Forensic Science International: Genetics 13 (2014) 217–223

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

An evidence based strategy for normalization of quantitative PCR data from miRNA expression analysis in forensic organ tissue identification

Eva Sauer [1], Iris Babion, Burkhard Madea, Cornelius Courts [1,*]

Institute of Legal Medicine, University of Bonn, Stiftsplatz 12, 53111 Bonn, Germany

ARTICLE INFO

Article history:
Received 11 June 2014
Received in revised form 1 August 2014
Accepted 6 August 2014

Keywords:
micro-RNA
Normalization
RT-qPCR
Organ tissue identification
Endogenous references

ABSTRACT

Messenger-RNA (mRNA)-based analysis of organ tissues and their differentiation in complex crime stains has recently been introduced as a potential and powerful tool to forensic genetics. Given the notoriously low quality of many forensic samples it seems advisable, though, to substitute mRNA with micro-RNA (miRNA) which is much less susceptible to degradation. However, reliable miRNA detection and quantification using quantitative PCR requires a solid and forensically relevant normalization strategy. In our study we evaluated a panel of 15 carefully selected reference genes for their suitability as endogenous controls in miRNA qPCR normalization in forensically relevant settings. We analyzed assay performances and expression variances in 35 individual samples and mixtures thereof integrating highly standardized protocols with contemporary methodologies and included several well-established computational algorithms. Based on these empirical results, we recommend *SNORD48*, *SNORD24*, and *RNU6-2* as endogenous references since these exhibit the most stable expression levels and the least expected variation among the evaluated candidate reference genes in the given set of forensically relevant organ tissues including skin. To account for the lack of consensus on how best to perform and interpret quantitative PCR experiments, our study's documentation is according to MIQE guidelines, defining the "minimum information for publication of quantitative real-time PCR experiments".

1. Introduction

While the inference of body fluids is a common task for instance in sexual crimes, the identification of organ tissues may be required less frequently but can produce crucial forensically relevant information. If the presence of organ tissue on an object, e.g. a weapon or a bullet, can be confirmed, this finding may be useful in crime reconstruction indicating the infliction and enabling the characterization of a particular traumatic injury of a person [1,2] and even more evidence can be gathered if the tissue can be linked conclusively to a person through DNA profiling [3].

Methods in forensic organ tissue identification used to be mostly based on immunohistological or enzymatic techniques [4–7] which may pose problems in terms of sensitivity, especially when only trace amounts of material are present, and might hamper DNA profiling.

In recent years, however, RNA based analytical methods are on the rise in forensic molecular biology [8] and several international trial exercises for forensic messenger-RNA (mRNA) analysis in the identification and differentiation of body fluids have already been conducted [9–12]. In addition, Lindenbergh et al. [13] recently presented an mRNA profiling method for the inference of organ tissues.

One drawback associated with the analysis of mRNA, however, is its susceptibility to degradation. Micro-RNAs (miRNA), due to their short length of 18–23 bp, are much less affected by degradation than mRNA. Therefore, in addition, feasibility and practicability of forensic miRNA analysis was discussed [14] and miRNA expression analysis based on quantitative PCR (qPCR) is being assayed in forensic settings by several groups since 2009 [15–18].

Quantitative PCR is widely considered as the gold standard for the quantification of miRNA expression but for qPCR to deliver a reliable and biologically meaningful report of target molecule numbers an accurate and relevant normalization of non-biological

* Corresponding author. Tel.: +49 228 738355; fax: +49 228 73 8339.
E-mail addresses: cornelius.courts@uni-bonn.de, rechtsmedizin.bonn@gmail.com (C. Courts).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.fsigen.2014.08.005
1872-4973/

variances is essential [19–22]. A robust normalization strategy that is specific for a particular experimental setup should encompass an individual and evidence based selection of one or a group of reference genes [23–25].

Previously, we introduced the first empirically based strategy for the normalization of qPCR derived miRNA expression data in forensically relevant body fluids [26]. This is to be complemented by the present study, wherein we present a group of endogenous reference genes selected out of 15 preselected markers on the basis of empirical evidence for the normalization of qPCR data from miRNA expression analysis in forensic organ tissue identification.

2. Materials and methods

2.1. Adherence to the MIQE guidelines

To facilitate reliable and unequivocal interpretation of the qPCR results reported herein, information that is rated 'essential' according to the "minimum information for publication of quantitative real-time PCR experiments" (MIQE) guidelines [27] is reported, where applicable and when available.

2.2. Organ specimens

Human organ specimens for each of the seven tested organ tissues, i.e. brain, heart muscle, kidney, liver, lung, skeletal muscle, and skin were collected from five individuals (3 males, 2 females, age 10–77 years, median: 41 years, postmortem interval ranging from ~3 h to 5 days) during medico-legal autopsies at the Institute of Forensic Medicine, Bonn. No diseased findings were admissible in the sampled organs to exclude potential interference with the analysis.

Within 1 h of sampling, the excised tissues were stored in RNAlater® solution (Life Technologies, Darmstadt, Germany) at −80 °C until further processing.

All samples were anonymized and discarded after use. The study design and experimental procedures had been approved by the ethics committee of the University Hospital of Bonn.

2.3. RNA extraction and quantification

To remove ambient RNases, all devices, machines, and surfaces utilized during the extraction procedure were thoroughly cleaned using RNase-Zap® (Life Technologies) and only RNase-free reagents and plastic consumables were used.

Tissue samples were subjected to extraction of total RNA using the *mir*Vana™ miRNA Isolation Kit (Life Technologies) following the manufacturer's protocol for fresh unfrozen tissue with slight modifications. An extraction negative control underwent the same procedure.

