ORIGINAL ARTICLE

mRNA-based skin identification for forensic applications

Mijke Visser · Dmitry Zubakov · Kaye N. Ballantyne · Manfred Kayser

Abstract Although the identification of human skin cells is of important relevance in many forensic cases, there is currently no reliable method available. Here, we present a highly specific and sensitive messenger RNA (mRNA) approach for skin identification, meeting the key requirements in forensic analyses. We examined 11 candidate genes with skin-specific expression, as ascertained from expression databases and the literature, as well as five candidate reference genes ascertained from previous studies, in skin samples and in other forensically relevant tissues. We identified mRNA transcripts from three genes *CDSN*, *LOR* and *KRT9*, showing strong over-expression in skin samples relative to samples from forensic body fluids, making them suitable markers for skin identification. Out of the candidate reference genes tested, only *ACTB* showed similarly high expression in skin and body-fluid samples, providing a suitable reference marker for quantitative real-time PCR (qPCR) analysis of skin. Analyses of palmar and thumbprint skin samples indicate that our qPCR approach for the three skin-targeted mRNA markers, as well as the reference mRNA marker ACTB, is highly sensitive, allowing successful detection of minute amounts of skin material including full, half and quarter thumbprints, albeit with decreased success in decreasing print material. Furthermore, thumbprints stored for 6.5 months provided similar results relative to freshly analysed samples implying reasonable time-wise stability of the three skin-targeted mRNAs as well as the *ACTB* reference mRNA. Our study represents the first attempt towards reliable mRNA-based skin identification in forensic applications with particular relevance for future trace/touched object analyses in forensic case work. Although the approach for skin identification introduced here can be informative when applied on its own, we recommend for increased reliability the integration of (one or more of) the skin-targeted mRNA markers presented here into multiplex assays additionally including mRNA markers targeting alternative cell types expected in forensic samples.

Keywords mRNA · Body-fluid identification · Skin · Forensic · *CDSN* · *LOR* · *KRT9* · *ACTB*

M. Visser · D. Zubakov · K. N. Ballantyne · M. Kayser (✉)
Department of Forensic Molecular Biology,
Erasmus MC University Medical Center Rotterdam,
P.O. Box 2040, 3000 CA, Rotterdam, The Netherlands
e-mail: m.kayser@erasmusmc.nl
URL: www.erasmusmc.nl/fmb/

Introduction

A plethora of methods are currently available in forensic biology to identify the cellular origin of biological traces on evidential material. Most commonly, presumptive tests are employed, such as the Phadebas test for saliva, the ABAcard HemaTrace method for blood, or prostate specific antigen detection for semen [1]. However, many of these tests can give false positive results, or their sensitivity level is too low compared to that of short tandem repeat (STR) profiling; thus, intensive research is being performed to find more accurate and more sensitive alternatives to these time- and sample-consuming tests.

Messenger RNA (mRNA) may provide the necessary specificity, sensitivity and automation capabilities that modern forensic biology laboratories require for cellular origin identification. Each cell type has its own transcriptome, and differences in expression patterns can be

utilised to identify mRNA markers with strong over-expression in one relative to the other forensically relevant cell types (such as body fluids). With the early identification of various mRNA markers with respective over-expression in venous blood, menstrual blood, saliva, semen and vaginal secretions, recent work has focused on refining these mRNA markers to increase sensitivity or degradation resistance, or on the development of multiplex assays to allow parallel testing of several mRNA markers in minute amounts of sample [2–9]. More recently, micro-RNAs (miRNAs) have been started to be explored for the forensic determination of cellular origin [10, 11].

However, at present, there are no mRNA markers specifically useful for skin cell determination included in published tissue identification multiplexes. Moreover, there is currently no widely used presumptive test for the positive identification of skin material, although alternative light sources, such as the Polilight, can indicate the presence of epithelial cells [12], and a histological method is available to distinguish between skin, buccal and vaginal epithelium [11, 13]. Hence, the ability to conclusively identify and confirm, on the molecular level, the presence of skin epithelial cells is still lacking. Given the success of mRNA profiling with other tissue types of forensic importance, this approach may also provide potential for skin identification. Being able to identify skin cells via mRNA profiling would be of considerable relevance to forensic trace DNA profiling, an area which has been controversial of late (for a recent summary, see [14]). Molecular proof that the evidentiary STR profile on an item originated from skin cells may allow the linking of the item and the STR profile to the suspect. This would remove some of the objections to the use of trace sample amounts for prosecution purposes, by allowing activity level statements to be made about trace DNA.

The development and maintenance of the outer epithelial layer of human skin has been comprehensively studied, and a considerable amount of knowledge has accumulated regarding the molecular control of epidermal cell differentiation. The large stem cell population in the basal layer of the epidermis continually renews the skin, differentiating into transit-amplifying cells before undergoing terminal differentiation [14, 15]. As the cells migrate upwards through the living spinous and granular layers, they become progressively keratinised as production of keratin proteins 1 and 10 increases, followed by the expression of the proteins filaggrin, involucrin and loricrin [16]. While undergoing apoptotic-like destruction, the nuclei are extruded, and the cells become cornified [15]. With the complete renewal of the epidermis every 40–56 days, humans shed considerable numbers of cells everyday, containing DNA and stripped nuclei, and presumably epidermal proteins and mRNA transcripts as well [17–19].

In the present study, we identified, through thorough database and literature searches combined with empirical laboratory testing of candidate markers, a set of mRNA transcripts that are highly over-expressed in skin epithelial cells relative to other cell types of forensic relevance, and performed appropriate tests to describe the sensitivity of the developed assays and of the time-wise stability of the proposed markers. We present here a highly specific and sensitive mRNA approach for skin identification for future applications in forensic analyses.

Materials and methods

Sample collection

For initial testing of candidate skin and reference genes, we used venous blood collected by finger prick ($n=15$) or venipuncture ($n=12$), saliva ($n=15$), semen ($n=3$), menstrual blood ($n=3$), vaginal secretions ($n=7$), skin biopsy ($n=2$) and palmar skin ($n=15$) samples collected from healthy volunteers (eight men, seven women, ranging from 22 to 42 years of age) under informed consent. All samples from each individual (except for skin biopsies) were collected within a 1-h period, with skin sample collection always preceding finger prick blood collection. For saliva collection, volunteers were asked to rinse the mouth with cold water 30 min before sampling. Saliva (3–5 ml) was collected in sterile 50 ml Falcon polyethylene tubes cooled on ice, with the RNA extracted immediately to prevent degradation. One saliva soaked cotton swab per individual was used for RNA isolation. For skin palmar samples, volunteers were asked to rinse hands with water (without soap) 30 min before sampling and to avoid touching other skin or often touched surfaces (keyboards, etc.) during this time. Cotton swabs were extensively rolled and moved over either the left or right hand (insides of both palms and fingers) of the volunteer. After collection of the different body fluids and tissues, all samples were immediately processed for RNA isolation. Semen, menstrual blood and vaginal secretion samples were collected and processed as described previously [10].

