Erforschung und Etablierung von microRNA als forensischer Biomarker zur Identifikation biologischer Spurenarten

Dissertation

zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

Vorgelegt von

Eva Katharina Sauer

aus Lennestadt

Bonn, Dezember 2016

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Burkhard Madea

2. Gutachter: Prof. Dr. Walter Witke

Tag der Promotion: 15.05.2017

Erscheinungsjahr: 2017

Inhaltsverzeichnis

Publikationen

Die der vorliegenden kumulativen Dissertationsschrift zugrunde liegenden Arbeiten, die am Institut für Rechtsmedizin des Universitätsklinikums Bonn sowie am Institut für Rechtsmedizin des Universitätsklinikums Schleswig-Holstein entstanden sind, wurden in den folgenden Publikationen veröffentlicht beziehungsweise zur Veröffentlichung eingereicht:

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 Science International: Genetics 11 (2014) 174–181 (doi.org/10.1016/j.fsigen.2014.03.011).

E. Sauer, I. Babion, B. Madea, C. Courts. An evidence based strategy for normalization of quantitative PCR data from miRNA expression analysis in forensic organ tissue identification. Forensic Science International: Genetics 13 (2014) 217–223 (doi.org/10.1016/j.fsigen.2014.08.005).

E. Sauer, A.K. Reinke, C. Courts. Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR. Forensic Science International: Genetics 22 (2016) 89–99 (doi.org/10.1016/j.fsigen.2016.01.018).

E. Sauer, A. Extra, P. Cachée, C. Courts. Identification of organ tissue types and skin from forensic samples by microRNA expression analysis. (Manuskript befindet sich derzeit im Revisionsprozess bei Forensic Science International: Genetics).

Die notwendigen Lizenzen zum Nachdruck der Publikationen in dieser Dissertation wurden vom Elsevier Verlag erteilt.

Zudem wurden Teile der Arbeit auf folgenden nationalen und internationalen Kongressen präsentiert und teilweise als ‚Extended Abstracts' veröffentlicht:

25. Weltkongress der International Society for Forensic Genetics (ISFG), 2013, Melbourne. Poster: Evidence based strategy for normalization of qPCR data in forensic miRNA analysis. Extended Abstract: E. Sauer, B. Madea, C. Courts. Forensic Science International: Genetics - Supplement Series 4 (2013) e148–e149.

92. Jahrestagung der Deutschen Gesellschaft für Rechtsmedizin (DGRM), 2013, Saarbrücken. Präsentation: Evidenzbasierte Strategie zur Normalisierung quantitativer PCR Daten in der forensischen miRNA-Analyse.

93. Jahrestagung der DGRM, 2014, Greifswald/Heringsdorf. Präsentation: Evidenzbasierte Strategie zur Normalisierung quantitativer PCR Daten in der forensischen miRNA-Analyse von Organgeweben.

26. Weltkongress der ISFG, 2015, Krakau. Poster: Validation of forensic body fluid identification based on empirically normalized miRNA expression data. Extended Abstract: E. Sauer, A.K. Reinke, C. Courts. Forensic Science International: Genetics - Supplement Series 5 (2015) e462–e464.

36. Spurenworkshop der DGRM und der Spurenkommission, 2016, Essen. Präsentation: miRNA-basierte Identifizierung fünf forensisch relevanter Körperflüssigkeiten auf Grundlage empirischer qPCR-Daten-Normalisierung.

Zusammenfassung

Spuren von forensisch-biologischem Interesse sind kleine Antragungen von Blut, Sekreten oder Gewebefragmenten an oder auf Personen, Oberflächen oder Gegenständen. Ihre Analyse ermöglicht nicht nur den Rückschluss auf spurenlegende Personen (beispielsweise Täter oder Opfer) durch Individualisierung anhand von DNA-Profilen, sondern kann auch Informationen zu einem Handlungsablauf liefern, etwa durch die Klärung der körperlichen Herkunft der Spur. Da die ‚klassischen' Verfahren zur Spurenartidentifikation eine Reihe von Nachteilen aufweisen, werden seit einigen Jahren alternative, nukleinsäurebasierte Herangehensweisen erforscht. MiRNAs besitzen neben der Grundvoraussetzung zelltypspezifischer Expression und der Co-Analysierbarkeit mit DNA weitere Charakteristika – allem voran ihre intrinsisch hohe Widerstandsfähigkeit gegenüber Degradation – durch die sie in besonderem Maße für die Anwendung in typischerweise nur in geringen Mengen vorhandenem, beeinträchtigtem forensischen Spurenmaterial geeignet sind.

Die in dieser Dissertationsschrift zusammengefassten Studien befassten sich daher mit der Identifikation robuster miRNA-Marker für die forensisch relevantesten Körperflüssigkeiten Blut, Speichel, Sperma, Vaginalsekret und Menstruationsblut sowie erstmalig mit der miRNA-basierten Identifikation der Organgewebe Gehirn, Lunge, Leber, Niere, Herzmuskel, Skelettmuskel und Haut im forensischen Kontext.

Um reliable Aussagen zur Ausgangsmenge einer untersuchten miRNA zu erhalten und verlässlich nur wenig unterschiedliche Expressionsniveaus erfassen zu können, erfordert die angewendete RT-qPCR-Methode eine stringente, für die gegebenen Versuchsbedingungen optimierte Datennormalisierung zur Eliminierung nicht-biologischer Varianzen. Hierfür wurden zunächst in zwei separaten Studien jeweils eine Gruppe von Referenzgenen (unter den gegebenen Bedingungen zwischen den Körperflüssigkeiten beziehungsweise Organgeweben möglichst stabil exprimierte snoRNAs oder snRNAs) ermittelt und validiert.

Eine unvoreingenommene Auswahl differentiell zwischen den untersuchten Körperflüssigkeiten beziehungsweise Organgewebe exprimierter miRNA-Kandidaten wurde mit Hilfe von Microarray-Experimenten getroffen und anschließend mittels RT-qPCR analysiert.

Die Marker, die dabei in vereinigten Proben mehrerer Individuen die besten Trenneigenschaften aufwiesen, also in ihrer Zielspurenart deutlich höher exprimiert waren als in den verbleibenden Spurenarten, wurden anschließend in Einzelproben evaluiert, wodurch die Analyse durch die Ergänzung um die interindividuellen Unterschiede vervollständigt wurde.