Approximately 100 mg per thawed tissue sample were cut into pieces <2 mm^3 and incubated with 1 ml Lysis/Binding Buffer at 56 °C for 2 h prior to extraction. Unlysed tissue was passed through QIAshredder™ columns (Qiagen, Hilden, Germany) and added to the lysate. Tissue debris remaining in the columns was discarded. Total RNA eluates were stored at −80 °C until further processing.

For the elimination of potential traces of genomic DNA, DNase I digestion was carried out with the TURBO DNA-free™ Kit (Ambion) as per manufacturer's protocol and was then repeated with identical conditions as this procedure optimized removal of genomic DNA (data not shown). No inhibition testing was performed.

For the determination of total RNA concentration and quality, represented as RNA integrity number (RIN)[28], the Quant-iT™ RNA Assay Kit on a Qubit fluorometer (both Life Technologies) and the RNA 6000 Nano Kit on an Agilent 2100 Bioanalyzer (both Agilent, Böblingen, Germany) were applied, respectively.

2.4. Preparation of samples

Based on quantification results, all individual samples were diluted to 2 ng/μl. Additionally, pooled samples per organ were prepared by combining identical volumes of the five diluted individual samples. For efficiency determination experiments, a mixture of all seven organ types was created, containing identical volumes of the above mentioned pooled samples.

2.5. Selection of candidate reference genes

Candidate reference genes were selected from the manufacturer's recommended panel focusing on those with a standard deviation of the average $C_q < 1$ across 38 human tissues [29] as well as based on a literature survey encompassing a study on miRNA normalization in human tissues [30] and forensic studies [15–17,31]. The following 15 potential reference genes were selected: *hsa-miR26b-5p*, *hsa-miR-92a-3p*, *hsa-miR-93-3p*, *hsa-miR-191-5p*, *RNU6-2*, *RNY3*, *SNORA74A*, *SNORD18A*, *SNORD24*, *SNORD44*, *SNORD47*, *SNORD48*, *SNORD49A*, *SNORD58B*, and *SNORD75* (Table 1).

2.6. Reverse transcription qPCR (RT-qPCR)

Complementary DNA (cDNA) was synthesized using target-specific stem-loop primers [33] (Table 1) and the TaqMan® MicroRNA Reverse Transcription Kit (Life Technologies), following the manufacturer's protocol. Each 15 μl reaction volume contained 10 ng total RNA, 1× RT primers, 50 U MultiScribe™ reverse transcriptase, 1 mM dNTPs, 3.8 U RNase inhibitor, and 1× reverse transcription buffer. RT cycling conditions consisted of 16 °C for 30 min, 42 °C for 30 min, and 85 °C for 5 min and were performed on a T3 Thermocycler (Biometra, Göttingen, Germany). Besides extraction negative and H_2O controls, RT (−)-controls were employed to detect potential contaminations with genomic DNA. Reverse transcriptions were conducted in duplicate for all individual and pooled samples. RT reaction products were stored at −20 °C.

Target-specific TaqMan® Assays (Table 1) and the TaqMan® Universal PCR Master Mix II, No AmpErase® UNG (Life Technologies) were used for qPCR reactions according to the manufacturer's protocol. Each 20 μl reaction volume contained 1× TaqMan Universal PCR Master Mix II, and 1× specific TaqMan® Assay to which 1.3 μl of the appropriate RT reaction product were added. QPCR reactions were run in duplicates for all RT reaction products, resulting in four technical replicates per sample, with the exception of the mixture containing all organ tissues for efficiency determination, in which case triplicates were performed, amounting to six technical replicates. The internal PCR control from the Quantifiler® Human DNA Quantification Kit (Life Technologies) was used as an inter plate calibrator. All qPCRs were conducted in MicroAmp® Optical 96-Well Reaction Plates (Life Technologies) and on an ABI Prism 7500 (Life Technologies) applying the following cycling conditions: 95 °C for 10 min and 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Data collection was performed during the 60 °C step by the SDS software version 1.2.3 (Life Technologies), a limit-of-detection analysis was not done. Along with the C_q-values calculated automatically by the SDS software (threshold value = 0.2, baseline setting: cycles 3 − 15), raw fluorescence data (R_n-values) were exported for further analyses.

Table 1
Specifications of the 15 candidate reference genes.