For sensitivity testing of the skin-targeted markers, left and right thumbprints of five individuals were collected on clean, sterilised glass slides. Samples were either immediately processed for RNA isolation or stored in dust-free, non-humid conditions for 6.5 months. Thumbprints were then collected either by wet–dry swabbing or the Pinpoint™ Slide RNA Isolation System II (Zymo Research, California, USA). In brief, swabbing was performed by rolling with simultaneous downward pressure a cotton or flocked swab moistened with lysis buffer over the thumbprint area, followed by collection of any excess moisture or

cells with a dry cotton or flocked swab (protocol modified from [20]). The Pinpoint™ Slide RNA Isolation System II utilises a proprietary solution, which is painted over the fingerprint, embedding the cellular material within the solution as it dries. Approximately 150 μl of Pinpoint™ Solution was spread in a thin layer across the print using an elongated 20 μl pipette tip. Slides were dried for 30 min at room temperature, with the dried solution and embedded print lifted and placed in a sterile micro centrifuge tube.

Ascertainment of candidate genes

The initial ascertainment of genes was performed using the built-in function of GNF SymAtlas database (now the BioGPS Gene Portal Hub) of gene expression in human tissues [21]. All candidate genes with expression 10-fold higher in skin than the median expression across all other tissues in this database were ascertained. This online database on gene function and structure based on human microarray expression data from the Genomics Institute of the Novartis Research Foundation. Users can view expression patterns across (currently) 84 tissues and can view detailed annotation information on many of the transcripts (http://www.biogps.gnf.org). In addition, KRT9 was ascertained from the literature as being specifically expressed in the palms and soles of the feet [22]. Commonly used reference genes were selected from the literature for examination of their amplification in skin and other forensically relevant tissues [4, 5, 7, 8, 23–26].

RNA isolation and cDNA synthesis

Half a millilitre of liquid blood (obtained by venipuncture), or one blood soaked cotton swab (obtained by finger prick), was first extracted by acid phenol-chloroform to remove the haem and greatly increase RNA yield. After centrifugation, total RNA was isolated from the upper phase with the Qiagen RNeasy Plus Mini kit (QIAGEN GmbH, Hilden, Germany) according to the manufacturer's instructions. The same kit was used for the saliva samples with the exception that during the lysis step with RLT buffer, the saliva soaked cotton swabs were incubated for 30 min at 56°C before RNA extraction. Total RNA was isolated from skin swabs using the Qiagen RNeasy Fibrous Tissue Mini kit (Qiagen) according to the manufacturer's instructions with minor adjustments; cotton swabs were incubated for 30 min at 56°C in RLT buffer, proteinase K and water, before the complete mixture including the swab was passed over a QiaShredder column (Qiagen), after which the flow-through was incubated for another 10 min at 56°C. RNA extractions and subsequent DNase I digestion from semen, menstrual blood and vaginal secretion samples was performed as described previously [10]. The quantity of total RNA was determined by a NanoDrop

ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

Total RNA from fingerprints was isolated using the Pinpoint™ Slide RNA Isolation System II (Zymo Research Corporation) according to the manufacturer's instructions with minor adjustments; following print collection with Pinpoint™ solution as described above, the material was incubated with 60 μl RNA digestion buffer and 15 μl proteinase K for 4 h at 55°C. The lysate was then applied to a Zymo-Spin Column, with the manufacturer's protocol in subsequent steps. An additional on-column DNase I treatment performed between the two wash steps with an incubation mixture of 10 μl DNase I and 70 μl RDD buffer (Qiagen), for 15 min at room temperature.

The reverse-transcriptase (RT) reaction was performed using RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas GmbH, St. Leon-Rot, Germany) according to the manufacturer's instructions for random hexamers protocol, in a total volume of 20 μl. Each reaction contained 5–11 μl of total RNA. "No-RT controls" (i.e. controls without reverse transcriptase) were performed to detect for genomic DNA contamination.

Quantitative PCR

PCR primers for the initial screening of target genes (CALML5, CDSN, DCD, DSC1, FLG, KRT1, KRT2A, KRT9, KRT10, LOR and SCGB2A2) and five commonly used reference genes (ACTB, B2M, GAPDH, PPIB and UBC) were designed using Primer3 software so that the forward and reverse primers were complementary to different exons of the respective genes and most closely located to the 3′ end of the corresponding RefSeq complementary DNA (cDNA) (Supplementary Table S1). Following marker selection, PCR primers and hydrolysis probes (TaqMan gene expression assays) for CDSN, KRT9, LOR and ACTB were commercially obtained (Applied Biosystems, The Netherlands). Probes were labelled with FAM (target genes) or VIC (reference gene) reporter dyes at the 5′ end and a non-fluorescent quencher at the 3′ end of the oligo in combination with MGB (Table 1). To ensure that only mRNA, and not genomic DNA, was detected by the tissue-specific PCR assays, PCR primers which spanned exon junctions were used, resulting in small amplicons for mRNA and very large or no product for genomic DNA. The only exception for this was the primer set used for LOR, as one of the two exons of this gene is too small for priming; thus, one single exon was targeted. Therefore, all samples and PCR conditions were also controlled so that genomic DNA could not be amplified. During the course of testing, the primer set for the reference gene ACTB had to be changed to ensure amplification of small amounts of low quality material. Initially, the amplicon size was 171 bp

Table 1 Three suitable skin-targeted mRNA markers and one reference mRNA marker, their Genbank accession number, PCR primer and probe sequences and assay IDs (Applied Biosystems) with the reporter dye and the sizes of the qPCR amplified products

Gene	Genbank accession number	Primer/Probe sequence (5′→3′)	Assay ID/reference	Reporter dye	Amplicon size (bp)
CDSN	NM_001264.3	–[a]	Hs 00169911 m1	FAM	79
LOR	NM_000427.2	–[a]	Hs 0189462. s1	FAM	105
KRT9	NM_000226.3	–[a]	Hs 00413861. m1	FAM	94
ACTB(171)	NM_001101.2	–[a]	4326315E (part number)	VIC	171
ACTB(74)	NM_001101.3	Forward: tgacccagatcatgtttgag Reverse: cgtacagggatagcacag Probe: HEX-caccccagccatgtacg -TAMRA	Primer3 software	HEX	74

[a] Primer/Probe sequences of the commercially available Taqman assays are undisclosed by the manufacturer

(ACTB-171 primer set, Applied Biosystems). However, this appeared too long for efficient amplification and thus was changed to a custom-designed primer/probe set ACTB-74, with a 74-bp amplicon. As both primer sets target the same transcript, reference gene quantitation was not affected.

