identification of vaginal samples by the tetraplex, it was intended to establish a relationship between the expression pattern of these markers and the hormonal cycle phase. The main objective was to alert to possible inter-individual variations of mRNA expression in this fluid and to prevent false-positive conclusions. It is known from the literature (described in Introduction, section 2.3.2.) that especially MUC4 varies depending on progesterone concentration. Thus, its expression would be expected to decrease in samples from donors in the lutein phase as well as, in samples from menopausal donors, since at these stages there is a marked decrease in female hormone production.

However, this study included only two representative samples of these phases and no differences in MUC4 expression were found. Thus, due to the scarcity of representative samples of different phases of the hormonal cycle, it was impossible to infer any conclusion. For a detailed study of this theme more representative samples of the different phases should be included (Attachment 5 – Supplementary Table 2).

In addition to the identification of all semen samples by the tetraplex, due to the high expression of PRM1 and SEMG1 in all samples, it can also be concluded that none of the samples come from an azoospermic or vasectomized men because PRM1 is only detectable in fertile individuals. This statement was confirmed by microscopic observation of semen samples stained by 'Christmas Tree' staining, which allowed the visualization of spermatozoa.

To evaluate the effectiveness of tetraplex in samples submitted to different extraction methods, RNA semen samples extracted using the manual extraction protocol were also analyzed.

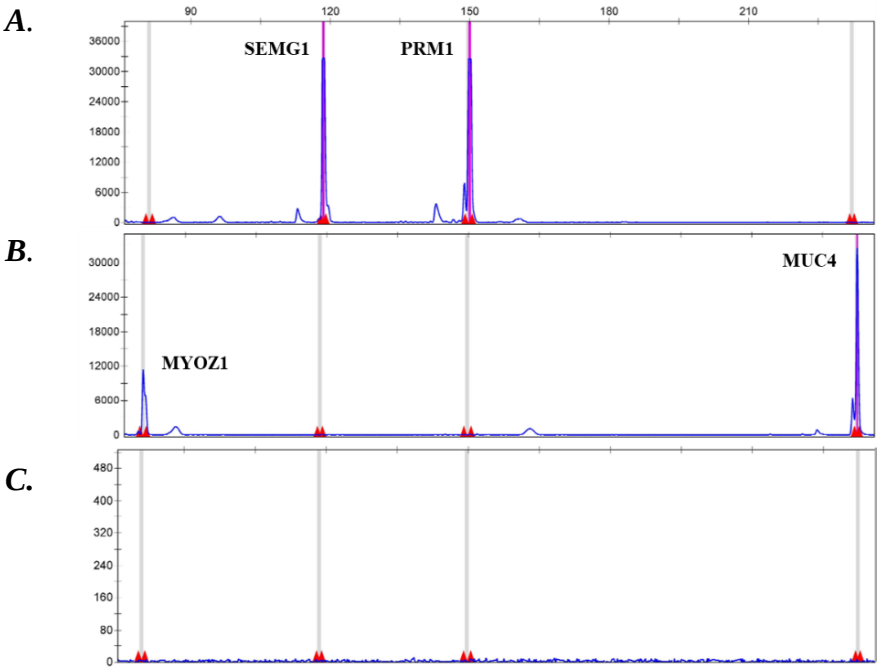


Figure 6. *Results of tetraplex assay in capillary electrophoresis format. Representative electropherograms of the:* ***A.*** *Tetraplex expression in semen samples extracted with the commercial kit,* ***B.*** *Tetraplex expression in vaginal secretion samples and* ***C.*** *Tetraplex expression in semen samples extracted with the manual protocol.*

Table 15. *Summarized results of tetraplex mRNA expression in semen and vaginal secretion samples. Results marked with (*) means that mRNA expression was not expected = possible contamination.*

Body Fluid samples	Extraction method	Tetraplex			
		MYOZ1	PRM1	SEMG1	MUC4
Vaginal secretions	Commercial kit	9/10			10/10
Semen	Commercial kit		14/14	14/14	
Semen	Manual protocol		0/13	0/13	2/13*
Semen	Manual & Purified protocol		1/4	1/4	1/4*

*No expected expression

3.2 Analysis of Tetraplex system in purified samples

Due to the strong hypothesis of the presence of inhibitors that impede the amplification of mRNA markers, the tetraplex was tested on 4 semen samples extracted using the manual method and submitted to the purification protocol (S4MP, S5MP, S6MP, S7MP). The capillary electrophoresis was performed and the electropherograms were analyzed.

In 3 out of 4 samples no expression of any marker was observed, as in the example of Figure 7. However, in the S7MP sample the expression of the semen specific markers was detected and contrary to what would be expected, a significant peak in the vaginal secretion-specific marker MUC4 was also observed (Table 15). As previously mentioned, it is believed that this unexpected result is an intra-laboratory contamination possibly arising from the use of the same laboratory material, since no co-expression results are observed in sample S7 in previously assays.

In this assay, the electropherograms corroborate the quantification results of the purified samples, since the sample S7MP corresponds to the sample with the highest RNA concentration and those with higher values of purity levels (Attachment 4 - Supplementary table 1). The remaining samples with low purity values even after purification process could not observe any biomarkers expression.

Due to the lack of positive results, this method of manual extraction, even after cleaning the RNA extracts by the purification procedure, proved to be a precarious method for mRNA profiling in forensic use. It is thought that the lack of results for the detection of biological fluids is because this co-extraction protocol is specific for extraction of DNA and RNA from animal cell cultures and no from body fluids.

Due to the lack of financial resources necessary to optimize the manual extraction method specific to body fluids, it was not possible to proceed with the study for the remaining samples.

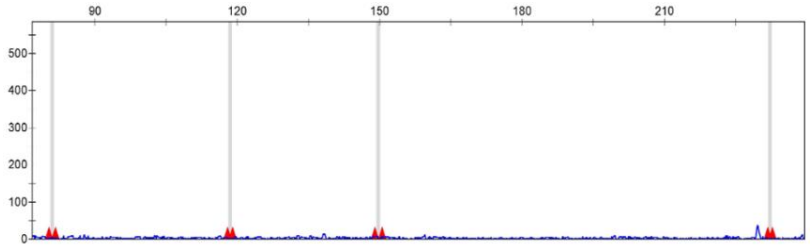


Figure 7. *Results of tetraplex assay in capillary electrophoresis format. Representative electropherograms of the Tetraplex expression in semen samples extracted with the manual protocol and submitted a purification process.*

3.3 Analysis of Tetraplex system in mix samples

A multiplex suitable for forensic use must be able to identify and distinguish different fluid simultaneously. In order to simulate the samples commonly found in scenarios of alleged Sexual Assault and to test the performance of the developed tetraplex system, semen/vaginal secretions mixtures were also analyzed.

To evaluate the co-extraction method in low quantities of complex mixtures two assays were performed. In the first assay, the tetraplex was tested using RNA samples extracted with the commercial kit following the protocol. To improve the performance of the entire procedure, in the second trial, the tetraplex was tested using RNA samples including spin baskets during the extraction process.

The tetraplex was tested on single samples from the same donors with reference quantities (Ref1, Ref2, Ref3, Ref4) and the electropherograms were analyzed. The results showed high expression, except for MYOZ1 which revealed weak expression as expected. The electropherograms of mixed samples from the two assays were compared with the results from the reference samples. These results are summarized in Table 16.

According to the results, the second assay showed an improvement in the expression of body fluid markers with a percentage of expressed markers of 63%. While only 38% of the markers showed positive signs in the first trial. This improvement was expected as spin baskets allow complete lysate recovery and therefore maximize DNA and RNA yields.

From a general perspective it is observed that in the mix samples, the specific markers of vaginal fluid have a weaker expression than the semen markers, since MUC 4 and MYOZ1 always show values lower than the values of the reference samples. Thus, it is recognized that in these mix samples, semen is the dominant fluid. This conclusion is supported by the results of the quantification study of RNA extracts from single samples, since according to Ingold *et al.*, 2018, the dominant fluid is usually the one with the highest DNA and RNA concentration values.

Analyzing the optimized method, the undetectability of MYOZ1 is observed in 3 out of 4 samples. These results are explained by the poor expression shown in previous studies. Thus, this marker showed only positive but low intensity results in Mix 2 (1 VS swab +

5µl S). It is believed that this expression was possible due to the decrease in the amount of dominant fluid.

Except for MYOZ1, tetraplex expression in Mix 1, 2 and 3 was similar, with no significant differences in the intensity of expression of the remaining markers. In Mix 4, whose ratios of each body fluid are the lowest (1/2 vaginal swab + 5µl semen) no expression of the markers was observed (Table 16).

In mixtures with traces of fluid samples, this technique would have to undergo further adjustments in order to obtain positive results. The addition of proteinase K during the extraction process would be one of the favorable options for acting on proteins that could prevent marker expression.

Table 16. *Results of mRNA expression in mixed samples in the first assay (1˚) and the second and optimized assay (2˚). Green represents the detection of mRNA expression equal to the expression of reference samples; Yellow color indicates expression values lower than the values of the reference samples; Red color means no detection of mRNA expression.*

mRNA Markers	Mix 1 1 VS swab + 10µl S		Mix 2 1 VS swab + 5µl S		Mix 3 1/2 VS swab + 10µl S		Mix 4 1/2 VS swab + 5µl S	
	1˚	2˚	1˚	2˚	1˚	2˚	1˚	2˚
MYOZ1								
SEMG1								
PMR1								
MUC4								

4. Analysis of sensitivity testing

The sensitivity of the tetraplex system was tested using different amounts of cDNA inputs, from 2pg to 5ng. The peak height results, shown in Table 17 represent the intensity of marker expression. According to the Table 17, the green color represents the intensity values similar to the values observed in the singleplex analysis (expected values). The yellow color shows the successive decrease of marker expression along the decrease of cDNA concentration. The red color indicates the absence of marker detection. In general,

in the serial dilutions of the two body fluids, the expression of the different markers decreases when the concentration of the template samples also decreases.

The two semen-specific markers have been shown to be equally sensitive since they are detected up to concentrations of 39pg/µl. However, PRM1 showed a higher intensity of expression, with practically no changes up to the concentration of 156pg/µl.

For vaginal secretions, MYOZ1 marker turned out to be the weakest expression tetraplex marker. As expected, other studies prove that MYOZ1 is the least sensitive mRNA marker for low sample quantities. MYOZ1 expression intensity decreases from 2.5ng/µl dilution (D3) (Table 17) and has a detection limit of 156pg/µl. MUC4 was found to be the most sensitive tetraplex marker since it was still detectable in 10pg/µl.

In forensic caseworks, most samples are found in very small quantities. It is therefore extremely important that all techniques are sensitive enough to obtain results from limited quantities. The sensitivity results of the tetraplex (four markets are detectable until a concentration of 156pg/µl de cDNA input) shown the system is applicable for forensic use since as described in the literature, the stains typically found in forensic cases with an average size of 50µl containing a high amount of total RNA (Juusola and Ballantyne, 2005).