Official gene symbol	NCBI-Alias	NCBI-/miRBase-mature sequence accession	TaqMan assay ID	Target sequence (amplicon length in base pairs)	References
hsa-miR-26b-5p		MIMAT0000083	000407	UUCAAGUAAUUCAGGAUAGGU (21)	[29]
hsa-miR-92a-3p		MIMAT0000092	000430	UAUUGCACUUGUCCCGGCCUG (21)	[29]
hsa-miR-93-3p		MIMAT0004509	002139	ACUGCUGAGCUAGCACUUCCCG (22)	[30]
hsa-miR-191-5p		MIMAT0000440	002299	CAACGGAAUCCCAAAAGCAGCUG (23)	[30]
RNU6-2	*U6*, *RNU6B*	NR_002752	001093	CGCAAGGATGACACGCAAATTCGTGAAGCGTTCCATATTTTT (42)	[15,17,29,31]
RNY3	*HY3*, *Y3*	AC005251	001214	CCAGTCACAGATTTCTTTGTTCCTTCTCCACTCCCACTGCATCACTT-AACTAGCCTT (57)	[29]
SNORA74A	*U19*, *RNU19*	X94290	001003	TTGCACCTCTGAGAGTGGAATGACTCCTGTGGAGTTGATCCTAGTC-TGGGTGCAAACAATT (61)	[29]
SNORD18A	*U18A*	AB061820	001204	CAGTAGTGATGAAATTCCACTTCATTGGTCCGTGTTTCTGAACCACA-TGATTTTCTCGGATGTTCTGATG (70)	[29]
SNORD24	*U24*, *RNU24*	NR_002447	001001	ATTTGCTATCTGAGAGATGGTGATGACATTTTAAACCACCAAGATCG-CTGATGCA (55)	[16,29,31]
SNORD44	*U44*, *RNU44*	NR_002750	001094	CCTGGATGATGATAGCAAATGCTGACTGAACATGAAGGTCTTAATTA-GCTCTAACTGACT (60)	[16,29,31]
SNORD47	*U47*, *RNU47*	X96647	001223	TAATGATTCTGCCAAATGAAATATAATGATATCACTGTAAAACCGTTCC-ATTTTGATTCTGAGGT (65)	[29]
SNORD48	*U48*, *RNU48*	NR_002745	001006	GATGACCCCAGGTAACTCTGAGTGTGTCGCTGATGCCATCACCGCAGCG-CTCTGACC (57)	[16,29]
SNORD49A	*U49*, *U49A*, *RNU49*	NR_002744	001005	CACTAATAGGAAGTGCCGTCAGAAGCGATAACTGACGAAGACTACTCC-TGTCTGATT (57)	[29]
SNORD58B	*U58b*, *RNU58B*	AB061824	001206	CTGCGATGATGGCATTTCTTAGGACACCTTTGGATTAATAATGAAAAC-AACTACTCTCTGAGCAGC (66)	[29]
SNORD75	*U75*	AF141346	001219	AGCCTGTGATGCTTTAAGAGTAGTGGACAGAAGGGATTTCTGAAATTCT-ATTCTGAGGCT (60)	[29]

NCBI, National Center for Biotechnology Information; miRBase, microRNA database [32].

2.7. Data analysis and software-based selection of endogenous reference genes

The LinRegPCR program version 2014.1 [34] was employed to compute C_q-values and amplification efficiencies from R_n-values. The arithmetic mean values of amplification efficiencies per triplicate repeats were used in further analysis, with efficiencies outside 5% of the group median being excluded from mean efficiency calculation. For C_q calculation, a common threshold value was set to −0.7 log (fluorescence). C_q-values deviating more than one cycle from the median of the technical replicates were excluded from subsequent pre-processing. For comparison, amplification efficiencies were computed analogously using the real-time PCR Miner algorithm [35].

Analysis of qPCR data including pre-processing was then performed using the GenEx software version 5.4.4 (multiD Analyses, Göteborg, Sweden) into which LinRegPCR and SDS spread sheet exported data was imported, respectively. Pre-processing of qPCR encompassed the following steps in the given order: interplate calibration, efficiency correction, and averaging of technical qPCR replicates.

To evaluate gene expression stability, we applied the following algorithms: NormFinder [36], geNorm [25,25], both implemented in the GenEx software, and the Excel-based BestKeeper [37]. NormFinder takes intra- and inter-group variances into account and provides a stability value per gene as a direct measure for the estimated expression stability, indicating the systematic error introduced when using the respective gene for normalization. Moreover, it is possible to assess the optimal number of reference genes by means of the accumulated standard deviation. GeNorm calculates and compares a so called gene stability measure (*M*-value) of all candidate genes, selecting an optimal pair of reference genes by stepwise exclusion of the gene with the highest *M*-value. BestKeeper uses pair-wise correlation analysis of the C_q-values of all pairs of candidate reference genes to determine the most stable gene. Calculations were performed separately for pooled samples per organ type and for individual tissue samples.

3. Results

3.1. Quantity and integrity

Quantity and integrity of total RNA varied among samples of the same organ tissue and between groups (Table 2). In terms of concentration, liver samples exhibited not only the highest values but also wide spread (160–999.5 ng/μl). Samples of skeletal muscle, heart muscle, and skin yielded comparatively low amounts of total RNA (120–200 ng/μl, 150–259.5 ng/μl, and 50.2–216 ng/μl, respectively). RIN values varied between 1.4 and 3.7, with skeletal and heart muscle samples displaying the

Table 2
Total RNA quantity and integrity per sample.

	Sex	Age (year)	Total RNA concentration [ng/μl] (RIN)						
			Brain	Heart	Kidney	Liver	Lung	SM	Skin
1	F	60	401 (2.5)	170 (3.7)	442.5 (2.2)	316.5 (2.9)	120 (2.9)	130 (3.1)	125 (2.2)
2	M	77	390 (2.4)	259.5 (3.1)	220 (2.4)	160 (2.1)	53.5 (1.9)	150 (2.3)	50.2 (1.4)
3	M	41	368.5 (2.4)	170 (2.3)	663 (2.0)	999.5 (2.1)	748.5 (2.3)	120 (2.8)	51.1 (1.5)
4	F	26	442.5 (2.2)	150 (3.0)	359.5 (2.2)	561.5 (2.1)	392.5 (2.2)	200 (2.3)	216 (2.3)
5	M	10	425 (2.3)	155 (3.3)	392.5 (2.3)	778 (2.3)	282 (2.6)	175 (3.6)	74.8 (2.1)
Mean			405.4 (2.4)	180.9 (3.1)	415.5 (2.2)	563.1 (2.3)	319.3 (2.4)	155 (2.8)	103.4 (1.9)

RIN, RNA integrity number; F, female; M, male; SM, skeletal muscle.

Table 3
Amplification efficiencies of candidate reference genes calculated by LinRegPCR software.

