J Forensic Sci, November 2007, Vol. 52, No. 6
doi: 10.1111/j.1556-4029.2007.00550.x
Available online at: www.blackwell-synergy.com

Jane Juusola,[1,2,3] *Ph.D. and Jack Ballantyne,*[1,2,3] *Ph.D.*

mRNA Profiling for Body Fluid Identification by Multiplex Quantitative RT-PCR*

ABSTRACT: An alternative approach to conventional protein-based body fluid identification is gene expression profiling analysis. In the present work, we report the development of sensitive and robust multiplex quantitative reverse transcriptase-PCR assays for the identification of blood, saliva, semen, and menstrual blood. Each body fluid assay comprises a triplex system that detects transcripts from two body fluid-specific genes and a housekeeping gene GAPDH. The body fluid-specific genes include erythroid δ-aminolevulinate synthase (ALAS2) and β-spectrin (SPTB) for blood, statherin (STATH) and histatin 3 (HTN3) for saliva, protamine 1 (PRM1) and protamine 2 (PRM2) for semen, and matrix metalloproteinase 7 (MMP7) and matrix metalloproteinase 10 (MMP10) for menstrual blood. Normalization of both body fluid-specific genes to the housekeeping gene by means of appropriate cycle threshold metrics ensures the high specificity of each assay for its cognate body fluid.

KEYWORDS: forensic science, RNA, mRNA profiling, body fluid identification, blood identification, saliva identification, semen identification, menstrual blood identification, multiplex real-time PCR

Messenger RNA profiling is likely to play a major role in the future of forensic biochemistry, not only for the identification of body fluids and tissues, which is the focus of the present work, but also in the determination of the age of an individual (a stain donor) (1), as well as the age of a stain (time since deposition) (2).

For body fluid and tissue identification, mRNA profiling is based upon the premise that each single tissue type is comprised of cells that have a unique transcriptome or gene expression (i.e., mRNA) profile. However, as many body fluids and tissues are made up of multiple different types of cells, forensic scientists will be mostly interested in the multicellular transcriptome concept, which refers to the collection of genes that are expressed within the constellation of differentiated cells that makes up a particular body fluid (3). These genes include ubiquitously expressed housekeeping genes, which are responsible for cell maintenance functions and typically are expressed in all cell types, and tissue-specific genes that are uniquely expressed in certain tissues or body fluids only.

Conventional biochemical and immunological tests for body fluid identification are protein based, performed in a series, not parallel, manner, and are therefore costly in terms of time and sample. In addition, there are no confirmatory tests, for example, for saliva and menstrual blood. The potential advantages of a mRNA-based assay include the ability to co-extract DNA and RNA, greater specificity compared with conventional methods, the ability to simultaneously analyze multiple markers and tissue types through a common assay format, and the potential for automation, all of which can save time and preserve sample.

Although a variety of other platforms are also available for mRNA analysis, such as microarray analysis and traditional polymerase chain reaction (PCR) coupled with capillary electrophoresis (CE) analysis, the present work describes assays for body fluid identification using multiplex real-time quantitative PCR (qPCR). qPCR technology has wide applications in DNA quantitation and gene expression analyses and many types of qPCR assays have been developed. The real-time PCR instrument is capable of simultaneously detecting four different fluorescent dyes, one of which can be used for an internal positive control, such as a housekeeping gene, and one of which is an internal reference dye, leaving two dyes for two tissue-specific genes of choice. As a result, multiplexes can be developed to detect up to two body fluid- or tissue-specific genes and one housekeeping control gene simultaneously. The ability of qPCR to quantitate target sequences is important in establishing the tissue-specificity of a gene product, particularly when the relative abundance of a number of different mRNA molecules can demonstrate a unique or restricted pattern of expression.

In this work, using genes that we previously characterized as possessing a restricted pattern of gene expression, we report the development of sensitive and robust quantitative reverse transcriptase-PCR assays (qRT-PCR) for the identification of blood, saliva, semen, and menstrual blood.

Materials and Methods

Body Fluid Samples

Body fluids were collected from healthy individuals using procedures approved by the University's Institutional Review Board. Blood was collected by venipuncture and 50 μL aliquots placed onto cotton cloth and dried at room temperature. Seventeen individuals donated blood samples, including nine females (age range: 14 months to 69 years, and umbilical cord blood) and eight males (19 to 80 years old). Saliva was collected in a microcentrifuge tube and 50 μL aliquots placed onto cotton cloth and dried at room temperature. Seven individuals donated saliva samples, including five females (2 to 33 years old) and two males (27 and 32 years old). Freshly ejaculated semen was collected in plastic cups, and then

[1]Graduate Program in Biomolecular Sciences, University of Central Florida, Orlando, FL.
[2]Department of Chemistry, University of Central Florida, Orlando, FL.
[3]National Center for Forensic Science, University of Central Florida, Orlando, FL.

*Presented in part at SAFS, MAFS, MAAFS, CFSF Joint Meeting in Orlando, FL, September 23, 2004, and Promega's 16th International Symposium on Human Identification in Grapevine, TX, September 29, 2005.

Received 24 Dec. 2006; and in revised form 4 June 2007, and 6 Jan. 2007; accepted 1 June 2007; published form 21 Dec. 2007.

aliquoted onto cotton cloth in known volumes and dried at room temperature or allowed to dry onto cotton swabs at room temperature. Six male individuals (24 to 37 years old) donated semen samples. Semen-free vaginal secretions (mid to late cycle, days 14–25) and menstrual blood (early cycle, days 1–3) were collected using sterile cotton-tipped swabs. Six female individuals (22 to 33 years old) donated menstrual blood swabs, including three individuals who were taking a hormonal oral contraceptive and three individuals who were not. The same six individuals and one pregnant individual donated semen-free mid-cycle vaginal swabs. In general and unless otherwise indicated, a 50 μL stain or a single cotton swab was used for RNA isolation.

Poly-A RNA from human skin, and total RNA from human muscle, adipose, and brain tissue samples were obtained from BD Biosciences–Clontech (Mountain View, CA).

RNA Isolation and Quantitation

Total RNA was extracted from blood, saliva, and semen stains, and vaginal swabs by previously described methods (3,4). RNA extracts were treated with DNase I (Ambion, Inc., Austin, TX) as previously described (3,4) and quantitated with Quant-iT™ RiboGreen® RNA Kit (Molecular Probes, Eugene, OR) according to the manufacturer's protocol for the high-range assay (5). Fluorescence was determined using a Wallac Victor^2 microplate reader (Perkin Elmer Life Sciences, Boston, MA).

cDNA Synthesis

For the reverse transcriptase (RT) reaction, total RNA template and nuclease-free water were combined to a final volume of 15 μL, and heated at 75°C for 3 min to eliminate target mRNA secondary structure then snap-cooled on ice. To the denatured RNA, 6 μL of dNTP mix (10 mM) (Applied Biosystems, Foster City, CA), 3 μL random decamers (50 μM) (Ambion, Inc.), 3 μL of 10X first-strand buffer (500 mM Tris-HCl pH 8.3, 750 mM KCl, 30 mM $MgCl_2$, and 50 mM dithiothrertol [DTT]) (Ambion, Inc.), 1.5 μL of SUPERase-In™ RNase Inhibitor (20 U/μL) (Ambion, Inc.), and 1.5 μL of Moloney Murine Leukemia Virus-RT (100 U/μL) (Ambion, Inc.) were added to yield a final reaction volume of 30 μL. This reaction mixture was incubated at 42°C for 1 h, and then at 95°C for 10 min to inactivate the RT. In general, the following amounts of RNA were used in the RT reaction: for blood, saliva, semen, and menstrual blood, 50 ng of total RNA, for vaginal secretions, 100 ng of total RNA, for muscle, adipose, and brain, 25 ng of total RNA, and for skin, 5 ng of Poly-A RNA. For the sensitivity studies, the total RNA input ranged from 20 fg to 500 ng into the 30 μL RT reaction.

