SHORT COMMUNICATION

M. Bauer · D. Patzelt

Protamine mRNA as molecular marker for spermatozoa in semen stains

Received: 21 May 2002 / Accepted: 2 October 2002 / Published online: 4 February 2003 © Springer-Verlag 2003

Abstract Cytological detection of spermatozoa with subsequent DNA analysis is the most important biological evidence in sexual crimes when suitable samples are available. Immunological and enzymatic detection of semenspecific proteins may be helpful but cannot replace specific identification of spermatozoa. We have recently shown that detection of cell-specific gene expression can be used to identify menstrual blood. In this paper we demonstrate that the basic nucleoproteins protamine 1 and 2 are suitable markers for spermatozoa because they are exclusively expressed in the haploid genome and that protamine mRNA can be detected in semen stains by the highly sensitive reverse transcriptase-polymerase chain reaction (RT-PCR). With semi-nested PCR, 10–100 spermatozoa are theoretically sufficient to provide positive amplification results, with hot-start PCR at least 100-1000 cells are required corresponding to an average semen volume of 0.01–0.1 µl. This new method thus allows specific identification of spermatozoa with molecular biology tools and may broaden the spectrum of investigations in the forensic laboratory.

Keywords Semen stains · mRNA · Protamine · Spermatozoa

Introduction

In rape cases evidence that ejaculation has occurred depends largely on the microscopic identification of spermatozoa. Enzyme activity-based methods (e.g. acid phosphatase, γ -glutamyltransferase) and antigen-antibody reactions (e.g. PSA, semenogelin, SVSA, MHS-5) which allow the detection of proteins present in seminal fluid may be useful

M. Bauer $(\mathbb{E}) \cdot D$. Patzelt

Institute of Legal Medicine, University of Würzburg, Versbacher Strasse 3, 97078 Würzburg, Germany Tel.: +49-931-20147020, Fax: +49-931-20147000, e-mail: reme005@mail.uni-wuerzburg.de as a preliminary test or in the absence of spermatozoa [1, 2, 3, 4, 5, 6, 7, 8] but are not specific for semen and show some cross-reactivity with other body fluids and biological samples [9]. Since the identification of the perpetrator by DNA analysis of semen stains has become the main purpose of forensic rape investigations, some scientists believe that detection of male DNA by Y-chromosomal STR analysis [10] might be sufficient evidence for the presence of semen in samples derived from vaginal swabs where cytology is negative [11].

However, since male DNA may have come from other sources than semen (e.g. contamination) and since the detection of male DNA on clothes or other non-vaginal samples does not allow any conclusion concerning the origin of the sample, the question whether spermatozoa are present or not remains crucial in many cases. Based on recent experiments with messenger RNA detection in dried blood stains [12], our objective in this study was to establish an alternative to cytology by developing a sensitive molecular assay which enables scientists to prove the presence of spermatozoa by detection of specific gene expression.

Materials and methods

Fresh ejaculates (n = 18) or semen stains (n = 48) were obtained from healthy, non-vasectomised volunteers. From each ejaculate at least 6 dilution series up to 10^{-5} were prepared with sterile phosphate buffered saline (PBS) following liquefaction at 37° C. Aliquots of 1µl of semen and of each dilution were pipetted onto a clean white cotton fabric, air-dried and stored for 2–30 days. Smears of each sample were made on glass slides. Further smears were generated using selected dried semen stains at different dilutions which were eluted in phosphate buffered saline.

Vaginal swabs without prior intercourse, buccal swabs, venous blood and menstrual blood were obtained from healthy volunteers, various tissue samples (vaginal swabs, blood, muscle, skin, lung, liver) from 16 autopsies (mean age 28.4 ± 7.3 years, range 16-39 years, mean postmortem time 21.3 ± 9 h). Brain tissue samples (parietal cortex) were taken from 10 different autopsies (mean age 42 ± 12.4 years, range 20-76 years, mean postmortem time 18.4 ± 4.3 h). The samples were air-dried (swabs) or immediately frozen and stored at -40° C.

Cytological smears were stained with hematoxylin-eosin and examined with $400 \times$ magnification for the presence and number of spermatozoa.