Gene symbol	Amplification efficiency of mixture	
	Mean[a]	SD
miR-26b	1.84	0.008
miR-92a	1.95	0.009
miR-93	1.85	0.012
miR-191	1.81	0.019
RNU6-2	1.84	0.012
RNY3	1.86	0.010
SNORA74A	1.58[b]	0.025
SNORD18A	1.85	0.020
SNORD24	1.85	0.029
SNORD44	1.86	0.017
SNORD47	1.86	0.015
SNORD48	1.77	0.014
SNORD49A	1.82	0.021
SNORD58B	1.81	0.021
SNORD75	1.82	0.017

SD, standard deviation.
[a] Efficiencies are given as values between 1 and 2, with 2 representing an amplification efficiency of 100%.
[b] Excluded from study.

highest (2.3–3.6 and 2.3–3.7, respectively) and skin sample the lowest values (1.4–2.3).

All extraction negatives, RT(−)- and H_2O controls were free of specific amplification (data not shown).

3.2. Amplification efficiency

Amplification efficiency per amplicon was derived from the mixture containing all seven organ tissues, including duplicated RT reactions and qPCR triplicates into the computation. Mean efficiencies per amplicon computed with LinRegPCR ranged from 95% (*miR-92a*) to 58% (*SNORA74A*) (Table 3). Due to its grossly outlying amplification efficiency, *SNORA74A* was excluded from further analysis.

3.3. Determination of most suitable reference genes

3.3.1. Pooled Samples

Analogous to our previous study [26], a first examination of the LinRegPCR spread sheet data for the pooled samples of each organ tissue including the remaining 14 candidate genes was conducted (Supplementary Table 1).

According to NormFinder results, *SNORD24* was the most stable gene (stability value = 0.3177) with only gradually increasing stability values for the eight most stable genes (Supplementary Fig. 1, upper panel). *RNY3* was least stable with a stability value of 1.6823. The accumulated standard deviation simultaneously calculated as an indicator of the optimal number of reference genes was lowest (0.1465) when eight reference genes were used (Supplementary Fig. 1, lower panel). There was, however, only a slight difference between the values when considering the combinations of between four and 13 genes.

GeNorm designated *SNORD18A* and *SNORD47* as the most stable pair of genes with an *M*-value of 0.1622 while *RNY3* displayed the highest *M*-value of 0.9036 (Supplementary Fig. 2).

BestKeeper calculated 13 out of 14 candidate genes to be stably expressed (i.e. standard deviation of C_q-value < 1.0) in the pooled sample set, with *SNORD49A* showing the least overall variation. *RNY3* was considered unstable with a standard deviation of 1.39 (Supplementary Fig. 3).

Candidate reference gene data put out by the three algorithms, respectively, were transformed into consecutively numbered ranks with 1 representing the most and 14 the least stable gene (Supplementary Table 2) and a comprehensive ranking order of gene stability was attained by calculation of the arithmetic mean ranking value per gene. In this comprehensive ranking *SNORD44* was top ranked in the pooled sample set, followed by *SNORD49A*, *SNORD24*, *SNORD18A*, and *SNORD47*. The least stable genes were *SNORD58B*, *SNORD75*, *miR-26b*, *miR-92a*, and *RNY3*.

Since these five markers were congruently ranked least stable with all three applied methods, they were improbable to be chosen for normalization purposes. *SNORD58B*, *SNORD75*, *miR-26b*, and *miR-92a* were excluded from further analyses, while *RNY3*, as the least stable candidate marker, was kept to confirm its relative instability compared to the other markers in the individual sample set and therefore, that the general outcome of the study was not biased by the exclusion of potential reference genes.

3.3.2. Individual samples

The remaining 10 candidate reference genes were subsequently analyzed in the 35 individual samples, i.e. five biological replicates per organ tissue (LinRegPCR spread sheet data, Table 4 and Supplementary Table 3).

NormFinder ranked *SNORD48* as the most stable gene with a stability value of 0.2918, followed by *SNORD24* and *RNU6-2*, whereas *RNY3* was the least stable gene (stability value = 1.4061) (Fig. 1, upper panel). The accumulated standard deviation was lowest (0.2039) when the use of eight reference genes was assumed (Fig. 1, lower panel). There emerged, however, only negligible differences between the values when considering the combinations of three to nine genes.

SNORD48 and *SNORD24* was the most stable pair of genes with an *M*-value of 0.4808 as per geNorm computations, followed by *RNU6-2* and *SNORD44* (Fig. 2). The least stable gene was *RNY3* (*M*-value = 0.9599).

Analyzed with BestKeeper, the candidate genes *miR-93*, *SNORD49A*, *SNORD48*, *miR-191*, *RNU6-2*, and *SNORD24* exhibited standard deviations of the C_q-values of <1.0 and were therefore considered to display stable expressions in the individual samples (Fig. 3).

The comprehensive ranking designated *SNORD48* as the most suitable reference gene for the given single sample set, followed by *SNORD24*, *RNU6-2*, and *SNORD49A* (Table 5). *RNY3* was evidently the least suitable candidate gene.

Table 4
Mean C_q-values and standard deviation per candidate reference gene of five individual organ tissue samples after pre-processing (amplification efficiency and C_q-values as per LinRegPCR software).