Mit der Entwicklung eines Entscheidungsalgorithmus unter Einsatz der Diskriminanzanalyse gelang es, alle fünf Körperflüssigkeiten anhand der Expressionsniveaus von vier miRNAs in Einzelquellproben zu unterscheiden. Der entwickelte Entscheidungsbaum wurde anschließend in bis zu 36 Jahre alten Spuren, verblindeten Proben und Mischungen mehrerer Körperflüssigkeiten getestet und zeigte auch hier, insbesondere für die Identifikation von Sperma und Blut im allgemeinen Sinne (venöses Blut und Menstruationsblut) gute Ergebnisse.

Eine statistisch valide Unterscheidung der Gehirn-, Leber-, Nieren-, Herzmuskel-, Skelettmuskel- und Hautproben von den jeweils anderen Organgewebeproben gelang durch die Anwendung binär logistischer Regressionsanalysen. Mit kleinen Einschränkungen konnte das erarbeitete Klassifikationsmodell im Anschluss auf gealterte Organabriebproben, verwesendes Organgewebematerial, Mischungen mehrerer Organgewebe und Asservate von simulierten Gewaltdelikten mit Einsatz von Stich- und Schusswaffen erfolgreich angewendet werden. Zudem wurde die vollständige Kompatibilität mit dem DNA-Arbeitsablauf gezeigt.

Die in dieser Dissertationsschrift zusammengefassten Arbeiten stellen mit ihren höchsten Standards genügenden, gemäß den MIQE-Richtlinien dokumentierten miRNA-Expressionsmessungen sowie dem bias-freien, statistisch reliablen Auswertungsgang die Grundlage für den Ausbau der Methode bis hin zu einer angestrebten Anwendung in der forensisch-genetischen Fallarbeit dar.

1 Allgemeine Einleitung

In der Strafverfolgung von Gewalt- und Sexualdelikten gilt es, einen gegebenen Sachverhalt hinsichtlich dreier hierarchischer Interpretationsebenen, der Quellen-, Handlungs- und Schuldebene, zu untersuchen [1–3]. Typische Fragestellungen lauten dabei etwa „Von welcher Person stammt eine bestimmte Spur?“ (Quellenebene), „Welche Handlungsabläufe können zu dem vorgefundenen Spurenbild geführt haben?“ (Handlungsebene) und „Ist der/die Angeklagte der ihm/ihr zur Last gelegten Tat schuldig?“ (Schuldebene). Während die Beantwortung der Schuldfrage allein dem Gericht obliegt, werden zur Bearbeitung von Fragestellungen auf den verbleibenden Ebenen Sachverständige unterschiedlicher Disziplinen herangezogen.

Seit ihrer konzeptionellen Einfassung durch die Erfindung der genetischen Identifikation auf Grundlage von Restriktionsfragmentlängenpolymorphismen („genetischer Fingerabdruck“) Mitte der 1980er Jahre [4,5] befasst sich die forensische Genetik in erster Linie mit der Individualisierung biologischen Materials anhand von DNA-Profilen zur Klärung des Ursprungs einer Spur (Quellenzuordnung) und ist zu einem unverzichtbaren Bestandteil der modernen Kriminaltechnik geworden [6].

Seit einigen Jahren werden zudem vermehrt auch nukleinsäurebasierte Methoden erforscht, mit deren Hilfe Fragen der Handlungsebene analysiert und so etwa wichtige Teilaspekte eines Tatgeschehens rekonstruiert werden können. Neben der Identifizierung von Spurenarten, die Gegenstand der vorliegenden Arbeit ist, sind hier vor allem Bestimmungen zeitlicher Zusammenhänge, wie die Abschätzung des Alters einer Spur [7–11], die Bestimmung der Tageszeit, zu der eine Spur deponiert wurde [12,13], die zeitliche Eingrenzung des Alters einer Wunde [11,14–18] oder die Einschätzung der Leichenliegezeit [11,19–22] zu nennen. Des Weiteren sind beispielsweise die Bestimmung des Alters eines Spurenverursachers [23–32], die Eingrenzung der Todesursache [11,33–44], die postmortale Feststellung von Schwangerschaften [45] und toxikogenetische Untersuchungen [46–51] Gegenstand aktueller Forschung.

1.1 Forensische Relevanz der Spurenartidentifikation

Spuren von forensisch-biologischem Interesse sind kleine Antragungen von Blut, Sekreten oder Gewebeteilen an oder auf Personen, Oberflächen oder Gegenständen, deren Analyse einen Rückschluss auf die handelnden Personen und/oder einen Handlungshergang gestatten [52]. Forensisch-genetische Methoden ermöglichen die Erstellung von DNA-Profilen mittels der Analyse von Short Tandem Repeats (STRs), wodurch sich die Identität spurenlegender Personen jenseits begründbarer Zweifel feststellen lässt. Der Nachweis der DNA einer Person in einer an einem Tatort gefundenen Spur sagt jedoch nicht notwendigerweise etwas über die Umstände der Entstehung dieser Spur aus, sodass im Zweifelsfall ein vermutetes Tatgeschehen von möglichen Alternativszenarien abzugrenzen ist [53,54].

Diese Beantwortung der Fragen auf Handlungsebene gewinnt seit einiger Zeit aus zwei Gründen an Bedeutung: zum einen ermöglichen die in Sensitivität und Robustheit verbesserten STR-Analyseverfahren die erfolgreiche Untersuchung immer kleinerer Spurenantragungen und komplexerer Mischungen, zum anderen wächst das allgemeine ‚forensische Bewusstsein', sodass Tatverdächtige häufig nicht mehr den Kontakt zu ihren Opfern leugnen, sondern die Anwesenheit ihrer DNA an einem Tatort oder an geschädigten Personen durch ein harmloses, jedenfalls nicht strafrechtlich relevantes Alternativszenario erklären [3,53]. Für die Rekonstruktion eines Tatherganges spielt daher die Auswertung und Interpretation des an einem Tatort vorgefundenen Spurenbildes eine entscheidende Rolle und bei Gewalt- und Sexualdelikten ist die Analyse biologischer Spuren hinsichtlich ihrer körperlichen Herkunft, insbesondere wenn es sich um Mischspuren handelt, häufig unerlässlich. Auch wenn die Herkunft einer Spur offensichtlich scheint, ist eine sichere Bestätigung notwendig, damit sie als objektiver Beweis gelten kann [55].