Table 17. *Results of sensitivity tests for the tetraplex markers in semen and vaginal secretions dilution series. Green color represents intensity values similar to the values observed in the singleplex analysis. Yellow color shows values lower than those observed in the singleplex analysis. Red color represents no detection of mRNA markers.*

Dilution	Concentration (ng/µl)	Peak height			
		Semen		Vaginal Secretion	
		SEMG1	PRM1	MYOZ1	MUC4
D2	5	32476	31991	11005	29335
D3	2.5	32494	32074	8799	32445
D4	1.25	32451	31709	6544	31532
D5	0.625	**– Not analyzed –**			
D6	0.313	17575	32596	935	9286
D7	0.156	11238	32488	4850	15080
D8	0.039	8671	7529	–	1526
D9	0.010	–	–	–	3293
D10	0.002	–	–	–	–

5. Analysis of specificity testing

In forensic casework resolution, to obtain reliable results and avoid false positive conclusions in the identification of biological fluids, the specificity of the developed multiplex must be ensured.

The specificity of this tetraplex for the identification of semen samples and vaginal secretion samples is guaranteed by the presence of marker expression in these samples and the absent in other body fluid samples. For this study, Blood (B), Menstrual Blood (MB), and Saliva (SAL) were included and the results are summarized in Table 18.

As expected, in peripheral blood sample none tetraplex marker was amplified. However, in menstrual blood sample, MUC4 positive signals were detected.

For saliva, the expression of all four candidate markers was observed. MYOZ1 and MUC4 detection revealed cross-reactivity with buccal cells consistent with previously studies which reported that it is usually found positive signals of vaginal secretion markers in saliva samples (Cossu *et al.*, 2009; Lindenbergh *et al.*, 2012).

The detection of MUC4 in saliva and menstrual blood samples can be explained since this marker is one of the main constituents of secretion that covers epithelial surfaces (such as cervix and buccal epithelium). During the sampling, some of epithelial cells were probably transferred and due to the similarity of epithelial constitution of the various tissues, false positive results were obtained.

The same happens with the detection of MYOZ1 in the saliva sample. According to data from the BIOGPS database, this marker also shows high expression in tongue cells.

As suggested by Song *et al.*, and in order to avoid the detection of false positives, specific markers of mucus or epithelial cells due to their high biochemical similarity should not be included in forensic fluid identification. Another alternative to avoid the common co-expression of vaginal markers in saliva samples would be to include microbial markers specific to the vaginal flora.

However, no reasons are known for the unexpected high expression of semen-specific markers (PMR1 and SEMG1) in saliva samples.

It is recognized that in order to improve the specificity testing of this tetraplex system, other body fluids as well as more samples from different donors should be included. For a complete specificity test, non-human species should also be included in this study. The amount of materials available as well as the financial limit did not allow to progress with a more rigorous study.

***Table 18**. Results of specificity tests for the tetraplex markers tested on five body fluids: Blood, Menstrual Blood, Saliva, Semen, Vaginal Secretions.*

Body Fluid samples	MYOZ1	SEMG1	PRM1	MUC4
Blood (B)	–	–	–	–
Menstrual Blood (MB)	–	–	–	✔
Saliva (SAL)	✔	✔	✔	✔
Semen (S)	–	✔	✔	–
Vaginal Secretions (VS)	✔	–	–	✔

6. Analysis of DNA profiles

In forensic investigations, obtaining genetic profiles is crucial as it may allow donor(s) identification. However, in this study, since the samples come from anonymous volunteers, it was impossible to compare these profiles with the reference profiles.

Achieving a good quality identification profile depends on the state of the sample as well as the extraction and amplification procedure. This part of the study aimed to evaluate the genetic profiles amplified with 'GlobalFiler PCR Amplification kit' (Thermo Fisher Scientific®) from the DNA samples extracted with different methodologies. Electropherograms of semen DNA samples extracted with the two different procedures, the manufactured kit ExtractMe (n=14) (BLIRT S.A., Poland) and the manual protocol (n=13) (adapted (Shojaie *et al.*, 2014), semen DNA samples extracted and purified with the adapted manual protocol (n=13), the vaginal secretion DNA samples extracted with manufactured kit ExtractMe (n=10) and mixed samples (n=4) were analyzed using the GeneMapper ID-X version 1.4 Software.

6.1 Semen

In DNA extracts from semen samples extracted by the commercial kit (ExtractMe, BLIRT S.A., Poland), all DNA markers were successfully amplified (Attachment 6 - Supplementary Table 3) obtaining complete identification profiles such as the example of Figure 8.

On the other hand, only 2 out of 13 samples extracted using the manual method (15% of the samples), resulted in full male DNA profile corresponding to the profiles obtained with the Extraction Kit. In the remaining samples (85%), despite amplification of some markers (Attachment 6 - Supplementary Table 3) these are not enough to obtain any identification profile, as demonstrated by the example of the electropherogram Figure 9.

Due to lack of satisfactory results, semen samples extracted by the manual method were subjected to a purification process. These purified extracts were again amplified in order to obtain complete genetic profiles. Although an improvement in the number of amplified labels after the purification process is noticeable, the results are still lower than in the extraction kit. Only in 6 out of 13 samples were possible to obtain a complete DNA profile and the remaining samples were classified as “absence of profile”.

Although the values obtained by the quantification assay show enough concentration of DNA, the low purity levels may be one of the impediment factors for obtaining the DNA profiles. Another possible reason may be the inhibition of the DNA markers amplification by some reagent used, since this manual protocol is not specific for DNA and RNA co-extraction from body fluids, as previously reported.

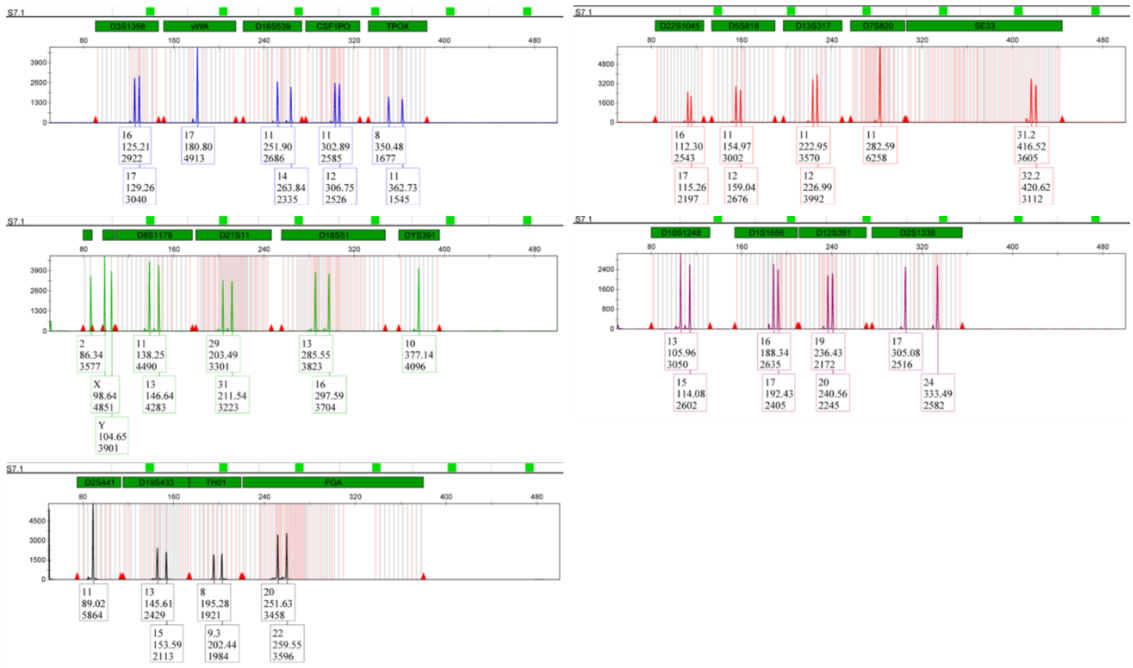


Figure 8. *Example of an electropherogram of a complete DNA profile from semen samples extracted using the commercial extraction (ExtractMe, BLIRT S.A., Poland).*

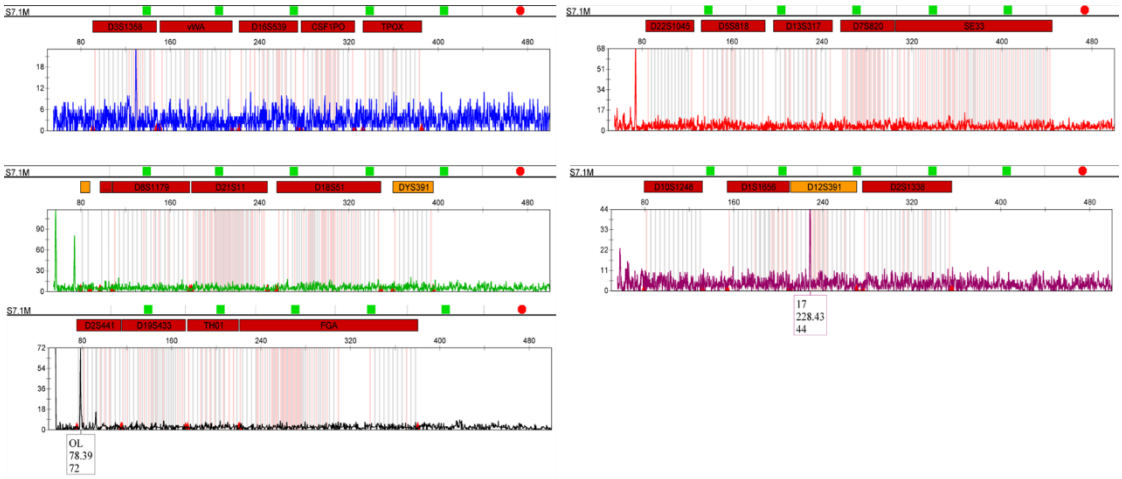


Figure 9. *Example of an electropherogram classified with 'absence of DNA profile' from semen samples extracted using the manual co-extraction protocol.*

These results may indicate the lack of reliability of the manual extraction process for DNA profiles analysis. Although this quick manual method is more economical, it has proved to be a non-robust method not suitable for forensic use. For the optimization of this manual method, more tests would be needed.

6.2 Vaginal Secretions

In vaginal secretion samples, all markers of the DNA female samples extracted using the commercial kit were successfully amplified (Attachment 6 - Supplementary Table 3) obtaining complete identification profiles such as the example of Figure 10.