Organ	miR-93	miR-191	RNU6-2	RNY3	SNORD18A	SNORD24	SNORD44	SNORD47	SNORD48	SNORD49A
Brain	26.05 ± 0.45	19.72 ± 0.23	24.69 ± 0.20	21.82 ± 0.57	21.83 ± 0.44	24.11 ± 0.43	23.75 ± 0.36	22.34 ± 0.51	20.45 ± 0.24	22.91 ± 0.28
Heart muscle	26.22 ± 1.07	20.92 ± 0.82	23.94 ± 0.73	21.93 ± 1.64	21.29 ± 1.20	23.19 ± 0.80	22.83 ± 1.09	21.77 ± 1.05	20.08 ± 0.69	23.27 ± 0.74
Kidney	26.43 ± 0.72	21.62 ± 1.31	24.70 ± 2.05	23.81 ± 1.29	23.24 ± 2.76	24.70 ± 1.81	24.43 ± 2.49	23.64 ± 2.32	21.37 ± 1.61	23.58 ± 1.95
Liver	27.11 ± 0.84	21.96 ± 0.60	25.30 ± 1.17	25.98 ± 1.44	22.55 ± 1.88	24.97 ± 1.10	23.95 ± 1.64	23.42 ± 2.08	21.44 ± 1.90	23.30 ± 0.82
Lung	25.77 ± 0.28	20.31 ± 0.42	23.99 ± 0.53	23.53 ± 0.62	21.46 ± 0.91	23.93 ± 0.55	23.18 ± 0.76	22.30 ± 1.15	20.60 ± 0.28	22.86 ± 0.41
Skeletal muscle	25.85 ± 0.27	20.44 ± 0.22	23.48 ± 0.79	22.04 ± 1.02	21.14 ± 0.58	22.80 ± 0.26	22.43 ± 0.51	21.70 ± 0.39	19.69 ± 0.28	22.88 ± 0.54
Skin	26.77 ± 0.45	21.30 ± 0.65	24.28 ± 0.74	24.48 ± 1.59	20.09 ± 0.54	23.25 ± 0.36	22.91 ± 0.84	21.08 ± 0.45	20.42 ± 0.58	22.56 ± 0.55

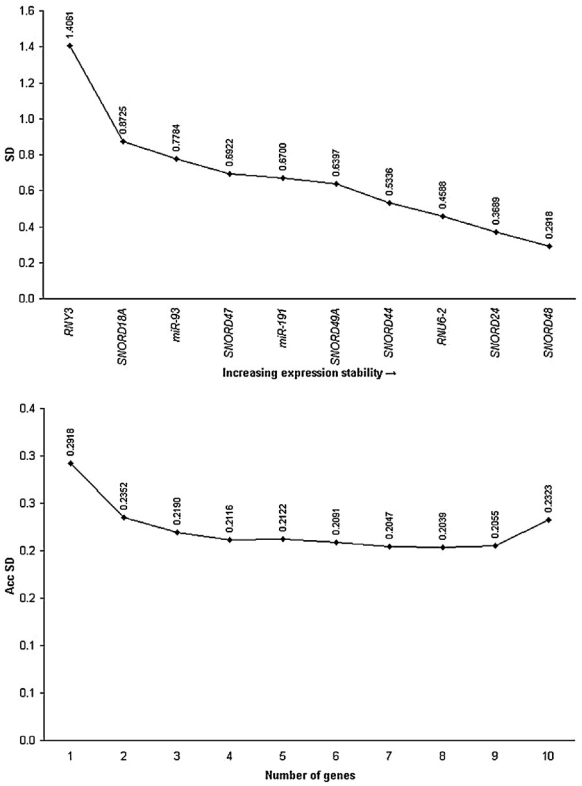

Fig. 1. NormFinder data analysis of 10 candidate genes in individual organ tissue samples. (Upper panel) Gene expression stability values of genes – from least (left) to most stable (right). (Lower panel) Determination of the optimal number of reference genes by computation of accumulated standard deviation values. Amplification efficiency and C_q-values as per LinRegPCR software.

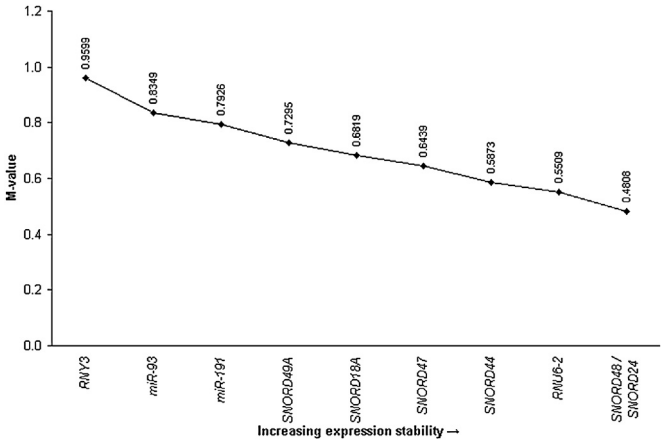

Fig. 2. GeNorm data analysis of 10 candidate genes in individual organ tissue samples. GeNorm proceeds by calculation of the gene stability measure (*M*-value) per gene – from least (left) to most stable (right); determination of the optimal pair of reference genes by stepwise exclusion of the gene with the highest *M*-value. Amplification efficiency and C_q-values as per LinRegPCR software.