Bei Verdacht auf Kindesmissbrauch etwa durch den Kindsvater oder den Lebensgefährten der Kindsmutter kann dessen DNA an der Kleidung oder Bettwäsche des Kindes durch die Übertragung von Hautzellen oder Speicheltropfen im normalen, unverdächtigen Umgang mit dem Kind erklärt werden. Der Nachweis, dass eine DNA-Spur aus einer Spermaantragung stammt, deutet dagegen auf eine andere, strafrechtlich relevante Handlung hin. Weitere Beispiele für das Gewicht, das die Identifizierung einer Körperflüssigkeit für strafrechtliche Ermittlungen von Sexualdelikten haben kann, sind etwa die Unterscheidung von venösem Blut und Menstruationsblut in Vaginalabstrichen des Opfers oder Penisab-

rieben eines Tatverdächtigen zur Feststellung vaginaler Verletzungen oder auch der Nachweis von Vaginalsekret auf einem Gegenstand, zum Beispiel einer Flasche, mit dem ein Opfer penetriert wurde. Das Vorhandensein von DNA des Opfers könnte durch die Verteidigung des Tatverdächtigen alternativ dadurch erklärt werden, dass das Opfer aus dieser Flasche getrunken hat und die DNA demnach aus einer Speichelspur stammt.

Im Rahmen von Gewaltdelikten kann, wenn es zu perforierenden oder klaffenden Verletzungen gekommen ist, an Tatorten und Tatmitteln neben Blut auch Organgewebematerial aufgefunden werden. Das etwa auf einer sichergestellten Stichwaffe gefundene Zellmaterial kann dann nicht nur der Identifizierung des Täters und des Opfers durch die Erstellung von DNA-Profilen dienen, sondern durch den differentiellen Nachweis humanen Organgewebes auf der Messerklinge auch die Einstichtiefe und damit Schwere und Gefährlichkeit einer Verletzung anzeigen. Bei Schusswaffendelikten kann die Identifikation von Organgeweben in Forwardspatter (‚Vorwärtsschleuderspuren', die durch den Austritt biologischen Materials aus einer Ausschusswunde in Flugrichtung des Projektils entstehen) oder Backspatter (‚Rückschleuderspuren', die durch den Austritt biologischen Materials aus einer Einschusswunde entgegen der Flugrichtung des Projektils entstehen) an Tatorten, auf der Kleidung eines Tatverdächtigen sowie an und in Waffen Obduktionsergebnisse und klassische kriminaltechnische Untersuchungen ergänzen und bestätigen. Zudem sind Aussagen über die Art und Schwere einer Verletzung auf diese Weise auch dann möglich, wenn die verletzte oder getötete Person vom Tatort entfernt wurde oder ausschließlich eine Waffe oder andere Gegenstände aufgefunden werden. Der Nachweis beispielsweise von Gehirngewebe auf der Kleidung eines Tatverdächtigen deutet darauf hin, dass sich dieser während der Tat (hier: ein Kopfschuss) in direkter räumlicher Nähe zu dem Opfer befand. Bei Beteiligung mehrerer Täter und Waffen kann die Identifikation von Organgeweben es außerdem ermöglichen, eine spezifische Wunde des Opfers einer bestimmten Waffe zuzuordnen und so aufzulösen, mit welcher Waffe eine tödliche Verletzung verursacht wurde [56,57].

Auch in Fällen von Leichenzerstückelung zum Zwecke der Leichenbeseitigung kann die Organgewebeidentifikation eine Handreichung für strafrechtliche Ermittlungen darstellen, da der Nachweis von Fragmenten innerer Organe, etwa in Abflussrohren, einen deutlich höheren Beweiswert enthält als der alleinige Nachweis der DNA des Opfers, wenn dieses berechtigten Zugang zum Tatort hatte, etwa weil es zu Lebzeiten dort wohnhaft war.

1.2 Methoden zur Spurenartidentifikation

Die Analyse einer Spur beginnt immer mit ihrer Identifikation als solche. Dieser erste Schritt des Erkennens einer Spur ist von grundsätzlicher Bedeutung, da eine nicht gesicherte Spur nicht untersucht wird und somit jegliche in dieser Spur enthaltene Information über den Spurenverursacher und ihre körperliche Herkunft verloren geht.

Während Haare bereits makroskopisch gut erkennbar sind, können selbst in den meisten Fällen ebenfalls gut sichtbare Blutflecke nicht unbedingt als solche erkannt werden, wenn sie sich etwa auf dunklem Textilgewebe befinden oder stark verdünnt vorliegen. Die Lokalisierung von Speichel-, Sperma- oder Vaginalsekretantragungen auf Spurenträgern ist aufgrund ihrer höchstens leicht gelblich-gräulichen Färbung optisch und ohne Hilfsmittel häufig problematisch [52]. Mittels alternativer Lichtquellen wie ultraviolettem Licht können Körperflüssigkeiten besser sichtbar gemacht und folglich besser lokalisiert werden [55,58]. Organgewebematerial ist häufig makroskopisch als solches zu erkennen, da es zumeist in Form von Gewebefragmenten vorliegt – die Zuordnung zu einem bestimmten Gewebetyp ist jedoch nicht immer ohne Weiteres möglich. Minimale Organgewebespuren, wie sie etwa von Stich- oder Schusswaffen oder Projektilen gesichert werden, sind zudem optisch nicht von dem in der Regel ebenfalls vorhandenen Blut zu unterscheiden.

Die Identifizierung der Spurenart im Sinne der Aufklärung eines Tathergangs bei Sexual- oder Gewaltdelikten erfolgt im nächsten Schritt der Probenanalyse.

