



EVALUATING AN MRNA BASED BODY FLUID IDENTIFICATION TEST USING SYBR GREEN FLUORESCENT DYE AND REAL-TIME PCR

Jo-Ann CONNOLLY, Graham WILLIAMS

Forensic and Analytical Research Centre, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, United Kingdom

Abstract

The requirement to have more definitive and wider ranging body fluid identification (BFI) tests has resulted in a range of mRNA based real-time PCR BFI assays utilising Taqman fluorescent dye. An attempt to make a reliable BFI test utilising the alternative SYBR Green fluorescent dye was carried out.

Samples were extracted from blood and saliva stains and then reverse transcribed using M-MLV and random hexamers. Using real-time PCR, relative quantitation of SPTB (blood), NCF2 (blood), KRT4 (saliva) and SPRR1A (saliva) was carried out on cDNA from the blood and saliva samples using the SYBR Green reagents. Melting curve analysis was also conducted following the amplification step. The RQ values were calculated using the formula $2^{-\Delta\Delta CT}$.

In all cases, the blood markers (SPTB and NCF2) were under expressed in saliva and over expressed in blood. The saliva markers (KRT4 and SPRR1A) were over expressed in saliva and under expressed in blood. Verification of the amplicons was carried out using melting curve analysis.

Relatively high levels of fluorescence were also detected in the reverse transcription blanks and negative controls; despite stringent anti-contamination procedures. This along with the melting curve analysis results suggest that the amplification may be a by-product of the test itself rather than contamination; due to the lack of specificity of the SYBR Green fluorescent dye. This was verified by contamination monitoring using negative controls along with SYBR Green and Taqman fluorescent dyes.

This shows that a SYBR Green based BFI test could be developed and it may be more appropriate to use than a Taqman based BFI test in a commercial forensic science environment. However, the amplification detected in the reverse transcription blanks and negative controls may render the results of a SYBR Green based BFI test unreliable and insufficiently robust for use in a court of law.

Keywords: - Forensic Genetics, Body fluid identification, mRNA, SYBR Green, Real-Time PCR

Introduction

In forensic casework, mRNA gene expression analysis is a new method for the identification of body fluids (BFI) from biological stains and is the subject of research by a number of groups [1-6]. In theory, mRNA gene expression is more specific than current body fluid identification methods that involve the

enzymatic and immunological detection of enzymes and proteins [3-7]; although it is more expensive and time consuming.

There are currently no immunological or enzymatic tests for vaginal secretions and menstrual blood. Development of a body fluid test using mRNA expression could lead to differentiations being made between vaginal secretions and saliva and

menstrual blood and blood due to trauma which could be crucial in forensic case work relating to sexual offences.

RNA has long been thought to be relatively unstable due to the presence of ribonucleases. However, several studies indicate that this is not the case, [6, 8-10]. The study by Zubakov *et al* shows that mRNA is present in blood and saliva stains up to 16 and 6 years, respectively [6].

This need to have more definitive and wider ranging body fluid identification tests has resulted in a range of mRNA based real-time PCR assays [1, 3, 6, 11]. The majority of these assays are on Applied Biosystems (ABI) real-time PCR machines and utilise Taqman[®] probes rather than SYBR Green[®] fluorescent dye which is an alternative dye that can be used.

Taqman[®] reagents are generally preferred by researchers as they have advantages over SYBR Green[®] reagents. Taqman[®] primers are sequence specific and will only bind to and consequently amplify DNA if the complementary sequence is present. SYBR Green[®] will bind to any double stranded DNA present [12].

Another advantage of Taqman[®] over SYBR Green[®] is that Taqman[®] primers can be multiplexed. As the Taqman[®] probes can be labelled with fluorescent dyes, each probe can be labelled with a different dye so that in data analysis, the fluorescence of each dye can be individually measured, and therefore showing which primer is binding to the DNA. SYBR Green[®] binds to all dsDNA and there is only one variant, therefore is unsuitable for multiplexing [12].

However, SYBR Green[®] does have advantages over Taqman[®]; in that the reagents are cheaper and that the primers do not require labelling. Melt-curve analysis can also be carried out on SYBR green[®] but not on Taqman[®]. This allows

for diagnostics function in case of a failed reaction.

