SCREENING OF GENE MARKERS FOR FORENSIC IDENTIFICATION OF VAGINAL SECRETION STAINS

إعداد
Muhammad Muqadas

الرياض
2-4/11/1428ـ الموافق (2-14/11/2007)
SCREENING OF GENE MARKERS FOR FORENSIC IDENTIFICATION OF VAGINAL SECRETION STAINS

Muhammad Muqadas, National centre of Excellence in Molecular Biology, Lahore, Pakistan.

ABSTRACT:

mRNA profiling methodology has proven to be a powerful approach for the molecular identification of different human body fluids including vaginal secretion stains being collected as an evidence material from the crime scene. For some frequently encountered body fluids like saliva and vaginal secretions there is no confirmatory test available. 13 genes were selected for this study from the public databases considering their expression pattern in vaginal secretions, biochemistry, immunology and physiological aspects. Reverse transcriptase-Polymerase chain reaction (RT-PCR) has been used to study gene expression. Two genes were taken from already published sources namely HBD-1 (DEFB1) and MUC4. Other eleven genes selected for this study are ADAMTS 5, CART 1 , CCL 20 , CNFN , DEFA 5 , HOXA 13 , INDO , KRT 7 , SPRR 2B , SERPIN B4 and SLPI. Out of these 13 genes, eight genes show expression in the vaginal secretions. This study describes a pathway for the molecular identification of the vaginal secretion stains for the forensic analysis that appears specific, sensitive and robust in terms of time and sample consumption. This technique is extremely useful especially for the rape and sexual assault cases.

Keywords: mRNA profiling; Vaginal secretions; Forensic Analysis; RT-PCR; cDNA.
INTRODUCTION:

Body fluids frequently encountered in forensic casework include blood, saliva, semen, and vaginal secretions. There are two key objectives in the examination of body-fluid stains or deposits; determining that a stain or deposit is in fact a body fluid, and that it has a human origin; and if of human origin, determining which person shed or deposited the fluid.

Biochemical tests for vaginal epithelial cells obtained during swabbing are Lugol’s Iodine or protein and ion detection, none show specificity to vaginal epithelial cells and are therefore unsuitable to use as evidential support in a court of law. For vaginal fluid, staining with glycogen may be performed; this is also not definite and reliable. For urine identification, tests for Urea and Creatinine may be performed. For feces identification tests for Urobilinogen may be performed. There has been no definitive test for human saliva. In forensic practice, the test for amylase has been a commonly used method as a presumptive test for the identification of saliva. Unfortunately, under certain circumstances, amylase can be detected in semen, blood, and plant sources, as well as saliva from other mammalian species, hence, the need for a more conclusive test.

With the current, advanced state of DNA technology, it may seem that the only necessary work is to obtain a DNA profile. However, DNA profiling is expensive, and is not useful if a stain is not actually a body fluid or of human origin. So testing for a particular body fluid present or not is a useful first step in any investigation.

The first generation of body fluids identification for vaginal secretion stains was carried out by biochemical, immunohistochemical and serological tests. But these tests have some limitations regarding the false positive results. It was the need of the hour to initiate a study that would replace these biochemical and serological tests and
increase authenticity and reliability of body fluids identification by using multiple markers for particular body fluids. There are no definitive tests for vaginal secretion and saliva. There is currently no accurate method to identify vaginal epithelial cells present in the vaginal secretions uniquely.

MATERIALS AND METHODS:

**Body fluid samples:**

Vaginal cotton swabs were used as biological material for body fluid analysis and were stored at -20 °C within the 6 hours after collection. Samples were transported to lab at controlled temperature. Swabs were used within two weeks for extraction of RNA after storage.