1.2.1 ‚Klassische' Methoden der Spurenartidentifikation

Die bisher verwendeten und auch derzeit noch weit verbreiteten Methoden zum Nachweis von Körperflüssigkeiten basieren auf chemischen, immunchromatographischen und enzymatischen Verfahren sowie auf histologisch-mikroskopischen Untersuchungen. In der Regel wird zuerst ein präsumtiver Test durchgeführt, welcher einen ersten Hinweis auf das Vorhandensein einer Körperflüssigkeit beziehungsweise darin enthaltener Substanzen gibt. Dessen Ergebnis wird anschließend, wenn verfügbar, durch einen konfirmatorischen Test bestätigt.

Für den Nachweis von Blut gibt es eine Reihe verschiedener präsumtiver Testverfahren. Die Luminol-Methode, welche darauf beruht, dass Hämoglobinderivate die Chemilumi-

neszenz des Luminols bei seiner Oxidation in alkalischer Lösung intensiveren [59–61], ist eine der sensitivsten präsumtiven Methoden. Der Test wird meist direkt am Tatort angewendet und ist in der Lage, Blut auch auf bereits gereinigten Oberflächen sichtbar zu machen [62,63]. Der größte Nachteil der Methode ist, dass für eine erfolgreiche Durchführung absolute Dunkelheit herrschen muss. Zudem können sich falsch-positive Ergebnisse in Gegenwart kupferhaltiger Chemikalien, einiger Bleichmittel und verschiedener tierischer und pflanzlicher Proteine ergeben und je nach angewandter Zusammensetzung kann Luminol die nachfolgenden DNA-Analysen beeinträchtigen [64,65]. Eine Reihe weiterer Vortests für Blut basiert auf der Pseudoperoxidaseaktivität des Hämoglobins, die den Abbau von Wasserstoffperoxid katalysiert, wodurch es in Lösung mit Benzidin oder Phenolphthalein zu einem Farbumschlag kommt [59,60]. Falsch-positive Ergebnisse ergeben sich hier in Anwesenheit chemischer Oxidantien und pflanzlicher Peroxidasen. Deutlich spezifischer sind immunologische Methoden, die allerdings aufgrund ihrer geringeren Sensitivität die Gefahr falsch-negativer Ergebnissen bergen [60,66]. Ein zusätzlicher bestätigender Nachweis von Blut wird in der forensischen Fallarbeit relativ selten geführt, kann aber beispielsweise durch eine Kristallprobe gelingen. Bei dieser wird ein getrockneter Blutfleck mit einem Halogenid und Eisessig versetzt und erhitzt, wobei sich braune rhombische, im Mikroskop nachweisbare Kristalle bilden [59,60,62]. Auch Enzyme Linked Immunosorbent Assay (ELISA)-basierte Methoden können zum Blutnachweis eingesetzt werden [67,68].

Der am weitesten verbreitete Vortest für Sperma ist der Test auf saure Phosphatase. Dabei wird durch das Enzym ein Phosphatsubstrat (z.B. α-Naphthylphosphat) hydrolysiert und das Reaktionsprodukt führt in Gegenwart eines Diazoniumsalzes zu einem Farbumschlag [69]. Die saure Phosphatase besitzt in der Samenflüssigkeit zwar eine um zwei bis drei Größenordnungen höhere Aktivität als in anderen Körperflüssigkeiten [58,70], trotzdem kann es zu falsch-positiven Ergebnissen kommen, etwa in Gegenwart von Vaginalsekret. Der bestätigende Spermanachweis wird in der Regel durch histologisch-mikroskopische Darstellung von Spermatozoen erbracht, etwa mittels einer Hämatoxylin-Eosin-Färbung [61]. Da diese Methode jedoch nach einer Vasektomie oder bei Azoospermie aus natürlicher Ursache zu einem falsch-negativen Ergebnis führt, ist der immunologische Nachweis des prostataspezifischen Antigens [69], welcher das Vorhandensein von Samenflüssigkeit auch in Abwesenheit von Spermatozoen nachweist ebenfalls weit verbreitet.

Für den Nachweis von Speichel, Vaginalsekret und Menstruationsblut liegen ausschließlich präsumtive ‚klassische' Methoden vor. Der am häufigsten angewendete Vortest für Speichel basiert auf der stärkespaltenden Aktivität des Enzyms α-Amylase: Stärke in Lösung mit Iod erzeugt eine starke tiefblaue Färbung, die Zugabe von Speichel führt zu einer Entfärbung der Lösung [55]. Obwohl α-Amylase in Speichel eine höhere Aktivität als in anderen Körperflüssigkeiten aufweist, können falsch-positive Befunde durch beispielsweise Sperma, Vaginalsekret, Schweiß oder Muttermilch entstehen [55,69]. Der präsumtive Nachweis von Vaginalsekret ist schwer zu führen – eine Möglichkeit besteht in der histologisch-mikroskopischen Darstellung glykogenhaltiger Zellen durch Anfärbung mit Lugol-Lösung oder der sogenannten Periodic-Acid-Schiff-Reaktion [55,61]. Beide Methoden sind jedoch anfällig für sowohl falsch-positive Ergebnisse durch Mundschleimhautzellen oder Zellen aus der männlichen Harnröhre als auch für falsch-negative Befunde in getrockneten Asservaten. Als Vortest auf Menstruationsblut wird seit Kurzem der immunologische Nachweis von bei der Fibrinolyse entstehenden Abbauprodukten, den sogenannten D-Dimeren, eingesetzt [71–73].

Für den Nachweis von Organgeweben in forensischem Spurenmaterial wurden, neben den gängigen histologischen Methoden, in einzelnen Fällen ELISA-basierte Ansätze erforscht [74–79]. Die Methoden haben vergleichsweise geringe Sensitivitäten und erweisen sich daher insbesondere dann als problematisch, wenn nur wenig Zellmaterial vorhanden ist [80].

