THE DEVELOPMENT OF AN RNA BASED ASSAY SYSTEM TO SUPPLANT CONVENTIONAL METHODS FOR BODY FLUID IDENTIFICATION

Jane Juusola[1], BS and Jack Ballantyne[1,2], PhD
[1]Department of Chemistry, University of Central Florida, Orlando, FL
[2]University of Central Florida Center for Forensic Science, Orlando, FL

The present work describes a prototype RNA based assay system to supplant conventional methods for body fluid identification. Before discussing the assay in detail, it is important to consider whether there is still a need for such testing. Prior to the advent of DNA analysis, it was standard practice to perform biochemical, serological, and immunological tests to identify body fluid stains prior to genetic analysis. Recently, however, a new trend has emerged in which body fluid identification is bypassed in favor of identification of human DNA in the sample extract. Proponents of this procedure argue that the presence of DNA from a particular individual is all that is necessary and it is inconsequential from which cell type (i.e. body fluid or tissue) the DNA originated. Nevertheless, is there still any forensic significance to ascertaining the source of a tissue or body fluid stain? We believe there is.

It may be useful to discuss examples of cases where body fluid identification could provide important probative evidence. In the first example, a sexual assault takes place involving vaginal intercourse whereby the female victim happens to be in menses. Blood is found on the suspect's clothing and, according to DNA testing, is consistent with having originated from the victim. The defense argues that the blood comes from the victim's nose when she was punched by the suspect, and any sexual act is denied. On the other hand, the prosecution argues that the blood is a result of a sexual act. The ability to identify blood as menstrual in origin, as opposed to circulating blood, would be significant. In another case, the identification of vaginal secretions, for which no test currently exists, would assist the investigation. An alleged rape occurs in a car, the DNA of the victim is found on the seat but the defense argues that it came from sweat or skin when the victim was sitting in the car. However, confirmation of the presence of vaginal secretions would be incontrovertible evidence of sexual activity. Additionally, if DNA from a victim is found on an implement believed to have been used in a sexual assault, the significance of this evidence would be enhanced by demonstrating that it originated from her vaginal secretions. The final example involves the sexual abuse of a young child by a person living in the same residence as the victim in which the finding of the suspect's DNA on the child's clothing or bed would not be surprising. In many circumstances, it would be more difficult for the suspect to provide an innocent explanation for the presence of his semen versus traces of his saliva on the victim. As exemplified by these case scenarios we consider it important, in many instances, to identify the nature of the body fluid present in a stain.

Conventional biochemical and immunological tests for body fluid identification are performed in a series, not parallel, manner and are therefore costly in terms of time and sample. In seeking to develop novel multiplex (i.e. parallel) analysis procedures for body fluid identification, we have considered assays based upon protein and messenger RNA (mRNA) since both are expressed in a tissue type specific manner. Multiplex analysis of complex protein mixtures such as those present in body fluids awaits further development in proteomics. Messenger RNA, the intermediate between DNA and proteins, is a better option because the technologies for massively parallel analysis continue to be developed in the post-genomic era.

We realize that there are multiple different tissue secretions present in body fluids. Each tissue type is comprised of cells that have a unique transcriptome, or gene expression (i.e. mRNA) profile. We term the collection of genes that are expressed within the constellation of differentiated cells that makes up a body fluid the multicellular transcriptome. These genes comprise ubiquitously expressed housekeeping genes, which are responsible for cell maintenance functions, and genes that are specifically expressed in certain tissues only. The mRNA molecules are present in different quantities depending upon the particular species of mRNA and the cell type, and can be classified as abundant, moderately abundant and rare. Eventually we hope to assay abundant or moderately abundant tissue-specific genes to automate the

identification of body fluids that are significant to forensic investigations, including blood, semen, saliva, vaginal secretions, skin, urine, muscle, adipose, brain, and fecal material. We hope to identify and characterize five to ten tissue specific genes per tissue type. Automation would be accomplished by using a 'body fluid identification chip' containing a microarray of cDNAs that would be able to recognize each of the candidate genes. In summary a mRNA based approach would allow for the facile identification of body fluids and could supplant the current battery of serological and biochemical tests. Some of the advantages of an mRNA-based approach are greater specificity, simultaneous analysis through a common assay format, improved timeliness, automation, and decreased sample consumption.