RNA from dried stains and swabs was extracted as described previously using the RNAgents total RNA isolation system (Promega, Madison, WI) [12]. The stains were cut out and placed into 150 µl denaturing solution containing guanidinium isothiocyanate. For swab extractions the whole cotton head of the swab was removed and put into a 1.5 ml tube with 500 µl denaturing solution. From undiluted semen stains, pieces with an average size of 1×1 mm were cut out. The same volumes of lysis solution were used with the Absolutely RNA Nanoprep kit (Stratagene, La Jolla, CA) and the extraction steps were performed as recommended by the manufacturer. For monitoring RNA isolation and reverse transcription and for enhancing RNA recovery, 10⁴ copies of a commercially available modified RNA homologous to the endoplasmic reticulum-bound form of mouse cyclophilin (RT-Check, Ambion, Austin, TX), a well conserved and ubiquitously expressed housekeeping gene [13] were added to the denaturing solution before phenolchloroform extraction (RNAgents total RNA isolation system) or before transfer to the spin cup (Absolutely RNA Nanoprep kit). With both protocols the final RNA solution had a volume of $10 \,\mu$ l.

For RNA isolation from tissue samples the RNAgents total RNA isolation system (Promega) was used. The samples were cut into small pieces (about 100–150 mg), placed into an adequate volume of denaturing solution, solubilised with a glass homogeniser and homogenised further by pipetting up and down through a 1-gauge syringe. After acid extraction with phenol/chloroform and precipitation with isopropanol, the precipitates were dissolved in ribonuclease-free water. The concentration was measured spectrophotometrically and the RNA integrity was assessed by running on non-denaturing 1.2% agarose gels. Sharp and distinct bands for 28 S and 18 S ribosomal RNA with higher intensity of the 28 S band indicated that no significant RNA degradation had occurred.

From the RNA solution obtained from dried stains and swabs, $6.8 \,\mu$ l was used for reverse transcription with Sensicript reverse transcriptase (Qiagen, Hilden, Germany), an enzyme adapted to low RNA concentrations, according to the manufacturers instructions in a total reaction volume of $10 \,\mu$ l. Of the total RNA extracted from tissue samples and body fluids 1 μ g was used for reverse transcriptase (Gibco Life Technologies, Karlsruhe, Germany) at 37°C for 60 min and in-activation at 95°C for 5 min in a total volume of 20 μ l.

PCR was performed with 1 μ l of the RT mix in a total reaction volume of 20 μ l. Primer sequences [14] were:

- GCC AGG TAC AGA TGC TGT CGC AG (protamine 1-forward)
- GTG TCT TCT ACA TCT CGG TCT G (protamine 1-reverse)
- TAC CTG GGG CGG CAG CAC C (protamine 1-nested)
- GTC CGA TAC CGC GTG AGG AGC CTG (protamine 2-forward)
- GCC TTC TGC ATG TTC TCT TCC TGG (protamine 2-reverse)
- CGC TCG CAC GAG GTG TAC AGG CAG (protamine 2nested)
- CCA TCG TGT AAT CAA GGA CTT CAT (cyclophilin forward)
- CTT GCC ATC CAG CCA GGA GGT CTT (cyclophilin reverse).

Protamine 1-forward, protamine 2-reverse and cyclophilin forward primers were modified with 6-FAM. The expected amplification product sizes were 150 bp (protamine 1), 128 bp (protamine 1 semi-nested), 300 bp (protamine 2), 270 bp (protamine 2 semi-nested) and 192 bp (cyclophilin). The primers for both protamine 1 and 2 were designed to span the intron contained in these genes so that DNA amplification would generate longer products (protamine 1: 242 bp, protamine 2: 463 bp).

For PCR 0.7 Units of AmpliTaq Gold (Applied Biosystems, Foster City, CA) was used. PCR conditions were as follows: 53 cycles, annealing temperature 60°C and magnesium concentration 1.5 mmol/l. The reactions were run on a Perkin Elmer 2400 cycler. For semi-nested PCR, protamine 1-forward and nested primers and protamine 2-reverse and nested primers were used, respectively in two amplification rounds, each with 35 cycles. Negative controls for the RNA extraction level included samples with sperm-free cotton fabric and samples containing only the isolation reagents, on the RT level samples without RNA and without reverse transcriptase and on the PCR level samples without cDNA and with low tissue concentration cDNA (e.g. 50 ng blood RNA in 10 µl RT reaction volume) for the semi-nested PCR.