Zusammengenommen ergeben sich für die konventionellen proteinbasierten Verfahren eine Reihe von Problemen bedingt durch die teilweise hohen Fehlerraten (falsch-positive und falsch-negative Ergebnisse), die stark variierenden Sensitivitäten, die nicht immer gegebene Kompatibilität mit nachfolgenden DNA-analytischen Methoden, den Mangel an definierten Grenzwerten, wodurch es zu einer subjektiven, anwenderabhängigen Auswertung kommt, und allem voran den hohen Materialverbrauch [55,58]: da für jede potentiell vorhandene Spurenart seriell ein separater Test durchgeführt werden muss, müssen von jedem Asservat/jeder Spur mehrere Proben genommen werden, etwa durch Ausschneiden eines Teils der Spur. Dies ist bei kleinen Spurenantragungen nur begrenzt und oft nicht möglich. Zudem gibt es derzeit nur wenige Ansätze zur Bestimmung von Organgewebe aus Spurenmaterial und keine konfirmatorischen Tests für die Identifizierung von Speichel, Vaginalsekret und Menstruationsblut.

Aufgrund der Einschränkungen sind viele forensische Labore dazu übergegangen, diese Vortests zu umgehen und der Probennahme direkt die DNA-Analyse anzuschließen. Befürworter dieser Vorgehensweise argumentieren, dass der Nachweis humaner DNA implizit auch das Vorhandensein biologischen Materials belege. Durch die fehlende Einordnung der Spurenart gehen jedoch potentiell entscheidende Informationen zur Kontextualisierung einer Spur verloren [81,82].

Neue Forschungsansätze zielen deshalb darauf ab, die Identifikation der Spurenart möglichst parallel für alle relevanten Körperflüssigkeiten oder Organgewebe und dabei non-invasiv (z.B. Raman-Spektroskopie [83]) oder kompatibel mit dem DNA-Analysearbeitsablauf (nukleinsäurebasierte Methoden, siehe unten) zu gestalten.

1.2.2 Nukleinsäurebasierte Ansätze zur Spurenartidentifikation

Technische und methodische Entwicklungen sowie neue Erkenntnisse in diversen molekularbiologischen Bereichen ermöglichen seit einigen Jahren die Erforschung neuer Methoden zur Identifizierung von Körperflüssigkeiten und Organgeweben, die entweder auf der Detektion differentieller Methylierungsmuster der Desoxyribonukleinsäure (DNA; englisch: deoxyribonucleic acid) oder der Messung differentieller Expression von Ribonukleinsäuren (RNA; englisch: ribonucleic acid) – genauer Boten-RNAs (mRNAs; englisch: messenger RNAs) oder microRNAs (miRNAs) – in unterschiedlichen Zellarten basieren.

Erforderlich für die Integration einer solchen Methode in die forensisch-genetische Fallbearbeitung ist eine vollständige Kompatibilität mit dem routinemäßigen DNA-Analysearbeitsablauf, das heißt sie darf weder die STR-Analyse zur Individualisierung der Spur stören, noch den Verbrauch zusätzlichen Probenmaterials erfordern. Außerdem sollte sie möglichst eine parallele Untersuchung mehrerer oder aller relevanten Körperflüssigkeiten und Organe erlauben sowie eine hohe Spezifität für die jeweiligen Spurenarten und ausreichende Sensitivität aufweisen [55,82].

Die Methylierung der DNA gehört neben Modifikationen von Histon-Proteinen und der Chromatin-Struktur zu den sogenannten epigenetischen Veränderungen, die die Genaktivität beeinflussen, ohne die Sequenz – also den Informationsgehalt – der DNA zu verändern. Für forensisch-genetische Analysen ist das Methylierungsmuster interessant, da dieses als einzige epigenetische Modifikation in extrahierter DNA erhalten bleibt.

DNA-Methylierung erfolgt an der Position C5 des Pyrimidinrings von Cytosinen, wobei Cytosine in CpG-Dinukleotiden bevorzugte Substrate für die DNA-Methyltransferasen darstellen [84]; etwa 70 bis 80% aller CpG-Dinukleotide der humanen DNA liegen methyliert vor [85,86]. Unmethylierte CpG-Dinukleotide sind häufig gruppiert in sogenannten CpG-Inseln, welche zumeist am 5‘-Ende von Genen in Promotorenbereichen lokalisiert sind [87–89]. DNA-Methylierung führt in der Regel zur Kondensation des Chromatins und damit zur Repression der Transkription durch sterische Hinderung des Bindens von Transkriptionsfaktoren [88,90,91]. Genomweite Studien haben gezeigt, dass bestimmte Regionen, sogenannte tDMRs (englisch: tissue-specific differentially methylated regions), zwischen unterschiedlichen Geweben differentielle Methylierung aufweisen, also abhängig vom Zelltyp und dessen Funktionen hypo- oder hypermethyliert vorliegen [92–94]. Die Analyse des Methylierungsstatus ausgewählter, sich in tDMRs befindender CpGs ermöglicht somit die Identifikation der Gewebe- oder Zellart einer DNA-Probe.

Zur Detektion werden derzeit zwei unterschiedliche Methoden angewendet: eine Herangehensweise basiert auf dem Verdau der DNA mit methylierungssensitiven Restriktionsenzymen, welche die DNA nur an Erkennungssequenzen mit unmethylierten CpGs schneiden, sodass bei einer anschließenden Polymerasekettenreaktion (PCR; englisch: polymerase chain reaction) mit die Schnittstelle flankierenden Primern nur unverdaute, also an dieser Stelle methylierte DNA amplifiziert und mithin ein messbares Signal erzeugt wird [95–99]; bei der anderen Methode wird die DNA mit Natriumbisulfit behandelt, wobei unmethylierte Cytosine durch Sulfonierung und hydrolytische Desaminierung in Uracile umgewandelt werden, welche in einer nachfolgenden PCR den komplementären Einbau von Adeninen statt Guaninen bedingen – die resultierenden Einzelnukleotid-Polymorphismen können dann mittels Sequenzierung [97,100–102] oder Primerextensionsreaktionen [97,103] detektiert werden.