A number of groups have used SYBR Green as part of their studies and have found problems with it, as expected, and prefer to use Taqman probes. It is generally accepted that Taqman is a superior probe than SYBR Green; however, the use of a multiplexed Taqman body fluid identification kit is likely to be very expensive. If a SYBR Green based test is capable of distinguishing between body fluids, then this would be a more cost-effective alternative, which would be of benefit to forensic scientists as well as researchers.

The aim of this study was to assess whether a cost-effective SYBR Green[®] based real-time PCR body fluid identification test can be reliably developed for use by commercial forensic science providers. This was initially carried out on human blood and saliva due to the ease of access to the samples.

Materials and methods

Sample collection and RNA isolation

In this study, blood and saliva samples were collected from a number of healthy volunteers. Fresh blood was obtained using the finger prick method then spotting the blood onto sterile filter paper. Saliva samples were collected in the form of sterile buccal swabs. All samples were extracted immediately after collection.

Total RNA was isolated from blood and saliva using the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol, with minor modifications as suggested by Zubakov *et al* (2009) [13]. These modifications included soaking the filter paper and buccal swabs in RLT buffer for 1 h at 4°C prior to extraction.

cDNA synthesis

Reverse transcription of the total RNA was performed with random hexamers and the M-MLV Reverse Transcriptase Kit (Ambion) according to the manufacturer's protocol. Reverse transcription blanks were also carried out for each body fluid and primer combination.

Real-Time PCR

Real-time polymerase chain reaction (PCR) was performed using the relative quantitation method on a 7500 Fast Real-Time PCR system (ABI). 1 µl of cDNA (or 1 µl of DNA/RNA free H₂O (Sigma) for negative controls) was amplified in a total reaction volume of 10 µl. The reaction mix contained 5 µl SYBR[®] Green master mix (ABI), 1 µl of forward primer, 1 µl reverse primer and 2 µl DNA/RNA free H₂O. The primer sequences were as follows: β-spectrin (SPTB – blood[3]), Forward 5'- AGG-ATG-GCT-TGG-CCT-TTA-AT-3' and Reverse 5'- ACT-GCC-AGC-ACC-TTC-ATC-TT-3'; Neutrophil Cytosolic Factor 2 (NCF2 – blood[13]), Forward 5'- ATT-ACC-TAG-GCA-AGG-CGA-CG-3' and Reverse 5'- TCT-GGG-TGG-AGG-CTC-AGC-T-3'; Keratin 4 (KRT4 – saliva[13]), Forward 5'- AGG-AGG-TCA-CCA-TCA-ACC-AG-3' and Reverse 5'-GCT-CAA-GGT-TTT-TGC-

TGG-AG-3'; Small Proline-Rich Protein 1A (SPRR1A – saliva[13]), Forward 5'- TGG-CCA-CTG-GAT-ACT-GAA-CA-3' and Reverse 5'- CCC-AAA-TCC-ATC-CTC-AAA-TG-3'. The housekeeping gene targeted in each reaction was 18s ribosomal RNA (18s rRNA), with the following primer sequences: Forward 5'- CTC-AAC-ACG-GGA-AAC-CTC-AC-3' and Reverse 5'- CGC-TCC-ACC-AAC-TAA-GAA-CG-3'. All PCR primer concentrations were 0.2ng/µl (All unlabelled primers were from MWG Eurofins).

The real-time PCR run conditions consisted of a denaturing step for 10 minutes at 95 °C, followed by 45 cycles of, 30 seconds at 95 °C, 45 seconds at 60 °C and 30 seconds at 72 °C. This was then followed by a melting curve stage that consisted of 15 seconds at 95 °C, 1 minute at 60 °C, the temperature was then ramped (at a rate of 1%) to 95 °C and held for 30 seconds and then 15 seconds at 60 °C.

Data analysis was performed using 7500 Fast Real-Time PCR software version 2.0.2.