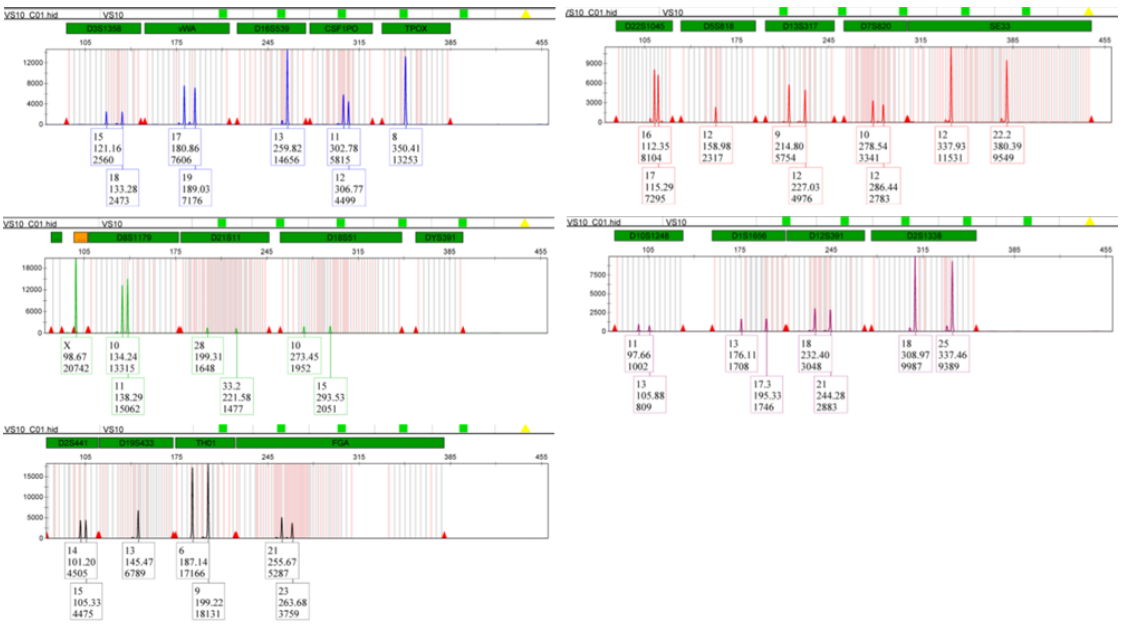


Figure 10. *Example of an electropherogram of a complete DNA profile from vaginal secretion samples extracted using the commercial extraction kit (ExtractMe, BLIRT S.A., Poland).*

6.3 Mixtures

Fluid mixtures are commonly found in forensic scenarios and even in small quantities can contribute to the resolution of these cases. So, in order to identify the donors, the genetic profiles of the mixed samples were also performed.

In these four mixed samples extracted with the commercial kit, all DNA markers were successfully amplified (Attachment 6 - Supplementary Table 3) obtaining complete identification profiles. As expected, the same genetic profile was obtained for the four mixes samples and it corresponds to the profiles of samples S7 and VS3 previously analyzed Figure 11.

However, in the electropherograms analysis, the shape of the peaks showed excess of total DNA input in the genotyping procedure. In order to facilitate the interpretation of the results, samples should be diluted to provide adequate DNA input.

The different ratios of semen and vaginal secretion did not reveal any implications since it was possible to obtain good genetic profiles of all the samples and no differences were observed. These results also prove the high sensitivity of the fundamental methods used to obtain these profiles and to identify the presence of different donors even in vestigial mix samples.

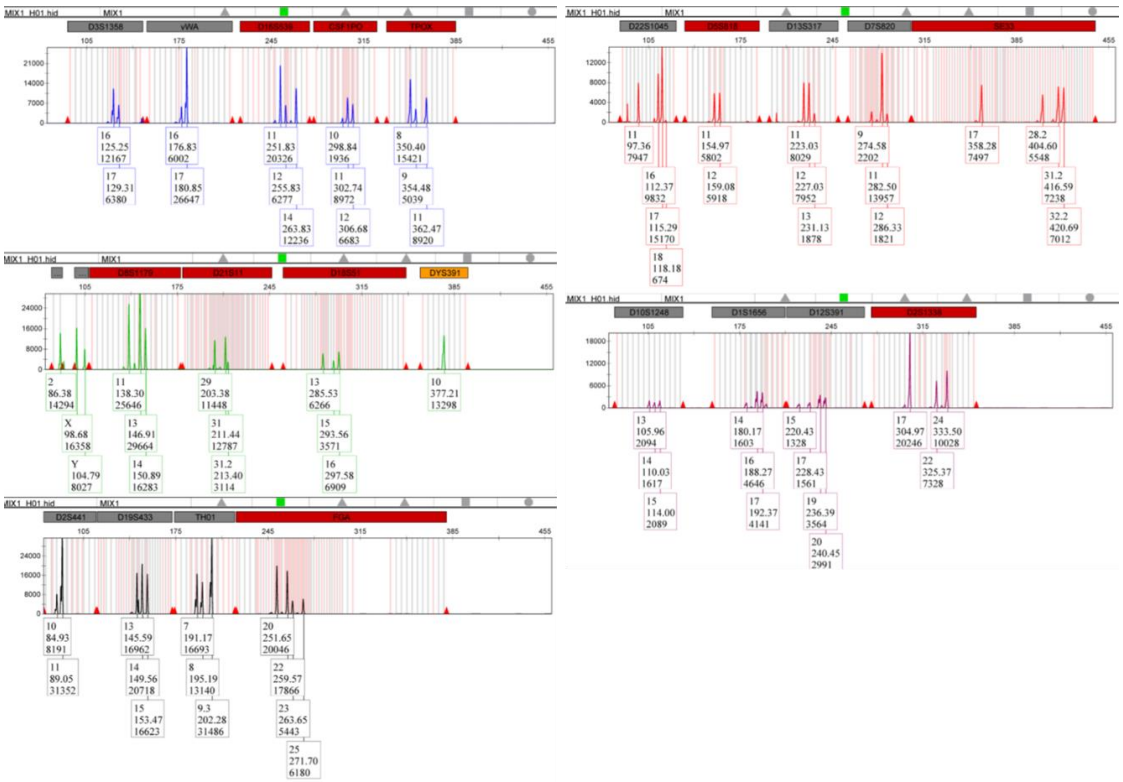


Figure 11. *Example of an electropherogram of a complete DNA profile from mixed samples extracted using the commercial extraction kit (ExtractMe, BLIRT S.A., Poland).*

CONCLUSIONS

This research work presents the following conclusions:

1. The use of the commercial extraction kit is the most reliable method for extracting nucleic acids from small amounts of forensic spots ensuring the quality of the DNA and RNA extracted samples.

2. Simultaneous DNA and RNA extraction has proven to be a reliable technique for limited quantity samples. This method allows also, the DNA extraction to identify donor samples as well as RNA extraction to identify the nature of the fluid.

3. The spectrophotometer Nanodrop is not a rigorous method for quantifying nucleic acids. For forensic applications, Real Time PCR quantification would be the most appropriate method. This method provides the exact concentration ensuring the use of exact DNA and RNA inputs in further analysis.

4. The selected markers were specific to identify these body fluids. However, MYOZ1 was found to have a much weaker expression than the others.

5. This developed multiplex has been found to be valid for identifying semen and vaginal secretions spots and it is enough sensibility for forensic use.

Thus, it is concluded that, due to frequent difficulty in analyzing forensic samples and the impossibility of applying conventional methods, the study of the mRNA expression profile, namely the developed tetraplex is the reliable method capable of identifying semen and vaginal secretions samples in forensic context. However, it is recognized that in future work more semen and vaginal secretions specific markers should be included in this multiplex to improve its specificity. To enhance the specificity test it would be advantageous to test this multiplex on non-human samples. It would also be interesting to extend this multiplex to make it able to detect more body fluids as saliva, menstrual blood, peripheral blood, skin, among others. In the future it is also expected to get the financial support to perform all the missing tests and treatments.

REFERENCES

Afolabia OA, Roederb AD, Iyengara A and Hadia S (2017) Evaluation of genetic markers for forensic identification of human body fluids. *Forensic Science International: Genetics Supplement Series* **6**:41-43.

Albani PP and Fleming R (2018) Novel messenger RNAs for body fluid identification. *Sci Justice* **58**:145-152.

An J-H, Shin K-J, Yang W-I and Lee H-Y (2012) Body fluid identification in forensics. *BMB Reports* **45**:545-553.

APAV (2017) Estatísticas APAV Crimes Sexuais a ≥18 entre 2013-2016.

APAV (2018) Estatísticas APAV Relatorio Anual 2017.

Aregger M and Cowling VH (2017) Regulation of mRNA capping in the cell cycle. *RNA Biol* **14**:11-14.

Bexon K and Williams G (2015) Characterising the fluctuation of microRNA expression throughout a full menstrual cycle. *Forensic Science International: Genetics Supplement Series* **5**:e264-e266.

Blackman S, Stafford-Allen B, Hanson EK, Panasiuk M, Brooker AL, Rendell P, Ballantyne J and Wells S (2018) Developmental validation of the ParaDNA((R)) Body Fluid ID System-A rapid multiplex mRNA-profiling system for the forensic identification of body fluids. *Forensic Sci Int Genet* **37**:151-161.

Cossu C, Germann U, Kratzer A, Bar W and Haas C (2009) How specific are the vaginal secretion mRNA-markers HBD1 and MUC4? *Forensic Science International: Genetics Supplement Series* **2**:536–537.

Cotton RW and Fisher MB (2015) Review: Properties of sperm and seminal fluid, informed by research on reproduction and contraception. *Forensic Sci Int* **18**:66-77.

Del Pino M, Svanholm-Barrie C, Torne A, Marimon L, Gaber J, Sagasta A, Persing DH and Ordi J (2015) mRNA biomarker detection in liquid-based cytology: a new approach in the prevention of cervical cancer. *Mod Pathol* **28**:312-320.

Dere E, Wilson SK, Anderson LM and Boekelheide K (2016) From the Cover: Sperm Molecular Biomarkers Are Sensitive Indicators of Testicular Injury following Subchronic Model Toxicant Exposure. *Toxicol Sci* **153**:327-340.

Dharmaraj N, Chapela PJ, Morgado M, Hawkins SM, Lessey BA, Young SL and Carson DD (2014) Expression of the transmembrane mucins, MUC1, MUC4 and MUC16, in normal endometrium and in endometriosis. *Hum Reprod* **29**:1730-1738.

Dorum G, Ingold S, Hanson E, Ballantyne J, Snipen L and Haas C (2018) Predicting the origin of stains from next generation sequencing mRNA data. *Forensic Sci Int Genet* **34**:37-48.

Etheridge A, Lee I, Hood L, Galas D and Wang K (2011) Extracellular microRNA: A new source of biomarkers. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **717**:85-90.

Evans GE, Phillipson GTM, Sykes PH, McNoe LA, Print CG and Evans JJ (2018) Does the endometrial gene expression of fertile women vary within and between cycles? *Hum Reprod* **33**:452-463.

Fleming RI and Harbison S (2010) The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids. *Forensic Sci Int Genet* **4**:244-256.