Die Verwendung des Methylierungsmusters für die Identifizierung von Spurenarten hat den Vorteil, dass diese Methode auch bei bereits extrahierter DNA, also auch in alten Fällen, angewendet werden kann. Ein erheblicher Nachteil der Herangehensweise besteht jedoch darin, dass DNA, die für die Analyse des Methylierungsmusters verwendet wird, nicht mehr für die Erstellung von STR-Profilen zur Verfügung steht und *vice versa*. Insbesondere die Bislufitkonvertierung bedingt den Einsatz einer relativ großen DNA-Menge, welche in forensischem Probenmaterial häufig nicht enthalten ist [3]. Hinzu

kommt, dass die Methode meist auf der Analyse des prozentualen Anteils der methylierten und unmethylierten Versionen bestimmter CpGs beruht, dieses Verhältnis sich in Gemischen mit anderen Spurenarten jedoch verschiebt und so die Analyse von Mischungen nur schwer möglich ist [3].

Während für die Identifizierung von Sperma bereits hoch spezifische Methylierungsmarker beschrieben werden konnten [98,99], ergaben sich für den Nachweis von Blut, Speichel, Vaginalsekret, Menstruationsblut und Haut unterschiedliche Ergebnisse [95–104]. Für die Identifizierung von Organgeweben anhand differentieller Methylierungsmuster im forensischen Kontext liegen bisher keine veröffentlichten Studien vor.

Messenger RNA stellt eine dynamische Zwischenstufe der Genexpression dar – als Transkript eines Gens fungiert ein mRNA-Molekül als temporärer Träger der Erbinformation und damit als Matrize für die Biosynthese der Proteine, welche die phänotypische, biologisch aktive Manifestation der Gene darstellen. Die Biosynthese der mRNA, die Transkription, erfolgt im Zellkern durch eine DNA-abhängige RNA-Polymerase an der Vorlage des komplementären Matrizenstrangs der DNA. Zunächst wird eine unreife Vorläuferform produziert, die sogenannte prä-mRNA. Durch den komplexen co-transkriptionellen Prozess des Spleißens werden nicht-codierende Sequenzen (Introns) aus der mRNA entfernt und, abhängig vom Bedarf der Zelle, die Anzahl und Kombination der verbleibenden codierenden Sequenzen (Exons) modifiziert. Außerdem erfolgen, ebenfalls co-transkriptionell, die Reifungsprozesse der Polyadenylierung (enzymatische Addition von Adenin-Nukleotiden an das 3'-Ende der prä-mRNA) und des Cappings (enzymatische Addition eines modifizierten Guanin-Nukleotids an das 5'-Ende der prä-mRNA) [105]. Diese stabilisieren die reife mRNA und ermöglichen einen gezielten Kernexport sowie die korrekte Initiation der Proteinbiosynthese, der Translation, durch Rekrutierung der Ribosomen im Zytoplasma.

Die Zusammensetzung des Transkriptoms, also die Gesamtheit aller zu einem bestimmten Zeitpunkt in einer Zelle vorhandenen mRNAs, bildet dabei präzise den aktuellen Bedarf an Genprodukten ab, welcher die Funktionen und den Zustand der Zelle repräsentiert. Das Prinzip der mRNA-basierten Spurenartidentifikation beruht darauf, möglichst spurenartspezifische Marker, also ausschließlich in den Zellen einer bestimmten Spurenart exprimierte mRNAs oder spurenartspezifische Signaturen, also in ihrer Kombinationen für die Spurenart spezifische Marker, nachzuweisen und so das Vorhandensein dieser Spurenart

festzustellen. Bei der Auswahl geeigneter Marker wird dabei nicht ein einzelner Zelltyp, sondern die Gesamtheit der durchaus heterogen zusammengesetzten Körperflüssigkeiten beziehungsweise Organgewebe betrachtet.

Seit der Publikation erster Ergebnisse zur Identifizierung von Körperflüssigkeiten mittels mRNA-Expressionsanalyse von Juusola und Ballantyne aus dem Jahr 2003 [81] wurden zahlreiche weitere einschlägige Studien veröffentlicht, die sich mit einer oder mehrerer der fünf forensisch relevantesten Körperflüssigkeiten Blut, Speichel, Sperma, Vaginalsekret und Menstruationsblut [106–122], aber auch mit anderen möglichen Spurenarten wie Schweiß, Urin und Nasensekret [122–124] befassten. In einer Studie wurden zudem mRNA-Marker für die Identifikation der Organgewebe Gehirn, Lunge, Leber, Niere, Herzmuskel, Skelettmuskel und Haut beschrieben [56]. Spezifität und Sensitivität der beschriebenen Kandidaten variieren stark, doch eine Reihe von Markern konnte in internationalen Ringversuchen durch mehrere Arbeitsgruppen bestätigt werden [125–130]. In einigen Ländern, wie etwa den Niederlanden, wird die mRNA-basierte Spurenartidentifikation bereits in der forensisch-genetischen Routine eingesetzt und bei Gerichtsverfahren eingebracht [131].

Überwiegend wurden in den veröffentlichten Studien und auch den Ringversuchen die Marker im Anschluss an die reverse Transkription der RNA-Moleküle in komplementäre DNA (cDNA; englisch: complementary DNA) mittels Multiplex-Endpoint-PCR angereichert und fluoreszenzmarkiert – und anschließend die PCR-Produkte durch kapillarelektrophoretische Auftrennung detektiert. Vorteile dieser Herangehensweise gegenüber den ‚klassischen' Methoden zur Identifizierung von Spurenarten sind die deutlich höhere Spezifität, die Möglichkeit, mehrere Marker für mehrere Spurenarten parallel in demselben Reaktionsansatz zu testen sowie die Anwesenheit einer Spurenart auch in komplexen Mischungen bestimmen zu können. Hinzu kommt, dass durch die Verwendung eines kombinierten Extraktionsverfahrens sowohl die genomische DNA als auch die mRNA aus derselben Spur erhalten werden kann [115,132] – so kann es gelingen, ohne zusätzlichen Probenverbrauch selbst bei kleinsten Tatortspuren den Spurenverursacher und die Spurenart zu identifizieren. Ein Nachteil des Endpoint-PCR-Ansatzes ist jedoch, dass die Marker so gewählt sind, dass sie möglichst ein binäres Expressionsverhalten aufweisen, also entweder ‚vorhanden' oder ‚nicht vorhanden' sind. Dadurch ist es nicht nur nicht möglich Mischungsanteile der unterschiedlichen Spurenarten zu bestimmen [133], sondern es ergeben sich auch Probleme dadurch, dass zum einen RNAs selten wirklich zelltypexklusiv

exprimiert werden, sondern vielmehr in vielen Zellarten schwache Ausprägung aufweisen – wodurch es bei unpräzise definierten Grenzwerten zu falsch-positiven Ergebnissen kommen kann [43] – und zum anderen diese rein qualitative Erhebung nur bedingt für die Entwicklung quantitativer statistischer Modelle, die die Einbindung einer Fehlerschätzung zur Einschätzung der Zuverlässigkeit des Modells erlauben, geeignet ist [43,134].