Results and discussion

Real-Time PCR

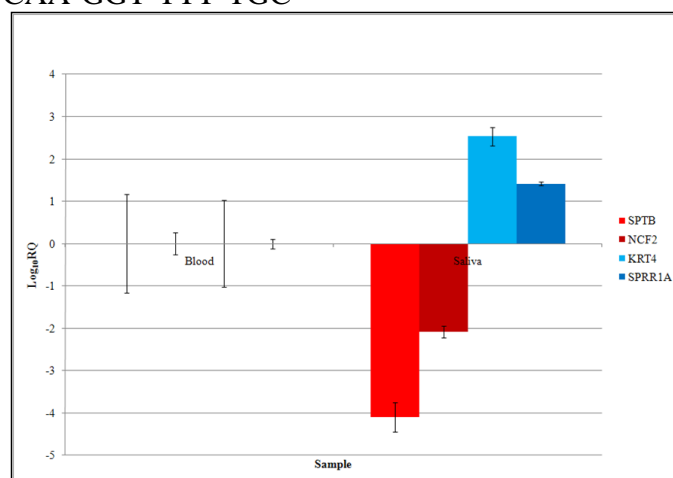


Figure 1: A gene expression plot of blood and saliva specific markers on blood and saliva samples using real-time PCR



In figure 1, real-Time PCR results were generated from a series of experiments for blood and saliva specific markers on blood and saliva samples. The calibrator sample in each experiment was blood and the formula used to create the gene expression plot was $2^{-\Delta\Delta CT}$.

The results from this study revealed high specificity for both the blood and saliva markers. SPTB and NCF2 were under

expressed in saliva signifying an over expression of these markers in blood. KRT4 and SPRR1A were over expressed in saliva this signified an under expression in blood. For each marker it can be seen that there is a significant difference in levels of expression between blood and saliva.

Reverse transcription blanks and negative controls

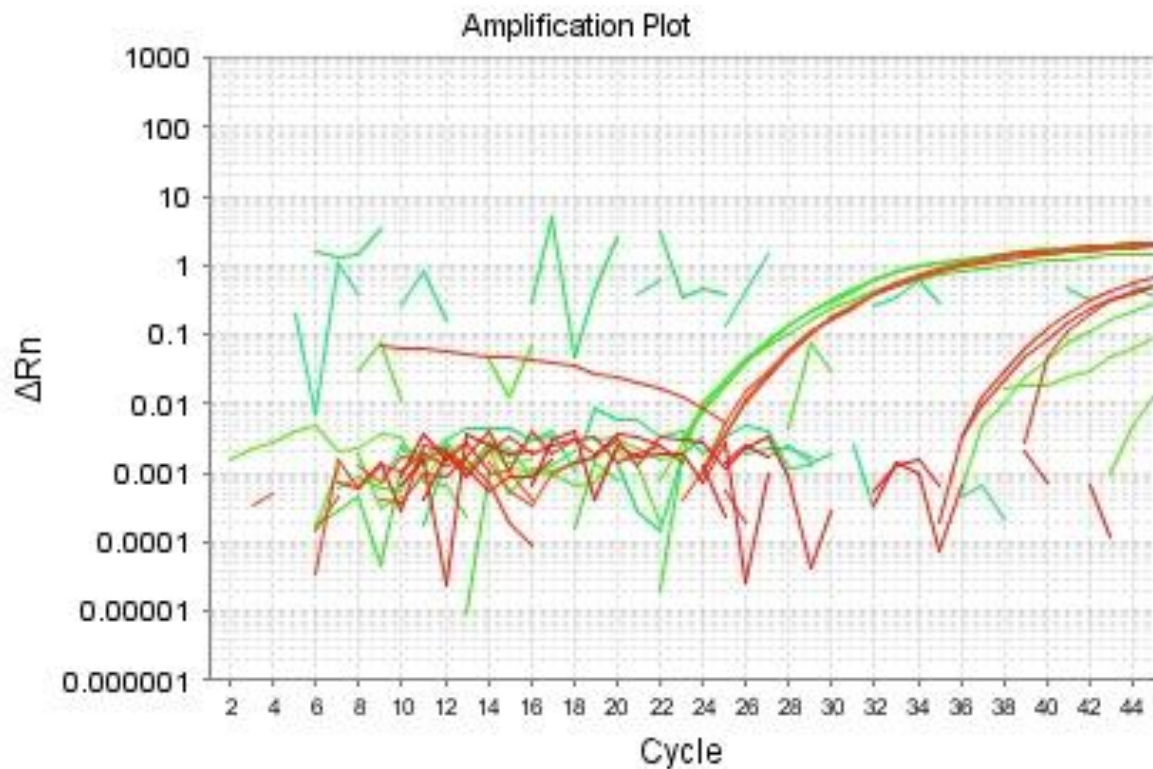


Figure 2: An amplification plot showing the amplification detected in the reverse transcription blanks and negative controls. All test sample amplifications have been omitted for clarity.