Fujimoto S, Manabe S, Morimoto C, Hamano Y and Tamaki K (2017) Effect of the absence of spermatozoa on microrna-based semen identification. *Forens Sci Int-Gen S* **6**:E238-E240.

Fujimoto S, Manabe S, Morimoto C, Ozeki M, Hamano Y and Tamaki K (2018) Optimal small-molecular reference RNA for RT-qPCR-based body fluid identification. *Forensic Science International: Genetics Supplement Series* **37**:135-142.

Garcia-Peiro A, Martinez-Heredia J, Oliver-Bonet M, Abad C, Amengual MJ, Navarro J, Jones C, Coward K, Gosalvez J and Benet J (2011) Protamine 1 to protamine 2 ratio

correlates with dynamic aspects of DNA fragmentation in human sperm. *Fertil Steril* **95**:105-109.

Goffinet F, Kayem G, Maillard F, Trebeden H, Cabrol D, Weill B and Batteux F (2005) Detection of interleukin 6 mRNA by RT-PCR in vaginal secretions: association with preterm delivery and neonatal infection in women with preterm labour and intact membranes. *Eur J Obstet Gynecol Reprod Biol* **123**:167-173.

Gomaa R, Salehi J and Behl S (2017) DNA Methylation as a Biomarker for Body Fluid Identification. *Arab Journal of Forensic Sciences and Forensic Medicine* **1**.

Gomes I, Strohbücker B, Rothschild MA and Schneider PM (2013) Evaluation of mRNA specific markers using a pentaplex system for the identification of skin and saliva from contact trace evidence. *Forensic Science International: Genetics Supplement Series* **4**:e180-e181.

Haas C, Hanson E, Anjos MJ, Ballantyne KN, Banemann R, Bhoelai B, Borges E, Carvalho M, Courts C, De Cock G, Drobnic K, Dotsch M, Fleming R, Franchi C, Gomes I, Hadzic G, Harbison SA, Harteveld J, Hjort B, Hollard C, Hoff-Olsen P, Huls C, Keyser C, Maronas O, McCallum N, Moore D, Morling N, Niederstatter H, Noel F, Parson W, Phillips C, Popielarz C, Roeder AD, Salvaderi L, Sauer E, Schneider PM, Shanthan G, Court DS, Turanska M, van Oorschot RA, Vennemann M, Vidaki A, Zatkalikova L and Ballantyne J (2014) RNA/DNA co-analysis from human menstrual blood and vaginal secretion stains: results of a fourth and fifth collaborative EDNAP exercise. *Forensic Sci Int Genet* **8**:203-212.

Haas C, Hanson E, Anjos MJ, Banemann R, Berti A, Borges E, Carracedo A, Carvalho M, Courts C, De Cock G, Dotsch M, Flynn S, Gomes I, Hollard C, Hjort B, Hoff-Olsen P, Hribikova K, Lindenbergh A, Ludes B, Maronas O, McCallum N, Moore D, Morling N, Niederstatter H, Noel F, Parson W, Popielarz C, Rapone C, Roeder AD, Ruiz Y, Sauer E, Schneider PM, Sijen T, Court DS, Sviezena B, Turanska M, Vidaki A, Zatkalikova L and Ballantyne J (2013) RNA/DNA co-analysis from human saliva and semen stains--results of a third collaborative EDNAP exercise. *Forensic Sci Int Genet* **7**:230-239.

Haas C, Muheim C, Kratzer A, Bär W and Maake C (2009) mRNA profiling for the identification of sperm and seminal plasma. *Forensic Science International: Genetics Supplement Series* **2**:534-535.

Hanson E, Ingold S, Haas C and Ballantyne J (2018) Messenger RNA biomarker signatures for forensic body fluid identification revealed by targeted RNA sequencing. *Forensic Sci Int Genet* **34**:206-221.

Hanson EK and Ballantyne J (2013) Highly specific mRNA biomarkers for the identification of vaginal secretions in sexual assault investigations. *Sci Justice* **53**:14-22.

Hara M, Nakanishi H, Yoneyama K, Saito K and Takada A (2016) Effects of storage conditions on forensic examinations of blood samples and bloodstains stored for 20 years. *Leg Med (Tokyo)* **18**:81-84.

Harbison S and Fleming R (2016) Forensic body fluid identification: state of the art. *Research and Reports in Forensic Medical Science*.

Harteveld J, Lindenbergh A and Sijen T (2013) RNA cell typing and DNA profiling of mixed samples: can cell types and donors be associated? *Sci Justice* **53**:261-269.

Heidary Z, Zaki-Dizaji M, Saliminejad K and Khorramkhorshid HR (2019) Expression Analysis of the CRISP2, CATSPER1, PATE1 and SEMG1 in the Sperm of Men with Idiopathic Asthenozoospermia. *J Reprod Infertil* **20**:70-75.

Ingold S, Dorum G, Hanson E, Berti A, Branicki W, Brito P, Elsmore P, Gettings KB, Giangasparo F, Gross TE, Hansen S, Hanssen EN, Kampmann ML, Kayser M, Laurent FX, Morling N, Mosquera-Miguel A, Parson W, Phillips C, Porto MJ, Pospiech E, Roeder AD, Schneider PM, Schulze Johann K, Steffen CR, Syndercombe-Court D, Trautmann M, van den Berge M, van der Gaag KJ, Vannier J, Verdoliva V, Vidaki A, Xavier C, Ballantyne J and Haas C (2018) Body fluid identification using a targeted mRNA massively parallel sequencing approach - results of a EUROFORGEN/EDNAP collaborative exercise. *Forensic Sci Int Genet* **34**:105-115.

Jakubowska J, Maciejewska A, Pawlowski R and Bielawski KP (2013) mRNA profiling for vaginal fluid and menstrual blood identification. *Forensic Sci Int Genet* **7**:272-278.

Jung SE, Cho S, Antunes J, Gomes I, Uchimoto ML, Oh YN, Di Giacomo L, Schneider PM, Park MS, van der Meer D, Williams G, McCord B, Ahn HJ, Choi DH, Lee YH, Lee SD and Lee HY (2016) A collaborative exercise on DNA methylation based body fluid typing. *Electrophoresis* **37**:2759-2766.

Juusola J and Ballantyne J (2005) Multiplex mRNA profiling for the identification of body fluids. *Forensic Sci Int* **152**:1-12

Kohlmeier F and Schneider PM (2012) Successful mRNA profiling of 23 years old blood stains. *Forensic Sci Int Genet* **6**:274-276.

Kulstein G, Schacker U and Wiegand P (2018) 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 Sci Int Genet* **36**:130-140.

Lindenbergh A, de Pagter M, Ramdayal G, Visser M, Zubakov D, Kayser M and Sijen T (2012) A multiplex (m)RNA-profiling system for the forensic identification of body fluids and contact traces. *Forensic Sci Int Genet* **6**:565-577.

Magalhaes T, Dinis-Oliveira RJ, Silva B, Corte-Real F and Nuno Vieira D (2015) Biological Evidence Management for DNA Analysis in Cases of Sexual Assault. *ScientificWorldJournal* **2015**:365674.

Matull WR, Andreola F, Loh A, Adiguzel Z, Deheragoda M, Qureshi U, Batra SK, Swallow DM and Pereira SP (2008) MUC4 and MUC5AC are highly specific tumour-associated mucins in biliary tract cancer. *Br J Cancer* **98**:1675-1681.

Oliva R (2006) Protamines and male infertility. *Hum Reprod Update* **12**:417-435.

OMS (2012) Prevenção da violência sexual e da violência pelo parceiro íntimo contra a mulher: ação e produção de evidência.

Park JL, Kwon OH, Kim JH, Yoo HS, Lee HC, Woo KM, Kim SY, Lee SH and Kim YS (2014) Identification of body fluid-specific DNA methylation markers for use in forensic science. *Forensic Sci Int Genet* **13**:147-153.

Park SM, Park SY, Kim JH, Kang TW, Park JL, Woo KM, Kim JS, Lee HC, Kim SY and Lee SH (2013) Genome-wide mRNA profiling and multiplex quantitative RT-PCR for forensic body fluid identification. *Forensic Sci Int Genet* **7**:143-150.

Pelley JW (2007) RNA Transcription and Control of Gene Expression, in *Biochemistry* pp 137-147.

Pinheiro MdFT (2008) *CSI Criminal*.

Pinheiro MF (2015) <*Princípios de Genética Forense.pdf*>.

Richard ML, Harper KA, Craig RL, Onorato AJ, Robertson JM and Donfack J (2012) Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis. *Forensic Sci Int Genet* **6**:452-460.

Roeder AD and Haas C (2013) mRNA profiling using a minimum of five mRNA markers per body fluid and a novel scoring method for body fluid identification. *Int J Legal Med* **127**:707-721.

Romero-Montoya L, Martínez-Rodríguez H, Pérez MA and Argüello-García R (2011) Relationship of spermatoscopy, prostatic acid phosphatase activity and prostate-specific antigen (p30) assays with further DNA typing in forensic samples from rape cases. *Forensic Sci Int* **206**:111-118.

Sauer E, Reinke AK and Courts C (2016) Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR. *Forensic Sci Int Genet* **22**:89-99.

Setzer M, Juusola J and Ballantyne J (2008) Recovery and stability of RNA in vaginal swabs and blood, semen, and saliva stains. *J Forensic Sci* **53**:296-305.

Shojaie N, Ghaffari SM, Ghaffari MS and Safari Z (2014) Simultaneous Extraction of DNA and RNA from Animal Cells by a Modified Laemmli Buffer. *Protocol Exchange*.

Sijen T (2015) Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers. *Forensic Sci Int Genet* **18**:21-32.

Sikirzhytskaya A, Sikirzhytski V and Lednev IK (2012) Raman spectroscopic signature of vaginal fluid and its potential application in forensic body fluid identification. *Forensic Sci Int* **216**:44-48.

Sirker M, Fimmers R, Schneider PM and Gomes I (2017) Evaluating the forensic application of 19 target microRNAs as biomarkers in body fluid and tissue identification. *Forensic Sci Int Genet* **27**:41-49.

Sirker M, Schneider PM and Gomes I (2016) A 17-month time course study of human RNA and DNA degradation in body fluids under dry and humid environmental conditions. *Int J Legal Med* **130**:1431-1438.

Song F, Luo H and Hou Y (2015) Developed and evaluated a multiplex mRNA profiling system for body fluid identification in Chinese Han population. *J Forensic Leg Med* **35**:73-80.

Sorbye SW, Fismen S, Gutteberg TJ, Mortensen ES and Skjeldestad FE (2016) Primary cervical cancer screening with an HPV mRNA test: a prospective cohort study. *BMJ Open* **6**:e011981.

Suttipasit P and Wongwittayapanich S (2018) Detection of prostate specific antigen and semenogelin in specimens from female rape victims. *Journal of Forensic and Legal Medicine* **54**:102-108.