Amplification products were visualised with ethidium bromide-stained 2.5% agarose gels and with laser-induced fluorescence capillary electrophoresis (ABI Prism 310 genetic analyzer) using Genescan analysis 2.0 software.

Results

The amplification of an artificial RNA homologous to the house-keeping gene cyclophilin served as control for RNA extraction, reverse transcription and amplification. Failure of cyclophilin amplification thus indicated loss of RNA during sample processing in the laboratory due to pipet-ting errors or exogenous ribonuclease contamination. This occurred in less than 5% of all samples. Only samples in which no in-vitro RNA degradation or RNA loss had occurred as demonstrated by positive cyclophilin PCR reaction were considered for this study.

All sperm-free samples, including blood, vaginal swabs, saliva and tissues taken at autopsy, were always negative for both protamine 1 and protamine 2 RT-PCR with hotstart PCR. The use of semi-nested PCR generally confirmed these results, however, a considerable percentage of reactions (up to 30%) showed protamine amplification without any tissue preference. When the amount of RNA used for reverse transcription was reduced from 1 μ g to 50 ng, the rate of positive protamine amplification results dropped below 10%.

The approximate sperm concentrations of the ejaculates were $40-90\times10^{6}$ /ml and therefore in the range of normospermia.



Fig. 1 Agarose gel showing amplification results for protamine 1 and 2 RT-PCR (*Lanes: L* ladder, $1 \ 1 \ \mu$ l of undiluted semen stain, 2 1 \ \mu l of a 1:10 dilution, $3 \ 1 \ \mu$ l of a 1:100 dilution, $4 \ 1 \ \mu$ l of a 1:1,000 dilution, $5 \ 1 \ \mu$ l of a 1:10,000 dilution, $6 \ negative \ control$)



Fig. 2 Agarose gel showing amplification results for protamine 1 hot-start (**A**) and semi-nested (**B**) PCR and for control PCR with cyclophilin primers (**C**) *Lanes*: *L* ladder, *I* undiluted semen stain, 2 semen stain diluted 1:10, 3 semen stain diluted 1:100, 4 semen stain diluted 1:2,500, 5 semen stain diluted 1:500, 6 semen stain diluted 1:1,000, 7 semen stain diluted 1:2,500, 8 negative control, 9 undiluted semen stain, 1 µl without cyclophilin mRNA, *11* positive control, undiluted semen stain, 1 µl with cyclophilin mRNA,

In undiluted semen stains both protamine mRNAs could always be detected even in 1 mm² pieces independent of the storage duration (2–30 days, room temperature).

In the diluted semen stains RT-PCR for protamine 1 and protamine 2 was always positive in samples diluted 1:10 (Fig. 1). With more diluted semen samples (1:100) results were inconsistent with amplification failure in about 30% (protamine 1) and 35% (protamine 2) of all reactions. Dilutions higher than 1:100 always tested negative.

No differences were seen between samples isolated with the traditional single-step technique (RNAgents total RNA isolation system, Promega) and with a spin columnbased isolation system adapted to low elution volumes (Absolutely RNA Nanoprep Kit, Stratagene).

With semi-nested PCR positive amplification results for both protamine 1 and 2 could be obtained with dilutions up to 1:1,000 (Fig. 2). However, contamination or artefacts due to high cycle numbers occurred in 25% of all experiments with bands appearing in both negative controls and highly diluted blood samples.

Capillary electrophoresis produced clear, distinct peaks even in cases where the bands on the agarose gels were only faintly discernible (Fig. 3).

Microscopic examination of the slides obtained by elution from dried semen stains revealed that spermatozoa could not be identified in stains at a dilution of 1:10 and less and in the majority of the undiluted stains with $1 \mu l$ sample amounts.