Standard End-Point PCR

End-point PCR was performed using the GeneAmp® PCR System 9700 (Applied Biosystems). Two microliters of the RT-reaction were amplified in a total reaction volume of 25 μL. The standard reaction mixture contained buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, and 1.5 mM $MgCl_2$), 0.5 mM dNTP mix (Applied Biosystems), 0.4 μM of each PCR primer, and 1.25 U AmpliTaq Gold® DNA polymerase (Roche Molecular Systems, Inc., Branchburg, NJ). PCR primer sequences for erythroid δ-aminolevulinate synthase (ALAS2) (forward primer, F: 5′-TGTGTCCGTCTGGTGTAGTA; reverse primer, R: 5′-AAACTTACTGGTGCCTGAGA), β-spectrin (SPTB) (4), histatin 3 (HTN3)

TABLE 1—*Real-time PCR primer and probe sequences and concentrations in triplex assays.*

Body Fluid	Gene	Primer (F and R) and Probe (P) Sequences/dyes	nM in triplex			
			BL	SA	SE	MB
Housekeeping Gene	GAPDH	F: 5′-ATGGAAATCCCATCACCATCTT	300	600	900	150
		R: 5′-CGCCCCACTTGATTTTGG	300	600	900	150
		P: 5′-NED-CAGGAGCGAGATCC	75	150	250	75
Blood	ALAS2	F: 5′-GCCGACACCCTCAGGTCTT	900	–	–	–
		R: 5′-GAAACTTACTGGTGCCTGAGATGTT	900			
		P: 5′-VIC-AAGCCACACAGGAGAC	250			
	SPTB	F: 5′-GCCTTTAATGCCCTGATACACAA	900	–	–	–
		R: 5′-GAGTCCTTCAGCTTATCAAAGTCGAT	900			
		P: 5′-FAM-CACCGGCCCGACCT	200			
Saliva	HTN3	F: 5′-CTTGGCTCTCATGCTTTCCAT	–	900	–	–
		R: 5′-TTTATACCCATGATGTCTCTTTGCA		900		
		P: 5′-FAM-ACTGGAGCTGATTCAC		250		
	STATH	F: 5′-TCTTGGCTCTCATGGTTTCCA	–	900	–	–
		R: 5′-CCAATTCTACGCAAAAATTTCTCTT		900		
		P: 5′-VIC-ATTGGAGCTGATTCATC		250		
Semen	PRM1	F: 5′-CAGATATTACCGCCAGAGACAAAG	–	–	900	–
		R: 5′-AATTAGTGTCTTCTACATCTCGGTCTGT			900	
		P: 5′-FAM-CAGCACCTCATGGCT			250	
	PRM2	F: 5′-GGCGCAAAAGACGCTCC	–	–	900	–
		R: 5′-GCCCAGGAAGCTTAGTGCC			900	
		P: 5′-VIC-TTCTGCAGCCTCTGCGAT			250	
Menstrual blood	MMP-7	F: 5′-GGGAGGCATGAGTGAGCTACA	–	–	–	900
		R: 5′-TGGCATTTTTTGTTTCTGAGTCATA				900
		P: 5′-FAM-AACAGGCTCAGGACTAT				250
	MMP-10	F: 5′-TGGTCACTTCAGCTCCTTTCCT	–	–	–	900
		R: 5′-AATGGCAGAATCAACAGCATCTC				900
		P: 5′-VIC-CACCTTACATACAGGATTG				250

F, forward primer; R, reverse primer; P, probe; BL, blood; SA, saliva; SE, semen; MB, menstrual blood; ALAS, erythroid δ-aminolevulinate synthase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HTN3, histatin 3; MMP, matrix metalloproteinase; PCR, polymerase chain reaction; PRM, protamine; SPTB, β-spectrin; STATH, statherin.

(3,4), statherin (STATH) (3,4), MMP-7 (4), and MMP-10 (F: 5′-ACAGGGAAGCTAGACACTGA; R: 5′-CTGGAGAATGTGAGTGGAGT) were designed using Primer 3 Online primer design software (6). PCR Primer sequences for PRM1 and PRM2 were obtained from published sources (4,7). Primers were custom synthesized by Life Technologies (Grand Island, NY). The standard PCR conditions used for all traditional PCR consisted of a denaturing step (95°C, 11 min) followed by 35 cycles (94°C, 20 sec; 55°C, 30 sec; and 72°C, 40 sec) and a final extension step (72°C, 5 min).

Polymerase chain reaction products were separated on 4% agarose gels and stained with SYBR® Gold nucleic acid stain (Molecular Probes), as previously described (4). The gels were visualized on the Omega10 Chemiluminescence Imaging System (ULTRA-LUM, Inc., Claremont, CA) and analyzed with ONE-Dscan 2.05, 1-D Gel Analysis Software for Windows (Scanalytics, Inc., Fairfax, VA).

Real-Time PCR Primer and Probe Design

Real-time PCR primer and minor-groove binding (MGB®) probe sequences were designed using Primer Express Software™ Version 2.0 (Applied Biosystems). Real-time PCR primer pairs were designed to span at least one exon–exon boundary and be 18–22 bases in length, but needed to function with high annealing temperatures (58–60°C) and short mRNA/cDNA amplicons (50–200 bp). The real-time PCR probe for each gene was designed to anneal at the exon–exon boundary enclosed by the primers, be 13–18 bases in length, and have an annealing temperature about 10°C higher than the respective primer pair (68–70°C). A TaqMan® MGB® probe from Applied Biosystems, which has a nonfluorescent quencher, acts as a Tm enhancer, requires fewer nucleotides, and offers higher specificity compared with the standard TAMRA probes, was selected as the probe type. Table 1 lists the real-time primer and probe sequences used during assay development.

Real-Time Polymerase Chain Reaction

Real-time PCR was performed using the Relative Quantitation protocol on an ABI PRISM 7000 Sequence Detection System (Applied Biosystems). Two microliters of the RT-reaction were amplified in a total reaction volume of 25 μL. The standard reaction mixture included proprietary 2X TaqMan® Universal PCR Master Mix, with optimized buffer conditions, dNTPs with dUTP, AmpErase UNG, passive reference, and AmpliTaq Gold DNA Polymerase (Applied Biosystems), and qPCR primers (Invitrogen, Grand Island, NY) and TaqMan® MGB probe (Applied Biosystems). Singleplex real-time PCR assays had a 900 nM concentration for each primer and a 250 nM probe concentration. For the body fluid identification triplex reactions, optimal primer and probe concentrations were determined experimentally and are listed in Table 1. Real-time PCR cycling conditions consisted of an incubation step (50°C, 2 min), a denaturation step (95°C, 10 min) followed by 50 cycles (95°C, 15 sec and 60°C, 30 sec). For data analysis, the threshold was set manually to 0.200 (the default setting).

Results and Discussion

Selection of Body Fluid-Specific Gene Candidates

The following genes were used to develop the multiplex real-time PCR assays for the positive identification of blood, saliva,

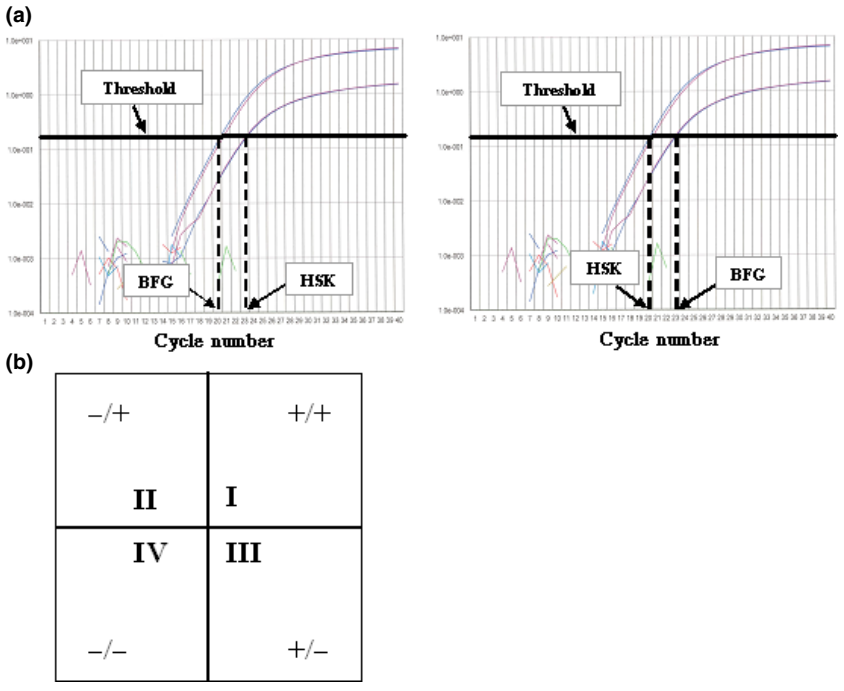

FIG. 1—*Real-time PCR data interpretation and dCt analysis. For the interpretation of multiplex real-time PCR results, we chose to use a delta Ct (dCt) calculation. (a) A Ct value for each gene is the cycle number at which the fluorescence signal, generated by the amplification of that gene, passes a preset threshold line (0.200). A dCt can be determined by subtracting the average body fluid gene (BFG) Ct value from the average housekeeping gene (HSK) Ct value. A positive dCt value indicated that a body fluid gene was present at a higher level than the housekeeping gene and that the body fluid being tested for was present (left panel). A negative dCt value indicated that the body fluid gene was present at a lower level than the housekeeping gene (right panel) or not detected at all, and therefore the body fluid was not detected. In cases when no Ct value was obtained (''undetermined''), a Ct value equal to the highest cycle number used in the assay, which was 50 cycles, was substituted into the calculation. (b) With two body fluid markers in a single reaction, the dCt values were plotted together in a quadrant plot. Possible results included double positive (+/+), positive/negative (+/−), negative/positive (−/+), and double negative (−/−).*