Takamura A, Watanabe K and Akutsu T (2017) Advanced forensic validation for human spermatozoa identification using SPERM HY-LITER Express with quantitative image analysis. *Int J Legal Med* **131**:933-939.

Tanioka H, Yokoi N, Komuro A, Shimamoto T, Kawasaki S, Matsuda A and Kinoshita S (2009) Investigation of the corneal filament in filamentary keratitis. *Invest Ophthalmol Vis Sci* **50**:3696-3702.

Tian Y, Li L, Zhang F and Xu J (2016) Seminal plasma HSPA2 mRNA content is associated with semen quality. *J Assist Reprod Genet* **33**:1079-1084.

van den Berge M, Carracedo A, Gomes I, Graham EA, Haas C, Hjort B, Hoff-Olsen P, Maronas O, Mevag B, Morling N, Niederstatter H, Parson W, Schneider PM, Court DS,

Vidaki A and Sijen T (2014) A collaborative European exercise on mRNA-based body fluid/skin typing and interpretation of DNA and RNA results. *Forensic Sci Int Genet* **10**:40-48.

Vennemann M and Koppelkamm A (2010) mRNA profiling in forensic genetics I: Possibilities and limitations. *Forensic Sci Int* **203**:71-75.

Vennemann M, Scott G, Curran L, Bittner F and Tobe SS (2014) Sensitivity and specificity of presumptive tests for blood, saliva and semen. *Forensic Sci Med Pathol* **10**:69-75.

Virkler K and Lednev IK (2009) Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene. *Forensic Sci Int* **188**:1-17.

Westring CG, Wiuf M, Nielsen SJ, Fogleman JC, Old JB, Lenz C, Reich KA and Morling N (2014) SPERM HY-LITER for the identification of spermatozoa from sexual assault evidence. Forensic Sci Int Genet 12:161-167.

Zhou QZ, Guo XB, Zhang WS, Zhou JH, Yang C, Bian J, Chen MK, Guo WB, Wang P, Qi T, Wang CY, Yang JK and Liu CD (2019) Expressions of miR-525-3p and its target gene SEMG1 in the spermatozoa of patients with asthenozoospermia. *Andrology* **7**:220-227.

ATTACHMENT 1

INFORMAÇÃO AO PARTICIPANTE

Identificação da natureza de fluídos biológicos provenientes de amostras colhidas no âmbito de Agressões Sexuais através de mRNA

Este documento contém informação sobre o estudo a ser realizado no âmbito da Tese de Mestrado de Ciências Forenses de Sara Duarte Loureiro, estudante da Universidade do Porto, intitulado "Identificação da natureza de fluídos biológicos provenientes de amostras colhidas no âmbito de Agressão Sexual através da análise de mRNA". Para a sua realização é necessária a colaboração de voluntários do sexo masculino e do sexo feminino. Após a leitura da folha de 'Informação ao Participante', na permanência de quaisquer dúvidas ou necessidade de mais esclarecimentos, esteja à vontade para colocar as suas questões.

Informação sobre o Estudo

Em investigações forenses, nomeadamente em agressões sexuais, é necessária a deteção e determinação da natureza biológica de manchas de fluídos, uma vez que estes possuem extrema importância na resolução jurídica dos casos, nomeadamente na reconstrução das circunstâncias do crime. Muitas vezes a identificação e análise destes vestígios tornam-se um obstáculo devido às limitações dos métodos convencionais comummente utilizados. Desta forma, os métodos moleculares, nomeadamente a análise da expressão génica diferencial, através de RNA mensageiro (mRNA), têm sido distinguidos como técnicas emergentes.

O presente estudo "Identificação da natureza de fluídos biológicos provenientes de amostras de Agressão Sexual através da análise de mRNA" tem como objetivo comprovar a eficácia de mRNA como biomarcador forense analisando as diferenças de expressão génica destas moléculas em amostras de sémen e secreções vaginais.

O que irei fazer se aceitar participar?

Será pedido aos participantes do sexo masculino, uma amostra de sémen que será recolhida pelo próprio num recipiente esterilizado e próprio para o efeito. Será igualmente pedido aos elementos do sexo feminino três amostras de secreção vaginal colhidas pela própria com recurso a uma zaragatoa devidamente esterilizados.

Sou obrigado(a) a participar?

A participação neste estudo é voluntária e não será remunerada.

Os meus dados serão confidenciais?

Os dados recolhidos destinam-se a ser usados exclusivamente nesta investigação e serão tratados de forma anónima. Não terá qualquer identificação pessoal, mantendo, desta forma, a máxima confidencialidade dos seus dados pessoais. Também a informação obtida não será utilizada para quaisquer outros fins. O material biológico colhido será destruído após o estudo e nunca será usado para qualquer outra finalidade. Este estudo poderá ser publicado em Revistas Científicas ou divulgado em Congressos, mas sempre sem revelar o nome dos participantes.

A participação no estudo poderá trazer riscos ou benefícios?

A sua participação no estudo não lhe trará nem riscos, nem benefícios diretos; todavia, a análise das amostras recolhidas, poderá ajudar a compreender melhor as novas metodologias, contribuindo para aperfeiçoar a identificação da natureza de vestígios biológicos forenses. Este conhecimento poderá, posteriormente, ser implementado em laboratórios de âmbito Forense Portugueses.

Se decidir participar, ser-lhe-á entregue esta folha de 'Informação ao Participante' para guardar e ser-lhe-á pedido que assine a declaração de Consentimento Informado.

Para mais informações ou esclarecimentos não hesite em contactar através do seguinte email: saraduarteloureiro@gmail.com

Muito obrigada pela atenção.

Data: ____/____/____

A Investigadora: _____

ATTACHMENT 2

INFORMAÇÃO SOBRE O PROCEDIMENTO DE COLHEITA (HOMENS)

É pedido aos participantes do sexo masculino, uma amostra de sémen que será recolhida pelo próprio num recipiente esterilizado e próprio para o efeito. O sémen deverá advir de uma ejaculação direta sem que antes tenha havido qualquer contacto sexual. As amostras de sémen deverão ser armazenadas no congelador até serem entregues ao investigador

INFORMAÇÃO SOBRE O PROCEDIMENTO DE COLHEITA (MULHERES)

Pede-se às voluntárias do estudo que introduzam a zaragatoa cerca de 1 cm no canal vaginal. De seguida, devem girar delicadamente cerca de 360º. Posteriormente, retirar a zaragatoa colocá-la na embalagem parcialmente aberta. Estas amostras devem permanecer num período de secagem cerca de 1h30m num local resguardado sem exposição solar direta e de seguida, colocadas no congelador até serem entregues ao investigador.

O processo deve ser repetido mais duas vezes de modo a entregar 3 amostras de secreção vaginal ao investigador.

Pede-se igualmente, que preencham a folha anexa com indicação da toma ou não de algum método contracetivo hormonal e com o dia do ciclo hormonal correspondente ao dia da colheira.

FORMULÁRIO PARA PARTICIPANTE FEMININO

Código: _____

Método contracetivo hormonal: S ☐ N ☐ Qual? _____

Fase do ciclo hormonal: Menstruação (1º-5º dia)

Fase Folicular (até 13º dia) ☐

Fase Ovulatória (14º-17º dia) ☐

Fase Luteínica (18º-28º dia) ☐

ATTACHMENT 3

DECLARAÇÃO DE CONSENTIMENTO

Considerando a "Declaração de Helsínquia" da Associação Médica Mundial

(Helsínquia 1964; Tóquio 1975; Veneza 1983; Hong Kong 1989; Somerset West 1996 e Edimburgo 2000)

Identificação da natureza de fluídos biológicos provenientes de amostras colhidas no âmbito de Agressões Sexuais através de mRNA

Eu, abaixo-assinado, _____,
tomei conhecimento do estudo em que serei incluído(a) e compreendi a explicação que me foi fornecida acerca da investigação que se tenciona realizar. Foi-me ainda dada oportunidade de fazer as perguntas que julguei necessárias e de todas obtive resposta satisfatória.

Foi-me dado a conhecer que, de acordo com as recomendações da Declaração de Helsínquia, a informação ou explicação que me foi prestada versou os objetivos, os métodos, os benefícios previstos, os riscos potenciais e o eventual desconforto da investigação em curso.

Foi-me ainda explicado que os registos dos resultados poderão ser consultados pelos responsáveis científicos e ser objeto de publicação, mas que os elementos da identidade pessoal serão sempre tratados de modo estritamente confidencial, uma vez que apenas o investigador principal terá acesso ao documento onde se encontram as concordâncias entre o código dado à amostra e os dados dos participantes.

Também me foi esclarecido que o material biológico colhido será destruído após o estudo e nunca será usado para qualquer outra finalidade. Por fim, foi-me afirmado que tenho o direito de recusar a todo o tempo a minha participação no estudo, sem que isso possa ter como efeito qualquer prejuízo.

Aceito participar de livre vontade no estudo acima mencionado.

Concordo que seja efetuada a colheita de amostras biológicas para realizar as análises e os estudos genéticos que fazem parte desta investigação.

Também consinto a divulgação dos resultados obtidos no meio científico, desde que seja garantido o seu anonimato.

Data: ____ / _____ / 201___

Assinatura do voluntário: **_____**

O Investigador responsável: **_____**

ATTACHMENT 4

Suplementary Table 1. *Quantification results by Nanodrop spectopgotometer for DNA and RNA co-extracted from semen and vaginal secretions samples.*