Discussion

The presence of mRNA in mature spermatozoa has been confirmed by many laboratories since 1988 [15]. It was suggested that the set of mRNAs for sperm-specific nucleoproteins, transition protein TP-1, ß-actin, c-myc, ß1-integrin and others in mature spermatozoa represents the remnants of stored mRNA from post-meiotically transcribed genes [16]. Recently short interspersed repeat-like sequences were described as part of the spermatozoal RNA [17]. During spermatogenesis histones are replaced by protamines which are the basic nuclear proteins of mature spermatozoa and are essential for the unique, highly condensed chromatin structure in the sperm nucleus [18]. Transcripts corresponding to human protamine 1 and protamine 2 which are among the most rapidly diverging proteins [19], were identified in epididymal as well as in ejaculated spermatozoa [20]. Due to the high interspecies variability, protamine gene sequencing was suggested for identification of the human origin of biological specimens [21] even if other methods are used for this purpose today. In humans no variations are present in the mRNA and protein sequence despite one polymorphic site in the coding region of the DNA. Thus protamine mRNA has to be considered as a suitable marker for the presence of spermatozoa in forensic samples.

Protamines are believed to be exclusively expressed from the haploid genome during spermatogenesis [20]. Although in experiments with expressed sequence tag analysis which are accessible via internet, cDNA derived from protamine 1 and 2 mRNA could be detected in brain tissue (http://genome-www5.stanford.edu/cgi-bin/SMD/source/

Fig. 3 Capillary electrophoresis electropherogramm showing peaks for protamine 1, cyclophilin and protamine 2 from diluted semen stains (same sample as on lane 2 in Fig. 2). Differences to the expected product size are due to the addition of fluorescent marker molecules to the primers (FAM)



sourceSearch), this was not confirmed by microarray gene expression analysis where weak protamine expression was found in the trachea but not in various brain areas [22]. Even if protamines are expressed in extragonadal tissues such as brain or trachea this would obviously not be relevant for forensic purposes. Our experiments with autopsy material including brain never showed any protamine gene expression even with semi-nested PCR when the samples were diluted prior to reverse transcription.

Since experimental data concerning protamine expression in extragonadal tissues and cells which might interfere with the examination of semen stains in forensic casework were lacking, we investigated protamine expression in various tissue samples and body fluids which included vaginal and buccal swabs, vaginal and endometrial tissue samples, saliva, blood and menstrual blood. The unequivocal result was that protamine expression could not be detected in these tissues and cells thus confirming that protamine mRNA is indeed exclusively expressed in spermatozoa and that cells which do not use protamine as nucleoprotein obviously do not transcribe the gene regularly.

In previous experiments we have shown that isolation of mRNA from dried blood stains suitable for RT-PCR is possible with a modified RNA extraction protocol [12] and that epithelial and endometrial cells may be identified in blood stains using cell-specific and tissue-specific mRNAs as markers [23]. In this study we were able to demonstrate that mRNA can be extracted from dried semen stains with slight modifications of this protocol. Recovery of RNA from dried stains could be enhanced by adding modified RNA molecules during the isolation procedure which improves the precipitation of RNA molecules present in low copy numbers by the mechanism of co-precipitation. Since this RNA was homologous to the ubiquitously expressed house-keeping gene cyclophilin and could be reverse-transcribed, PCR with primers derived from this cyclophilin sequence could be used as control for the efficacy of RNA isolation and reverse transcription.

RT-PCR with primers specific for protamine 1 and protamine 2 resulted in successful amplification of samples containing only 100-1,000 spermatozoa. The use of seminested PCR lowered the detection limit to 10-100 cells which corresponds to 0.01–0.001 µl of semen and is comparable with the PSA test [6]. Although in forensic casework this sensitivity level might not always be realised due to the significantly higher contamination risk of seminested PCR techniques, simple, non-nested hot-start PCR should be sufficient to handle routine cases with semen stains on clothes or other fabrics because our experiments have shown that undiluted semen stains as small as 1 mm² could be consistantly identified by RT-PCR. Since laserinduced fluorescence capillary electrophoresis has a lower detection limit for labeled DNA than ethidium bromide staining the use of this visualisation method can improve sensitivity by providing clearly distinguishable peaks in cases in which agarose gel electrophoresis did not allow the diagnosis of positive amplification.

A set of negative controls covering every step is necessary to ensure contamination-free reactions if semi-nested PCR is used. This control set should include highly diluted cDNA from sources other than semen, for example blood. With simple hot-start PCR such extensive precautions are not required since false-positive results were never seen in our experiments.