TABLE 2A—*Ct and dCt data for body fluid identification triplexes.*

	Blood Triplex								Saliva Triplex							
	Ct						dCt		Ct						dCt	
	SPTB		ALAS2		GAPDH				HTN3		STATH		GAPDH			
RT	+	−	+	−	+	−	GAPDH-SPTB	GAPDH-ALAS2	+	−	+	−	+	−	GAPDH-HTN3	GAPDH-STATH
BL A	28.4	U	26.3	U	39.8	45.4	**11.4**	**13.5**	U	U	U	U	30.4	44.3	−19.6	−19.6
BL B	27.0	U	25.6	U	44.3	U	**17.3**	**18.7**	U	U	U	U	30.0	U	−20.0	−20.0
BL C	27.6	U	26.1	U	37.4	U	**9.8**	**11.3**	U	U	U	U	28.5	42.3	−21.5	−21.5
BL D	27.8	U	27.1	U	U	47.1	**22.2**	**22.9**	U	U	U	U	29.6	43.1	−20.4	−20.4
BL E	26.2	U	24.2	U	U	U	**23.8**	**25.8**	U	U	U	U	29.7	44.9	−20.3	−20.3
BL F	26.5	U	26.4	U	43.6	48.7	**17.1**	**17.2**	U	U	U	U	28.5	41.3	−21.5	−21.5
BL G	24.3	U	23.0	U	U	48.0	**25.7**	**27.0**	U	U	U	U	28.6	U	−21.4	−21.4
BL H	29.0	U	26.8	U	U	U	**21.0**	**23.2**	U	U	U	U	29.7	U	−20.3	−20.3
BL I	27.5	U	26.8	U	U	45.0	**22.5**	**23.2**	U	U	U	U	28.6	41.3	−21.4	−21.4
BL J	27.1	U	25.7	U	37.1	U	**10.0**	**11.4**	U	U	U	U	29.4	43.3	−20.6	−20.6
BL K	28.0	U	26.7	U	U	U	**22.0**	**23.3**	U	U	U	U	29.7	U	−20.3	−20.3
BL L	28.1	U	26.6	U	44.7	U	**16.6**	**18.1**	U	U	U	U	29.9	48.5	−20.1	−20.1
BL M	27.5	U	27.0	U	U	U	**22.5**	**23.0**	U	U	U	U	30.2	U	−19.8	−19.8
BL N	27.2	U	26.5	U	U	U	**22.8**	**23.5**	U	U	U	U	29.5	U	−20.5	−20.5
BL O	26.1	U	24.6	U	U	49.1	**23.9**	**25.4**	U	U	U	U	32.0	U	−18.0	−18.0
BL P	28.4	U	27.7	U	U	U	**21.6**	**22.3**	U	U	U	U	29.6	43.1	−20.4	−20.4
BL Q	27.9	U	25.2	U	U	44.7	**22.1**	**24.8**	U	U	U	U	30.2	41.6	−19.8	−19.8
SA A	U	U	U	U	45.6	U	−4.4	−4.4	34.2	U	36.7	U	39.8	42.7	**5.6**	**3.1**
SA B	U	U	U	U	45.0	48.8	−5.0	−5.0	33.1	U	35.9	U	40.1	43.5	**7.0**	**4.2**
SA C	U	U	U	U	42.3	U	−7.7	−7.7	35.9	U	37.4	U	36.4	U	**0.5**	−1.0
SA D	U	U	U	U	43.5	47.4	−6.5	−6.5	35.5	U	37.2	U	38.5	42.0	**3.0**	**1.3**
SA E	U	U	U	U	43.5	47.4	−6.5	−6.5	31.0	U	33.7	U	45.4	41.2	**14.4**	**11.7**
SA F	U	U	U	U	45.1	48.4	−4.9	−4.9	31.8	U	34.1	U	45.5	44.7	**13.7**	**11.4**
SA G	U	U	U	U	U	U	0.0	0.0	34.3	U	37.5	U	43.2	U	**8.9**	**5.7**
SE A	37.2	U	U	U	38.5	42.5	1.3	−11.5	U	U	U	U	32.3	35.8	−17.7	−17.7
SE B	U	U	U	U	41.3	U	−8.7	−8.7	U	U	U	U	34.5	U	−15.5	−15.5
SE C	U	U	U	U	49.9	U	−0.1	−0.1	U	U	U	U	38.9	43.0	−11.1	−11.1
SE D	U	U	U	U	45.5	U	−4.5	−4.5	U	U	U	U	39.3	45.7	−10.7	−10.7
SE E	U	U	U	U	U	U	0.0	0.0	U	U	U	U	36.3	37.3	−13.7	−13.7
SE F	U	U	U	U	41.9	48.9	−8.1	−8.1	39.6	U	43.2	U	34.1	39.8	−5.5	−9.1
VS A	U	U	U	U	35.3	48.5	−14.7	−14.7	U	U	U	U	29.8	41.7	−20.2	−20.2
VS B	37.8	U	U	U	36.3	46.5	−1.5	−13.7	U	U	U	U	31.4	41.2	−18.6	−18.6
VS C	38.5	U	U	U	35.5	48.3	−3.0	−14.5	U	U	U	U	30.2	38.9	−19.8	−19.8
VS D	U	U	U	U	35.5	42.6	−14.5	−14.5	41.1	U	39.8	U	31.1	38.0	−10.0	−8.6
VS E	U	U	U	U	37.3	U	−12.7	−12.7	U	U	U	U	31.6	42.7	−18.4	−18.4
MB A	32.4	U	30.0	U	42.7	U	10.3	12.7	42.4	U	40.5	U	27.7	40.3	−14.7	−12.8
MB B	35.6	U	31.6	U	39.4	46.1	3.8	7.8	U	U	44.6	U	27.3	39.0	−22.7	−17.3
MB C	33.1	U	32.4	U	32.4	42.8	−0.7	**0.0**	U	U	48.7	U	27.3	38.2	−22.7	−21.4
MB D	32.6	U	31.2	U	34.5	44.1	1.9	3.3	U	U	49.2	U	28.4	39.6	−21.6	−20.8
MB E	31.5	U	28.3	U	32.6	47.5	1.1	4.3	U	U	U	U	26.6	38.8	−23.4	−23.4
MB F	33.4	U	30.5	U	32.5	45.4	−0.9	**2.0**	U	U	U	U	25.7	43.0	−24.3	−24.3
SK	32.9	U	40.1	U	33.5	U	**0.6**	−6.6	36.8	U	38.6	U	26.6	U	−10.2	−12.0
MU	26.4	U	U	U	31.1	U	**4.7**	−18.9	U	U	U	U	25.4	U	−24.6	−24.6
AD	31.2	U	34.1	U	33.8	48.5	**2.6**	−0.3	32.0	U	34.4	U	29.7	46.0	−2.3	−4.7
BR	27.8	U	45.2	U	32.8	U	**5.0**	−12.4	U	U	U	U	28.1	U	−21.9	−21.9
DNA	U	U	U	U	40.8	U	−9.2	−9.2	U	U	U	U	31.0	U	−19.0	−19.0

Note: "U" was substituted with a value of 50, the number of amplification cycles, into the dCt calculations.