Das grundlegende Prinzip der Identifikation von Spurenarten mittels miRNA-Expressionsanalyse ähnelt dem der bereits beschriebenen mRNA-basierten Herangehensweise. Seine Anwendung im forensisch-genetischen Kontext wird bisher allerdings fast ausschließlich für Körperflüssigkeiten erforscht [82,134–145]. Die biologischen und methodischen Grundlagen dieses Ansatzes werden im nachfolgenden Kapitel 1.3 erläutert.

1.3 MicroRNA

MiRNAs sind eine Klasse einzelsträngiger, nicht-codierender RNA-Moleküle (ncRNAs) mit einer Länge von etwa 18 bis 24 Nukleotiden (nt), die eine wichtige Rolle in der post-transkriptionellen Genexpressionsregulation vieler zellulärer Prozesse spielen [146–148].

Die ersten miRNAs, bezeichnet als lin-4 und let-7, wurden 1993 und 2000 im Fadenwurm *Caenorhabditis elegans* entdeckt [149,150]. Doch erst nachdem let-7 in nahezu allen Metazoen nachgewiesen werden konnte [151] und daraufhin zahlreiche weitere miRNAs identifiziert wurden, setzten sich die Bezeichnung „microRNA“ und das Konzept von miRNAs als große, evolutionär konservierte Klasse von Riboregulatoren durch [152–154]. In der sogenannten miRBase-Datenbank, in der neu entdeckte miRNAs sequentiell nummeriert und katalogisiert werden, sind zum Zeitpunkt dieser Niederschrift 2.588 reife humane miRNAs registriert [155–157].

1.3.1 Biogenese und Funktion

MiRNA-codierende Gene finden sich über das gesamte Genom verteilt, sowohl in Introns protein-codierender Gene, in Introns und Exons ncRNA-codierender Gene als auch in intergenischen Bereichen [152–154,158,159]. Etwa die Hälfte der bekannten miRNA-Gene

befinden sich dabei in Genclustern, die in polycistronische Primärtranskripte umgeschrieben werden [152,153,159,160].

Der kanonische Biogeneseweg der miRNAs in Säugerzellen beginnt mit der Synthese primärer Transkripte, sogenannter pri-miRNAs, unterschiedlicher Länge durch RNA-Polymerasen vom Typ II [161–163]. In einer pri-miRNA, die polyadenyliert und mit 5'-Kappe vorliegt, bilden die miRNA-codierenden Bereiche etwa 60 bis 100 nt lange Haarnadelstrukturen mit unvollkommener Basenpaarung. Diese Strukturen werden noch im Nukleus durch den die Typ III Ribonuklease Drosha und das doppelsträngige RNA-bindende Protein DGCR8 enthaltenden Mikroprozessor-Komplex aus den pri-miRNAs herausgeschnitten, wodurch sogenannte Vorläufer-miRNAs (pre-miRNAs; englisch: precursor miRNAs) entstehen [164–169]. Nach dem Exportin-5-vermittelten Kernexport [170–172] werden die pre-miRNAs von einem zytoplasmatischen Proteinkomplex aus der Typ III Ribonuklease Dicer und den RNA-bindenden Proteinen TRBP und PACT in etwa 22 Basenpaare lange miRNA-Duplexe gespalten [173–178]. Eine solche nunmehr reife, noch doppelsträngige miRNA wird anschließend in einen Ribonukleoproteinkomplex mit von TRBP und Dicer rekrutierten Proteinen der Argonaut-Familie (Ago-Proteine) aufgenommen [179–181]. Liegt eine vollständige Basenpaarung der beiden miRNA-Stränge vor, wird einer der beiden Stränge durch die Endoribonuklease-Domäne der Ago2-Proteine aktiv gespalten; bei unvollständiger Basenpaarung oder Vorliegen eines der anderen drei humanen Ago-Proteine, welche keine Ribonukleaseaktivität besitzen, werden die Stränge passiv entwunden, durch die entstehende strukturelle Spannung in zwei Einzelstränge aufgetrennt und ein Strang aus dem Ribonukleoproteinkomplex entlassen, welcher dann frei im Zytoplasma vorliegt und abgebaut wird [182–187]. Nach der Dissoziation von Dicer, TRBP und PACT und der Assoziation des Proteins GW182 wird der entstandene Komplex als miRISC (englisch: miRNA-induced silencing complex) bezeichnet. Welcher miRNA-Strang als sogenannter Führungsstrang im miRISC verbleibt, ist hauptsächlich von der thermodynamischen Stabilität der miRNA-Duplexenden abhängig: der Strang mit dem instabileren 5'-Ende bleibt erhalten [188,189]. Zudem selektieren die Ago-Proteine miRNAs, die mit einem Adenin oder Uracil beginnen [190,191]. Da es sich dabei jedoch nicht um eine strikte Selektion handelt, bleibt auch der komplementäre Gegenstrang mit variierender Häufigkeit erhalten [192].

Der reife miRISC-Komplex bindet seine Zielmoleküle durch Basenpaarung zwischen der enthaltenen miRNA und den einzelsträngigen mRNA-Molekülen, wobei insbesondere die Nukleotide an den Positionen 2 bis 7 des 5'-Endes der miRNA, die sogenannte Seed-Sequenz, von Bedeutung sind [193–195]. Die Bindestellen innerhalb der mRNA liegen dabei hauptsächlich in der 3' untranslatierten Region [194], aber auch in der codierenden Sequenz [196–198] und der 5' untranslatierten Region [199,200]. Der genaue Modus der Genexpressionsregulation ist abhängig vom Ausmaß der Komplementarität der im miRISC enthaltenen miRNA zur Ziel-mRNA sowie dem enthaltenen Ago-Protein: einige wenige humane miRNAs sind fast vollständig komplementär zu ihren Zielmolekülen und induzieren in Anwesenheit von Ago2 die endonukleolytische Spalten der mRNA [201,202]; der Großteil humaner miRNAs ist jedoch nur teilweise komplementär zu den Zielmolekülen, wodurch es zur sterischen Inhibierung des Initiations- oder Elongationsschrittes der Translation [203–207] und/oder zu Deadenylierung der mRNA und damit zu deren beschleunigtem Abbau kommt [208,209].