Figure 2 demonstrates that the reverse transcription blanks showed some significant amplification and, to a lesser extent, there was some amplification

detected in the negative controls; thus indicating a possible issue with contamination or with the test itself.

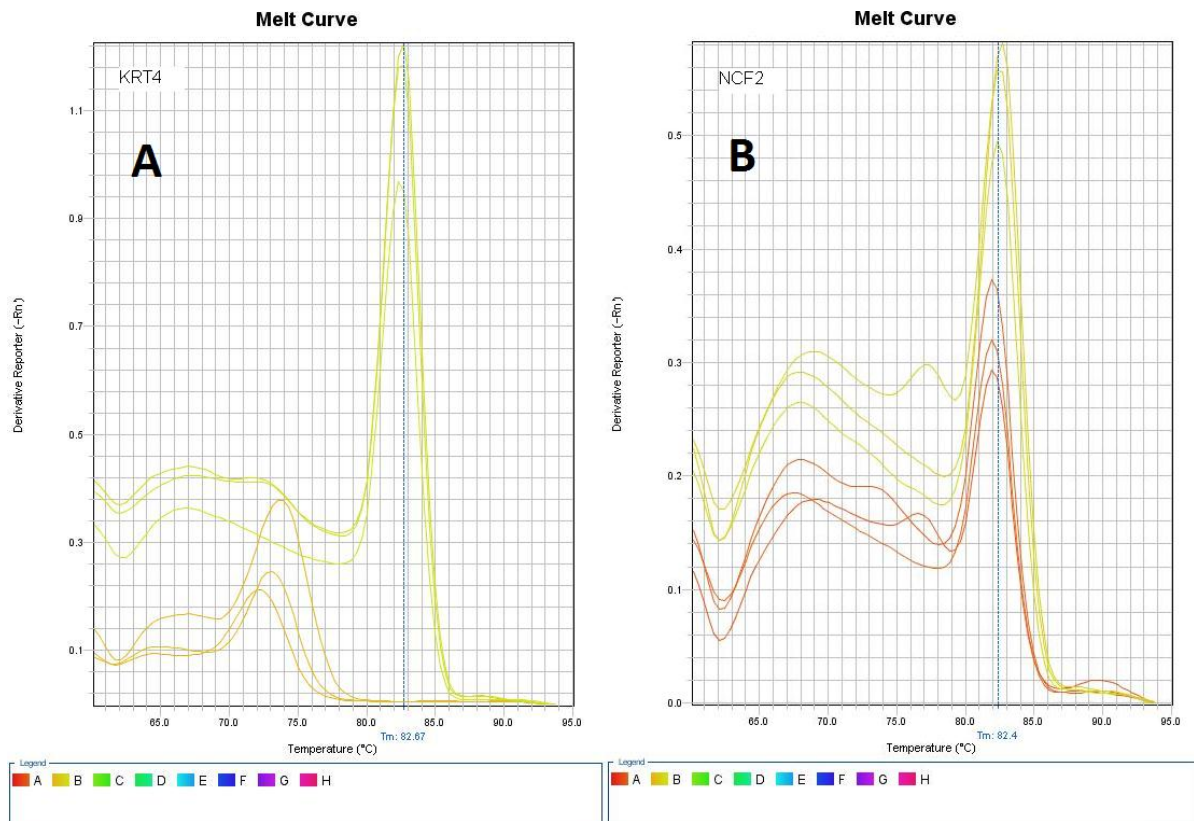
Melting curve analysis

Figure 3: A representative sample of melting curves of amplified products.