RT, Reverse Transcriptase; U, undetermined; BL, blood; SA, saliva; SE, semen; VS, vaginal secretions; MB, menstrual blood; SK, skin; MU, muscle; AD, adipose; BR, brain; ALAS, erythroid δ-aminolevulinate synthase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HTN3, histatin 3; SPTB, β-spectrin; STATH, statherin; Ct, cycle threshold; dCt, delta Ct.

semen, and menstrual blood: erythroid δ-aminolevulinate (ALAS2) (8–11) and β-spectrin (SPTB) (4,12,13) for blood, statherin protamine 2 (STATH) (3,4) and histatin 3 (HTN3) for saliva (3,4), protamine 1 (PRM1) (4,7,14) and (PRM2) (4,7,14) for semen, and matrix metalloproteinase 7 (MMP-7) (4,15) and matrix metalloproteinase 10 (MMP-10) (15) for menstrual blood. Specificity was established by demonstrating that the mRNA for the candidate gene was present in one type of body fluid stain but absent from all others (data not shown and [3,4]). The body fluids tested included blood, saliva, semen, mid-cycle vaginal secretions, and menstrual blood. The testing involved the 'monoplex' amplification of gene transcripts by RT-PCR from total RNA isolated from body fluid stains and a subsequent separation and detection of the amplimers by agarose gel electrophoresis (data not shown and [4]). ALAS2 and SPTB were detected in RNA isolated from blood stains and, to a lesser extent, in RNA isolated from menstrual blood, but were undetectable in RNA isolated from saliva, semen, and vaginal secretions stains (data not shown and [4]). STATH and HTN3 were amplified from RNA extracted from saliva stains, but were undetectable in RNA extracts from blood, semen, vaginal secretions, and menstrual blood (3,4). PRM1 and PRM2 were amplifiable from semen stain RNA extracts, but were undetectable in blood, saliva, vaginal secretions, and menstrual blood (4,7,14). MMP-7 and MMP-10 were amplified from menstrual blood RNA, but were

TABLE 2B—*Ct and dCt data for body fluid identification triplexes.*

	Semen Triplex								Menstrual Blood Triplex							
	Ct						dCt		Ct						dCt	
	PRM1		PRM2		GAPDH				MMP-7		MMP-10		GAPDH			
RT	+	−	+	−	+	−	GAPDH-PRM1	GAPDH-PRM2	+	−	+	−	+	−	GAPDH-MMP-7	GAPDH-MMP-10
BL A	U	U	U	U	33.1	U	−16.9	−16.9	U	U	39.6	U	33.3	U	−16.7	−6.3
BL B	U	U	U	U	32.9	U	−17.1	−17.1	U	U	44.1	U	34.1	U	−15.9	−10.0
BL C	U	U	U	U	30.2	U	−19.8	−19.8	U	U	U	U	41.9	U	−8.1	−8.1
BL D	U	U	49.6	U	32.1	U	−17.9	−17.5	U	U	U	U	38.4	U	−11.6	−11.6
BL E	49.9	U	41.6	U	29.5	28.9	−20.4	−12.1	U	U	U	U	37.5	U	−12.5	−12.5
BL F	U	U	U	U	28.0	43.2	−22.0	−22.0	U	U	U	U	37.4	U	−12.6	−12.6
BL G	43.5	U	U	U	27.7	U	−15.8	−22.3	40.7	U	U	U	38.4	U	−2.3	−11.6
BL H	U	U	U	U	30.4	U	−19.6	−19.6	U	U	U	U	42.1	U	−7.9	−7.9
BL I	U	U	44.3	U	28.6	41.2	−21.4	−15.7	40.9	U	U	U	37.9	U	−3.0	−12.1
BL J	U	U	U	U	29.7	U	−20.3	−20.3	41.0	U	U	U	38.7	U	−2.3	−11.3
BL K	U	U	U	U	33.1	U	−16.9	−16.9	U	U	U	U	47.5	U	−2.5	−2.5
BL L	U	U	U	U	33.1	U	−16.9	−16.9	U	U	U	U	39.4	U	−10.6	−10.6
BL M	U	U	U	U	31.0	U	−19.0	−19.0	U	U	U	U	38.0	U	−12.0	−12.0
BL N	U	U	U	U	33.1	U	−16.9	−16.9	U	U	U	U	33.1	U	−16.9	−16.9
BL O	U	U	U	U	36.6	U	−13.4	−13.4	U	U	49.1	U	34.2	U	−15.8	−14.9
BL P	U	U	U	U	30.2	U	−19.8	−19.8	U	U	U	U	33.9	U	−16.1	−16.1
BL Q	U	U	U	U	30.1	U	−19.9	−19.9	U	U	U	U	34.7	46.2	−15.3	−15.3
SA A	U	U	U	U	42.9	U	−7.1	−7.1	U	U	U	U	42.7	U	−7.3	−7.3
SA B	U	U	U	U	45.0	U	−5.0	−5.0	44.0	U	U	U	41.5	46.1	−2.5	−8.5
SA C	U	U	U	U	39.5	U	−10.5	−10.5	U	U	U	U	43.9	U	−6.1	−6.1
SA D	U	U	U	U	39.8	U	−10.2	−10.2	U	U	U	U	42.0	U	−8.0	−8.0
SA E	U	U	U	U	41.2	U	−8.8	−8.8	U	U	46.2	U	41.1	U	−8.9	−5.1
SA F	U	U	U	U	40.4	U	−9.6	−9.6	U	U	41.8	U	44.0	U	−6.0	**2.2**
SA G	U	U	U	U	41.5	U	−8.5	−8.5	U	U	U	U	42.9	U	−7.1	−7.1
SE A	28.7	U	30.4	U	33.1	35.6	**4.4**	**2.7**	U	U	U	U	45.4	45.6	−4.6	−4.6
SE B	38.2	U	34.0	U	34.0	U	−4.2	**0.0**	41.7	U	U	U	37.8	46.9	−3.9	−12.2
SE C	28.4	U	30.0	U	46.2	U	**17.8**	**16.2**	U	U	41.7	U	42.4	U	−7.6	**0.7**
SE D	45.4	U	44.0	U	U	U	4.6	6.0	U	U	U	U	43.6	U	−6.4	−6.4
SE E	32.2	U	34.3	U	35.0	35.3	2.8	0.7	U	U	U	U	47.2	U	−2.8	−2.8
SE F	31.9	U	31.0	U	39.6	48.5	**7.7**	**8.6**	U	U	U	U	45.5	47.2	−4.5	−4.5
VS A	U	U	U	U	30.3	42.1	−19.7	−19.7	U	U	U	U	34.2	43.7	−15.8	−15.8
VS B	U	U	U	U	31.7	41.9	−18.3	−18.3	37.3	U	U	U	36.7	44.4	−0.6	−13.3
VS C	37.9	U	37.3	U	30.5	39.1	−7.4	−6.8	42.2	U	U	U	34.6	44.6	−7.6	−15.4
VS D	46.9	U	41.3	U	30.4	38.4	−16.5	−10.9	35.1	U	42.2	U	35.8	42.7	**0.7**	−6.4
VS E	U	U	U	U	32.6	44.5	−17.4	−17.4	U	U	U	U	35.2	U	−14.8	−14.8
MB A	U	U	U	U	28.1	U	−21.9	−21.9	27.1	U	25.5	U	U	U	**22.9**	**24.5**
MB B	U	U	U	U	27.3	U	−22.7	−22.7	36.3	U	33.1	U	31.1	41.8	−5.2	−2.0
MB C	U	U	U	U	27.9	U	−22.1	−22.1	33.4	U	24.7	U	U	45.9	**16.6**	**25.3**
MB D	U	U	U	U	29.5	48.2	−20.5	−20.5	24.3	U	26.6	U	U	48.2	25.7	23.4
MB E	U	U	U	U	27.0	49.5	−23.0	−23.0	28.3	U	24.3	U	U	48.3	21.7	25.7
MB F	47.2	U	U	U	30.6	U	−16.6	−19.4	29.7	U	25.6	U	U	U	**20.3**	**24.4**
SK	U	U	U	U	31.8	U	−18.2	−18.2	29.3	U	30.2	U	27.9	U	−1.4	−2.3
MU	U	U	U	U	28.9	U	−21.1	−21.1	U	U	41.6	U	28.2	U	−21.8	−13.4
AD	42.8	U	42.9	U	34.2	U	−8.6	−8.7	37.4	U	U	U	33.5	U	−3.9	−16.5
BR	U	U	U	U	34.0	U	−16.0	−16.0	U	U	U	U	34.3	U	−15.7	−15.7
DNA	U	U	U	U	U	U	**0.0**	**0.0**	U	U	U	U	U	U	**0.0**	**0.0**

Note: "U" was substituted with a value of 50, the number of amplification cycles, into the dCt calculations.