Quantitative real-time PCR reactions (qPCR) for the examination of candidate genes were performed with custom-designed primers on an ABI 7300 PCR machine (Applied Biosystems, The Netherlands) with the Maxima™ SYBR Green/ROX qPCR Master Mix (2X) (Fermentas GmbH) using the following parameters: initial denaturation at 94°C for 10 min, followed by 45 cycles of 94°C for 15 s and 60°C for 30 s. In each PCR reaction, 2 μl of 5-fold diluted cDNA was used as input. Specificity of the primers and the absence of genomic DNA contamination were checked with melting profiling and agarose gel electrophoresis.

TaqMan gene expression assays were performed for CDSN, KRT9, LOR and ACTB to test marker specificity across different tissues and body fluids. qPCR reactions were performed on a LightCycler® 480 Real-time PCR System (Roche Diagnostics Nederland B.V.) with the Maxima™ Probe/ROX qPCR Master Mix (Fermentas GmbH) using the following parameters: initial denaturation at 95°C for 10 min, followed by 50 cycles of 95°C for 10 s, 60°C for 30 s and 72°C for 1 s and a final cooling step to 40°C. qPCR amplifications were performed in a total volume of 20 μl, with the reaction mix containing 10 μl of PCR master mix, 1× primer/probe mix, 1.5 μl of undiluted cDNA and 7.5 μl of double-distilled water (ddH$_2$O). PCR reactions were performed in Lightcycler 480 Multiwell PCR plates (Roche). LightCycler 480 System Software v1.5 (Roche) was used to analyse the qPCR data. Threshold cycle (Ct) measurements were automatically determined by the software. The reference gene ACTB was used to normalise the amplification signal between the samples of different cellular origin, individuals

and amount of input cDNA. Comparative analyses between skin, saliva, blood, semen, menstrual blood and vaginal secretions were done with the delta Ct (ΔCt) method, where $ΔCt = Ct_{(target gene)} - Ct_{(reference gene)}$. Low ΔCt values indicate high expression of a gene, whereas higher ΔCt values indicate very low expression of a target gene in relation to the expression of ACTB. If no signal was detected from either the reference gene or target gene, the sample was removed from analysis, although the number of failures was noted.

Sensitivity and stability testing was performed with commercial TaqMan primers and probes, with the exception of ACTB-74. As the amplicon given by the commercial primers was too large to be suitable, custom primers were designed as described above, and listed in Table 1. ACTB-74 was duplexed with CDSN in the qPCR reactions. qPCR reactions were performed on a LightCycler® 480 Realtime PCR System (Roche) with the Maxima™ Probe/ROX qPCR Master Mix (Fermentas GmbH). Amplifications were performed in a total volume of 20 μl, with each PCR reaction containing 5 μl of undiluted cDNA as input. For the singleplex LOR and KRT9 amplifications, and CDSN within the duplex, 1× primer/probe mix was used. For ACTB-74 within the duplex, 1 μM forward and reverse primers and 0.25 μM probe were used. Cycling parameters were as listed above for the specificity testing. ΔCt values for the target genes were calculated as listed above. If no signal was detected from either the reference or target gene, the sample was removed from analysis, although the number of failures was noted.

All qPCR assays, including the commercially obtained Taqman assays, for which the manufacturer states 100% efficiency, as well as the endogenous controls ACTB-171 and ACTB-74, were tested and verified to perform near 100% PCR efficiency by performing standard curve analysis (LightCycler 480 System Software v1.5) of cDNA dilutions using skin biopsy samples (data not shown).

Statistical analyses

All statistical analyses were performed in SPSS v15.0 (SPSS Inc.)

Results

Testing of candidate genes

Ten skin candidate genes *CALML5*, *CDSN*, *DCD*, *DSC1*, *FLG*, *KRT1*, *KRT2A*, *KRT10*, *LOR* and *SCGB2A2* were identified from the GNF SymAtlas database [21] as having expression levels in skin 10-fold higher than the median expression across all other tissues in this database. In addition, *KRT9* was chosen from the literature for being specifically and solely expressed in the terminally differentiated epidermis of palms and soles of hand and feet [22]. All candidate genes were experimentally examined via qPCR in skin samples as well as in a variety of forensically relevant body fluids, such as blood, menstrual blood, semen, saliva and vaginal secretions for detectable levels of expression. Five of the ascertained candidate genes showed non-specific amplification (*CALML5, DSC1, FLG, KRT1* and *KRT2A*) among the samples tested, while three (*DCD, KRT10* and *SCGB2A2*) revealed expression signals in skin samples expected to be too low for practical forensic applications (data not shown). This initial screening resulted in a panel of three mRNA markers that showed considerably high expression in skin samples together with over-expression in skin samples relative to all other forensically relevant cell types tested (in the following called "skin-targeted mRNA markers"), which we used in further investigations: corneodesmosin (*CDSN*), loricrin (*LOR*) and keratin 9 (*KRT9*).

In addition, five commonly used reference genes (*ACTB*, *B2M*, *GAPDH*, *PPIB* and *UBC*) [4, 5, 7, 8, 23–26] were tested in skin samples. Of those, the expression of *ACTB* appeared to be most abundant in skin material (data not shown). When further tested in the various forensically relevant body fluids, *ACTB* showed similarly high levels of expression in all cell types including skin (Fig. 1) and was therefore chosen as reference mRNA marker in all further qPCR testings.

Using *ACTB* as reference gene, we analysed the three skin-targeted mRNA markers in palmar skin samples serving as an appropriate skin sample type for the envisioned future forensic application in the "touched object" scenario. We observed high levels of expression for all three mRNA markers in most individual samples tested. Because the reference gene *ACTB* (i.e. its 171-bp amplicon) was not detected in five of the 30 palmar samples, we excluded these five samples from further

analyses. Using *ACTB* as the reference gene, we obtained average ΔCt values of −4.46 for *CDSN*, −3.62 for *LOR* and −0.67 for *KRT9*. No significant differences in ΔCt values were observed between samples from left and right palms within the same individual (Kolmogorov–Smirnov $Z=0.685$, $p=0.736$); however, there was substantial variation in ΔCt values between individuals (Kruskal–Wallis $\chi^2=39.38$, $p<0.0001$). As the Ct values of the reference gene *ACTB* showed significant variation between but not within individuals (i.e. between technical replicates), the observed variation in ΔCt is most likely due to variation in the amount of material deposited. *CDSN* and *LOR* were detected in all samples, whereas *KRT9* was not detected in three of the 30 samples tested. However, given the high Ct values observed in these three samples for the other two skin-targeted mRNA markers, it appears that very little material was collected initially with these samples, and thus, *KRT9* fell below the level of detection.

Specificity of skin-targeted mRNA markers

To further establish the specificity of the three skin-targeted mRNA markers *CDSN*, *LOR*, and *KRT9*, we tested their expression levels in non-skin samples, namely, blood (venous and menstrual), saliva, vaginal secretions and semen, as well as in palmar skin samples, using an enlarged sample set (Fig. 2). Raw Ct values were normalised relative to the expression of the identified reference gene *ACTB* with primer set *ACTB*-171 for skin, saliva, finger prick and venous blood, and primer set *ACTB*-74 for vaginal secretions, menstrual blood and semen samples (see methods for different *ACTB* assays). All three skin-targeted mRNA markers showed significantly lower ΔCt values in palmar skin samples, as compared to the non-skin samples (ANOVA, $F=442.27$, $p<0.0001$). Only *LOR* was detected in all tissues, but was at much higher levels in skin than in non-skin samples, and these differences were statistically significant (ANOVA, $F=298.1$, $p<0.0001$). *CDSN* was below the detection threshold in semen, while *KRT9* was below the detection threshold in venous and menstrual blood, as well as in semen. All other post hoc, Bonferroni-adjusted comparisons between skin and non-skin tissues were found to be highly significant, with p values ranging from 1.27×10^{-5} (*LOR*; skin vs semen) to 2.44×10^{-48} (*CDSN;* skin vs fingerprick blood).

As both saliva and vaginal secretions contain epithelial cells, it was unsurprising to find slightly higher levels of expression of the mRNA markers within these samples, compared to the expression of the mRNA markers in other non-epithelial samples. However, skin vs vaginal secretions and skin vs saliva expression levels were significantly different for all skin-targeted mRNA markers (all $p<0.00001$). In addition, higher expression

Fig. 1 Ct values of the reference gene *ACTB* from biopsied skin samples and those from forensically relevant body fluids. Equal amounts of total RNA (100 ng) were used for cDNA synthesis in each sample. *Error bars* represent the maximum and minimum observed Ct values from two to four individuals for each sample type
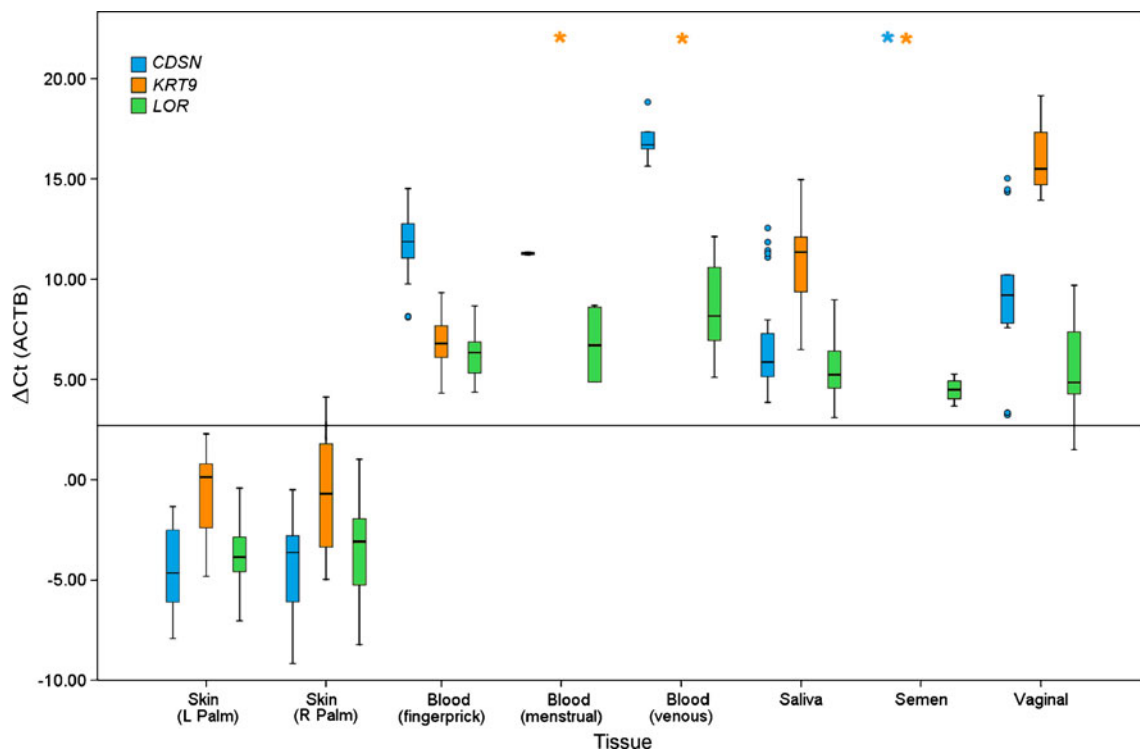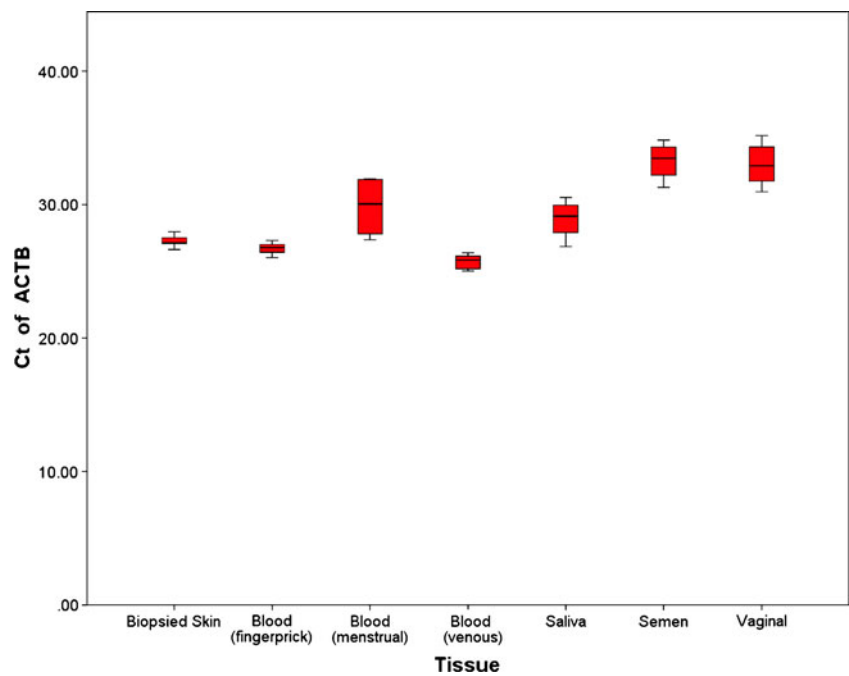
Fig. 2 ΔCt values for three skin-targeted mRNA markers *CDSN, KRT9* and *LOR* in various skin samples and across forensically relevant body fluids: Palmar skin (L), *n*=15; Palmar skin (R), *n*=15; Blood (fingerprick), *n*=15; blood (menstrual), *n*=3; blood (venous), *n*=12; saliva, *n*=15; semen, *n*=2; vaginal, *n*=7. The *box* for each gene represents the interquartile range (25–75th percentile), and the *line* within this box is the median value. *Bottom* and *top bars* indicate the minimum and maximum value of all the data, respectively, or 1.5 times the interquartile range of the data, whichever is smaller. Outlier values are indicated by *closed circles*. Note the significantly lower ΔCt values observed in skin compared to non-skin samples. An *asterisk* indicates that the target gene was not detected in that tissue

levels of the skin-targeted mRNAs were detected in blood collected via fingerpricks, compared with venous blood. This is presumably due to low-level mixing of skin cells with the blood in the fingerprick samples. However, despite this, a clear distinction could be made between ΔCt values of skin samples and non-skin samples for all three skin-targeted mRNA markers. There are only two exceptions: KRT9, where a single right palmar sample displayed a ΔCt more characteristic of non-skin samples, and LOR, where two samples have ΔCt values considerably lower than the other vaginal secretion samples. Excluding this one skin sample from the analysis allows a ΔCt threshold value of 2.7 to be set for differentiating skin from non-skin samples with our approach; any sample below this value was attributed as being from skin, while any sample with a ΔCt>2.7 was correctly called as non-skin.

Sensitivity of skin-targeted mRNA markers

The sensitivity of the qPCR assays in detecting the three skin-targeted mRNA markers, respectively, in low amounts of skin material was tested using thumbprints of several individuals that were set on glass slides. However, an efficient method of recovering the cellular material and mRNA had to be investigated first, to ensure that total mRNA could be retrieved in sufficient quantity from fingerprints left on the glass slides. Therefore, three collection methods—wet/dry cotton swabbing [20], wet/dry swabbing with flocked swabs and the Pinpoint™ Slide RNA isolation System II kit (Zymo Research)—were tested, and the results were compared. To ensure consistency between samples, three donors (good, medium and poor shedders, as determined by the overall success rate of amplification of the target and reference genes, data not shown) provided 5–10 thumbprints each and for each of the sample collection/RNA isolation protocol. Overall, the Pinpoint™ method proved to be the only one that allowed reproducible detection of skin-targeted mRNA markers (data not shown). With the other two methods, only the reference gene ACTB amplified occasionally and only from samples obtained from the good shedder, whereas the skin-targeted mRNA markers could not be amplified successfully, most likely due to insufficient yield from the thumbprint (data not shown). Therefore, the Pinpoint™ method was used for skin cell collection in all further experiments, particularly during the sensitivity testing due to the low amount of material expected.

All three skin-targeted mRNA markers proved to be exceptionally sensitive for the detection of skin cells given the assays applied (Fig. 3). From full thumbprints, all three mRNA markers in all samples except one were detected with a Ct value below 40, with average Ct values of 34.72 for CDSN, 36.38 for LOR, and 36.45 for KRT9 while the average Ct for the reference gene ACTB was 35.8 (primer set ACTB-74). Dropout of KRT9 was observed in one full print, and the donor of this particular print appeared to be a poor shedder based on the expression levels detected in other genes and thumbprints examined. The ΔCt values were generally below the 2.7 threshold as determined in the previously described analysis, with the exception of the LOR ΔCt (3.1) obtained from the poor shedder and the KRT9 ΔCt at 2.8 of one other full thumbprint.

Furthermore, when half and quarter thumbprints were analysed, it was noted that the Cts increased more than may have been predicted based on sample size where an increase in Ct of 1 would be expected for full-half print differences and two for full-quarter prints. This may reflect differences in the amount of mRNA deposited between prints tested, or alternatively that the stochastic limit of qPCR detection was reached, resulting in excessive variation between samples. The average Cts observed for half prints were all below 40 cycles, but dropout was observed in one fifth CDSN and LOR and three fifths KRT9 amplifications. The generally lower expression of KRT9 resulted in the average ΔCt ratio being above the 2.7 threshold (3.2) from half prints, although CDSN and LOR were still below this level. The quarter print samples showed considerably more dropout (20% for ACTB, 40% for CDSN and LOR and 70% for KRT9), as may be expected from such low amounts of material collected from quarter prints. However, all samples showed positive amplification for CDSN and, therefore, were able to be correctly assigned as skin based on this marker (average ΔCt 0.32), as well as three of the LOR amplifications. All remaining LOR and all KRT9 amplifications from quarter prints would not have been identified correctly as skin samples on the basis of the 2.7 ΔCt threshold. Thus, as could be expected, as the level of material present drops, there is an increasing risk of misclassification of the sample as originating from a non-skin source. Stochastic effects may result in inconsistent ratios of the three target genes to the reference ACTB gene. To prevent incorrect sample assignment, the limit of detection threshold for each of the three skin-targeted mRNA markers may be set at 40 cycles, beyond which the ΔCt value may be inaccurate. This is particularly the case when the ACTB Ct is above 35, indicating only minute traces of material are present. In the future, a larger dataset, consisting of many more samples, should be used to precisely define more specific limits.

Time wise stability of skin-targeted mRNA markers

To obtain an indication about the time-wise stability of the expression signals of the three skin-targeted mRNA markers, we analysed full thumbprints stored at room

Fig. 3 Detection sensitivity of qPCR assays for three skin-targeted mRNA markers *CDSN, KRT9* and *LOR*, with full (*n*=5), half (*n*=5) and quarter (*n*=10) thumbprints. All sample points below the line would be correctly identified as originating from skin, when applying a threshold of 2.7 (see specificity results for further details). An *asterisk* indicates that dropout was observed, with the number specified for each marker and fingerprint size
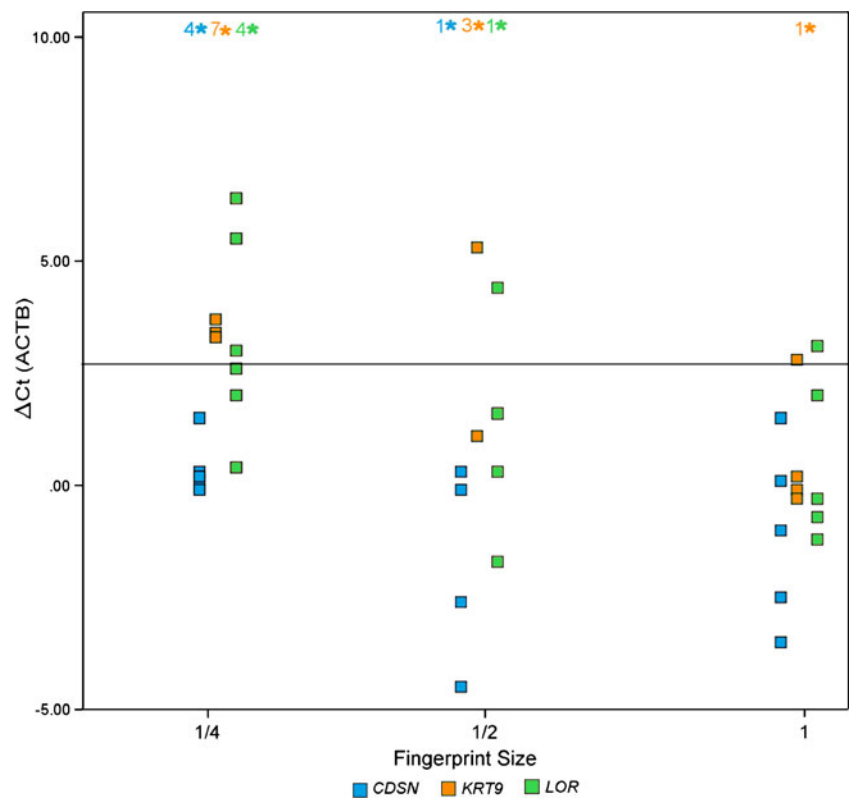

temperature for 6.5 months. Compared to thumbprints processed immediately after deposition, no significant difference in expression signals were observed for any of the three skin-targeted mRNA markers, with ΔCt values remaining relatively stable (Fig. 4; ANOVA; *CDSN*, $F=0.002$, $p=0.965$; *KRT9*, $F=3.992$, $p=0.081$; *LOR*, $F=0.911$, $p=0.357$). Notably, the reference mRNA marker *ACTB* also remained stable, with a non-significant increase of 0.1 in Ct ($F=0.021$, $p=0.886$). Only *KRT9* showed a (albeit not statistically significant) trend towards decreasing detection levels with increasing thumbprint storage time, with average Ct values increasing by 1.8 cycles and ΔCt by 1.6. These results indicate that the *KRT9* transcript may be more susceptible to time-wise degradation than the two other skin-targeted mRNA markers, as *KRT9* was not detected in 40% of samples stored for 6.5 months.

Discussion

Although touched objects and fingerprints have been used in forensic biology to generate DNA profiles for over 10 years [27], there is not a definitive test for determining the presence of skin epithelial cells in a biological trace. The inability to prove that a suspect's DNA originated specifically from skin cells can lead to challenges against DNA evidence in court. For instance, a suspect who has been identified by DNA profiling may allege that his/her

DNA profile is present on an item found at a crime scene due to speaking or sneezing over it, but that he/she did not touch or use the item in question from which his/her DNA profile was obtained. Alternatively, a person may claim sexual assault with a specific object, while an alleged
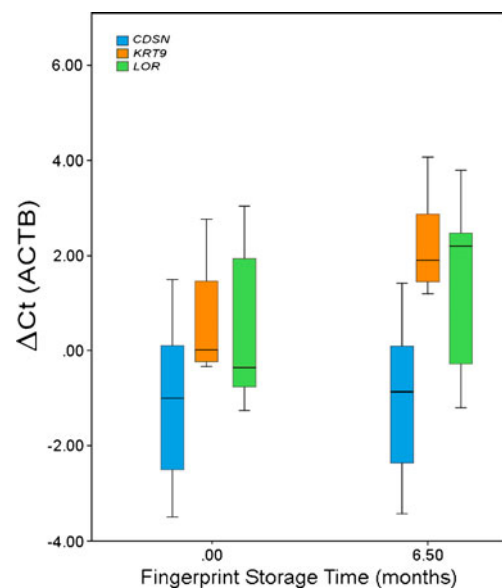
Fig. 4 Time-wise stability of three skin-targeted mRNA markers *CDSN, KRT9* and *LOR* after 6.5 months of thumbprint storage. No significant differences were observed between freshly analysed samples (*n*=5) and thumbprints stored for 6.5 months at room temperature (*n*=10) for all three mRNA markers

suspect whose DNA profile was collected from such object may claim that he/she simply touched the object instead of sexually assaulting the person. While histological staining or alternative light source methods can be used, these do not provide conclusive evidence for the presence of skin cells, lack the required sensitivity or reduce the efficiency of DNA profiling and, therefore, are not in common usage. Instead, an mRNA assay targeting genes that are highly over-expressed in skin relative to other cell types that can be expected at a crime scene, such as forensically relevant body fluids, would allow the conclusive determination of the presence of epithelial cells in an extract and is compatible with automation and parallel analysis with conventional DNA markers, such as forensic STRs.

There are various mRNA transcripts with putative over-expression in skin available in public databases and in the literature that are usually derived from genome-wide gene expression data. However, expression signals obtained with microarray probes do not always agree with those derived from specific qPCR experiments, and some of the cell types highly relevant to forensic analysis are usually not covered by expression databases. Furthermore, sensitivity limits between array-based methods and qPCR are different, which may lead to further discrepancies. Hence, careful experimental validation of candidate mRNA markers obtained from expression databases is needed using a methodology suitable for forensic applications (such as qPCR), as well as testing forensically relevant cell types.