**RNA isolation:**

Total RNA was extracted from the vaginal cotton swabs using the AllPrep DNA/RNA Mini Kit (Qiagen) according to the manufacturer’s protocol. Four stored vaginal cotton swabs were taken and placed directly into two 1.5 ml eppendorf tubes already containing the 500 µl of Buffer RLT Plus (provided by the kit), placing two swabs into each tube. Sterile razor blade was used to cut the cotton swab just up from the swab cotton area. Incubated these swabs in eppendorf tubes for 15 minutes on ice and then for 15 minutes at room temperature with occasional vortexing and flicking. Vortex for one minute for the complete lysis and homogenization of the vaginal epithelial cells. Then transferred into AllPrep DNA spin column and centrifuge for 30 seconds at 13000 rpm at room temperature. Take flow through and add equal volume of freshly prepared 70 % ethanol then transfer into RNeasy spin column and centrifuge for 30 seconds at 13000 rpm at room temperature. Flow through was discarded and added 700 µl Buffer RW1 to the RNeasy spin column then centrifuge and discard the flow through. Now add 500 µl Buffer RPE, centrifuge and discard the flow through. This step was repeated twice. After
this, RNA was precipitated with 50 µl of the RNase-free water and stored at -70 °C until needed.

**Quantification and purity of RNA:**

UV absorption measurements were used as an indication of RNA purity and quantity. Bio-Rad SmartSpec™ Spectrophotometer were used for purity and quantification studies. Protocol given by the AllPrep DNA/RNA Mini kit (Qiagen) was opted for RNA quantification and purity estimation. Formaldehyde agarose gel electrophoresis was also used for purity and quantification studies.

**cDNA Synthesis:**

cDNA was synthesized using the RevertAid™ H Minus First Strand cDNA Synthesis Kit according to the following protocol:

PCR tube was taken and placed on ice. Now add 9 µl of total extracted RNA (10ng), 1 µl of oligo (dT) ₁₈ primer (0.5µg/µl) and 2 µl of DEPC-treated water. Mixed gently and incubated at 70°C for 5min. Then add these components respectively, 4 µl of 5 X reaction buffer, 1 µl of Ribolock™ Ribonuclease Inhibitor (20u/µl) and 2 µl of 10mM dNTP mix. Now incubated at 37°C for 5min. Finally added 1 µl of RevertAid™ H minus M-MuLV Reverse Transcriptase (200u/µl) and incubated at 42°C for 60min then stopped the reaction by heating at 70°C for 10min on PCR Machine. Synthesized cDNA was stored at -20 °C until needed.

**Gene selection parameters:**

Genes for the vaginal secretions has been found on basis of physiology, biochemistry and immunology of the vaginal secretion stains. These genes have been taken from sequenced human genome data available in public databases. For selection of genes having expression in the vaginal secretions, to be used as marker for molecular
identification of vaginal secretions stains for forensic use, following parameters were considered,

1- Genes having more than one exon.
2- Genes having no any reported pseudo genes.
3- Genes having transcript variants with common exon.
4- Genes having expression in the vaginal secretions / vaginal fluid / vaginal mucus / vaginal epithelial cells / cervix / cervical mucus according to the Human Protein Reference Database (HPRD).
5- Genes having expression according to the CGAP (Cancer Genome Anatomy Project).

PCR Primer Sequences:

For amplification GAPDH was used as house keeping gene and served as an internal control (Positive Control). Specific forward and reverse primers for genes that were expressed in the vaginal secretions were designed with the help of computer-assisted program “Primer 3”.

Table No. 1; List of Primers used for amplification:

<table>
<thead>
<tr>
<th>Sr. NO.</th>
<th>GENE NAME</th>
<th>SEQUENCE 5’→ 3’</th>
<th>PRIMER LENGTH</th>
<th>PCR PRODUCT</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MUC4</td>
<td>Forward GGA CCA CAT TTT ATC AGG AA 20 235</td>
<td>(Ballantyne,2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse TAG AGA AAC AGG GCA TAG GA 20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>HBD-1</td>
<td>Forward TTC CTG AAA TCC TGA GTG TT 20 213</td>
<td>(Ballantyne,2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse TAA CAG GTG CCT TGA ATT TT 20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>SLPI</td>
<td>Forward GTG GAA GGC TCT GGA AAG T 19 265</td>
<td>Primer 3 (New Designed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse ACT GGC CAT CCA TCT CAC 18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>HOXA 13</td>
<td>Forward ACG AAC CCT TGG GTC TTC 18 193</td>
<td>Primer 3 (New Designed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse GTA TAA GGC ACG CGC TTC 18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Primer 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(New Designed)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ADAMTS 5</td>
<td>Forward CTA CTG CAC AGG GAA GAG G</td>
<td>19</td>
<td>168</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse AGG ACA CCT GCA TAT TTG G</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>CART 1</td>
<td>Forward CTT TTA CAG CAA AGC GTG TG</td>
<td>20</td>
<td>239</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse CAC ATT TAT CCC CAA GTT CA</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>CCL 20</td>
<td>Forward GTG CTG TAC CAA GAG TTG TC</td>
<td>20</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse TCC AAG ACA GCA GTC AAA GT</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>CNFN</td>
<td>Forward ATG TCC TAC CTC GTG ACC AG</td>
<td>20</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse GCA CAG AGG AGC AAA AGT G</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>DEFA 5</td>
<td>Forward ATG AGG CTA CAA CCC AGA AG</td>
<td>20</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse CCT TGT ACT GAA TCT TTG AC</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>INDO</td>
<td>Forward TCC TGG TCT CTC TAT TGG TG</td>
<td>20</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse GGA TTT GGT GAA ACA CTT GA</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>KRT 7-I</td>
<td>Forward ACA CAT CTG TGG TGC TGT C</td>
<td>19</td>
<td>215</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse CCG GTT CAT CTC TGA AAT C</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>SPRR 2B</td>
<td>Forward GAT CTG CTT TGG AGA AAC TG</td>
<td>20</td>
<td>519</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse CTC TTT CTG GTG AAG CTC TG</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>SERPIN B4</td>
<td>Forward GTG TGG AAT CTA CTG ATT TTG C</td>
<td>22</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse CCC ATC AGG AAA TAG GTT TT</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>GAPDH</td>
<td>Forward ACC ACA GTC CAT GCC ATC A</td>
<td>19</td>
<td>452</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse TCC ACC ACC CTG TTG CTG TA</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>HOXA13 (II)</td>
<td>Forward CTG GAT ATG CCA GTG GTG</td>
<td>18</td>
<td>313</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reverse GCC TCC GTT TGT CCT TAG</td>
<td>19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Primers dilution:**

Primers made by DNA Synthesis Lab were in fluid form. These include SLPI, HOXA13, MUC4 and HBD-1. Dilution was made 20 times with the T.E Buffer for primers and used directly. SLPI was used 2.5µl each (F10µg/µl;R6.27µg/µl), MUC4 6.5 µl each (F10.42µg/µl;R8.32µg/µl), DEFB1 (HBD-1) 4.0 µl each (F9.99µg/µl;R9.75µg/µl) and HOXA13 4.5 µl (F4.21µg/µl;R3.21µg/µl) was used for PCR amplification.
**Polymerase chain reaction:**

Bio-Rad iCycler Thermal Cycler PCR Machine and GeneAmp® PCR system 2700 (Applied Biosystems) were used in this study.

**PCR Amplification:**

For this purpose PCR reactions were set with 2.5 µl of 10X Taq Buffer with (NH₄)₂SO₄ and 15 mM MgCl₂ (Fermentas) [750 mM Tris-HCl (pH 8.8 at 25°C), 200 mM (NH₄)₂SO₄, 15 mM MgCl₂, 0.1 % (v/v) Tween 20], 1 µl of 0.125 mM dNTP Mix [0.125 mM each dATP, dCTP, dGTP, dTTP], 1 µl of Taq Polymerase (2 units /µl), Primers were used as KRT7 1.2 µM each primer, CART1 2.4 µM each, ADAMTS5 1.4 µM each, SPRR2B each primer 1.2 µM, 0.5 mM Spermine tetrahydochloride 0.30 µl and annealing temperatures mentioned below. Total PCR reaction volume was kept 25 µl.