	DNA			RNA		
Sample ID	ng/µL	260/280	260/230	ng/µl	260/280	260/230
Semen - Commercial Kit						
S1	1.72	1.59	0.02	14.99	2.47	0.07
S2	1.30	1.79	0.26	3.99	6.22	0.17
S3	1.54	1.02	0.43	9.37	4.68	0.05
S4	2.60	1.77	0.09	4.43	*n.d.*	0.06
S5*	31.04	1.97	1.33	35.25	1.25	0.83
S6*	45.31	1.05	0.14	104.95	1.87	2.36
S7*	66.12	1.83	1.35	140.23	1.89	1.02
S8*	35.81	2.04	2.14	39.83	1.96	0.92
S9*	19.91	1.73	1.44	12.13	2.21	0.21
S10*	13.63	1.89	1.08	6.04	2.06	0.12
S11*	37,72	1.68	1.62	56.58	1.93	0.96
S12*	45.79	1.86	1.77	62.38	1.93	1.50
S13*	84.14	1.84	2.39	52.23	1.88	0.97
S14*	80.12	1.77	1.43	160.99	1.89	1.38
Semen - Manual Method						
S1 M	4092.00	0.93	0.18	212.6	0.86	0.87
S2 M	213.39	1.73	1.3	87.95	0.74	0-27
S3 M	26.6	1.78	0.97	44.11	0.85	0.20
S4 M	73.28	0.96	0.21	112.3	0.93	0.36
S5 M	46.41	0.91	0.19	323.11	0.87	1.98
S6 M	103.65	1.14	0.32	265.45	0.94	1.17
S7 M	50.58	1.18	0.31	143.68	0.99	0.45
S8 M	15.07	1.11	0.35	45.32	0.82	0.20
S9 M	7.09	0.88	0.26	83.59	0.88	0.27
S10 M	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*
S11 M	49.86	1.28	0.56	164.17	0.88	0.51
S12 M	14.12	0.95	0.2	167.53	0.91	0.52
S13 M	12.68	1.04	0.48	98.14	0.78	0.29
S14 M	32.10	1.24	0.36	156.62	0.86	0.48
Semen - Purifed Manual Method						
S4 MP	28.83	1.67	0.43	12.25	1.45	0.49
S5 MP	72.10	1.95	0.97	5.94	1.36	*n.d.*
S6 MP	163.08	1.76	1.53	37.44	1.66	0.84
S7 MP	48.88	1.46	1.86	62.98	1.62	0.55
Vaginal secretions						
VS1	13.00	1.64	1.63	26.78	1.75	0.6
VS2	12.18	1.53	0.06	18.71	2.15	0.18
VS3	12.23	1.38	0.11	12.42	2.18	0.06
VS4	8.93	1.11	0.25	2.49	1.71	0.01
VS5	15.56	1.37	0.33	5.77	1.37	0.17
VS6	16.95	2.02	1.19	2.02	1.7	0.01
VS7	7.44	1.36	0.06	15.61	1.73	0.19
VS8	14.10	1.69	0.25	33.28	2.15	0.50
VS9	12.38	1.82	0.43	33.79	2.12	0.10
VS10	19.33	1.57	0.29	7.95	1.98	0.03

*Optimized assay

n.a.- not analyzed

n.d.- not determined

ATTACHMENT 5

Supplementary Table 2. *Information about the hormonal cycles and the intake of birth control medication of female volunteers.*

Code	Birth control method	Hormonal cycle phase
VS1	Pill	Luteal
VS2	Pill	Follicular
VS3	None	Menopause
VS4	None	Follicular
VS5	None	Follicular
VS6	Pill	Follicular
VS7	*n.a.*	*n.a.*
VS8	Pill	Follicular
VS9	Pill	Follicular
VS10	*n.a.*	*n.a.*