RT, Reverse Transcriptase; U, undetermined; BL, blood; SA, saliva; SE, semen; VS, vaginal secretions; MB, menstrual blood; SK, skin; MU, muscle; AD, adipose; BR, brain; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; MMP, matrix metalloproteinase; PRM, protamine; Ct, cycle threshold; dCt, delta Ct.

undetectable in blood, saliva, semen, and vaginal secretions [(4) and data not shown]. In summary, body fluid-specificity of the selected gene candidates was confirmed by the presence of appropriately sized RT-PCR products in one body fluid exclusive of the others tested (data not shown and [3,4]).

Development of Multiplex Real-Time Assays for Body Fluid Identification

Multiplex PCR development for the real-time PCR platform involves a similar strategy as that employed with standard end-point PCR, where proper primer design and optimization of the component primer concentrations are key to successful assay development (4). With qPCR assays, however, proper probe design and optimal probe concentrations are also critical. In the methods section, we have detailed our approach to real-time PCR primer and probe design. Instrument manufacturer (Applied Biosystems) recommended primer and probe concentrations were used for all singleplex reactions, with a final concentration of 900 nM for each primer and a concentration of 250 nM for the probe. These suggested conditions were also the starting point for multiplexing the selected genes. We decided to develop triplexes for the

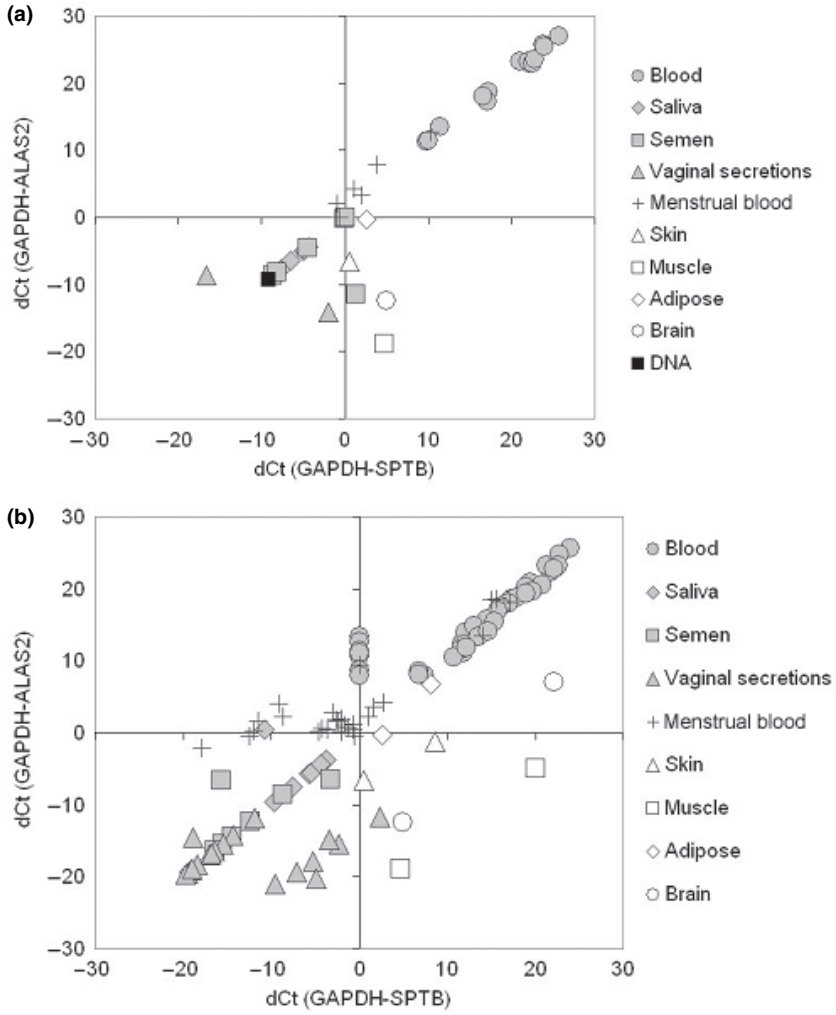

FIG. 2—*Specificity of real-time PCR triplex for blood identification. (a) To test the specificity of the qRT-PCR assay for the identification of blood, body fluid samples from multiple individuals (blood, n = 17; saliva, n = 7, semen, n = 6, vaginal secretions, n = 5; and menstrual blood, n = 6) were analyzed with the blood triplex assay. Other samples tested included RNA samples from skin (n = 1), muscle (n = 1), adipose tissue (n = 1), and brain (n = 1) and DNA (n = 1). The results are displayed as scatter plots in which each sample's dCt (GAPDH-SPTB) and dCt (GAPDH-ALAS2) are plotted. (b) To determine whether specificity was maintained over a range of RNA input concentrations, the blood triplex assay was tested using different input quantities of body fluid RNA (20 fg–500 ng) and tissue RNA (25 and 100 ng). The results are displayed as scatter plots in which each sample's dCt (GAPDH-SPTB) and dCt (GAPDH-ALAS2) are plotted.*

identification of body fluids, which would include two body fluid-specific markers and a housekeeping gene, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), as an internal control. The probe for each gene was labeled with a distinguishable reporter dye. As with the multiplex RT-PCR assay for the CE format that we previously published (4), including two markers for each body fluid would provide redundancy and, at least in principle, reduce the occurrence of false positive or false negative results. For reasons described below, the goal of optimizing the primer and probe concentrations was to bring the cycle threshold (Ct) values for the two body fluid genes as close to each other as possible (within approximately three Ct values of each other), and for the housekeeping gene to have a Ct value higher than the body fluid genes in a positive sample and a Ct value lower than the body fluid genes in a negative sample.

Delta Ct Metric for qRT-PCR Data Interpretation

For the interpretation of multiplex real-time PCR results, we employed a delta Ct (dCt) metric which measures the concentration of body fluid-specific genes in relation to an internal positive housekeeping gene control. The Ct value for each gene is the cycle number at which the fluorescence signal generated by the amplification of that gene passes a preset threshold (Fig. 1*a*). For our assays, we chose to use a constant threshold value of 0.200, which permitted easier comparison of results within and between assays. The dCt is determined by subtracting the average body fluid gene Ct value from the average housekeeping gene Ct value, with the average Ct obtained as a result of at least two separate amplifications of the same sample. A positive dCt value in a sample indicates that a body fluid gene is present at a higher level than the housekeeping gene, and that the body fluid being tested for is present (Fig. 1*a*, left panel). A negative dCt value indicates that the body fluid gene is present at a lower level than the housekeeping gene or not detected at all, and therefore that particular body fluid is not detectable in that sample (Fig. 1*a*, right panel). In cases when the amplification plot does not cross the threshold and no Ct value is obtained (“undetermined”), a Ct value equal to the highest cycle number used in the assay, which is 50 cycles, is used. The high specificity of the markers used for these assays often resulted in an

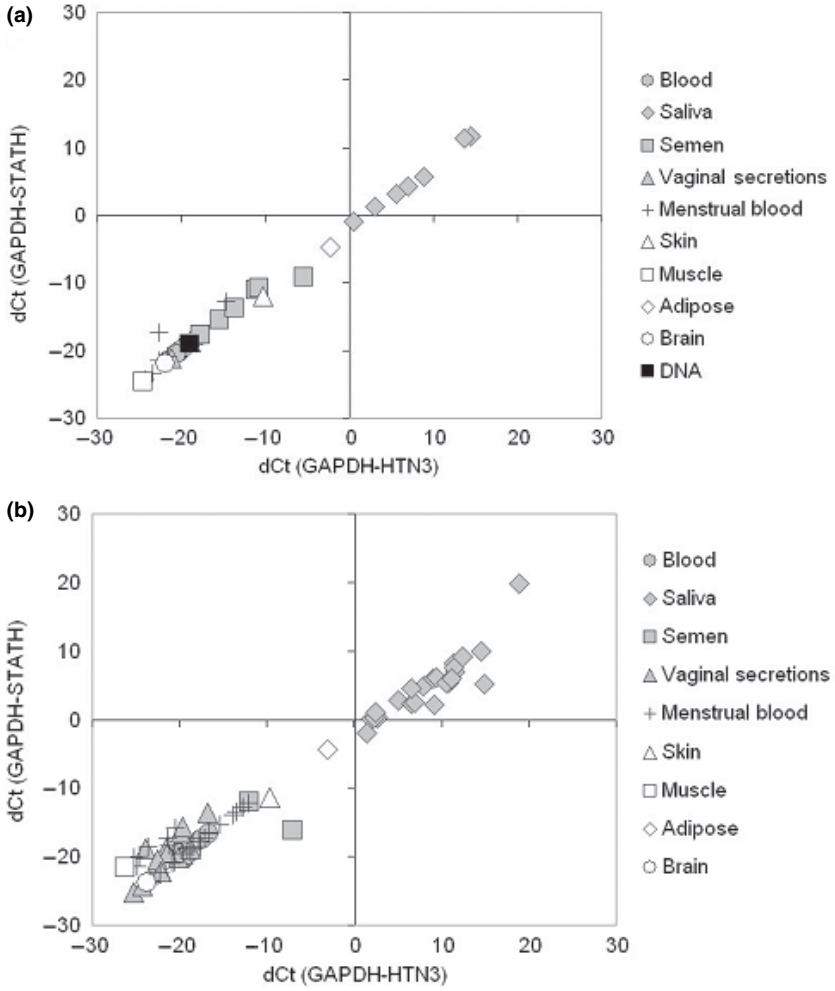

FIG. 3—*Specificity of real-time PCR triplex for saliva identification. (a) To test the specificity of the qRT-PCR assay for the identification of saliva, body fluid samples from multiple individuals (blood, n = 17; saliva, n = 7, semen, n = 6, vaginal secretions, n = 5; and menstrual blood, n = 6) were analyzed with the saliva triplex assay. Other samples tested included RNA samples from skin (n = 1), muscle (n = 1), adipose tissue (n = 1), and brain (n = 1) and DNA (n = 1). The results are displayed as scatter plots in which each sample's dCt (GAPDH-HTN3) and dCt (GAPDH-STATH) are plotted. (b) To determine whether specificity was maintained over a range of RNA input concentrations, the saliva triplex assay was tested using different input quantities of body fluid RNA (20 fg–500 ng) and tissue RNA (25 and 50 ng). The results are displayed as scatter plots in which each sample's dCt (GAPDH-HTN3) and dCt (GAPDH-STATH) are plotted.*

"undetermined" for at least one of the body fluid genes in samples that were known not to contain the body fluid gene assayed (Tables 2*a* and *b*).

As each body fluid-specific assay comprises two body fluid genes and one housekeeping gene, two dCt's are produced for each sample. The two dCt values for each sample are used to create a ΔdCt scatter plot and, as each of the two dCt's can be positive or negative in value, the resulting data points can reside in one of four quadrants (Fig. 1*b*). Thus possible ΔdCt results include double positive (+/+), positive/negative (+/−), negative/positive (−/+), and double negative (−/−) (Fig. 1*b*). A double positive ΔdCt result, with the data point in the upper right quadrant (quadrant I), indicates that both body fluid genes are expressed at higher levels than the housekeeping gene, showing that the body fluid being tested is present (a positive result). A double negative result, with the data point in the lower left quadrant (quadrant IV), indicates that neither body fluid gene is detected (or that the body fluid genes are expressed at lower levels than the housekeeping gene), indicating that the body fluid being tested is not detectable (a negative result). A housekeeping gene signal (as indicated by a Ct value <50) in a sample that gives a double negative result indicates the presence of sufficient nondegraded RNA in the sample for analysis and, consequently, the absence of the body fluid gene mRNAs in the sample. A positive/negative or negative/positive result, with the data point in the lower right (quadrant III) or upper left quadrants (quadrant II), respectively, indicates that one body fluid gene is expressed higher than the housekeeping gene, while the other body fluid gene is not detectable or has a lower level of expression than the housekeeping gene. Data points in Quadrants II and III would be termed "inconclusive," indicating that further testing may be necessary or that results from the other real-time body fluid PCR assays may need to be taken into account before a confirmation of body fluid source could be made.

Body Fluid Multiplex qRT-PCR Assay Specificity

Real-time PCR triplexes were developed for the identification of dried stains originating from blood, saliva, semen, and menstrual blood. Specificity testing of each of the qRT-PCR multiplexes used RNA extracted from commonly encountered body fluid samples (blood, saliva, semen, and vaginal secretions) from multiple individuals (blood, $n = 17$; saliva, $n = 7$; semen, $n = 6$; vaginal

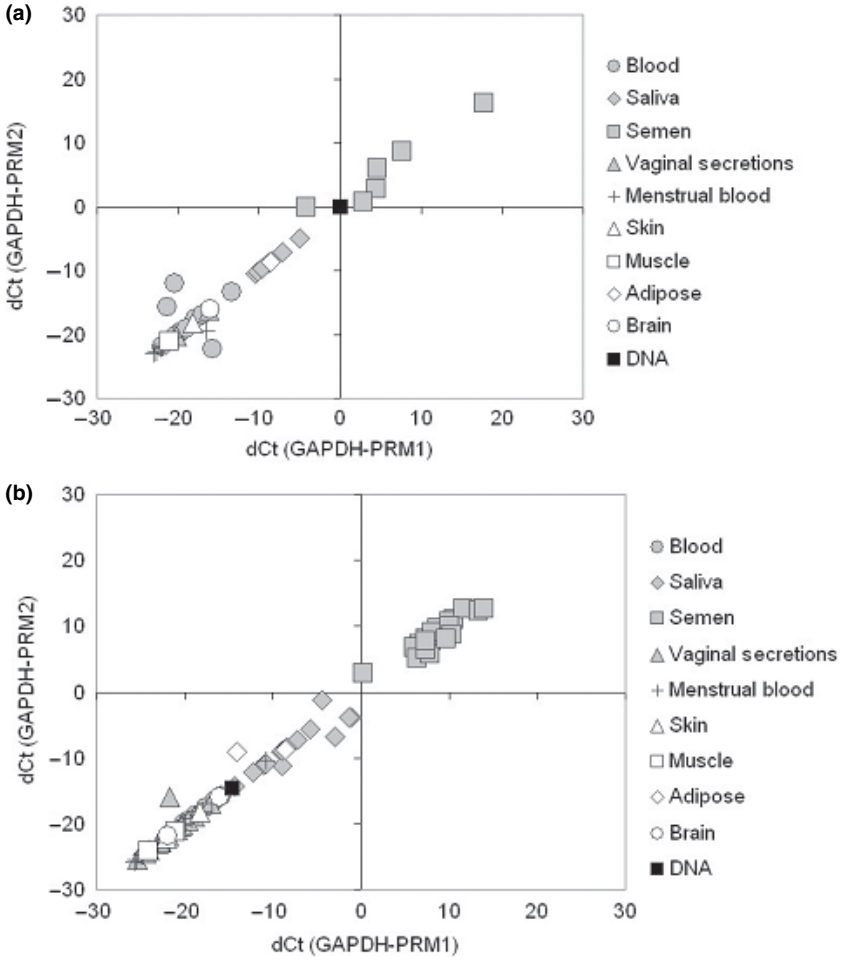

FIG. 4—*Specificity of real-time PCR triplex for semen identification. (a) To test the specificity of the qRT-PCR assay for the identification of semen, body fluid samples from multiple individuals (blood, n = 17; saliva, n = 7, semen, n = 6, vaginal secretions, n = 5; and menstrual blood, n = 6) were analyzed with the semen triplex assay. Other samples tested included RNA samples from skin (n = 1), muscle (n = 1), adipose tissue (n = 1), and brain (n = 1) and DNA (n = 1). The results are displayed as scatter plots in which each sample's dCt (GAPDH-PRM1) and dCt (GAPDH-PRM2) are plotted. (b) To determine whether specificity was maintained over a range of RNA input concentrations, the semen triplex assay was tested using different input quantities of body fluid RNA (20 fg–500 ng) and tissue RNA (25 and 100 ng). The results are displayed as scatter plots in which each sample's dCt (GAPDH-PRM1) and dCt (GAPDH-PRM2) are plotted.*

secretions, $n = 5$; and menstrual blood, $n = 6$). Other samples tested included RNA samples from skin ($n = 1$), muscle ($n = 1$), adipose tissue ($n = 1$), and brain ($n = 1$) and DNA ($n = 1$). The results of the specificity testing are summarized in Tables 2*a* and *b*. All four triplexes were also tested using different input quantities of RNA (20 fg–500 ng) to determine whether specificity was maintained over a range of RNA input concentrations (Figs. 2–5).