In figure 3, image A is from targeting KRT4 and image B is from NCF2. The larger peak in both images represents the main source of the fluorescence generated during real-time PCR. In image A (KRT4), there is a series of large peaks corresponding to a T_m of $\sim 82^\circ\text{C}$ and a series of smaller peaks ($T_m = \sim 73^\circ\text{C}$). In image B, the larger peaks corresponding to a T_m of $\sim 82^\circ\text{C}$. Primer dimers can be observed by the presence of a smaller hump or shoulder to the left of the main peaks.

The melting curve analysis was carried out as there were concerns that the

fluorescence detected did not correspond to the targeted amplicons given the lack of specificity of the SYBR Green dye. If all of the amplification was from the targeted markers, then a series of peaks with the same T_m value would be expected. If all of the amplification was from contamination then a series of peaks with a variety of T_m values would be obtained. However, all peaks exhibited are expected given that the amplification is from the markers targeted rather than from PCR or non-targeted products.

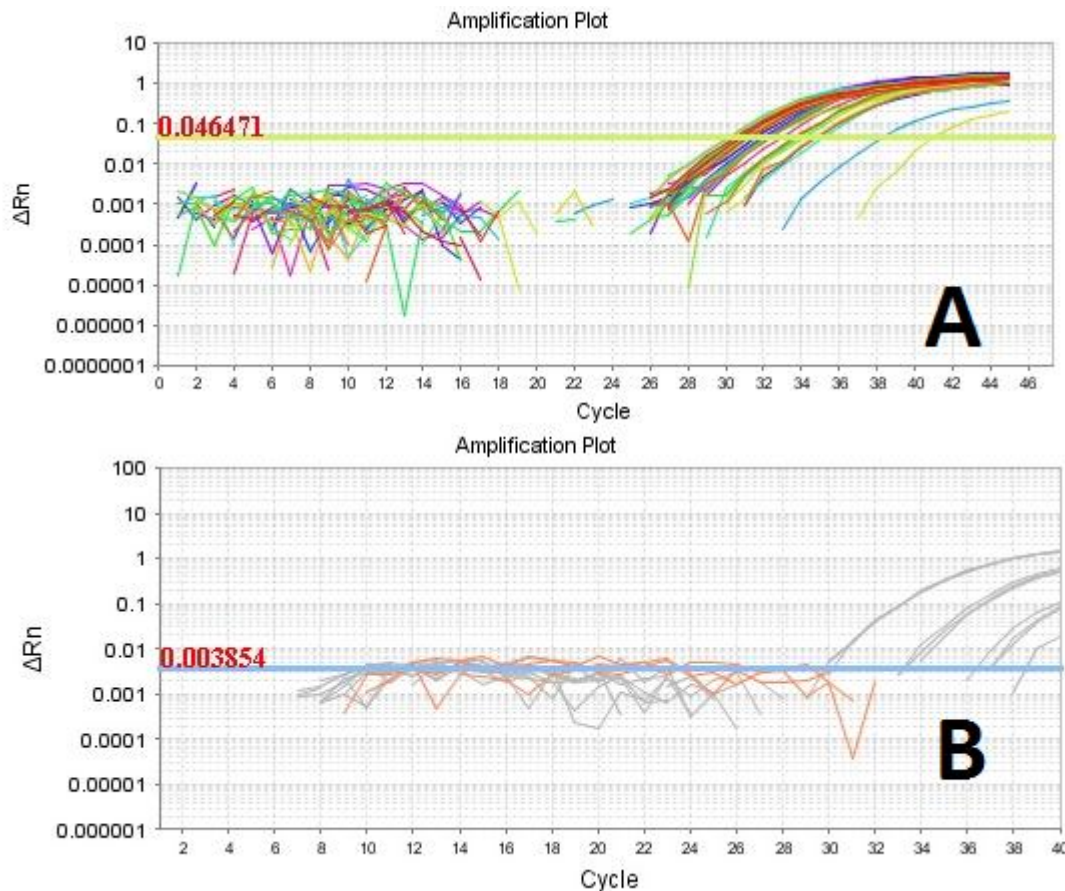
Contamination

Figure 4: Real-time PCR results on negative controls using SYBR Green (A) and Taqman (B) fluorescent dyes. The amplification in image A is from negative and positive controls whilst the amplification in image B is from positive controls alone. The threshold was determined automatically by the 7500 real-time PCR software V2.0.2.