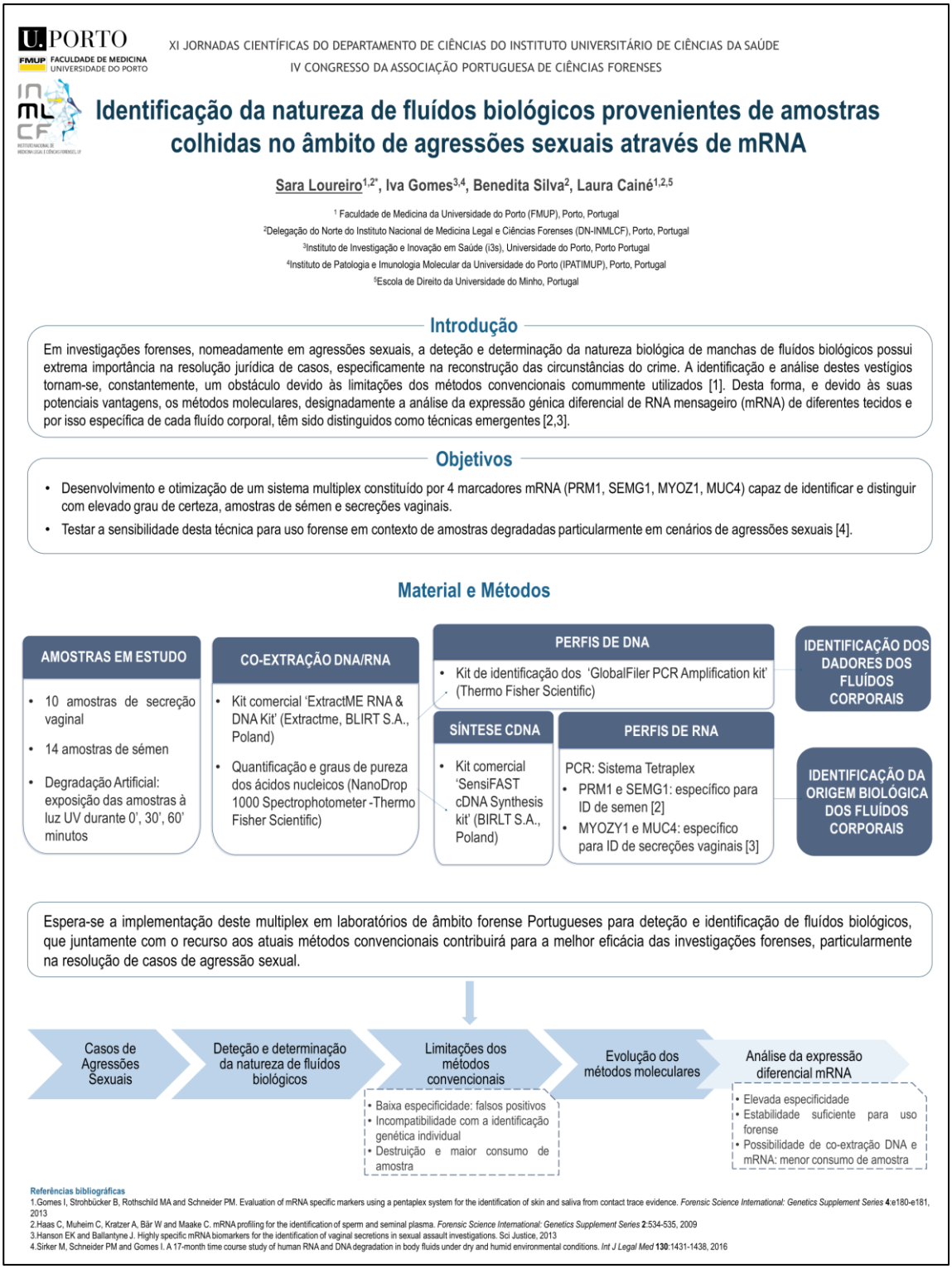
n.a. - no answer

Supplementary Table 3. *DNA profiles. (-) means no markers amplification.*

Sample ID	D3S1348	vWA	D16S539	CSF1PO	TIPOX	AMELOGENINA	D8S1179	D21S11	D18S51	DYS391	D2S441	D19S433	TH01	FGA	D22S1045	D5S818	D13S317	D7S820	SE33	D10S1248	D1S1656	D12S391	D2S1338
Semen - Commercial Kit																							
S1	15, 18	16, 18	9	10, 12	8	2 XY	11, 13	28, 3	13, 19	10	10, 11	14, 16	8; 9.3	19, 22	16	12	8	10, 11	14; 29,2	13, 14	16	22, 23	17
S2	16	14, 19	12	10, 11	9	2 XY	10, 12	28, 29	12, 14	11	11, 11.3	14, 15.2	6, 9.3	21	11, 16	11, 12	11, 12	9, 11	19, 30.2	13, 15	15, 17	19.3, 22	17, 20
S3	16, 17	16	9, 12	11, 12	8	2 XY	13, 14	30, 31.2	12, 15	11	10, 15	14	9, 9.3	21, 22.2	15, 16	10, 12	10, 11	8, 9	20, 28.2	13, 17	15, 17.3	18, 20	20, 24
S4	16, 18	15, 19	12, 13	10, 11	8, 11	2 XY	12, 14	31.2, 32.2	17	11	11, 14	13	6	19, 23	16, 17	12	8, 12	10, 12	18.2, 29.2	12,15	12, 15	17, 24	18, 23
S5	16	16, 18	11, 12	10, 15	8	2 XY	14	27, 30.2	13, 14	11	11.3, 14	12, 13	6, 9.3	20, 21	16, 17	12, 14	11	9, 10	21, 28.2	14, 15	14, 18.3	18	18, 23
S6	15	18, 19	11, 12	9, 10	8, 10	2 XY	10, 15	29, 31.2	10, 18	11	10, 14	13, 14	8, 9	20, 24	15	11, 12	11, 12	8, 10	28.2, 29.2	14, 16	14.3, 16	15, 18	20, 24
S7	16, 17	17	11, 14	11, 12	8, 11	2 XY	11, 13	29, 31	13, 16	10	11	13, 15	8, 9.3	20, 22	16, 17	11, 12	11, 12	11	31.2, 32.2	13, 15	16, 17	19, 20	17, 24
S8	14, 18	15, 18	12, 13	10	11	2 XY	12, 13	29, 31.2	13, 19	11	11.3, 14	13, 14	9, 9.3	20	15	11, 12	11, 12	11	15, 26.2	13, 14	14, 16	15, 18	17, 26
S9	17, 18	17, 18	12, 14	10, 12	9, 10	2 XY	13, 14	28, 29	12, 13	11	10, 14	14	8, 9.3	21, 22	11, 16	12	10, 12	10,12	22.2, 27.2	14, 17	15, 15.3	19, 21	19, 20
S10	15, 17	15, 17	9, 11	11	9, 11	2 XY	12, 14	29	17, 18	10	10	14, 15	9.3	22, 22.2	14, 16	12	11, 13	11, 13	29.2, 31.2	14	12, 15	18.3, 23	24, 25
S11	16, 18	16, 20	12, 13	11	8	2XY	13, 14	29, 31	12, 15	11	11	13, 15	9.3	24, 25	15	10, 12	8, 11	8, 11	18, 30.2	14	15, 15.3	19, 22	17, 19
S12	14, 15	15	11, 13	10, 11	8, 10	2XY	13, 15	30, 31.2	14, 15	10	11	14	6, 9.3	19, 21	15, 16	10, 13	11	10, 11	30.2, 31.2	13, 14	12, 17	16, 20	17, 25
S13	14, 17	15, 19	9, 11	11	8	2XY	13, 14	27, 30.2	14, 18	11	14	13, 14	9, 9.3	21, 22	11, 16	9, 12	9, 12	8, 11	17, 22.2	13, 15	12	16, 18	17, 18
S14	14, 15	16	11	10, 12	8	2XY	12, 15	27, 30.2	12, 13	9	11	13, 14	9.3	20, 26	16	13	8, 12	12, 14	18, 31.2	13, 14	14, 16.3	20, 24	19, 20
Semen - Manual Method																							
S1 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S2 M	16	14, 19	12	10, 11	9	2 XY	10, 12	28, 29	12, 14	11	11, 11.3	14, 15.2	6, 9.3	21	11, 16	11, 12	11, 12	9, 11	19, 30.2	13, 15	15, 17	19.3, 22	17, 20
S3 M	16, 17	16	9, 12	11, 12	8	2 XY	13, 14	30, 31.2	12, 15	11	10, 15	14	9, 9.3	21, 22.2	15, 16	10, 12	10, 11	8, 9	20, 28.2	13, 17	15, 17.3	18, 20	20, 24
S4 M	–	–	–	–	–	–	–	31.2, 32.2	17	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S5 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S6 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S7 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S8 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S9 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S11 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S12 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S13 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S14 M	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Semen - Manual Purified Method																							
S1 MP	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S2 MP	16	14, 19	12	10, 11	9	2 XY	10, 12	28, 29	12, 14	11	11, 11.3	14, 15.2	6, 9.3	21	11, 16	11, 12	11, 12	9, 11	19, 30.2	13, 15	15, 17	19.3, 22	17, 20
S3 M	16, 17	16	9, 12	11, 12	8	2 XY	13, 14	30, 31.2	12, 15	11	10, 15	14	9, 9.3	21, 22.2	15, 16	10, 12	10, 11	8, 9	20, 28.2	13, 17	15, 17.3	18, 20	20, 24
S4 M	–	–	–	–	–	–	–	31.2, 32.2	17	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S5 MP	16	16, 18	11, 12	10, 15	8	2 XY	14	27, 30.2	13, 14	11	11.3, 14	12, 13	6, 9.3	20, 21	16, 17	12, 14	11	9, 10	21, 28.2	14, 15	14, 18.3	18	18, 23
S6 MP	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S7 MP	16, 17	17	11, 14	11, 12	8, 11	2 XY	11, 13	29, 31	13, 16	10	11	13, 15	8, 9.3	20, 22	16, 17	11, 12	11, 12	11	31.2, 32.2	13, 15	16, 17	19, 20	17, 24
S8 MP	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S9 MP	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S11 MP	16, 18	16, 20	12, 13	11	8	2XY	13, 14	29, 31	12, 15	11	11	13, 15	9.3	24, 25	15	10, 12	8, 11	8, 11	18, 30.2	14	15, 15.3	19, 22	17, 19
S12 MP	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
S13 MP	–	–	–	–	–	2XY	13, 14	30.2	14		14	13, 14	9, 9.3	–	–	–	–	–	–	15	12	16	18
S14 MP	14, 15	16	11	10, 12	8	2XY	12, 15	27, 30.2	12, 13	9	11	13, 14	9.3	20, 26	16	13	8, 12	12, 14	18, 31.2	13, 14	14, 16.3	20, 24	19, 20
Vaginal Secretions - commercial Kit																							
VS1	14, 17	14, 16	10, 11	11, 12	8, 11	XX	11, 13	29, 32.2	13, 18		14, 15	14	6, 8	20, 24	11, 15	12	11	10, 11	17, 21	14, 15	12, 17.3	19, 24	17, 24
VS2	15, 16	14,16	11, 13	10, 11	8, 11	XX	12, 14	28, 29	13		14, 15	13, 14	6,9	20,22	15, 16	12	11, 14	10, 12	14, 21	14, 15	12, 17.3	19, 20	17
VS3	16	16, 17	11,12	10, 11	8, 9	XX	13, 14	31, 31.2	15		10, 11	14	7, 9.3	23, 25	11, 17	11, 12	12, 13	9, 12	17, 28.2	14	14, 17.3	15, 17	17, 22
VS4	14, 16	14, 18	10, 13	10, 12	8	XX	11, 14	28, 32.2	13		11, 15	13, 14	8, 9	22, 24	15	12	11, 14	11, 12	17, 21	15	17.3	20, 24	17, 24
VS5	17, 19	16, 18	11	10, 11	8, 11	XX	13, 14	29, 30	14, 21		11.3, 14	12, 14	6, 9	20, 24	16	11, 12	11	8, 10	15, 29.2	12	16, 16.3	17.3, 18	18, 19
VS6	14, 16	15, 19	9, 12	10, 11	8	XX	12	29	13, 14		11.3, 14	14	6, 8	23	16, 17	9, 12	9, 10	9, 14	26.2, 27.2	15, 16	15, 15.3	18, 19	17, 25
VS7	16, 17	16	11, 12	12, 13	11, 12	XX	12, 14	29, 30	12, 13		11, 11.3	13, 16	6, 8	20, 21	16	11, 13	9, 11	10	13, 15	13	16, 16.3	17, 19.3	16, 23
VS8	16, 17	15, 17	11, 12	10, 12	9	XX	13, 14	30, 31	12, 15		10, 11	10, 14	7, 9.3	25	11, 15	11	12, 13	11, 12	17	13, 14	16, 17.3	15, 23	22, 24
VS9	15, 16	17, 19	11	11, 13	11	XX	11, 13	30, 31.2	14, 15		11	12, 16	6, 9	25, 26	15, 16	12, 13	11, 13	10, 12	29.2, 31.2	16	12, 16.3	18, 25	18, 19
VS10	15, 18	17, 19	13	11, 12	8	XX	10, 11	28, 33.2	10, 15		14, 15	13	6, 9	21, 23	16, 17	12	9, 12	10, 12	12, 22.2	11, 13	13, 17.3	18, 21	18, 25
Mixtures - Commercial Kit																							
Mix 1	16, 17	16, 17	11, 12, 14	10, 11, 12	8, 9, 11	2 XY	11, 13, 14	29, 31, 31.2	13, 15, 16	10	10, 11	13, 14, 15	7, 8, 9.3	20, 22, 23, 25	11, 16, 17, 18	11, 12	11, 12, 13	9, 11, 12	17, 28.2, 31.2, 32.2	13, 14, 15	14, 16, 17	15, 17, 19, 20	17, 22, 24
Mix 2	16, 17	16, 17	11, 12, 14	10, 11, 12	8, 9, 11	2 XY	11, 13, 14	29, 31, 31.2	13, 15, 16	10	10, 11	13, 14, 15	7, 8, 9.3	20, 22, 23, 25	11, 16, 17, 18	11, 12	11, 12, 13	9, 11, 12	17, 28.2, 31.2, 32.2	13, 14, 15	13.3, 16, 17	15, 17, 19, 20	17, 22, 24
Mix 3	16, 17	16, 17	11, 12, 14	10, 11, 12	8, 9, 11	2 XY	11, 13, 14	29, 31, 31.2	13, 15, 16	10	10, 11	13, 14, 15	7, 8, 9.3	20, 22, 23, 25	11, 16, 17, 18	11, 12	11, 12, 13	9, 11, 12	17, 28.2, 31.2, 32.2	13, 14, 15	14, 16, 17	15, 17, 19, 20	17, 22, 24
Mix 4	16, 17	16, 17	11, 12, 14	10, 11, 12	8, 9, 11	2 XY	11, 13, 14	29, 31, 31.2	13, 15, 16	10	10, 11	13, 14, 15	7, 8, 9.3	20, 22, 23, 25	11, 16, 17, 18	11, 12	11, 12, 13	9, 11, 12	17, 28.2, 31.2, 32.2	13, 14, 15	14, 16, 17	15, 17, 19, 20	17, 22, 24

ATTACHMENT 7

Poster presented at the 4th Congress of the Associação Portuguesa de Ciências Forenses (APCF, 2019).



U.PORTO
FMUP FACULDADE DE MEDICINA UNIVERSIDADE DO PORTO

XI JORNADAS CIENTÍFICAS DO DEPARTAMENTO DE CIÊNCIAS DO INSTITUTO UNIVERSITÁRIO DE CIÊNCIAS DA SAÚDE
IV CONGRESSO DA ASSOCIAÇÃO PORTUGUESA DE CIÊNCIAS FORENSES

Identificação da natureza de fluídos biológicos provenientes de amostras colhidas no âmbito de agressões sexuais através de mRNA

Sara Loureiro[1,2*], Iva Gomes[3,4], Benedita Silva[2], Laura Cainé[1,2,5]

[1] Faculdade de Medicina da Universidade do Porto (FMUP), Porto, Portugal
[2]Delegação do Norte do Instituto Nacional de Medicina Legal e Ciências Forenses (DN-INMLCF), Porto, Portugal
[3]Instituto de Investigação e Inovação em Saúde (i3s), Universidade do Porto, Porto Portugal
[4]Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Porto, Portugal
[5]Escola de Direito da Universidade do Minho, Portugal

Introdução

Em investigações forenses, nomeadamente em agressões sexuais, a deteção e determinação da natureza biológica de manchas de fluídos biológicos possui extrema importância na resolução jurídica de casos, especificamente na reconstrução das circunstâncias do crime. A identificação e análise destes vestígios tornam-se, constantemente, um obstáculo devido às limitações dos métodos convencionais comummente utilizados [1]. Desta forma, e devido às suas potenciais vantagens, os métodos moleculares, designadamente a análise da expressão génica diferencial de RNA mensageiro (mRNA) de diferentes tecidos e por isso específica de cada fluído corporal, têm sido distinguidos como técnicas emergentes [2,3].

Objetivos

- Desenvolvimento e otimização de um sistema multiplex constituído por 4 marcadores mRNA (PRM1, SEMG1, MYOZ1, MUC4) capaz de identificar e distinguir com elevado grau de certeza, amostras de sémen e secreções vaginais.
- Testar a sensibilidade desta técnica para uso forense em contexto de amostras degradadas particularmente em cenários de agressões sexuais [4].

Material e Métodos

AMOSTRAS EM ESTUDO
- 10 amostras de secreção vaginal
- 14 amostras de sémen
- Degradação Artificial: exposição das amostras à luz UV durante 0', 30', 60' minutos

CO-EXTRAÇÃO DNA/RNA
- Kit comercial 'ExtractME RNA & DNA Kit' (Extractme, BLIRT S.A., Poland)
- Quantificação e graus de pureza dos ácidos nucleicos (NanoDrop 1000 Spectrophotometer -Thermo Fisher Scientific)

PERFIS DE DNA
- Kit de identificação dos 'GlobalFiler PCR Amplification kit' (Thermo Fisher Scientific)

IDENTIFICAÇÃO DOS DADORES DOS FLUÍDOS CORPORAIS

SÍNTESE CDNA
- Kit comercial 'SensiFAST cDNA Synthesis kit' (BIRLT S.A., Poland)

PERFIS DE RNA

PCR: Sistema Tetraplex
- PRM1 e SEMG1: específico para ID de semen [2]
- MYOZY1 e MUC4: específico para ID de secreções vaginais [3]

IDENTIFICAÇÃO DA ORIGEM BIOLÓGICA DOS FLUÍDOS CORPORAIS

Espera-se a implementação deste multiplex em laboratórios de âmbito forense Portugueses para deteção e identificação de fluídos biológicos, que juntamente com o recurso aos atuais métodos convencionais contribuirá para a melhor eficácia das investigações forenses, particularmente na resolução de casos de agressão sexual.