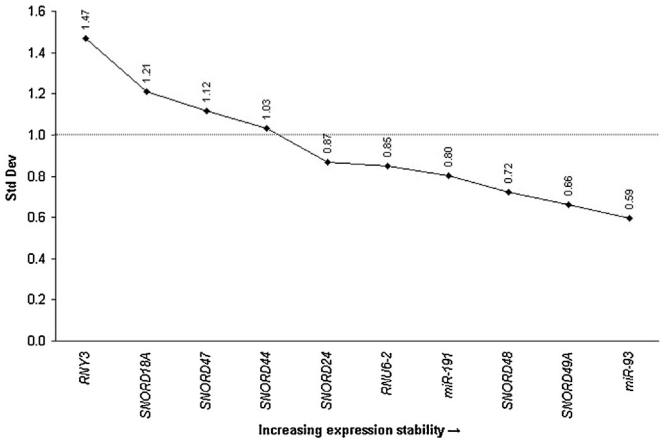

Fig. 3. BestKeeper data analysis of 10 candidate genes in individual organ tissue samples. BestKeeper proceeds by pair-wise correlation analysis of the C_q-values of all pairs of candidate reference genes – from least (left) to most stable (right). Amplification efficiency and C_q-values as per LinRegPCR software.

3.3.3. *Computations based on efficiencies calculated by Real-time PCR Miner*

Analyses of the simultaneously computed C_q-values as per SDS software, efficiency corrected according to Real-time PCR Miner (data not shown), resulted in slightly different ranking orders, but a similar overall outcome (Supplementary Table 4). This has already been observed in previous work [26] and thus and to declutter the results the separate data set based on Real-time PCR Miner calculations is not reported herein.

4. Discussion

Quantitative PCR can deliver reliable and biologically meaningful results only if an accurate and relevant normalization of non-biological variances is applied [19–22]. Non-biological variances can include variations in PCR efficiency, amount of starting material by sample-to-sample variation, RNA integrity, RT efficiency, and cDNA sample loading [38–40]. To compensate for internal non-biological variances, the use of appropriate endogenous reference genes is essential, which have to be selected based on their empirically proven suitability in a given experimental setting. This is particularly important for qPCR based miRNA analysis for which a general agreement on methodological standardization has not been achieved yet.

Table 5
Comprehensive ranking order of the candidate reference genes for individual organ tissue samples, derived by integrating rankings of NormFinder, geNorm, and BestKeeper. Computation with amplification efficiency and C_q-values as per LinRegPCR software.

Ranking order	NormFinder	geNorm	BestKeeper	Comprehensive ranking (mean rank value)
1	*SNORD48*	*SNORD48* and *SNORD24*	*miR-93*	*SNORD48* (1.67)
2	*SNORD24*		*SNORD49A*	*SNORD24* (3.00)
3	*RNU6-2*	*RNU6-2*	*SNORD48*	*RNU6-2* (3.67)
4	*SNORD44*	*SNORD44*	*miR-191*	*SNORD49A* (4.67)
5	*SNORD49A*	*SNORD47*	*RNU6-2*	*SNORD44* (5.00)
6	*miR-191*	*SNORD18A*	*SNORD24*	*miR-191*, *miR-93* (6.00)
7	*SNORD47*	*SNORD49A*	*SNORD44*	
8	*miR-93*	*miR-191*	*SNORD47*	*SNORD47* (6.67)
9	*SNORD18A*	*miR-93*	*SNORD18A*	*SNORD18A* (8.00)
10	*RNY3*	*RNY3*	*RNY3*	*RNY3* (10.00)

Previously, we established a normalization strategy for five forensically relevant body fluids, including highly standardized protocols in terms of handling, storage, and extraction of samples to minimize external variances and empirically determined *SNORD24*, *SNORD38B*, and *SNORD43* as the most suitable endogenous references from a preselected panel of candidate genes in a given sample set [26]. Analogously, the aim of the present study was to establish a reliable and empirically derived reference framework for normalization of qPCR data in the analysis of miRNA expression in six forensically relevant organ tissues and skin. The selection of 15 candidates for the starting panel was based upon a recommended panel of reference genes [29] and a literature survey [15–17,30,31].

Another important aspect that has to be accounted for in qPCR data analysis is PCR efficiency, which is known to vary in samples across different tissues, e.g. due to inhibitors and with variations on the total RNA fraction pattern extracted. Thus, a separate determination of qPCR efficiency for each performed transcript is necessary [19,21,41] as omission of correction for differential PCR efficiencies may bias the expression results [24,42–45].

As yet, several algorithms for computing C_q-values and PCR efficiencies from raw fluorescent data have been presented in the literature and Ruijter et al. [46] recently published a first comprehensive benchmark study of the evaluation of nine of these methods. They reported that the LinRegPCR method [34] was top ranked for precision and resolution and also showed high linearity without introducing excessive bias [46].

To account for LinRegPCR's tendency to underestimate PCR efficiencies compared to the standard curve analysis method we did not apply the commonly used efficiency criteria (90–110% of efficiency), but accepted efficiencies down to 70% as calculated per LinRegPCR software.

In addition, we employed the Real-time PCR Miner method for efficiency computation [35], which was highly ranked in the benchmark study by Ruijter et al. [34] as well. In accordance with our previous work, no meaningful differences were seen between the reference gene rankings put out by computations with LinRegPCR or Real-time PCR Miner efficiency data, respectively. We used the amplification efficiencies as computed from a mixture of all seven organ tissues for all subsequent calculations, which is in our view the most conservative approach since it is usually unknown which tissues types are present in a given casework sample.

There are several algorithms available for the identification of the most suitable endogenous reference out of a set of candidate genes. As yet, there is no consensus as to which of these performs best, we employed NormFinder, geNorm, and BestKeeper, representing three well-established and frequently used algorithms, and reported both the results for each algorithm individually and in combination as described by Wang et al. [47].

We further took into account that the composition of a sample encountered in forensic casework is usually unknown in terms of the number and types of organ tissues present as well as the identity of the contributor(s). In our view, it was the most conservative approach to compromise on the recommendation of a group of reference genes for a normalization procedure that assumes that all relevant types of organ tissue may be present in a given specimen. It is possible though, to apply a more specific normalization strategy if, for whatever reason, a sample's composition is less unknown, e.g. if the presence of one organ tissue can reliably be excluded. This would have to be validated in the given setting.