Blood—The blood triplex consists of SPTB, ALAS, and GAPDH. SPTB was detectable in the highest amounts in the blood samples (Ct range: 24.3–29.0), slightly lower in menstrual blood and the four tissues tested (Ct: 31.5–35.6), and even lower or undetectable after 50 cycles in the other body fluids (Table 2*a*). Similarly, ALAS2 was detectable in highest levels in blood (Ct range: 24.2–27.7), slightly lower in menstrual blood and adipose (Ct: 28.3–32.4), and significantly lower or undetectable after 50 cycles in the other body fluids and tissues (Table 2*a*). In some instances no detectable amplification of GAPDH ("U," undetermined) was observed. This may be attributable to reagent titration as a result of the particularly high abundance of the two blood markers in blood. For these samples, as well as for the body fluid samples that produced an undetermined Ct value, the number equal to the final cycle used (50) was substituted into the dCt calculation. The two dimensional ΔdCt plots show that the blood stain samples and some menstrual blood samples were located in the upper right quadrant (+/+), with saliva, semen, and vaginal secretions samples located in the lower left quadrant (−/−) (Figs. 2*a* and *b*). The quadrant location of the four tissues samples tested was dependent on the RNA input into the RT reaction. The input of a relatively high RNA concentration (100 ng per 30 μL reaction) shifted the adipose and brain samples into the upper right quadrant (+/+), whereas a lower total RNA input (25 ng per 30 μL reaction) kept them in the lower right quadrant (+/−) (Fig. 2*b*).

Saliva—The saliva triplex consists of HTN3, STATH, and GAPDH. With a few exceptions, both HTN3 and STATH demonstrated a high level of specificity for saliva (Ct ranges: 31.0–35.9 and 33.7–37.5, respectively) (Table 2*a*). The most notable exception was adipose, which provided Ct values within the same range (32.0 and 34.4, respectively); however, its Ct value for GAPDH was lower and therefore gave two negative (−/−) dCt values. Overall, for the saliva triplex, only the saliva data points (closed diamonds)

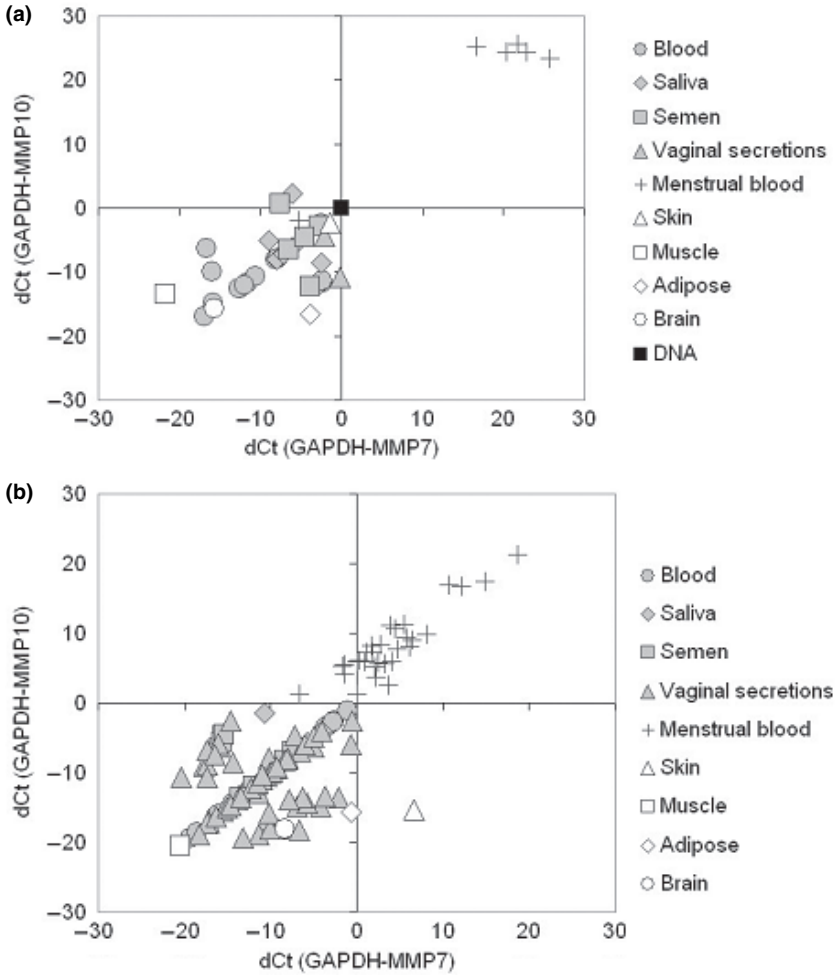

FIG. 5—*Specificity of real-time PCR triplex for menstrual blood identification. (a) To test the specificity of the qRT-PCR assay for the identification of menstrual blood, body fluid samples from multiple individuals (blood, n = 17; saliva, n = 7, semen, n = 6, vaginal secretions, n = 5; and menstrual blood, n = 6) were analyzed with the menstrual blood triplex assay. Other samples tested included RNA samples from skin (n = 1), muscle (n = 1), adipose tissue (n = 1), and brain (n = 1) and DNA (n = 1). The results are displayed as scatter plots in which each sample's dCt (GAPDH-MMP7) and dCt (GAPDH-MMP10) are plotted. (b) To determine whether specificity was maintained over a range of RNA input concentrations, the menstrual blood triplex assay was tested using different input quantities of body fluid RNA (20 fg–500 ng) and tissue RNA (25 ng). The results are displayed as scatter plots in which each sample's dCt (GAPDH-MMP7) and dCt (GAPDH-MMP10) are plotted.*

were present in the upper right quadrant (+/+), whereas the data points for blood, semen, vaginal secretions, menstrual blood, skin, muscle, adipose, and brain were all present in the lower left quadrant (−/−) (Figs. 3*a* and *b*). One of the known saliva samples (SA-C) exhibited a ΔdCt of (0.5, −1.0) and thus nominally breaks the +/+ rule (i.e., it appears in quadrant III instead of quadrant I). However, detailed inspection of the ΔdCt plot (Fig. 3*a*) shows that the sample is clearly clustered along with all of the other saliva samples and easily distinguished from non-saliva containing samples. Normal biological variation in gene expression is possibly the reason for the lower levels of HTN3 and STATH mRNA in this sample.

Semen—The semen triplex, which includes PRM1, PRM2, and GAPDH, demonstrated a similar high level of specificity, with relatively high or undetectable Ct values in non-semen samples after 50 cycles of amplification (Table 2*b*). Five of the six semen samples gave (+/+) dCt results as expected. One of the semen samples (SE-B), however, produced unexpectedly high, and reproducible, Ct values for PRM1 and/or PRM2 that resulted in the sample not being located in the upper right quadrant as its ΔdCt value was (−/0). This variation may be the result of inter-individual variation in gene expression levels and underscores the importance of recognizing the limitations of the ΔdCt data analysis method. In the case of SE-B, all three genes demonstrate moderate levels of expression, with a Ct value of 38.2 for PRM1 and 34.0 for both PRM2 and GAPDH. While this is certainly not indicative of a negative result for semen (typically the PRM1 and PRM2 dCt's are undeterminable in a non-semen sample or >40 [for blood] [Table 2*b*]), the dCt values of −4.2 and 0.0 at first blush seem to be indicative thereof. Thus we would recommend examining the raw data (Ct values and amplifications plots) obtained with the qRT-PCR assays before designating a sample as positive or negative for the body fluid. With regard to the specificity testing for the semen triplex when ΔdCt values were examined, only the semen samples were located in the upper right quadrant (+/+), whereas the other body fluids and tissues were located in the lower left quadrant (−/−) (Figs. 4*a* and *b*).