Casos de Agressões Sexuais → Deteção e determinação da natureza de fluídos biológicos → Limitações dos métodos convencionais → Evolução dos métodos moleculares → Análise da expressão diferencial mRNA

Limitações dos métodos convencionais:
- Baixa especificidade: falsos positivos
- Incompatibilidade com a identificação genética individual
- Destruição e maior consumo de amostra

Análise da expressão diferencial mRNA:
- Elevada especificidade
- Estabilidade suficiente para uso forense
- Possibilidade de co-extração DNA e mRNA: menor consumo de amostra

Referências bibliográficas

1.Gomes I, Strohbücker B, Rothschild MA and Schneider PM. Evaluation of mRNA specific markers using a pentaplex system for the identification of skin and saliva from contact trace evidence. *Forensic Science International: Genetics Supplement Series* **4**:e180-e181, 2013
2.Haas C, Muheim C, Kratzer A, Bär W and Maake C. mRNA profiling for the identification of sperm and seminal plasma. *Forensic Science International: Genetics Supplement Series* **2**:534-535, 2009
3.Hanson EK and Ballantyne J. Highly specific mRNA biomarkers for the identification of vaginal secretions in sexual assault investigations. Sci Justice, 2013
4.Sirker M, Schneider PM and Gomes I. A 17-month time course study of human RNA and DNA degradation in body fluids under dry and humid environmental conditions. *Int J Legal Med* **130**:1431-1438, 2016

ATTACHMENT 8

Submitted abstract that will lead to a conference proceeding paper at the 26th Congress of the International Society for Forensic Genetics (ISFG, 2019, Prague).

Evaluation of two DNA/RNA co-extraction methods for body fluid identification in forensics

Loureiro S [1], Amorim A [2, 3, 4], Cainé L [1, 5,6], Silva B [5], Gomes I [2, 3]

[1] Department of Public Health,, Faculty of Medicine, University of Porto, Portugal.

[2] i3S, Institute for Research and Innovation in Health, Porto, Portugal.

[3] IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, Portugal.

[4] Department of Biology, Faculty of Sciences, University of Porto, Portugal.

[5] National Institute of Legal Medicine and Forensic Sciences, Porto, Portugal.

[6] Escola de Direito, Universidade do Minho, Portugal

Body fluid identification has become a field of interest in forensic casework as it can add value to particular investigative scenarios. Identifying the source of the biological material is not always an upfront task using conventional methods; therefore, profiling of specific mRNA markers can provide the answer. The implementation of RNA based analyses in forensic casework must focus on the quality and sensitivity of methods, starting with nucleic acid extraction, and without loss of DNA for STR profiling. In this work, two methods for DNA and RNA co-extraction were tested and compared. The kit ExtractME RNA & DNA Kit (Extractme, BLIRT S.A., Poland) and a quick, simple manual based protocol, adapted from Shojaie N. *et al*. [Protocol Exchange (2014), doi:10.1038/protex.2014.036], were used to simultaneously extract both types of nucleic acids from semen samples. As a first quality control (QC) check, DNA and RNA quantifications as well as purity assessment ratios (260/230 and 260/280) were estimated by Nanodrop 1000 Spectrophotometer (Thermo Fisher Scientific, USA). High quantities of both DNA and RNA were obtained by both methods allowing for the use of extracts in downstream applications such as PCR. When comparing the ratios of QC obtained by spectrophotometry readings, the commercial kit provided better results overall. Results suggest that the manual method allowed for more carryover of contaminants absorbing at 230nm and 280nm as the purity ratios were below the accepted range. Nevertheless, these ratios could be further improved by purification of both DNA and RNA. To finally asses, the use of the tested protocols in body fluid identification the gene expression profiles of two semen mRNA specific markers were compared between methods.

Poster presented at the 26th Congress of the International Society for Forensic Genetics (ISFG, 2019, Prague).

26th World Congress of the International Society for Forensic Genetics

Evaluation of two DNA/RNA co-extraction methods for body fluid identification in forensics

Loureiro S [1], Amorim A [2, 3, 4], Cainé L [1, 5,6], Silva B [5], Gomes I [2, 3]

[1] Department of Public Health, Forensic Sciences and Medical Education, Faculty of Medicine, University of Porto, Portugal. [2] i3S, Institute for Research and Innovation in Health, Porto, Portugal. [3] IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, Portugal. [4] Department of Biology, Faculty of Sciences, University of Porto, Portugal. [5] National Institute of Legal Medicine and Forensic Sciences, Porto, Portugal. [6] Escola de Direito, Universidade do Minho, Portugal

Introduction

In most scenarios of forensic investigations, as for example in cases of sexual aggressions, it may be necessary to detect and identify the biological nature of fluid stains. This is because they play an extremely important role in the resolution of legal cases, as for example, in the reconstruction of the crime circumstances. The identification and analysis of these biological traces are often very difficult due to the limitations of the conventional methods currently used. Due to its potential advantages, molecular methods, namely the analysis of differential gene expression patterns through messenger RNA (mRNA) of different tissues and, therefore, specific to each body fluid have been distinguished as emergent techniques (1). This research has been increasing over the past years focusing on the interest in identifying body fluids/tissues relevant for forensic applications (1-5).

Aim: **In this work, two methods for co-extraction of DNA and RNA were tested and compared. On one hand, a commercial kit that uses a spin, mini column based methodology, and on the other, a quick, simple isopropanol and nucleic acid ethanol precipitation based protocol. Both methods were used to extract nucleic acids from semen samples and the performance of methodologies were tested by mRNA profiling of two semen specific markers, PRM1 and SEGM1.**

Materials & Methods

Samples: 200 µl semen from healthy male volunteers (n=14).

Commercial co-extraction kit: ExtractME RNA & DNA Kit (BLIRT S.A., Poland) following the manufacturer instructions, with slight modifications to the protocol: addition of 20 µl of DTT 1mM and an incubation step of 30 min at 56ºC, in the lysis step.

Manual co-extraction protocol: adapted from Shojaie N. *et al.* (2) based on a modified Laemelli buffer, with addition of isopropanol to cell lysate, RNA and DNA precipitation, followed by purification of nucleic acids using absolute ethanol. Modifications were introduced to the lysis step: 20 µl DTT 1mM, 20 µl Prot K and incubation 30 min, 56ºC).

mRNA semen specific markers and primer references: PRM1 (3) and SEMG1 (4).

cDNA synthesis, end-point PCR and CE: SensiFAST cDNA Synthesis kit (Blirt S.A., Poland) following manufacturer's protocol and amplification conditions according to Gomes I. *et al.* (5). CE (capillary electrophoresis) for detection of PCR products on an ABI3500 platform (Applied Biosystems).

Nucleic acid quantifications and quality control ratios (QC): Nanodrop 1000 Spectrophotometer (ThermoFisher Scientific).

Results & Discussion

Nucleic Acid Concentrations: RNA concentrations were much higher for the majority of samples (12 out of 13) in extracts obtained with the non-commercial kit (Fig. 1). As for the DNA samples (Fig. 2) at least half of the extracts presented also higher concentrations for the "manual" protocol tested in this study.

QC Assessment of ExtractME DNA/RNA kit: the majority of extracts presented standard acceptable purity quality assessment ratios (260/280): ~2.0 for 9 RNA samples and ~1.8 for 11 out of the 14 DNA samples. As for the possible presence of inhibitors (260/230), only 4 samples (out of the total 28 DNA and RNA extracts) presented acceptable values of purity (~1.8-2.2). This means that this ratio was appreciably lower, indicating the presence of co-purified contaminants.

QC Assessment of Manual protocol: only 2 DNA samples presented purity acceptable values of ~1.8 (260/280). For the RNA extracts all samples presented values of 260/230 <1.0 again indicating the presence of inhibitors (i.e., co-purified contaminants) in all samples.

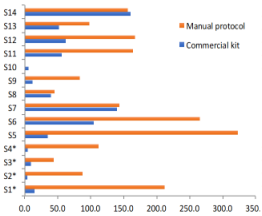


Fig. 1 **RNA concentrations** obtained for all semen samples from both methodologies (n=14 ExtractMe kit and n=13 manual protocol). *First tested assay - without DTT and no incubation.

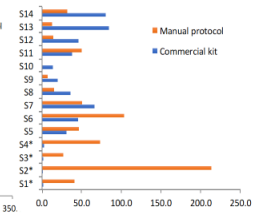


Fig. 2 **DNA concentrations** obtained for all semen samples from both methodologies (n=14 ExtractMe kit and n=14 manual protocol). *First tested assay - without DTT and no incubation.

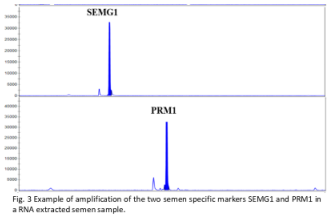


Fig. 3 Example of amplification of the two semen specific markers SEMG1 and PRM1 in a RNA extracted semen sample.

mRNA semen markers: PRM1 and SEGM1 markers were amplified in all RNA samples to further test the performance of each method. All semen extracts (n=14) obtained from the ExtractMe kit had positive expression for both mRNA markers (Fig. 3). For the non-commercial kit, no expression of the semen markers was detected in any of the samples (n=13). If considering only the RNA quantification data this result would be rather unexpected due to high concentrations estimated. However, when contemplating the QC assessment ratios, the lack of detection of expression in RNA extracts (manual protocol) could be due to a high concentration of contaminant carryover that probably inhibit PCR. In addition, one of the critical steps is the separation of DNA from RNA: in this method after isopropanol addition to lysate DNA precipitates forming a DNA coil which is then removed with, for e.g., a spatula. Most likely some DNA still remains in the RNA extract which would justify the high observed nucleic acid quantities absorbing at 260 nm.

The aim of this study was to compare the efficiencies of a column-based extraction commercial kit (ExtractMe, BLIRT S.A., Poland) and a simple, low cost manual based procedure (2). The objective was to test if the latter protocol would perform at the same level as a commercially developed and optimized kit. This would allow implementation of a straightforward and low cost DNA/RNA co-extraction protocol in any forensic laboratory but where many times, budget limitations can determine the implementation of RNA based methods. Also, the manual method (2) was developed for animal cell culture purification and therefore adjustments as well as refinements are needed to extract successfully DNA/RNA from body fluids or other tissue types, particularly relevant in forensics.

References

[1] Sijen T (2015) FSI Gen 18:21-32. [2] Shojaie N, *et al.* (2014) Protocol Exchange Nat Res. DOI: 10.1038/protex.2014.036. [3] Juusola J & Ballantyne J (2005) FSI 152:1-12. [4] Haas C, *et al.* (2009) FSI Gen Suppl Ser 2:534-535. [5] Gomes I, *et al.* (2013) FSI Gen Suppl Ser 4:e180-e181.

Acknowledgments: I.Gomes is funded by CEECIND/02609/2017.

SEDE ADMINISTRATIVA

FACULDADE DE **MEDICINA**

FACULDADE DE **CIÊNCIAS**

FACULDADE DE **DIREITO**

FACULDADE DE **FARMÁCIA**

FACULDADE DE **MEDICINA DENTÁRIA**

FACULDADE DE **PSICOLOGIA E DE CIÊNCIAS DA EDUCAÇÃO**

INSTITUTO DE **CIÊNCIAS BIOMÉDICAS ABEL SALAZAR**