Although the melting curve analysis supported the view that the amplification detected was from the markers targeted rather than from contamination or PCR product, further verification was required. This was done by running a plate with negative controls using SYBR Green and GAPDH as shown in Figure 4 in image A. Significant amplification can be seen with an approximate Ct value of 32. The same negative controls were then run on a plate using Taqman and GAPDH (as per manufacturer's instructions) as seen in Figure 4 in image B. In this case, no amplifications were detected from the negative controls, thus showing that there was no contamination and that the

amplification detected in image A is a by-product of the SYBR Green fluorescent dye.

The main objective of this study was to develop a test capable of distinguishing between blood and saliva using the SYBR Green[®] fluorescent dye rather than the Taqman[®] probe. This was achieved by comparing the expression levels of specific mRNA markers for both blood (NCF2 and SPTB) and saliva (KRT4 and SPRR1A) in blood and saliva samples. Although it has been reported that SPTB is expressed in other body fluids [3, 15], it was found to be sufficient to differentiate between blood and saliva. Over expression of a specific



marker in the test sample indicated there was more marker present in the test sample compared to the calibrator sample and *vice-versa*. In figure 1, the results illustrate a lower expression of the NCF2 and SPTB markers in saliva than in blood and a higher expression of KRT4 and SPRR1A in saliva than blood. For each marker there was a significant difference in levels of expression between blood and saliva.

The main advantage of a SYBR Green[®] test over a Taqman[®] based test is that it can be more cost-effective in certain circumstances. If mRNA profiling is to be utilised in forensic casework then the requirements of the police forces and other law enforcement agencies needs to be taken into consideration. By utilising Bayes' theorem based case assessment and interpretation (CAI) [14], it can be seen that there is no requirement for a routine assay as in most cases, there is usually an idea of which body fluids the DNA is thought to have come from, i.e. vaginal material or saliva, saliva or sweat, blood or an underlying saliva stain. In such cases, having a multiplexed BFI test that targets a full panel of body fluid specific markers is not cost effective nor is it a good use of resources. If a high-throughput mRNA profiling process were to be required, then a Taqman[®] based assay would be the most cost-effective for a forensic science provider in a competitive market. However, a body fluid identification test of this nature would not be routinely required as the current enzymatic based tests are usually sufficient.

There are times when such enzyme based tests may not be sufficient or available; such as when it is necessary to distinguish between saliva and vaginal material; i.e. from a penile swab in cases of consensual oral intercourse versus non-consensual vaginal intercourse. In such cases, mRNA profiling can then be utilised. If a forensic

science provider encounters such cases on an infrequent basis, it may be more cost effective to utilise a SYBR Green[®] based BFI test. The inclusion of unlabelled primers means that the assay can be tailored to suit the needs of the individual case and provide a more fit-for-purpose solution. How an mRNA based BFI test can be applied within a CAI model has yet to be investigated and will be the subject of further work in the future, but it is thought that such a test will merely augment current enzymatic/immunological tests to increase confidence in associating a DNA profile with a particular body fluid rather than replace them.

The issue of possible contamination as shown by a relatively high fluorescence being detected in the reverse transcription blanks and the negative controls (figure 2) gave some concerns and caused the experiments to be repeated several times with increasingly rigorous anti-contamination procedures; including the wearing of face masks, the use of several DNase based wipes and sprays, the incorporation of a DNA digestion step during the RNA extraction stage and the use of DNA/RNA free H₂O. Despite this, such fluorescence has never been eliminated or even significantly reduced. If these relatively high levels of fluorescence were due to contamination; then we would expect some significant variations from experiment to experiment. It is therefore possible that such fluorescence may actually be a by-product of the test itself. This hypothesis is supported by the melting curve analysis data (Figure 3); if there was DNA contamination taking place, then multiple T_m peaks would be detected and the plot would not exhibit such consistent peaks. One possible reason for the presence of amplification could be due to the non-specificity of the SYBR Green[®] reagent. This was then confirmed by running negative controls using both

SYBR Green and Taqman fluorescent dyes. The amplifications detected, as shown in Figure 4, were markedly different. Clear amplification has been detected in the negative controls using SYBR Green and GAPDH, but no amplification was detected using Taqman and GAPDH. If there was a contamination issue, then amplification would have been detected within the negative controls on the Taqman plate.