An important difference between this and previous studies on miRNA expression data normalization [30] lies in the properties of herein enclosed samples that due to the forensic scope of this study do not represent two conditions of the same tissue or cell type (e.g. healthy/cancerous or treated/untreated) but up to six distinct organ tissues and skin. The more types of tissue are included in a thus increasingly complex mixture, the higher the variances of expression values for any one reference gene are expected to be.

While it was unlikely in the first place to identify one or even a group of reference genes that exhibit no or very low expression variances between samples, we did, however, observe generally lower variances among the expression values of the different markers between tissue types and within the biological replicates per organ tissue as compared to those found in body fluids [26]. This may be based on the inherent complexity of body fluids, which consist of multiple and different cell types, while organ tissues appear to be more homogenic. This difference is also exemplified by the outputs of the three applied algorithms, e.g. 13 out of 14 and 6 out of 10 candidate genes displaying a stable expression as per BestKeeper in the pooled and individual sample set, respectively.

An initial screening of the candidate genes' suitability as endogenous references in forensic organ tissue identification was performed using pooled samples per tissue type, counterbalancing the expected inter-individual differences. The hereby determined ranking order was the basis for the exclusion of the markers *SNORD58B*, *SNORD75*, *miR-26b*, and *miR-92a*, which were less stable, compared to the other candidates. The consistent last rank of *RNY3* in the pooled as well as in the individual sample sets legitimates this exclusion retrospectively.

The final selection of the most suitable reference genes was then based on the performance of the nine best ranked candidate reference genes from the pooled sample set in a single sample set, since by this means both the differences between organ tissues as well as among different individuals were taken into account.

The comprehensive ranking designated *SNORD48* and *SNORD24* as the most stable reference genes followed by *RNU6-2*. All three markers were coherently top ranked by NormFinder and geNorm, and were considered as stably expressed by the BestKeeper algorithm. Additionally, the accumulated standard deviation computations by NormFinder as an indicator of the optimal number of reference genes did not improve greatly when considering more than three genes.

We are aware of the relatively small sample size and aim to further assess the reference gene panel in terms of its value for data normalization in future studies in which miRNA candidates for the identification of organ tissues and skin will have to be validated with the presented normalization strategy. This validation also will have to encompass degraded and/or compromised as well as mixed samples to represent realistic forensic conditions.

5. Conclusion

Herein we analyzed 15 potential reference genes and empirically determined *SNORD48*, *SNORD24*, and *RNU6-2* to be the most stable endogenous reference genes for a reliable normalization of qPCR data from forensic analysis of organ tissues in a given set of organ tissue samples.

Acknowledgement

We would like to thank the forensic pathologists at the Institute of Legal Medicine Bonn for their help with sample acquisition.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2014.08.005.