The amplification was carried out at conditions with initial denaturation at 95°C for 5 minutes, melting 94°C for 20 seconds, annealing at 55°C for 30 seconds, Extension at 72°C for 40 seconds, and final extension at 72 °C for 7 minutes for a total of 32 cycles.

**Agarose gel electrophoresis:**

2 % Agarose gel was run at 100 Volts for 1 hour. When the gel was run completely, the DNA fragments were visualized directly under the UV trans- illuminator.

**RESULTS:**

**Genes having expression in vaginal secretions:**

From the above given gene primers sets, following genes have shown expression in vaginal secretion stains. HBD-1, MUC4, ADAMTS5, SLPI, CART1, KRT7, SPRR2B, HOXA13.
PCR AMPLIFICATION RESULTS:

Amplification of HBD-1, MUC4, HOXA13 and SLPI

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
</table>

Lane 1: DNA Ladder 100 bp

Lane 2: Negative Control

Lane 3: HBD-1 (213 bp)

Lane 4: MUC4 amp (235 bp)

Lane 5: SLPI amplified (265 bp)

Lane 6: HOXA13 (193 bp)

Fig 1: Shows expression of four gene markers in the vaginal secretions. Non amplification of the negative control shows that reagents are free from contamination.
HOXA13 Amplification and Non Specificity

Lane 1: DNA Ladder 100 bp
Lane 2: Negative Control
Lane 3: HOXA13 from Sample number 1.
Lane 4: HOXA13 from Sample number 2.
Lane 5: HOXA13 from Sample number 3.

Fig 2: Shows non specificity of HOXA13 gene marker for its binding at two different sites within the genome.

Possible reason for the HOXA13 Non-specificity:

This was a surprising finding since the primers had been designed to anneal either across exon boundaries or within separate exons, so amplification products from
any contaminating DNA should have been non-existent or significantly longer. The most reasonable explanation was that unreported exon or processed pseudo genes were being detected by the sensitive RT-PCR procedure. To control this, sequencing strategy was used. Sequencing of the 193 bp amplicon was done to know the original sequence so that we can make modified primers to avoid this non-specificity. But from later experiments it was revealed that this gene is not constitutively expressing in all the samples. To remove this non specificity a modified primer (HOXA13-II) was also made but this primer set did not show any amplification at all.

**Amplification of HBD-1 and MUC4**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>

Lane 1: DNA Ladder 100 bp  
Lane 2: Negative Control  
Lane 3: HBD-1 (213 bp)  
Lane 4: GAPDH (Positive Control) 452 bp  
Lane 5: MUC4 (235 bp)
Fig 3: Shows expression of the two gene markers taken from already published sources. Amplification of positive control shows that RNA is correctly isolated and cDNA is correctly synthesized.

**Amplification of KRT7 and SPRR2B**

Lane 1: DNA Ladder 100 bp  
Lane 2: ADAMTS 5  
Lane 3: KRT7 (215 bp)  
Lane 4: SPRR2B (519 bp)  
Lane 5: CCL20  
Lane 6: CNFN  
Lane 7: Negative Control  
Lane 8: GAPDH (Positive Control) 452 bp

Fig 4: Shows expression of the two novel gene markers in the vaginal secretion stains.
Amplification of SPRR2B and ADAMTS 5

Lane 1: DNA Ladder 100 bp
Lane 2: CNFN
Lane 3: SERPINB4
Lane 4: SPRR2B (519 bp)
Lane 5: ADAMTS5
Lane 6: ADAMTS5 (168 bp)
Lane 7: GAPDH (Positive Control) 452 bp

Fig 5: Shows expression of the one more novel gene marker namely ADAMTS5.
Amplification of KTR7, CART1, MUC4 and HBD-1

Lane 1: DNA Ladder 100 bp
Lane 2: CART1 (239 bp)
Lane 3: KRT7 (215 bp)
Lane 4: MUC4 (235 bp)
Lane 5: HBD1 (213 bp)
Lane 6: GAPDH (Positive Control) 452 bp
Lane 7: Negative Control.