Menstrual Blood—The menstrual blood triplex is comprised of MMP-7, MMP-10, and GAPDH. Similar to the other body fluid-specific assays, both MMP-7 (Ct range: 24.3–36.3) and MMP-10 (Ct range: 24.3–33.1) demonstrate a higher level of expression in the menstrual blood samples compared with the other body fluids

and tissues (Table 2*b*). As with the blood triplex, the relatively high levels of expression for the menstrual blood markers prevent the GAPDH signal from reaching threshold, resulting in an "undetermined" Ct value for most of the samples. On the other hand, as with the semen triplex, one of the samples (MB-B) gave a false negative result because of an unexpectedly high level of expression for the GAPDH. Clearly, based on the Ct values for MMP-7 (36.3) and MMP-10 (33.1), the sample should be deemed positive. With that one exception, for the menstrual blood triplex, the menstrual blood ΔdCt data points were present in the upper right quadrant (+/+), while all of the blood, saliva, semen, vaginal secretions, skin, muscle, adipose, and brain were present in the lower left quadrant (−/−) (Figs. 5*a* and *b*).

Multiplex qRT-PCR Assay Sensitivity

The sensitivities of the real-time PCR triplex analyses were measured using varying quantities of total RNA input, ranging from 20 fg to 500 ng (data not shown). The limits of detection for the triplexes were established by the ability to obtain a positive dCt value for both of the body fluid-specific genes in samples comprising the particular body fluid. The sensitivities are expressed in total RNA input into a 30 μL RT reaction. As with our previously developed multiplex RT-PCR body fluid assay for the CE platform (4), the limits of detection for the triplexes varied in that the semen genes were the most sensitive (detection limit of 1 pg of total RNA input for the qRT-PCR assay), the saliva genes the least sensitive (2 ng) whereas the blood (150 pg) and menstrual blood (100 pg) genes were found to be of intermediate sensitivity.

Conclusions

The present work describes the development of sensitive and specific multiplex qRT-PCR assays for the identification of blood, saliva, semen, and menstrual blood. Advantages of a real-time PCR platform compared with a CE platform (4) include a more quantitative approach to gene expression analysis, increased sensitivity and virtually no post-amplification sample processing. Indeed, the qRT-PCR assays described herein are 5- to 40-fold more sensitive than a CE-based system for body fluid identification (4) and maintain specificity over a wide range of RNA input concentrations (20 fg–500 ng) (Figs 2*a*, 3*a*, 4*a*, and 5*a*). Recently, Nussbaumer et al. reported the development of a qRT-PCR assay for body fluid identification (16) using the HBA, KLK/PSA, and MUC genes. Although the sensitivity of the real-time assay was not reported, the HBA and KLK/PSA gene transcripts appeared to be specific for blood and semen, respectively, and the MUC gene was found to be variably expressed within and between different nonblood body fluids (16). In contrast, we have chosen to incorporate two body fluid-specific markers into each qRT-PCR assay in an attempt to improve assay specificity and provide analytical redundancy. As with the CE-based assay, the inclusion of two body fluid markers for each assay is expected to result in more analytical accuracy and precision (fewer false positives or negatives), by taking into account possible biological variation in gene expression levels between individuals (4).

The dCt and ΔdCt metrics for data analysis are measures of the relative expression of tissue-specific genes with respect to a housekeeping gene, GAPDH, and are expected to suffice as predictors for most samples. However, due to individual biological variation in gene expression, reliance on dCt and ΔdCt values alone may occasionally result in a false negative result, as was found with one each of the saliva, semen, and menstrual blood samples analyzed (Figs 3*a*, 4*a*, and 5*a*). For the semen result, inspection of the Ct values indicated that the sample's PRM1 and PRM2 genes were expressed at a moderate level in relation to the housekeeping gene (i.e., lower than most semen stains but higher than other non-semen body fluids). The outlier sample's ΔdCt value, while not located in the upper right quadrant as expected, nevertheless clustered separately from all of the other non-semen body fluid samples (Fig. 4*a*). Thus, we recommend examining Ct values as well as the other metrics especially for those samples appearing at the tail end of the ΔdCt distribution.

It appears to be important to control the amount of analyte (total RNA) input into the assay at the RT stage. For example, a false positive result for blood (ΔdCt in the upper right quadrant) could be obtained from relatively pure RNA isolated from adipose and brain tissue if 100 ng was added to the RT reaction whereas an input of 25 ng resulted in a lower right quadrant ΔdCt value. In addition, the presence of some blood-derived RNA in adipose and brain tissue isolates is not entirely unexpected due to the presence of blood vessels in these tissues. Nevertheless, akin to DNA profiling, it is important to control the amount of total RNA added to the mRNA profiling assay.

The species specificity of the assays was not empirically determined during the course of this work, although the primers used do not show significant homology with any other species in GenBank, including non-human primates.

Our assays incorporate the commonly used housekeeping gene GAPDH to normalize the expression of the body fluid-specific genes. However, GAPDH, like most other housekeeping genes used as controls in gene expression analysis, also exists in the genome as a processed pseudogene (17). Consequently, amplification products from contaminating genomic DNA can appear in the RT control sample, albeit at a lower concentration than products from bona fide transcripts (Tables 2a and b). Amplified product formation in the RT sample is not currently considered when performing the dCt calculations, and DNase I pretreatment of the RNA sample prior to RT-PCR is intended to eliminate or, at least significantly reduce the amplification of pseudogenes. Thus although data interpretation is relatively facile with the current assay formulations, we are currently evaluating a number of alternative non-pseudogene containing housekeeping genes for possible incorporation therein.

Acknowledgment

This work was financially supported by the Federal Bureau of Investigation through contract #J-FBI-03-287.

References

1. Alvarez M, Ballantyne J. The identification of newborns using messenger RNA profiling analysis. Anal Biochem 2006;357(1):21–34.
2. Anderson S, Howard B, Hobbs GR, Bishop CP. A method for determining the age of a bloodstain. Forensic Sci Int 2005;148(1):37–45.
3. Juusola J, Ballantyne J. Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification. Forensic Sci Int 2003;135(2):85–96.
4. Juusola J, Ballantyne J. Multiplex mRNA profiling for the identification of body fluids. Forensic Sci Int 2005;152(1):1–12.
5. Jones LJ, Yue ST, Cheung CY, Singer VL. RNA quantitation by fluorescence-based solution assay: RiboGreen reagent characterization. Anal Biochem 1998;265(2):368–74.
6. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365–86.
7. Steger K, Pauls K, Klonisch T, Franke FE, Bergmann M. Expression of protamine-1 and -2 mRNA during human spermiogenesis. Mol Hum Reprod 2000;6(3):219–25.

8. Bishop DF. Two different genes encode delta-aminolevulinate synthase in humans: nucleotide sequences of cDNAs for the housekeeping and erythroid genes. Nucleic Acids Res 1990;18(23):7187–8.
9. Bishop DF, Henderson AS, Astrin KH. Human delta-aminolevulinate synthase: assignment of the housekeeping gene to 3p21 and the erythroid-specific gene to the X chromosome. Genomics 1990;7(2):207–14.
10. Surinya KH, Cox TC, May BK. Identification and characterization of a conserved erythroid-specific enhancer located in intron 8 of the human 5-aminolevulinate synthase 2 gene. J Biol Chem 1998;273(27):16798–809.
11. Cox TC, Sadlon TJ, Schwarz QP, Matthews CS, Wise PD, Cox LL, et al. The major splice variant of human 5-aminolevulinate synthase-2 contributes significantly to erythroid heme biosynthesis. Int J Biochem Cell Biol 2004;36(2):281–95.
12. Amin KM, Scarpa AL, Winkelmann JC, Curtis PJ, Forget BG. The exon-intron organization of the human erythroid beta-spectrin gene. Genomics 1993;18(1):118–25.
13. Chu ZL, Wickrema A, Krantz SB, Winkelmann JC. Erythroid-specific processing of human beta spectrin I pre-mRNA. Blood 1994;84(6):1992–9.
14. Bauer M, Patzelt D. Protamine mRNA as molecular marker for spermatozoa in semen stains. Int J Legal Med 2003;117(3):175–9.
15. Bauer M, Patzelt D. Evaluation of mRNA markers for the identification of menstrual blood. J Forensic Sci 2002;47(6):1278–82.
16. Nussbaumer C, Gharehbaghi-Schnell E, Korschineck I. Messenger RNA profiling: a novel method for body fluid identification by Real-Time PCR. Forensic Sci Int 2006;157(2-3):181–6.
17. Zhang Z, Harrison PM, Liu Y, Gerstein M. Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. Genome Res 2003;13(12):2541–58.

Additional information and reprint requests:
Jack Ballantyne, Ph.D.
Associate Professor
Department of Chemistry
University of Central Florida
Building 5, 4000 Central Boulevard
Orlando, FL 32816-2366
E-mail: jballant@mail.ucf.edu