Negative results have to be interpreted cautiously because unfavourable storage conditions could lead to RNA degradation and to loss of protamine mRNA below the detection limit even when spermatozoa were initially present. This problem however is present in other forensic methods and even negative Y-chromosomal DNA amplification results e.g. from vaginal swabs, do not exclude the possibility that spermatozoa were initially present although DNA generally is more stable than RNA. When storage conditions are known and monitoring of the analysis steps by cyclophilin RT-PCR has confirmed that during RNA isolation and reverse transcription no degradation has occurred, then negative amplification results for protamine RT-PCR can be considered as strong evidence for the absence of spermatozoa.

Although more sophisticated techniques for the analysis of gene expression such as microarray gene expression analysis [22] or real-time RT-PCR, which was recently introduced into forensic science [24] are available, we feel that these methods do not suit the needs of forensic analysis at present because they are expensive, require timeconsuming data analysis and do not provide clear-cut and easily interpretable results. As long as the detection of one or two mRNAs is sufficient and no background gene expression in extragonadal cells has to be considered, simple RT-PCR is the method of choice for this purpose.

The molecular detection of sperm-specific mRNA might be useful in cases in which identification of spermatozoa is required but was not possible or not successful using microscopy due to limited sample amount or insufficient morphological preservation. When small dried semen stains with low sperm concentration are eluted to prepare a smear, spermatozoa might not be detectable by microscopy because they are morphologically not discernible from artefacts such as cotton fibre fragments or because the sperm heads were destroyed during storage or elution. We do not know how long mRNA is detectable when it is released from the cell but the sample rehydration in guanidinium isothiocyanate-containing solution prevents any further RNA degradation. This might explain in part why protamine mRNA in our study was detectable in samples which, after elution in phosphate buffered saline, did not allow microscopic identification of spermatozoa.

An important benefit of this method is the possibility of simultaneous RNA and DNA analysis from the same specimen (publication in preparation) so that in the case of very small sample amounts, no valuable material has to be sacrificed for microscopy.

The sensitivity of the test is high due to the amplification techniques involved. This might also be interesting for the analysis of vaginal swabs for the presence of spermatozoa in sexual assaults. In some cases microscopic examination is not successful because the sperm concentration is low and the smear which represents only part of the material present on the swab might fail to contain a sufficient number of identifiable sperm heads. In such cases the combined RNA and DNA analysis using the swab might be helpful. However, further investigations are necessary to show the practical relevance of this technique.

The shortcomings of this new method include the higher degradation sensitivity of RNA compared with DNA and the need for time-consuming laboratory work. Therefore it rather represents a molecular alternative or supplement to cytology and cannot replace immunological or enzymatic membrane tests such as PSA or acid phosphatase which are easy to handle and provide immediate results. However, the prerequisites of RNA work are present in every DNA laboratory and since RNA research is emerging in many fields of forensic science such as wound healing [25], there will be no general obstacles to establishing RNA analysis as a routine method in the future.

In conclusion our results show that the detection of protamine gene expression in dried semen stains provides evidence for the presence of spermatozoa. Due to the nucleic acid structure of RNA, simultaneous DNA analysis is possible according to preliminary results obtained in our laboratory. Identification of spermatozoa as well as genetic profiling of the perpetrator would thus be possible in a very elegant one-step molecular analysis.