Fig 6: Shows expression of four gene markers in the vaginal secretions.

**DISCUSSION:**

Initially vaginal secretions-specific genes were identified through a combination of literature and database searches and consideration of their physiology, biochemistry and immunology. Subsequently, searched genes having pseudo genes
present in the genome were rejected due to the difficulty of distinguishing mRNA derived PCR products from contaminating genomic DNA. Primers for each gene were then designed to span at least two different exons of the same gene in such a manner that the genomic DNA would produce amplicons that were either larger than that expected for that specific mRNA or would not be amplifiable under the specific PCR cycling conditions used.

In body fluid identification, mRNA profiling approach has been used on variety of body fluids like Semen, Saliva, Vaginal secretions and Blood. But all studies conducted so far used two or three gene markers for specific body fluid identification. The novelty of our work is the exploration of six new markers for molecular identification of vaginal secretion stains.

The reasons for using multiple gene markers for vaginal secretions identification are,

1- To account for any possible biological variation in gene expression levels.

2- An individual may possess a mutation that reduces or eliminates the expression.

3- Some genes may show expression in a specific state for example pregnancy, sexual stimulation and vaginal fungal infection.

4- The resulting multi gene marker identification system was able to reproducibly identify vaginal secretion stains as single or mixed body fluid stains.

5- This increases the specificity and authenticity of the results.

Saliva and vaginal secretions are unique in the fact they contain other existing flora, such as microbes or yeast, within the mouth and reproductive tract. It was struggled hard to keep primers only human specific but we have no experimental data. Therefore, there may be the amplification of RNA recovered from this microbial flora. Additional biological factors i.e. enzymes and hormones may also be present. Hormonal contents may vary the composition of vaginal secretions and also enhance the subsequent expression of
number of other unrelated genes. Results of the individual genes indicate that certain mRNAs may be more prone to degradation than others during sampling and then subsequent transportation and storage conditions, but this disadvantage may be overcome by using multiple markers per body fluid.

Current study used thirteen gene markers namely MUC4, HBD1, ADAMTS 5, CART 1, CCL 20, CNFN, DEFA 5, HOXA 13 (I&II), INDO, KRT 7, SPRR 2B, SERPIN B4 and SLPI. Eight genes out of these thirteen genes show expression in vaginal secretions. These genes are HBD1, MUC4, ADAMTS 5, SLPI, CART1, KRT7, SPRR2B and HOXA13 (I). Other five genes namely SERPIN B4, INDO, CCL 20, CNFN and DEFA 5 show no expression at all. MUC4 and HBD1 were the gene markers which were already reported in the literature but eleven gene markers are new and are reported first time for molecular identification of vaginal secretion stains. If there is no amplification at all another possible reason may be that primers designed, may be from a particular exon that might not be expressing but another exon of the same gene may be expressing so it is recommended to make multiple primers for a single gene to cover all the reported exons of a single gene.

This study was a preliminary investigation in order to gain an understanding of the usage of RNA in body fluid identification using mRNA profiling approach. Future experiments on mRNA profiling approach in biological stains identification will need to focus on three main areas to make this approach more applicable and powerful. Firstly, alternative protocols for RNA extraction, as well as cDNA synthesis and detection, should be tested to assess the isolation and quantitation of the isolated RNA and synthesized cDNA. Secondly, species specificity for these gene markers and validation studies for RNA isolation needs to be conducted. Thirdly, additional high abundance tissue-specific genes must be found and evaluated using Real Time PCR and Micro
array. This technique of Micro array analysis may be useful in evaluating the more abundant genes expressed in a particular body fluid. Those genes must have constitutive expression throughout the life, in every state and in every stage of life to be analyzed with authenticity and reproducibility.
REFERENCES:


