Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Towards simultaneous individual and tissue identification: A proof-of-principle study on parallel sequencing of STRs, amelogenin, and mRNAs with the Ion Torrent PGM

D. Zubakov [a], I. Kokmeijer [a], A. Ralf [a,1], N. Rajagopalan [b], L. Calandro [b], S. Wootton [b], R. Langit [b], C. Chang [b], R. Lagace [b], M. Kayser [a,*]

[a] Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands
[b] Thermo Fisher Scientific/Life Technologies, South San Francisco, CA, USA

ARTICLE INFO

ABSTRACT

DNA-based individual identification and RNA-based tissue identification represent two commonly-used tools in forensic investigation, aiming to identify crime scene sample donors and helping to provide links between DNA-identified sample donors and criminal acts. Currently however, both analyses are typically performed separately. In this proof-of-principle study, we developed an approach for the simultaneous analysis of forensic STRs, amelogenin, and forensic mRNAs based on parallel targeted DNA/RNA sequencing using the Ion Torrent Personal Genome Machine® (PGM™) System coupled with the AmpliSeq™ targeted amplification. We demonstrated that 9 autosomal STRs commonly used for individual identification (CSF1PO, D16S539, D3S1358, D5S818, D7S820, D8S1179, TH01, TPOX, and vWA), the AMELX/AMELY system widely applied for sex identification, and 12 mRNA markers previously established for forensic tissue identification (ALAS2 and SPTB for peripheral blood, MMP10 and MMP11 for menstrual blood, HTN3 and STATH for saliva, PRM1 and TGM4 for semen, CYP2B7P1 and MUC4 for vaginal secretion, CCL27 and LCE1C for skin) together with two candidate reference mRNA markers (HPRT1 and SDHA) can all be successfully combined. Unambiguous mRNA-based tissue identification was achieved in all samples from all forensically relevant tissues tested, and STR sequencing analysis of the tissue sample donors was 100% concordant with conventional STR profiling using a commercial kit. Successful STR analysis was obtained from 1 ng of genomic DNA and mRNA analysis from 10 ng total RNA; however, sensitivity limits were not investigated in this proof-of-principle study and are expected to be much lower. Since dried materials with noticeable RNA degradation and small DNA/RNA amplicons with high-coverage sequencing were used, the achieved correct individual and tissue identification demonstrates the suitability of this approach for analyzing degraded materials in future forensic applications. Overall, our study demonstrates the feasibility of simultaneously obtaining multilocus STR, amelogenin, and multilocus mRNA information for combined individual and tissue identification from a small sample of degraded biological material. Moreover, our study marks the first step towards combining many DNA/RNA markers for various forensic purposes to increase the effectiveness of molecular forensic analysis and to allow more forensically relevant information to be obtained from limited forensic material.

© 2015 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Human identification from DNA via comparative STR profiling has become a standard tool in forensic laboratories around the world [1]. However, the DNA-identified donor of a crime scene sample not necessarily is the perpetrator. One piece of information that helps linking (or not) identified sample donors with criminal acts, is to find out which type of cells or tissue, such as saliva, peripheral and menstrual blood, semen, vaginal secretion, and skin, was used for STR profiling and thus individual identification [2]. In fact, after STR-based individual identification, establishing the cellular origin of biological traces found at crime scenes is considered the next crucial task in forensic casework, as it allows for the reconstruction of the sequence of crime events and evaluation of the relevance of forensic evidences. Therefore,

* Corresponding author. Tel.: +31 10 703 80 73; fax: +31 10 704 45 75.
E-mail address: m.kayser@erasmusmc.nl (M. Kayser).
1 Current affiliation: DDL Diagnostic Laboratory, Rijswijk, The Netherlands.

forensic investigators require tools for unequivocal, confirmatory proof of the origin of biological stains used for STR profiling, which can be defended in court.

In recent years, progress in forensic molecular biology has led to the development of RNA-based methods for forensic tissue identification [3–5]. Multiple RNA markers were proposed for all forensically relevant tissues, some of them were evaluated in collaborative trials to identify the most informative and robust mRNA markers out of many suggested previously [6–13]. Notably, unlike most of the previous forensic methods used for forensic tissue identification that are often presumptive, prone to false positive or false negative results, and destructive to biological samples, the RNA approach allows for sensitive, specific, non-destructive and parallel identification of the cellular origin of biological traces [3]. Whereas the information upon the cellular origin of crime scene biological samples is typically obtained from RNA, individual identification of the donors of these samples is established from DNA, typically using STRs and sometimes SNPs. Currently, these two types of molecular analyses, DNA and RNA testing, are typically performed separately. However, in the future both analyses may want to be joined, especially with the perspective to combine with more DNA/RNA markers for additional forensic purposes to increase the effectiveness of molecular forensic analysis and to allow more forensically relevant information to be obtained from limited crime scene material.

In principle, massive parallel sequencing (MPS), also referred to as next generation sequencing (NGS) or second generation sequencing (SGS), allows the targeted analysis of large numbers of nucleic acid markers including DNA and (after a reverse transcription step) RNA markers, which is not feasible with alternative technologies that only allow limited multiplexing. Currently, NGS platforms such as the Ion Torrent[TM] Personal Genome Machine® (PGM[TM]) and others have been already applied to forensic STR [16–18] and SNP analysis [19–21] for human identification purposes. They are also used for forensic research purposes and suitability testing for future case work applications of mitochondrial DNA analysis for high-resolution maternal lineage identification [14,15]. Recently, we demonstrated that the parallel analysis of >500 Y-SNPs covering the entire currently known Y-chromosome phylogenetic tree is feasible in a single analysis with the Ion Torrent[TM] PGM[TM] for providing ultra-high resolution paternal lineage and ancestry information [16]. The combined analysis of different types of forensic biomarkers (DNA and RNA) for different forensic purposes with NGS platforms has not been demonstrated as of yet and was noted to be difficult in a recent review article on forensic NGS [17].

Here, we present a proof-of-principle study on parallel analysis of multiple STR markers, amelogenin, and multiple mRNA markers, and demonstrate that in principle combined individual and tissue identification is feasable using targeted NGS, particularly with the Ion Torrent PGM.

2. Materials and methods

2.1. Sample collection

A total of 7 volunteers including 4 male and 3 female individuals donated their body fluids for this study under informed consent. Native peripheral blood samples were obtained with finger pricks, and blood drops were absorbed by sterile cotton swabs. Unstimulated saliva was obtained by spitting into 50 mL Falcon tubes, aliquots were pipetted onto cotton swabs. Freshly ejaculated semen was collected in plastic cups, and then aliquotted onto cotton swabs. Semen-free vaginal secretions and menstrual blood were collected using sterile cotton-tipped swabs from volunteers who had abstained from sexual contact for a minimum of three days. Menstrual blood was collected at 3rd day of menstruation. The same donors contributed the menstrual blood and vaginal secretion samples. All cotton swabs were allowed to slowly dry at room temperature for at least overnight and stored at −80 °C until the isolation of nucleic acids. Skin samples were obtained from surgical left-over biopsy material (for sample description see [18]). The epidermal layer of the skin was removed from the dermal layer by first cutting the skin into pieces of ~1 mm wide and 7 mm long and incubating the skin pieces for 30 min at room temperature in ammonium thiocyanate solution (3.8% ATC in PBS) [19].

2.2. Preparation of nucleic acids

RNA and DNA were co-isolated from the same biological sample using AllPrep Kit (Qiagen) according to the manufacturer's protocol. RNA fraction was treated with Turbo DNase Kit from Ambion (Applied Biosystems, Netherlands) to remove traces of genomic DNA. The quality and quantity of RNA were assessed by capillary electrophoresis on Agilent 2100 Bioanalyzer automated analysis system (Agilent Technologies, Palo Alto, CA, USA). Genomic DNA concentrations were established with Quantifiler® Human DNA Quantification Kit (Applied Biosystems, Netherlands).

2.3. mRNA and STR sequencing

Twelve mRNA markers (2 per each tissue) previously suggested for tissue identification, as well as 2 house-keeping genes serving as positive controls were analyzed (Table 1). Ion AmpliSeq Designer software tool (Thermo Fisher) was used to design a panel of custom PCR primers producing the amplicons of average size 150 bp. Ion AmpliSeq[TM] RNA Library Kit 2.0 was used for RNA

Table 1
Primers for mRNA markers used for tissue identification in combination with STR/amelogenin sequencing with the Ion Torrent PGM. All primers were automatically designed by AmpliSeq Designer software to span the intron, but not exon–exon boundaries.

mRNA marker	Targeted tissue	Forward	Reverse	Reference
HPRT1	House-keeping	TTCCTCCTCCTGAGCAGTCA	ACCCTTTCCAAATCCTCAGCAT	[20]
SDHA	House-keeping	TTGAGATTTGCTGATGGAAGCATAAGA	GTGCTTTAGGTCTCCATAGAGCTT	[20]
MMP10	Menstrual blood	CCATAAGGAAAATTGATGCAGCTGTT	CAACTCCTGGAAAGTCATCAGCTA	[7]
MMP11	Menstrual blood	GACGCCACTCACCTTTACTGA	TCCCCTTCTCGGTGAGTCTT	[7]
ALAS2	Peripheral blood	GACCCTGACCACCTAAAGAAACTT	GAAGGTCAGGGCCCCATAC	[9]
SPTB	Peripheral blood	CTAGTGGACACGGCGGATAA	GATGCCCTGGTGATACTTCATGAG	[9]
HTN3	Saliva	TCAACGTGTTTAAATACATCAGCCATCT	AGCATGAGAGCCAAGATTAAAGCAA	[8]
STATH	Saliva	GTATGGCCCTTATCAGCCAGTT	TGGATGTAGAAGTCGTATTTGCCAATC	[8]
PRM1	Semen	GCAGATATTACCGCCAGAGACA	AGGAGTTTGGTGGATGTGCTATTTT	[8]
TGM4	Semen	AAATGTCCTGGACTGCTGCAT	CCCCAGTCCAATTCCCAATGAG	[8]
CCL27	Skin	GACTGTCACCTCCAGGCTTT	TCCTTAGCATCCCAAAATTCAGCTT	[13]
LCE1C	Skin	GTCTCTTCCTGCTGCAGTGT	TCTGTGGCAGTGGGACCTA	[13]
CYP2B7P1	Vaginal secretion	GAATTCAGCCACCAGAACCTCAT	CACCTGTTCAATCTCCTTGTAGACT	[7]
MUC4	Vaginal secretion	GAATGACAAGCCCTACCTCTGT	CCCAGCCCATTGAAGGTGTA	[7]

library preparation according to manufacturer's instructions. Standard protocol for RNA library preparations includes 20 cycles of PCR amplification, which was used here. Before combination in multiplex analysis, all mRNA markers were tested via singleplex analysis using all tissue samples (targeted and non-targeted ones). Singleplex analysis revealed much higher number of sequencing reads for the target mRNA markers in the targeted tissue samples (i.e., for blood mRNA markers in blood samples etc.) than in all non-targeted samples (data not shown), as we expected from previously established marker knowledge. Because it is well known that menstrual blood and vaginal secretion samples contain large amounts of non-human mRNA, which compromises total RNA measurements, we also tested 30 cycles of amplification in these samples to overcome the problem of insufficient human RNA content in these samples. Although we found that with 30 cycles the number of sequencing reads was increased, as expected, in the final multiplexing analyses we used the standard protocol of 20 cycles.

Nine forensic autosomal STRs commonly used for individual identification (CSF1PO, D16S539, D3S1358, D5S818, D7S820, D8S1179, TH01, TPOX, and vWA) as well as AMELX/AMELY locus for sex determination (Table 2), i.e., the Ion Torrent™ HID STR 10-plex (Life Technologies), were amplified by PCR in 20-ul reactions containing 8 ul of 2.5X Thermo Fisher Proprietary Reaction Mix 2 ul of 10X Primer Mix (STR 10-plex Ion Ampliseq™ primer panel) and 1 ng DNA unless specified differently. The PCR conditions were 95 °C for 1 min, followed by 30 cycles of 94 °C for 10 s and 59 °C for 90 s with a final hold time of 5 min at 60 °C. Primers were partially digested by adding 2 ul of FuPa reagent (Ion AmpliSeq™ Library Kit 2.0) to the PCR reaction, and incubated at 50 °C for 10 min, 55 °C for 10 min followed by 60 °C for 20 min. Samples were then purified using 40 ul AMPure XP Reagent (Agencourt®) following the recommended protocol and was then eluted in 74 ul EB buffer (10 mM Tris–HCl pH 8.0). DNA library was prepared using Ion Plus Fragment Library Kit based on the manufacturer's instructions. Libraries were quantified using the Ion Quantitation Kit then diluted, and finally pooled in equimolar ratios. The libraries were processed with the Ion OneTouch™ 2 System using Ion PGM™ Template OT2 200 Kit to produce 200 base-read libraries. Sequencing of the libraries was performed with Ion Torrent PGM using 318 v2 chips and Ion PGM™ 200 Sequencing Kit v2. Sequencing was performed as suggested by the manufacturer with the small adjustment of the increased number of flows from 500 to 750.

Firstly, separate sequencing runs were performed for mRNA and STR marker panels to test their performance for each sample. For the subsequent final multiplex experiments RNA and DNA libraries from the same sample were individually barcoded with Ion Xpress™ Barcode Adapters 1-96 Kit. Both cDNA and genomic DNA libraries were pooled in equal amounts and simultaneously processed with the Ion OneTouch™ 2 System using Ion PGM™ Template OT2 200 Kit to produce 200 base-read libraries. Sequencing of the libraries was performed with Ion Torrent PGM using ver. 318 chips v2 and Ion PGM™ 200 Sequencing Kit v2. In single machine run up to 16 barcoded samples (8 RNA and 8 STR) were sequenced simultaneously. In accordance to manufacturer's protocols, 1 ng genomic DNA and 10 ng total RNA were used as templates. The only difference between singleplex and multiplex experiments was that for the latter we combined differently barcoded samples (DNA and RNA) from different individuals in one sequencing run. Therefore, multiplex sequencing may be considered as a technical replication of singleplex experiments.

STR sequencing results obtained with Ion Torrent PGM were verified for concordance with classical capillary electrophoresis (CE) based genotyping using AmpFLSTR® Identifiler® PCR Amplification Kit (Thermo Fisher), which contained all 9 STR markers and the amelogenin system used here for Ion Torrent PGM sequencing.

2.4. Data analysis

Sequencing results were analyzed with Ion Torrent Suite v.4.0.1. Mapped sequencing reads for mRNA targets were counted with Coverage Analysis plugin v. 4.2. STR profiling was performed with HID_STR_Genotyper v.2.0. RNA-based tissue classification was performed with binary logistic regression analysis using IBM SPSS Statistics v. 21 software.

3. Results and discussion

3.1. Sample issues: mRNA integrity

In vitro stability of mRNA markers is an essential prerequisite for their use in forensic applications [5]. Current methods of mRNA-based tissue identification that mostly include end-point PCR followed by capillary electrophoresis are capable of successful analysis of even highly degraded samples [4]. To motivate future forensic use on degraded samples, in this proof-of-principle study, we used body fluid samples that were allowed to air-dry slowly on cotton swabs under room temperature, and are therefore expected to have undergone some RNA degradation. The result of the Bioanalyzer testing of RNA integrity indeed confirm that the samples we used for Ion Torrent PGM sequencing were compromised as shown in Fig. 1. However, despite the noticeable RNA degradation, tissue identification was successful for all samples tested. Our preliminary data suggest that the use of small cDNA amplicons of 150 bp for all 14 mRNA markers provides suitable prerequisites for successful analysis of degraded RNA, raising reasonable expectations that our approach will be suitable for the analysis of forensic casework samples, which will need to be tested in detailed future studies.

3.2. Sample issues: human mRNA quantity

Particular biological fluids, i.e., vaginal secretion and menstrual blood, are naturally contaminated with bacteria. Therefore, total RNA concentration measurements with Bioanalyzer, or any other conventional methods that are all based on the overall content of RNA (human and non-human), do not accurately reflect the content of human RNA in such samples. Likely because of this, we found in the singleplex mRNA pre-testing that some of the vaginal secretion and menstrual blood samples tested revealed lower numbers of sequencing reads than samples from other tissues (data not shown). One way to compensate for this problem would be increasing the total input amount of the RNA template; however, this is not always possible in forensic casework.

Table 2
GRCh37/hg19 genomic coordinates of forensic STRs and amelogenin (BED file) included in the Ion Torrent HID 10plex system used for human identification in combination with mRNA sequencing with the Ion Torrent PGM.

Locus	Chromosome	Start	End
AMELX	chrX	11315050	11315050
AMELY	chrY	6737935	6737935
CSF1PO	chr5	149455886	149455937
D16S539	chr16	86386308	86386351
D3S1358	chr3	45582229	45582292
D5S818	chr5	123111250	123111293
D7S820	chr7	83789542	83789593
D8S1179	chr8	125907105	125907156
TH01	chr11	2192318	2192345
TPOX	chr2	1493425	1493456
vWA	chr12	6093143	6093210
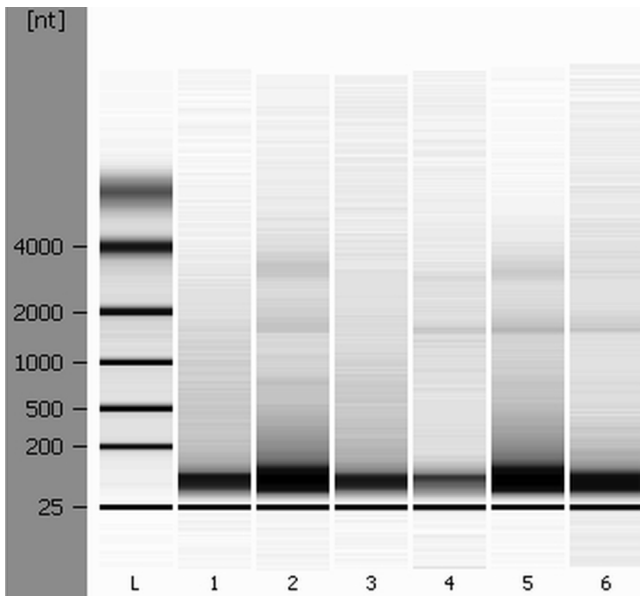
Fig. 1. BioAnalyzer gel electrophoresis exemplifying the degree of RNA degradation of the samples used in the current study. The #2–6 – RNA samples were isolated from different body fluids that were air-dried under room temperature: peripheral blood, vaginal secretion, menstrual blood, semen, and saliva, respectively, while the #1 RNA sample was isolated from skin biopsy material. L: size standard.

Therefore, in the preliminarily mRNA singleplex testing we increased the number of AmpliSeq PCR cycles during sequencing library preparation. Instead of 20 PCR cycles recommended in the standard manufacturer's protocol, we used 30 cycles of amplification for vaginal secretion and menstrual blood samples. As expected, this small change in protocol resulted into an increase of sequencing reads without noticeable negative effect on sequencing accuracy (data not shown). However, in the final multiplex analysis and regardless of the sample used (see below) the application of 20 cycles did provide accurate tissue identification for all samples tested, including those from vaginal secretion

and menstrual blood. Nevertheless, our pretesting results suggest that for the future practical application of the method, any unknown sample which fails to amplify during the library preparation using the standard protocol (i.e., 20 cycles) shall be repeated with an increased number of amplification cycles (i.e., 30). Notably, failure with the standard protocol in future applications could be because of the presence of vaginal and menstrual blood samples due to the human/non-human RNA effect described above, or due to any other human forensic samples with limited amounts of human RNA for simple quantity reasons.

3.3. Sequencing quality issues

Quality control of PGM sequencing runs (data not shown) indicated that only about half of the sequencing reads were usable. Apparently this indicates that our sequencing runs were not ideal as compared to theoretically achievable number of usable reads. However, this did not prevent the successful tissue identification with the mRNA markers used or obtaining the correct STR profiles (see below) likely because of the relatively small number of DNA/RNA markers and of pooled barcoded samples included in our experiments relative to the overall sequencing capacity of the chips used. If many more DNA and RNA markers will be included in the multiplex sequencing including those for other forensic purposes, this may become an issue. Objective quality metrics have yet to be established via additional future work.

3.4. Ion Torrent PGM mRNA profiling for tissue identification purposes

The results of mRNA sequencing using multiplex analysis of all RNA markers and together with amelogenin and 9 STRs revealed that the tissue-targeting mRNA markers constitute the vast majority of sequencing reads in each of the target tissues in all samples tested (Fig. 1), except for certain cases of naturally mixed body fluids, e.g., one of the two menstrual blood samples used. The tissue specificity of all mRNA markers tested was very high; almost all of the markers in non-target tissues were either not expressed or expressed extremely low compared to the target tissues (Suppl. Table 1).
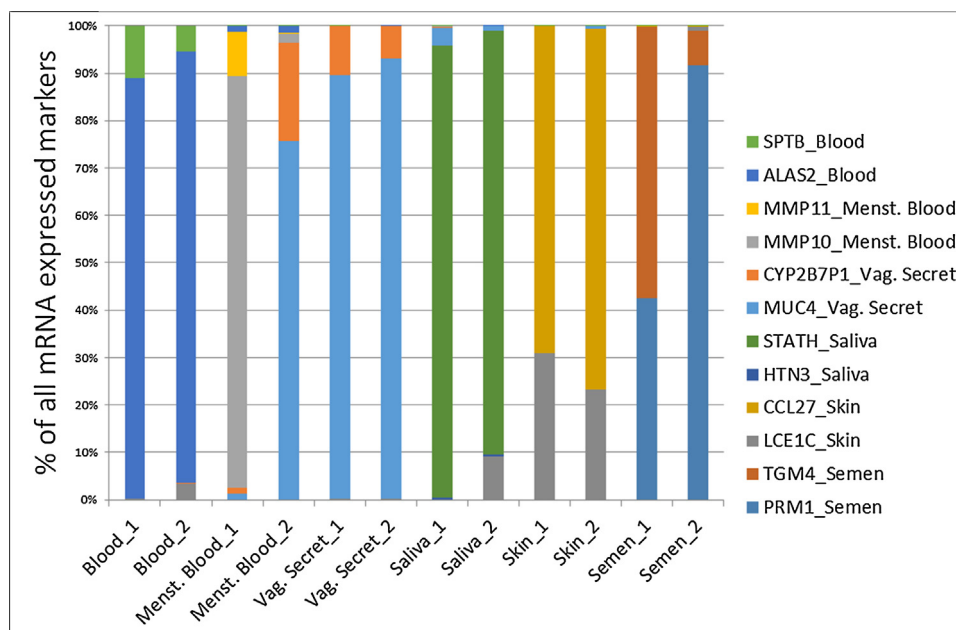
Fig. 2. Results of Ion Torrent PGM sequencing of tissue-specific mRNA markers (expressed as relative number of sequencing reads per each sample) as performed simultaneously together with STR/amelogenin sequencing.

Table 3
Binary logistic regression analysis of 12 mRNA markers used for tissue identification as carried out in combination with STR/amelogenin sequencing with the Ion Torrent PGM. Significant markers are highlighted in bold.

Marker	Blood		Menstrual blood		Saliva		Semen		Skin		Vaginal secretion	
	Score	*p*-value	Score	*p*-value	Score	*p*-value	Score	*p*-value	Score	*p*-value	Score	*p*-value
SPTB	**9.274**	**.002**	.366	.545	.373	.541	.373	.542	.369	.544	.374	.541
ALAS2	**11.583**	**.001**	.400	.527	.480	.488	.480	.488	.479	.489	.480	.488
MMP11	.185	.667	**5.461**	**.019**	.229	.632	.208	.648	.234	.629	.238	.626
MMP10	.230	.631	**5.728**	**.017**	.225	.636	.230	.631	.230	.631	.230	.631
STATH	.445	.505	.445	.505	**11.124**	**.001**	.445	.505	.445	.505	.445	.505
HTN3	.422	.516	.420	.517	**10.537**	**.001**	.422	.516	.422	.516	.422	.516
TGM4	.329	.566	.329	.566	.329	.566	**8.190**	**.004**	.322	.571	.329	.566
PRM1	.333	.564	.335	.563	.334	.563	**8.355**	**.004**	.335	.563	.335	.563
CCL27	.394	.530	.394	.530	.394	.530	.390	.532	**9.836**	**.002**	.395	.530
LCE1C	.209	.647	.889	.346	.006	.936	.598	.439	**9.880**	**.002**	.793	.373
CYP2B7P1	.757	.384	3.077	.079	.696	.404	.757	.384	.762	.383	2.866	.090
MUC4	.761	.383	.159	.690	.648	.421	.760	.383	.752	.386	**9.110**	**.003**

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.04.002.

Body fluids that naturally contain cells of mixed biological origin such as menstrual blood demonstrated the presence of multiple makers; along with menstrual blood specific genes (*MMP10* and *MMP11*), blood and vaginal secretion markers were also present. Given the intrinsic heterogeneity of this particular body fluid, the observed variability in the relative abundance of markers in different samples is to be expected. One of two analyzed menstrual blood samples (#2) revealed a higher vaginal secretion component (likely picked-up during sampling procedure) than the other menstrual blood sample tested (#1) as indicated by the high number of sequencing reads for *MUC4* and *CYP2B7P1*. This not surprising as finding vaginal secretion as well as peripheral blood mRNA markers in menstrual blood was observed before [7,21–23]. However, the presence of specific menstrual blood markers *MMP11* and *MMP10* (Fig 2, Suppl. Table 1) does allow its correct assignment to menstrual blood. In general, unambiguous, 100% accurate tissue identification was achieved in all cases as confirmed by binary regression analysis. As shown in Table 3, every individual mRNA marker used appeared to be highly specific for its target tissue with the exception of CYP2B7P1, which seems to be not significantly specific for vaginal secretion due to the natural presence of vaginal mucosa in menstrual blood. However, when vaginal secretion and menstrual blood are combined together and compared versus other tissues, CYP2B7P1 appeared as highly significant classifier (score 7.421, *p*-value 0.006).

Given the small sample size in our proof-of-principle study, the statistical power of the classification with logistic regression should not be overestimated. Important to stress here, our experiments corroborate existing knowledge upon the respective tissue specificity of the selected mRNA markers, which was established in previous works based on larger number of samples using alternative technologies. This includes the mRNA markers *LCE1C* and *CCL27* and their specificity for skin samples obtainable from forensically relevant touched objects/contact traces [24], while in the present study we used skin biopsy samples for technical reasons. However, the limited samples used in this proof-

of-principle study, together with previously established marker knowledge, proved the suitability of the Ion Torrent PGM platform for forensic tissue identification purpose, which should be tested on more samples in more detailed future studies.

In addition to the 12 tissue-specific mRNA markers, two housekeeping genes, *SDHA* and *HPRT1,* were included. These genes were selected from the published list of reference gene candidates recommended for PCR normalization [20]. However, they were never thoroughly tested in the forensic context as of yet. Therefore, using these genes for normalization in our work is only provisional. At the current stage of research, we consider using these housekeeping genes as positive controls rather than proper normalizers, as no difference was observed between the normalized and raw data, in terms of specificity of any particular marker (Suppl. Table 1). Missing normalization protocol potentially may complicate the interpretation of the results of mixed stains analysis, which however appears to be very challenging with any currently available method of tissue identification using RNA markers [25,26]. Future additional work on the housekeeping genes used here will determine their use as reference genes for normalization purposes.

Theoretically, the sensitivity of the mRNA profiling with Ion Torrent PGM sequencing should not be much different from the regular qRT-PCR analysis, since AmpliSeq amplification of mRNA targets works by the same principle. Thorough sensitivity testing of the method is the topic for future studies; however, even the currently used 10 ng of total RNA per reaction is in the range of usability in practical forensic casework and can possibly be lowered.

3.5. Ion Torrent PGM STR and amelogenin profiling for individual identification purposes

In total 7 individuals, who provided their tissue samples for this project, were sequenced at 9 forensic STRs and AMELY/AMELX simultaneously (Table 2) and together with the 14 mRNAs described above using the Ion Torrent PGM. In our hands and for the samples tested, the Ion Torrent™ HID STR 10-plex protocol

Table 4
STR microvariants detected by STR sequencing carried out in combination with amelogenin and mRNA sequencing with the Ion Torrent PGM.

STR marker/allele	Variant 1	Variant 2	Variant 3
D3S1358-16	[TATC]2[TGTC]1[TATC]13	[TATC]2[TGTC]2[TATC]12	[TATC]2[TGTC]3[TATC]11
D3S1358-17	[TATC]2[TGTC]2[TATC]13	[TATC]2[TGTC]3[TATC]12	
D8S1179-15	[TATC]2[TGTC]1[TATC]12	[TATC]3[TGTC]1[TATC]11	
vWA-14	[TAGA]9[CAGA]4[TAGA]1	[TAGA]3[TGGA]1[TAGA]3[CAGA]4[TAGA]1[CAGA]1[TAGA]1	
vWA-15	[TAGA]10[CAGA]4[TAGA]1	[TAGA]11[CAGA]3[TAGA]1	

used here worked without any technical problems. Data analysis plugin HID_STR_Genotyper allowed fast and easy calling of STR alleles as well as of amelogenin alleles. Concordance testing of Ion Torrent STR profing with conventional capillary electrophoresis performed with AmpFLSTR® Identifiler® PCR Amplification Kit which includes all 10 loci, revealed no length allele differences between the two methods (Suppl. Fig. 1).

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.04.002.

The added value of sequencing STRs such as with the Ion Torrent PGM lies in the detection of repeat motif variations that do not result in length variation, as previously demonstrated for the Ion Torrent[TM] HID STR 10-plex protocol and for the Ion Torrent PGM we used here [27]. Such microvariants have the potential to provide additional resolution needed, for example in complex mixed stain analysis. In our limited set of 7 samples, we found such microvariants at 3 out of the 9 STR markers analyzed, namely at D3S1358, D8S1179, and vWA (Table 4). Notably, there was one sample that when considering only the repeat length appeared to be homozygote for D3S1358 with 16 repeats, but in fact it was heterozygote because of a single point mutation in one of the repeat blocks. When more samples will be sequenced, many more repeat motif variants are expected to be found, which adds power to the discriminating capacity of STRs in general.

For STR/amelogenin sequencing, we used 1ng of genomic DNA as recommended by the manufacturer. Performing detailed sensitivity testing goes beyond the purpose of the presented proof-of-principle study; however, it is expected that the sensitivity of STR profiling with Ion Torrent PGM will be higher than shown here. Recently [27], a sensitivity threshold of 50 pg was reported for the Ion Torrent[TM] HID STR 10-plex protocole including the same 9 forensic STRs and amelogenin that we applied here in combination with mRNA markers. Future sensitivity testing will determine if this high STR sequencing sensitivity level is also achievable when combined with mRNA sequencing as done here.

3.6. Simultaneous mRNA and STR profiling with Ion Torrent PGM

In the current protocol, we started from the very same biological sample and finished with the very same sequencing run for STR, amelgoenin, and mRNA markers simultaneously; however, we had to split the DNA and RNA sample preparation steps in between. This was necessary because the current version of the RNA AmpliSeq pipeline used here does not allow to completely avoid the amplification of genomic DNA with the mRNA/cDNA primers. In fact, for some mRNA markers tested we noted amplification from genomic DNA (data not shown) with the primers used, which motivated us to separate the DNA and RNA sample preparation steps. However, ongoing developments are aimed to solve this problem via intelligent primer design, so that in the future all steps are expected to be done simultaneously.

In our experiments, we found that simultaneous sequencing of up to 16 barcoded samples (8 DNA and 8 RNA samples) caused no deleterious effect on both mRNA or STR profiling; apparently, the capacity of the Ion Torrent chip 318 v.2 allows much higher level of multiplicity, which shall be explored more in future studies.

All experiments in this study were carried out from single source samples and the investigation of mixed samples will be a matter of future detailed studies. As discussed elsewhere [28], correct tissue identification from mixed stain material is challenging with any method currently used, especially as long as appropriate methods to quantify human RNA prior to analysis are missing. This in principle also applies to NGS-based methods that however, have the advantage that the results are quantified directly (in contrast to currently used CE-based methods) via counting of sequencing reads. This is expected to improve mixed stain analysis in general and has already been demonstrated for forensic STR and SNP analysis in preliminary studies [27,29] with more work needed.

Previously, the analysis of STRs and a small number of miRNAs amplified from the same sample via separated PCR via a combined CE-based approach was proposed [30,31]. However, limitations in the type and number of miRNA markers used did not allow the identification of all forensically relevant tissues as we achieved here via combined NGS. Furthermore, the traditional CE-approach provides strong limitations to the total number of D/RNA markers combinable in the end-point reading due to limitations in fragment-size length spacing and the limited number of available fluorescence dyes. Unlike previous approaches, NGS such as via the Ion Torrent PGM used here, in principle allows simultaneous analysis of large numbers (many hundreds up to several thousands) of genetic markers as for instance we previously demonstrated for >500 Y-SNPs [16]. The PGM-based NGS tool for combined STR/amelogenin/mRNA analysis for simultaneous individual and tissue identification that we introduced with this proof-of-principle study serves as a starting point and we envision adding many more genetic markers for the same and additional forensic purposes in the future.

4. Conclusions

For the first time, we demonstrate the feasibility of parallel sequencing analysis of various nucleic acid markers of different types (i.e., STRs, amelogenin insertion/deletion, and mRNAs) for different forensic purposes (i.e., individual, sex, and tissue identification) simultaneously, and present the Ion Torrent PGM as a suitable platform to achieve this goal. Given the proof-of-principle nature of our study, it is up to future studies to investigate more forensically relevant aspects such as sensitivity, robustness, and reliability, and other elements of a formal forensic developmental validation study. Future work is aimed to increase the number of forensic STR loci included in the system, and to modify the RNA/cDNA primers allowing a fully parallel analysis protocol. Moreover, in the future, the suitability of the Ion Torrent PGM platform in combining targeted sequencing of different types of nuclear acid markers for different forensic purposes shall be further expanded, such as by additionally adding appearance and ancestry predicting SNPs for investigating purposes for cases where the generated STR profile does not match any known person. Eventually, we envision a targeted NGS tool that combines nucleic acid markers of different types for all relevant forensic purposes in a single analysis to provide a highly effective molecular tool to improve forensic investigation, especially in cases where the evidence material is highly limited and not enough for a battery of independent tests. Future work will show if the currently available NGS technologies and machineries, such as Ion Torrent PGM are already suitable to achieve such all-in-one forensic tool, or if further developments in sequencing technologies have to be awaited to achieve this ambitious goal.

References

[1] M.A. Jobling, P. Gill, Encoded evidence: DNA in forensic analysis, Nat. Rev. Genet. 5 (2004) 739–751.

[2] M. Kayser, P. de Knijff, Improving human forensics through advances in genetics, genomics and molecular biology (vol. 12, p179, 2011), Nat Rev Genet. 13 (2012) 754.

[3] D. Zubakov, M. Kayser, Forensic tissue identification with nucleic acids, in: D. Primorac, M. Schanfield (Eds.), Forensic DNA Applications: An Interdisciplinary Perspective, CRC Press, 2014, pp. 385–416.

[4] M. Setzer, J. Juusola, J. Ballantyne, Recovery and stability of RNA in vaginal swabs and blood, semen, and saliva stains, J. Forensic Sci. 53 (2008) 296–305.

[5] M. Bauer, RNA in forensic science, Forensic Sci. Int. Genet. 1 (2007) 69–74.

[6] D. Zubakov, E. Hanekamp, M. Kokshoorn, W. van Ijcken, M. Kayser, RNA stable markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples, Int. J. Legal Med. 122 (2008) 135–142.

[7] C. Haas, E. Hanson, M.J. Anjos, K.N. Ballantyne, R. Banemann, B. Bhoelai, et al., RNA/DNA co-analysis from human menstrual blood and vaginal secretion stains: results of a fourth and fifth collaborative EDNAP exercise, Forensic Sci. Int. Genet. 8 (2014) 203–212.

[8] C. Haas, E. Hanson, M.J. Anjos, R. Banemann, A. Berti, E. Borges, et al., RNA/DNA co-analysis from human saliva and semen stains—results of a third collaborative EDNAP exercise, Forensic Sci. Int. Genet. 7 (2013) 230–239.

[9] C. Haas, E. Hanson, M.J. Anjos, W. Bar, R. Banemann, A. Berti, et al., RNA/DNA co-analysis from blood stains—results of a second collaborative EDNAP exercise, Forensic Sci. Int. Genet. 6 (2012) 70–80.