Conclusions

In summary, an mRNA based body fluid identification utilising SYBR Green[®] dye capable of distinguishing between blood and saliva has been established. This has shown that a SYBR Green[®] based BFI test could be just as distinguishing as a Taqman[®] based BFI test. However, due to the non-specific binding properties of

SYBR Green, this test is less reliable than using Taqman, but does not exclude it as being an alternative. This is supported by a small number of research groups who appear to have used SYBR Green to screen for primers albeit not as a test in its own right [6, 11, 13].

To move this forward, further validation studies will be carried out using SYBR green by carrying out BFI tests on blind samples as well as mixtures of body fluids. Additional studies will also be carried out to further verify the amplified product. Such studies will include sequencing and capillary electrophoresis base fragment analysis.

Further blood and saliva specific primers will be incorporated into the test. Further work will also be required in order to expand this test to include menstrual blood, vaginal material and sweat.

References

- [1]. Ballantyne J, Juusola J. *Messenger RNA profiling: body fluid identification using multiplex reverse transcription-polymerase chain reaction (RT-PCR)*, **Google Patents**; 2007.
- [2]. Bauer M, Patzelt D. *Identification of menstrual blood by real time RT-PCR: Technical improvements and the practical value of negative test results*, **Forensic Science International** 174, 1, 55-9, 2008
- [3]. Haas C, Klessner B, Maake C, Bär W, Kratzer A. *mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR*, **Forensic Science International: Genetics** 3, 2, 80-8, 2009
- [4]. Hanson EK, Lubenow H, Ballantyne J. *Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs*, **Analytical Biochemistry** 387, 2, 303-14, 2009
- [5]. Sakurada K, Ikegaya H, Fukushima H, Akutsu T, Watanabe K, Yoshino M. *Evaluation of mRNA-based approach for identification of saliva and semen*, **Legal Medicine** 11, 3, 125-8, 2009
- [6]. Zubakov D, Kokshoorn M, Kloosterman A, Kayser M. *New markers for old stains: stable mRNA markers for blood and saliva identification from up to 16-year-old stains*, **International journal of legal medicine** 123, 1, 71-4, 2009



- [7]. Fleming RI, Harbison S. *The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids*, **Forensic Science International: Genetics**, 4, 4, 244-256, 2010
- [8]. Bauer M, Polzin S, Patzelt D. *Quantification of RNA degradation by semi-quantitative duplex and competitive RT-PCR: a possible indicator of the age of bloodstains?* **Forensic Science International**, 138, 1-3, 94-103, 2003
- [9]. Fleige S, Pfaffl MW. *RNA integrity and the effect on the real-time qRT-PCR performance*, **Molecular Aspects of Medicine**, 27, 2-3, 126-39, 2006
- [10]. Setzer M, Juusola J, Ballantyne J. *Recovery and Stability of RNA in Vaginal Swabs and Blood, Semen, and Saliva Stains*, **Journal of Forensic Sciences**, 53, 2, 296-305, 2008
- [11]. Hanson E, Lubenow H, Ballantyne J. *Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs*, **Forensic Science International: Genetics Supplement Series**, 2, 1, 503-4, 2009
- [12]. Wong ML, Medrano JF. *Real-time PCR for mRNA quantitation*, **Biotechniques**, 39, 1, 75, 2005
- [13]. Zubakov D, Hanekamp E, Kokshoorn M, van Ijcken W, Kayser M. *Stable RNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples*, **International journal of legal medicine**, 122, 2, 135-42, 2008
- [14]. Cook R, Evett IW, Jackson G, Jones PJ, Lambert JA. *A model for case assessment and interpretation*, **Science & Justice**, 38, 3, 151-6, 1998
- [15]. Haas C, Hanson E, Bär W, Banemann R, Bento AM, Berti A, et al. *mRNA profiling for the identification of blood--Results of a collaborative EDNAP exercise*, **Forensic Science International: Genetics**, 5, 1, 21-26 2011