There is a paucity of studies involving the detection of RNA in biological stains. We have sought to answer the following questions. Is it possible to isolate total RNA, which comprises messenger, ribosomal, transfer and small RNA from body fluid stains? Moreover, since mRNA only comprises 1-3% of total RNA in stains (with any particular transcript being present in 0.001-1.000 % of mRNA) is it possible to detect sufficient, relatively non-degraded mRNA from various housekeeping genes in the isolates using RT-PCR? How stable is such RNA and how sensitive is the RT-PCR assay? Finally, can we identify and develop suitable assays for a number of candidate tissue-specific mRNAs for use in body fluid identification? In this report, we present data indicating that it is possible to isolate RNA from blood, saliva, and semen stains and to detect mRNA from housekeeping genes in these stains. The RNA appears to be stable since it is possible to isolate RNA of sufficient quantity and quality for analysis in ten-week-old stains. Lastly, we have identified a number of candidate tissue-specific genes for saliva.

Total RNA was extracted from blood, semen and saliva stains using guanidine isothiocyanate and phenol:chloroform, followed by isopropanol precipitation and DNase I treatment of the re-solubilized pellet. The RNA was easily visualized as a DNase I resistant smear on a SYBR Gold stained agarose gel. We have also been able to detect mRNA for three housekeeping genes (□-actin, GAPDH and S15) in blood, saliva, and semen stains, using both oligo-dT and random decamers as first strand cDNA primers in a RT-PCR reaction. Amplified product was found in the negative control samples of the semen stain extracts, to which no reverse transcriptase was added. This was a surprising finding since the primers had been judiciously designed to anneal either across exon boundaries or within separate exons, and so amplification products from any contaminating DNA should have been non-existent or significantly larger. The most reasonable explanation was that processed psudogenes were being detected by the sensitive RT-PCR procedure from traces of contaminating genomic DNA in the RNA extracts. This was confirmed by analyzing genomic DNA from blood, saliva, and semen stains. This DNA contamination problem in the semen RNA extracts encouraged the adoption of the DNase I digestion step now standard in our protocol.

Total RNA was quantified using a RiboGreen® fluorescence assay. The amount of total RNA that can be extracted from stains can be in the order of several hundred nanograms. The average amount of total RNA extracted from a 50µl bloodstain was 230ng (range 170-260ng), from a buccal swab was 360ng (210-610ng), from 10µl, 25µl, and 50µl saliva stains was 29ng, 50ng, and 120ng respectively, and from 10µl, 25µl, and 50µl semen stains was 59ng, 110ng, and 190ng, respectively. This is sufficient RNA for the extremely sensitive RT-PCR method. As little as 0.16ng of total RNA isolated from a bloodstain was enough to detect the S15 housekeeping gene using oligo-dT as the first strand cDNA primer.

The stability of RNA was studied using bloodstains that had been left at room temperature for 0, 1, 2, 7, 14, 27, and 73 days. Even after 10 weeks, significant quantities of total RNA were recovered. However, when oligo-dT was used as the first strand cDNA primer to detect mRNA from housekeeping genes no product was obtained from 7-day-old stains. However, using random decamers as the first strand primer permitted the detection of mRNA even with the 10-week-old stains. These results are probably due to partial degradation of mRNA whereby the single stranded RNA becomes fragmented resulting in the 3' poly (A) tail becoming disassociated from the rest of the molecule. The oligo-dT primers, which bind to the poly (A) tail, would be ineffective under these circumstances whereas the random decamer primers would still be able to anneal and produce a product. Accordingly, we recommend the latter primers for routine use for stain analysis in which some degradation of the mRNA is expected.
Five saliva-specific genes have been identified so far, including statherin, histatin, proline-rich proteins BstNI subfamily 1 (PRB1), proline-rich proteins BstNI subfamily 2 (PRB2), and proline-rich proteins BstNI

subfamily 3 (PRB3). These gene products participate in a non-immune host defense system in the oral cavity. Primers have been designed so that the RT-PCR would amplify products in the 100-200 base pair range. Processed pseudogenes do not confound the analysis since the PCR products from genomic DNA are significantly larger than those obtained from mRNA. Specificity for saliva has been demonstrated by the absence of these genes in RNA extracted from semen and blood stains. Additionally, these genes were found to be detectable in all individuals tested.

To summarize, we have been able to isolate RNA from biological stains in sufficient quantity and quality for analysis. We have identified a number of candidate tissue-specific genes that may be useful for the positive identification of saliva. We are in the process of seeking other candidate tissue specific genes for the identification of other body fluids and tissues of forensic importance.