[10] J. Juusola, J. Ballantyne, Multiplex mRNA profiling for the identification of body fluids, Forensic Sci. Int. 152 (2005) 1–12.

[11] E. Hanson, C. Haas, R. Jucker, J. Ballantyne, Specific and sensitive mRNA biomarkers for the identification of skin in 'touch DNA' evidence, Forensic Sci. Int. Genet. 6 (2012) 548–558.

[12] M. Visser, D. Zubakov, K.N. Ballantyne, M. Kayser, mRNA-based skin identification for forensic applications, Int. J. Legal Med. 125 (2011) 253–263.

[13] E. Hanson, C. Haas, R. Jucker, J. Ballantyne, Specific and sensitive mRNA biomarkers for the identification of skin in 'touch DNA' evidence, Forensic Sci. Int. Genet. 6 (2012) 548–558.

[14] W. Parson, C. Strobl, G. Huber, B. Zimmermann, S.M. Gomes, L. Souto, et al., Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM), Forensic Sci Int-Gen. 7 (2013) 543–549.

[15] V. Codd, M. Mangino, P. van der Harst, P.S. Braund, M. Kaiser, A.J. Beveridge, et al., Common variants near TERC are associated with mean telomere length, Nat. Genet. 42 (2010) 197–199.

[16] A. Ralf, M. van Oven, K. Zhong, M. Kayser, Simultaneous analysis of hundreds of Y-chromosomal SNPs for high-resolution paternal lineage classification using targeted semiconductor sequencing, Hum. Mutat. 36 (January (1)) (2015) 151–159.

[17] C. Borsting, N. Morling, Next generation sequencing and its applications in forensic genetics, Forensic Sci. Int. Genet. (February (14)) (2015), doi:http://dx.doi.org/10.1016/j.fsigen.2015.02.002 pii: S1872-4973(15)00031-9.

[18] M. Visser, R.J. Palstra, M. Kayser, Human skin color is influenced by an intergenic DNA polymorphism regulating transcription of the nearby BNC2 pigmentation gene, Hum. Mol. Genet. 23 (2014) 5750–5762.

[19] A. Trost, J.W. Bauer, C. Lanschutzer, M. Laimer, M. Emberger, H. Hintner, et al., Rapid, high-quality and epidermal-specific isolation of RNA from human skin, Exp. Dermatol. 16 (2007) 185–190.

[20] J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, et al., Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, Genome Biol. 3 (2002) RESEARCH0034.

[21] P. Danaher, R.L. White, E.K. Hanson, J. Ballantyne, Facile semi-automated forensic body fluid identification by multiplex solution hybridization of NanoString® barcode probes to specific mRNA targets, Forensic Sci. Int. Genet. 14 (2015) 18–30.

[22] A. Lindenbergh, M. de Pagter, G. Ramdayal, M. Visser, D. Zubakov, M. Kayser, et al., A multiplex (m)RNA-profiling system for the forensic identification of body fluids and contact traces, Forensic Sci. Int. Genet. 6 (2012) 565–577.

[23] A.D. Roeder, C. Haas, mRNA profiling using a minimum of five mRNA markers per body fluid and a novel scoring method for body fluid identification, Int. J. Legal Med. 127 (2013) 707–721.

[24] C. Haas, E. Hanson, R. Banemann, A.M. Bento, A. Berti, Á. Carracedo, et al., RNA/DNA co-analysis from human skin and contact traces – results of a sixth collaborative EDNAP exercise, Forensic Sci. Int. Genet. 16 (2015) 139–147.

[25] J. Harteveld, A. Lindenbergh, T. Sijen, RNA cell typing and DNA profiling of mixed samples: can cell types and donors be associated? Sci. Justice 53 (2013) 261–269.

[26] M. van den Berge, A. Carracedo, I. Gomes, E.A.M. Graham, C. Haas, B. Hjort, et al., A collaborative European exercise on mRNA-based body fluid/skin typing and interpretation of DNA and RNA results, Forensic Sci. Int. Genet. 10 (2014) 40–48.

[27] S.L. Fordyce, H.S. Mogensen, C. Borsting, R.E. Lagace, C.W. Chang, N. Rajagopalan, et al., Second-generation sequencing of forensic STRs using the Ion Torrent HID STR 10-plex and the Ion PGM, Forensic Sci. Int. Genet. 14 (2015) 132–140.

[28] T. Sijen, Molecular approaches for forensic cell type identification: on mRNA, miRNA, DNA methylation and microbial markers, Forensic Sci. Int. Genet. (2014).

[29] C. Borsting, S.L. Fordyce, J. Olofsson, H.S. Mogensen, N. Morling, Evaluation of the Ion Torrent HID SNP 169-plex: a SNP typing assay developed for human identification by second generation sequencing, Forensic Sci. Int. Genet. 12 (2014) 144–154.

[30] Y. Li, J. Zhang, W. Wei, Z. Wang, M. Prinz, Y. Hou, A strategy for co-analysis of microRNAs and DNA, Forensic Sci. Int. Genet. 12 (2014) 24–29.

[31] D. van der Meer, M.L. Uchimoto, G. Williams, Simultaneous analysis of micro-RNA and DNA for determining the body fluid origin of DNA profiles, J. Forensic Sci. 58 (2013) 967–971.