Performing such a study, we have identified three mRNA markers, namely KRT9, CDSN and LOR, as being suitable for skin identification in the forensic context. That we found eight of the 11 candidate skin mRNA markers ascertained from a large and widely used expression database either not specific or not sensitive may be explained by the factors listed above. All three genes identified here as representing suitable skin mRNA markers are involved in the differentiation or maintenance of the keratinisation and cornification of the skin. KRT9 belongs to the superfamily of intermediate filament proteins expressed in all different epithelial cell types. There are 20 human epithelial keratins, all functioning as important building blocks composing the cytoskeleton of epithelial cells, and they all also have their own specific function or timing in the cellular dynamics. They are expressed, mostly in pairs or subsets, during terminal skin cell differentiation, in the different stages of development and in different epithelia [22]. KRT9, a type I keratin, is expressed only in the suprabasal cells of the epidermis and has previously been found to be specifically expressed in palmar and plantar skin [22]. CDSN, a 52- to 56-kDa basic glycoprotein, is specific to the cornified epithelia and the inner root sheath of hair follicles [29]. LOR is initially expressed in the granular layer of the epidermis during cornification and comprises of about 80% of the total protein mass of the cornified envelope [29, 30]. Both CDSN and LOR are involved in the assembly of the epidermal cornified cell envelope, with CDSN mainly detected in uppermost spinous and granular layers [29, 32, 33, 34].

Recently, Schulz et al. [28], in a study focussing on antibodies against various cytokeratins, found KRT10 to be specifically expressed in human skin samples. While we were able to confirm its specificity to skin, we found the expression level within palmar and fingerprint skin samples too low to be practically useful for trace samples in forensic applications (data not shown).

We have observed high over-expression of the mRNA markers KRT9, CDSN and LOR in skin, compared to blood, saliva, menstrual blood, semen and vaginal secretions (Fig. 2). Although low levels of expression were observed for all three genes in some non-skin samples, skin samples showed considerably higher levels of expression of the three skin-targeted mRNA markers. This observed strong difference in expression intensity between skin and non-skin samples allows the three mRNA markers to differentiate skin from non-skin samples based on the ΔCt measure. In particular, the clear distinction in ΔCt values between skin samples on one side and vaginal and saliva samples on the other indicates that the three mRNA markers are expressed predominately in cornified, keratinised epithelium. Oral and vaginal epithelium and skin epidermis are very similar, but differ in the proportion of cornified epithelial cells, which could be an explanation for the slightly higher level of expression of the target genes we detected in vaginal secretion and saliva compared to the other non-skin tissues. The results from fingerprick blood samples also indicates that low-level skin cell contamination of a body tissue or fluid, such as in blood obtained using fingerpricks, will not cause the test to falsely indicate skin as the major component. However, the lower ΔCt observed, compared to a non-mixed blood sample such as venous blood, may indicate that skin cells are present in the minority in the fingerprick blood samples we used. For completely reliable mixture determination, we advise the incorporation of the skin-targeted mRNA markers introduced here into multiplex assays additionally including mRNA markers targeting all alternative forensically relevant cell types.

The reference gene ACTB (β-actin), used here for normalising expression signals of skin-targeted mRNA markers, was chosen from five commonly used genes (ACTB, B2M, GAPDH, PPIB and UBC, [4, 5, 7, 8, 23–26]), which were all tested for ubiquitous expression across forensically relevant samples including skin. Of all reference candidate genes tested, ACTB was the most abundantly expressed in skin cells, with the least variation across all

cell types in this study. Actins are highly conserved proteins that are involved in cell motility, structure and integrity. *ACTB* is a major constituent of the contractile apparatus and one of the two non-muscle cytoskeletal actins. Although 18S rRNA is often used in forensic gene expression studies and in many other gene expression studies [4, 7, 24, 26], caution should be taken when using rRNA as reference for mRNA targets, due to observed imbalances between mRNA and rRNA fractions [26]. Furthermore, rRNA molecules are not present in purified mRNA samples, and the high abundance of 18S rRNA compared to target mRNA markers can create difficulties in the normalisation procedure [26]. Ideally, more than one reference gene would be applied, and additional candidate genes may be screened in future studies for suitability in trace samples. However, it should be kept in mind that the low amount of material present with touched objects and thus available for skin cell determination significantly limits the number of tests that can be performed on a sample, also limiting the number of reference genes that can be analysed from such sources (unless multiplexing can be employed successfully).

To ensure that the protocols would be suitable for practical forensic applications, such as to analyse samples taken from touched objects, primer sets were chosen to produce small amplicons ranging from 79 to 105 bp for the three skin-targeted mRNA markers. Initially, a commercially available assay for the reference gene *ACTB* was used, with an amplicon size of 171 bp. However, it was found out that this commercial assay was unable to amplify samples of low quantity/quality (such as fingerprints), and therefore, we developed and applied to such samples a new *ACTB* assay with a reduced amplicon size of only 74 bp. Sensitivity testing of *CDSN*, *LOR*, *KRT9* and reference gene *ACTB*-74 revealed that all four mRNA markers were stably detected from full thumbprints (Fig. 3). The sensitivity decreased with the size of the fingerprint sample decreasing, resulting in seven of 15 amplification/threshold failures for half thumbprints and 21 of 30 for quarter thumbprints. However, given the similar specificity of all three skin-targeted mRNA markers used, a single marker is expected to be sufficient for skin identification. Consequently, using a single marker returning a ΔCt below the 2.7 threshold, 80% of half thumbprints could be correctly classified, as well as 70% of the quarter prints. The observed high level of sensitivity of the assays employed for the suitable three skin-targeted as well as the ACTB-74 reference markers ensures that skin deposits can be detected in samples, which may contain only trace amounts and for which currently forensic scientists cannot postulate an origin for. As a further prerequisite for future applications to forensic case work, the three skin-targeted mRNA markers (as well as the *ACTB* reference mRNA) were shown to be remarkably stable over sample storage time, with no significant signal

loss observed over 6.5 months of fingerprint sample storage (although longer time intervals should be investigated in future studies). Despite a slight decrease in *KRT9* expression after 6.5 months, half of the prints still displayed positive amplification with ΔCt below the 2.7 threshold.

Finally, we would like to note that co-extraction of both DNA and RNA from the same sample followed by parallel STR and mRNA amplification is possible, and a suitable method has been demonstrated recently [35]. Thus, applying our approach for the identification of skin to co-extracted case work material would allow the conclusive linking within the same sample of the cellular origin of the STR profile obtained while limiting reagent cost and sample consumption.

Conclusion

From our results, it can be concluded that three mRNA markers *CDSN*, *LOR* and *KRT9*, in combination with the *ACTB* reference mRNA marker, and analysed via the described qPCR assays, are highly suitable for identifying skin cells, including small traces of skin material. These mRNA markers fill a gap in the current portfolio of mRNA markers for cellular origin identification in forensic applications, allowing the rapid and accurate identification of minute quantities of skin materials. Although the approach for skin identification introduced here can be informative when applied on its own, we recommend the inclusion of (one or more of) the skin-targeted mRNA markers presented here in multiplex systems, which additionally include mRNA markers targeting other forensically relevant cell types (as well as more than one reference gene). In general, we recommend, as emphasised before [8], that in forensic applications, a cell type/tissue should be identified by the presence of (e.g. mRNA) markers specific for the relevant cell type/tissue together with the absence of markers specific for alternative cell types/tissues, which is most efficiently achieved by combining all relevant markers in a sensitive and reliable multiplex assay. Therefore, our study not only provides suitable mRNA markers for identifying skin cells, including minute amounts of skin material but also contributes to the future completion of suitable multiplex mRNA systems for the reliable determination of all cellular types usually expected in forensic samples including skin.

References

1. Virkler K, Lednev IK (2009) Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene. Forensic Sci Int 188:1–17
2. Bauer M, Patzelt D (2002) Evaluation of mRNA markers for the identification of menstrual blood. J Forensic Sci 47:1278–1282
3. Fleming RI, Harbison S (2010) The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids. Forensic Sci Int Genet 4:244–256
4. Haas C, Klesser B, Maake C, Bär W, Kratzer A (2009) mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR. Forensic Sci Int Genet 3:80–88
5. Juusola J, Ballantyne J (2003) Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification. Forensic Sci Int 135:85–96
6. Juusola J, Ballantyne J (2007) mRNA profiling for body fluid identification by multiplex quantitative RT-PCR. J Forensic Sci 52:1252–1262
7. Sakurada K, Ikegaya H, Fukushima H, Akutsu T, Watanabe K, Yoshino M (2009) Evaluation of mRNA-based approach for identification of saliva and semen. Leg Med (Tokyo) 11:125–128
8. Zubakov D, Hanekamp E, Kokshoorn M, van IJcken W, Kayser M (2008) Stable RNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples. Int J Leg Med 122:135–142
9. Zubakov D, Kokshoorn M, Kloosterman A, Kayser M (2009) New markers for old stains: stable mRNA markers for blood and saliva identification from up to 16-year-old stains. Int J Leg Med 123:71–74
10. Zubakov D, Boersma AW, Choi Y, van Kuijk PF, Wiemer EA, Kayser M (2010) MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation. Int J Leg Med 124:217–226
11. Hanson EK, Lubenow H, Ballantyne J (2009) Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. Anal Biochem 387:303–314
12. Vandenberg N, van Oorschot RA (2006) The use of Polilight in the detection of seminal fluid, saliva, and bloodstains and comparison with conventional chemical-based screening tests. J Forensic Sci 51:361–370
13. French CE, Jensen CG, Vintiner SK, Elliot DA, McGlashan SR (2008) A novel histological technique for distinguishing between epithelial cells in forensic casework. Forensic Sci Int 178:1–6
14. van Oorschot RAH, Ballantyne KN, Mitchell RJ (2010) Forensic trace DNA: a review. Investigative Genetics 1:14
15. Mack JA, Anand S, Maytin EV (2005) Proliferation and cornification during development of the mammalian epidermis. Birth Defects Res C Embryo Today 75:314–329
16. Lippens S, Hoste E, Vandenabeele P, Agostinis P, Declercq W (2009) Cell death in the skin. Apoptosis 14:549–569
17. Blanpain C, Lowry WE, Pasolli HA, Fuchs E (2006) Canonical notch signaling functions as a commitment switch in the epidermal lineage. Genes Dev 20:3022–3035
18. Halprin KM (1972) Epidermal "turnover time"—a re-examination. Br J Dermatol 86:14–19
19. Kita T, Yamaguchi H, Yokoyama M, Tanaka T, Tanaka N (2008) Morphological study of fragmented DNA on touched objects. Forensic Sci Int Genet 3:32–36
20. Alessandrini F, Cecati M, Pesaresi M, Turchi C, Carle F, Tagliabracci A (2003) Fingerprints as evidence for a genetic profile: morphological study on fingerprints and analysis of exogenous and individual factors affecting DNA typing. J Forensic Sci 48:586–592
21. Pang BC, Cheung BK (2007) Double swab technique for collecting touched evidence. Leg Med (Tokyo) 9:181–184
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) Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci USA 99:4465–4470
23. Chu PG, Weiss LM (2002) Keratin expression in human tissues and neoplasms. Histopathology 40:403–439
24. Heinrich M, Lutz-Bonengel S, Matt K, Schmidt U (2007) Real-time PCR detection of five different "endogenous control gene" transcripts in forensic autopsy material. Forensic Sci Int Genet 1:163–169
25. Nussbaumer C, Gharehbaghi-Schnell E, Korschineck I (2006) Messenger RNA profiling: a novel method for body fluid identification by real-time PCR. Forensic Sci Int 157:181–186
26. Stamova BS, Apperson M, Walker WL et al (2009) Identification and validation of suitable endogenous reference genes for gene expression studies in human peripheral blood. BMC Med Genomics 2:49
27. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3:RESEARCH0034
28. van Oorschot RA, Jones MK (1997) DNA fingerprints from fingerprints Nature 387:767
29. Schulz MM, Buschner MGD, Leidig R, Wehner HD, Fritz P, Häbig K, Bonin M, Schütz M, Shiozawa T, Wehner F (2010) A new approach to the investigation of sexual offenses—cytoskeleton analysis reveals the origin of cells found on forensic swabs. J Forensic Sci 55:492–498
30. Jonca N, Guerrin M, Hadjiolova K, Caubet C, Gallinaro H, Simon M, Serre G (2002) Corneodesmosin, a component of epidermal corneocyte desmosomes, displays homophilic adhesive properties. J Biol Chem 277:5024–5029
31. Kalinin A, Marekov LN, Steinert PM (2001) Assembly of the epidermal cornified cell envelope. J Cell Sci 114:3069–3070
32. Yoneda K, Hohl D, McBride OW, Wang M, Cehrs KU, Idler WW, Steiner PM (1992) The human loricrin gene. J Biol Chem 267:18060–18066
33. Meyer-Hoffert U (2009) Reddish, scaly, and itchy: how proteases and their inhibitors contribute to inflammatory skin diseases. Arch Immunol Ther Exp (Warsz) 57:345–354
34. McGrath JA (2005) Inherited disorders of desmosomes. Australas J Dermatol 46:221–229
35. Bowden A, Fleming R, Harbison S (2010) A method for DNA and RNA co-extraction for use on forensic samples using the Promega DNA IQ system. Forensic Sci Int Genet. doi:10.1016/j.fsigen.2009.11.007