References

[1] W. Weimann, Über das Verspritzen von Gewebsteilen aus Einschussöffnungen und seine kriminalistische Bedeutung, Dtsch. Z. Gerichtl. Med. 17 (1931) 92–105.
[2] A. Brüning, F. Wiethold, Die Untersuchung und Beurteilung von Selbstmörderschusswaffen, Dtsch. Z. Gerichtl. Med. 23 (1934) 71–82.
[3] M. Kleiber, D. Stiller, P. Wiegand, Assessment of shooting distance on the basis of bloodstain analysis and histological examinations, Forensic Sci. Int. 119 (2001) 260–262.
[4] T. Takata, S. Miyaishi, T. Kitao, H. Ishizu, Identification of human brain from a tissue fragment by detection of neurofilament proteins, Forensic Sci. Int. 144 (2004) 1–6.
[5] A. Kimura, H. Ikeda, S. Yasuda, K. Yamaguchi, T. Tsuji, Brain tissue identification based on myosin heavy chain isoforms, Int. J. Legal Med. 107 (1995) 193–196.
[6] Y. Seo, E. Kakizaki, K. Takahama, A sandwich enzyme immunoassay for brain S-100 protein and its forensic application, Forensic Sci. Int. 87 (1997) 145–154.
[7] K. Takahama, Forensic application of organ-specific antigens, Forensic Sci. Int. 80 (1996) 63–69.
[8] M. Bauer, RNA in forensic science, Forensic Sci. Int. Genet. 1 (2007) 69–74.
[9] M. van den Berge, A. Carracedo, I. Gomes, E.A. 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.
[10] C. Haas, E. Hanson, W. Bar, R. Banemann, A.M. Bento, A. Berti, et al., mRNA profiling for the identification of blood – results of a collaborative EDNAP exercise, Forensic Sci. Int. Genet. 5 (2011) 21–26.
[11] 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.
[12] 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.
[13] A. Lindenbergh, B.M. van den, R.J. Oostra, C. Cleypool, A. Bruggink, A. Kloosterman, et al., Development of a mRNA profiling multiplex for the inference of organ tissues, Int. J. Legal Med. 127 (2013) 891–900.
[14] C. Courts, B. Madea, Micro-RNA – a potential for forensic science? Forensic Sci. Int. 203 (2010) 106–111.
[15] E.K. Hanson, H. Lubenow, J. Ballantyne, Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs, Anal. Biochem. 387 (2009) 303–314.
[16] D. Zubakov, A.W. Boersma, Y. Choi, P.F. van Kuijk, E.A. Wiemer, M. Kayser, MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation, Int. J. Legal Med. 124 (2010) 217–226.
[17] C. Courts, B. Madea, Specific Micro-RNA signatures for the detection of saliva and blood in forensic body-fluid identification, J. Forensic Sci. 56 (2011) 1464–1470.
[18] 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.
[19] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Res. 29 (2001) e45.
[20] S.A. Bustin, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems, J. Mol. Endocrinol. 29 (2002) 23–39.
[21] M.W. Pfaffl, G.W. Horgan, L. Dempfle, Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR, Nucleic Acids Res. 30 (2002) e36.
[22] S. Bustin, T. Nolan, Data analysis andinterpretation, in: S. Bustin (Ed.), A–Z of Quantitative PCR, first ed., International University Line, La Jolla, CA, 2004, pp. 439–492.
[23] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method, Methods 25 (2001) 402–408.
[24] M.W. Pfaffl, M. Hageleit, Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR, Biotechnol. Lett. 23 (2001) 275–282.
[25] 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.
[26] E. Sauer, B. Madea, C. Courts, An evidence based strategy for normalization of quantitative PCR data from miRNA expression analysis in forensically relevant body fluids, Forensic Sci. Int. Genet. 11 (2014) 174–181.
[27] S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, et al., The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments, Clin. Chem. 55 (2009) 611–622.
[28] A. Schroeder, O. Mueller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, et al., The RIN: an RNA integrity number for assigning integrity values to RNA measurements, BMC Mol. Biol. 7 (2006) 3.
[29] L. Wong, K. Lee, I. Russell, C. Chen, Endogenous Controls for Real-Time Quantitation of miRNA Using TaqMan MicroRNA Assays, Application Note (Applied Biosystems), 2010, pp. 1–8 (127AP11-01).
[30] H.J. Peltier, G.J. Latham, Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues, RNA 14 (2008) 844–852.
[31] Z. Wang, J. Zhang, H. Luo, Y. Ye, J. Yan, Y. Hou, Screening and confirmation of microRNA markers for forensic body fluid identification, Forensic Sci. Int. Genet. 7 (2013) 116–123.
[32] S. Griffiths-Jones, The microRNA registry, Nucleic Acids Res. 32 (2004) D109–D111.
[33] C. Chen, D.A. Ridzon, A.J. Broomer, Z. Zhou, D.H. Lee, J.T. Nguyen, et al., Real-time quantification of microRNAs by stem-loop RT-PCR, Nucleic Acids Res. 33 (2005) e179.
[34] J.M. Ruijter, C. Ramakers, W.M. Hoogaars, Y. Karlen, O. Bakker, M.J. van den Hoff, et al., Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data, Nucleic Acids Res. 37 (2009) e45.
[35] S. Zhao, R.D. Fernald, Comprehensive algorithm for quantitative real-time polymerase chain reaction, J. Comput. Biol. 12 (2005) 1047–1064.
[36] C.L. Andersen, J.L. Jensen, T.F. Orntoft, Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets, Cancer Res. 64 (2004) 5245–5250.
[37] M.W. Pfaffl, A. Tichopad, C. Prgomet, T.P. Neuvians, Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – excel-based tool using pair-wise correlations, Biotechnol. Lett. 26 (2004) 509–515.
[38] L. Wong, H. Pearson, A. Fletcher, C.P. Marquis, S. Mahler, Comparison of the efficiency of Moloney Murine Leukaemia Virus (M-MuLV) reverse transcriptase, RNase H-M-MuLV reverse transcriptase and Avian Myeloblastoma Leukaemia Virus (AMV) reverse transcriptase for the amplification of human immunoglobulin genes, Biotechnol. Tech. 12 (1998) 485–489.
[39] W.H. Karge III, E.J. Schaefer, J.M. Ordovas, Quantification of mRNA by polymerase chain reaction (PCR) using an internal standard and a nonradioactive detection method, Methods Mol. Biol. 110 (1998) 43–61.
[40] C. Mannhalter, D. Koizar, G. Mitterbauer, Evaluation of RNA isolation methods and reference genes for RT-PCR analyses of rare target RNA, Clin. Chem. Lab. Med. 38 (2000) 171–177.
[41] P.Y. Muller, H. Janovjak, A.R. Miserez, Z. Dobbie, Processing of gene expression data generated by quantitative real-time RT-PCR, Biotechniques 32 (2002) 1372–1379.
[42] J. Hellemans, G. Mortier, P.A. De, F. Speleman, J. Vandesompele, qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data, Genome Biol. 8 (2007) R19.
[43] S. Fleige, V. Walf, S. Huch, C. Prgomet, J. Sehm, M.W. Pfaffl, Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR, Biotechnol. Lett. 28 (2006) 1601–1613.
[44] Y. Karlen, A. McNair, S. Perseguers, C. Mazza, N. Mermod, Statistical significance of quantitative PCR, BMC Bioinform. 8 (2007) 131.
[45] S.E. Larkin, S. Holmes, I.A. Cree, T. Walker, V. Basketter, B. Bickers, et al., Identification of markers of prostate cancer progression using candidate gene expression, Br. J. Cancer 106 (2012) 157–165.
[46] J.M. Ruijter, M.W. Pfaffl, S. Zhao, A.N. Spiess, G. Boggy, J. Blom, et al., Evaluation of qPCR curve analysis methods for reliable biomarker discovery: bias, resolution, precision, and implications, Methods 59 (2013) 32–46.
[47] Q. Wang, T. Ishikawa, T. Michiue, B.L. Zhu, D.W. Guan, H. Maeda, Stability of endogenous reference genes in postmortem human brains for normalization of quantitative real-time PCR data: comprehensive evaluation using geNorm, NormFinder, and BestKeeper, Int. J. Legal Med. 126 (2012) 943–952.