References

- Tamaki K, Fujisawa K, Okajima H, Sato K, Katsumata J (1989) Identification of semen in stain by determination of the specific activity of L-tartrate-inhibitable acid phosphatase. Z Rechtsmed 102:429–436
- Abe S, Gunji H, Kunii S, Kuraya M, Fujita T, Hiraiwa K (1999) Semen specific γ-glutamyltransferase carries ABH antigens: a sandwich ELISA for simultaneous semen detection and its AB0 blood typing. Clin Chim Acta 283:183–194
- 3. Sato I, Yoshiike M, Yamasaki T, Yoshida K, Takano S, Mukai T, Iwamoto T (2001) A dot-blot-immunoassay for semen identification using a polyclonal antibody against semenogelin, a powerful seminal marker. Forensic Sci Int 122:27–34
- Abe S, Kunii S, Fujita T, Hiraiwa K (1998) Detection of human seminal γ-glutamyl transpeptidase in stains using sandwich ELISA. Forensic Sci Int 91:19–28
- Simich JP, Morris SL, Klick RL, Rittenhouse-Diakun K (1999) Validation of the use of a commercially available kit for the identification of prostate specific antigen (PSA) in semen stains. J Forensic Sci 44:1229–1231
- Hochmeister MN, Budowle B, Rudin O, Gehrig C, Borer U, Thali M, Dirnhofer R (1999) Evaluation of prostate-specific antigen (PSA) membrane test assays for the forensic identification of seminal fluid. J Forensic Sci 44:1057–1060
- Keil W, Bachus J, Tröger HD (1996) Evaluation of MHS-5 in detecting seminal fluid in vaginal swabs. Int J Legal Med 108: 186–190
- Steinman G (1995) Rapid spot tests for identifying suspected semen specimens. Forensic Sci Int 72:191–197

- Chen JT, Hortin GL (2000) Interferences with semen detection by an immunoassay for a seminal vesicle-specific antigen. J Forensic Sci 45:234–235
- Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 110:118–124
- Sibille I, Duverneuil C, Lorin de la Grandmaison G, Guerrouache K, Teissiere F, Durigon M, Mazancourt P de (2002) Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa. Forensic Sci Int 125:212–216
- Bauer M, Kraus A, Patzelt D (1999) Detection of epithelial cells in dried blood stains by reverse transcriptase-polymerase chain reaction. J Forensic Sci 44:1232–1236
- Hasel KW, Glass JR, Godbout M, Sutcliffe JG (1991) An endoplasmic reticulum-specific cyclophilin. Mol Cell Biol 11: 3484–3491
- 14. Domenjoud L, Nussbaum G, Adham IM, Greeske G, Engel W (1990) Genomic sequences of human protamines whose genes, PRM1 and PRM2, are clustered. Genomics 8:127–133
- Kramer JA, Krawetz SA (1997) RNA in spermatozoa: implications for the alternative haploid genome. Mol Hum Reprod 3: 473–478
- 16. Miller D (1997) RNA in the ejaculate spermatozoon: a window into molecular events in spermatogenesis and a record of the unusual requirements of haploid gene expression and post-meiotic equilibration. Mol Hum Reprod 3:669–676
- Miller D, Briggs D, Snowden H, Hamlington J, Rollinson S, Lilford R, Krawetz SA (1999) A complex population of RNAs exists in human ejaculate spermatozoa: implications for understanding molecular aspects of spermiogenesis. Gene 237:385– 392
- Steger K, Pauls K, Klonisch T, Franke FE, Bergmann M (2000) Expression of protamine-1 and -2 mRNA during human spermiogenesis. Mol Hum Reprod 6:219–225
- Retief JD, Winkfein RJ, Dixon GH, Adroer R, Queralt R, Ballabriga J, Oliva R (1993) Evolution of protamine P1 genes in primates. J Mol Evol 37:426–434
- Wykes SM, Visscher DW, Krawetz SA (1997) Haploid transcripts persist in mature human spermatozoa. Mol Hum Reprod 3:15–19
- Queralt R, Fabregues-Boixar O de, Adroer R, Gene M, Gomez-Catalan J, Huguet E, Oliva R (1993) Direct sequencing of the human protamine P1 gene and application in forensic medicine. J Forensic Sci 38:1491–1501
- 22. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, Patapoutian A, Hampton GM, Schultz PG, Hogenesch JB (2002) Largescale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 99:4465–4470
- Bauer M, Patzelt D (2002) Evaluation of mRNA markers for the identification of menstrual blood. J Forensic Sci 47:1278– 1282
- 24. Ishida K, Zhu BL, Maeda H (2000) Novel approach to quantitative reverse transcription PCR assay of mRNA component in autopsy material using the TaqMan fluorogenic detection system: dynamics of pulmonary surfactant apoprotein A. Forensic Sci Int 113:127–131
- 25. Sato Y, Ohshima T (2000) The expression of mRNA of proinflammatory cytokines during skin wound healing in mice: a preliminary study for forensic wound age estimation (II). Int J Legal Med 113:140–145