Da für die Genexpressionsregulation eine unvollständige Basenpaarung zwischen miRNA und Ziel-mRNA ausreichend ist, kann jede miRNA mehrere – bis hunderte – Zielmoleküle haben und jede mRNA kann Ziel diverser miRNAs sein – entweder, weil das mRNA-Molekül mehrere miRNA-Bindestellen besitzt, oder weil sich mehrere miRNAs dasselbe unvollständig übereinstimmende Bindemotiv teilen [195,210,211].

Die miRNA-Biogenese, -Reifung und -Stabilität unterliegen wiederum jeweils eigenen distinkten, noch nicht vollständig verstandenen Kontrollmechanismen [160,182,212]. Die Transkriptionsregulation erfolgt, wie bei mRNAs, durch RNA Polymerase II-assoziierte Transkriptionsfaktoren und epigenetische Modifikationen [161,162,180,212]. Während intergenische miRNAs dabei offensichtlich eigene Promotoren besitzen, scheinen in Introns codierte miRNAs sich teilweise die Promotorregionen mit ihren ‚Wirtsgenen' zu teilen, teilweise aber auch eigene Promotoren zu besitzen [182,213,214]. Die posttranskriptionelle Regulation der miRNA-Reifung sei hier am Beispiel des Mikroprozessorkomplexes aufgeführt, dessen Abundanz und Effizienz auf mehreren Ebenen beeinflusst wird: eine Autoregulation des Komplexes ergibt sich durch die stabilisierende Wirkung von DGCR8 auf Drosha einerseits und die Destabilisierung der *DGCR8*-mRNA durch Spaltung einer darin enthaltenen Haarnadelstruktur durch Drosha andererseits [212,215]; die nukleare Lokalisierung und Stabilität der Proteine des Mikroprozessor-

komplexes werden dann durch Proteinmodifikationen wie Phosphorylierung oder Acetylierung reguliert [216–218]. Hinzu kommen Interaktionen mit anderen RNA-bindenden Proteinen wie Lin28, dessen Bindung an die terminale Schleife einer Haarnadelstruktur sowohl die Prozessierung einer pri-miRNA durch Drosha als auch einer pre-miRNA durch Dicer verhindert [182,212]. Lin 28 initiiert zudem die Bindung von Uridinyltransferasen, die enzymatisch Uridine ans 3'-Ende einer pre-miRNA anfügen und so den miRNA-Abbau durch 3'-5'-Exonukleasen einleiten [219]. Ein weiterer Modifikationsprozess, der die Reifung von miRNA beeinflusst, ist die sogenannte RNA-Editierung, bei der Adenosine der pri- und pre-miRNAs durch Desaminierung in Inosine umgewandelt werden und die eine nur noch ineffiziente Bindung von Drosha und Dicer zur Folge hat [182,220]. Seit Kurzem werden zudem zirkuläre Formen alternativ gespleißter mRNAs erforscht, die als „miRNA-Schwämme" fungieren können [221].

Sobald sie mit Ago-Proteinen assoziiert vorliegen, sind miRNA-Moleküle sehr stabil [222]. Ob und wie miRNAs aus dem miRISC-Komplex entlassen werden, damit sie von zytoplasmatischen Exonukleasen abgebaut werden können ist noch nicht hinreichend verstanden [182]. Einige Studien deuten auf eine durch die Zielmoleküle eingeleitete Degradation hin: bei (fast) vollständiger Basenpaarung mit der Ziel-mRNA wird die miRNA uridinylisiert und so der exonukleäre Abbau initiiert [220,223,224].

1.3.2 Gewebespezifische Expression

Die Ausprägung und Verteilung von Molekülen zu kennen ist essentiell für deren funktionelle Charakterisierung, das Verständnis physiologischer und pathologischer Prozesse, an denen sie beteiligt sind, sowie ihre Erforschung als potentielle Biomarker, etwa zur Erkennung krankhafter Veränderungen. Die entscheidende Rolle von miRNAs im komplexen System der Genexpressionsregulation – es wird geschätzt, dass über 60 % der proteincodierenden Gene beim Menschen miRNA-basierter Regulation unterliegen [211] – und damit für die phänotypische Ausprägung einer Zelle, sowie die Tatsache, dass Veränderungen in der Expression hunderter mRNAs potentiell durch nur eine oder wenige miRNAs gesteuert werden können [195,210,211], legen nahe, dass die Zusammensetzung der Gesamtheit aller zu einem bestimmten Zeitpunkt in einer Zelle vorhandenen miRNAs, das sogenannte miRNom, reichhaltige biologische Information codiert [225].

Schon kurz nach ihrer Entdeckung als weit verbreitete Klasse von Riboregulatoren, wurde die spatiotemporale Expression der miRNAs und die damit verbundene Bedeutung für die embryonale Entwicklung [147,149,226–228], die Ausbildung und Aufrechterhaltung der Zelldifferenzierung [229–232] und die Zellproliferation und Apoptose [233–235] beschrieben. Eine Dysfunktion der miRNA-vermittelten posttranskriptionellen Genregulation wurde bereits mit zahlreichen pathologischen Vorgängen und Krankheitsbildern assoziiert [236–244]. Ähnlich der Transkriptomanalyse, kann also auch die Analyse des miRNoms, Aufschluss über den Zustand und Typus einer Zelle geben.

Eine Reihe von Studien haben sich bereits mittels unterschiedlicher methodischer Ansätze mit der Untersuchung der differentiellen Expression humaner miRNAs über zahlreiche Gewebetypen hinweg befasst [245–252]. Hanson et al. waren 2009 die erste Arbeitsgruppe, die die Hypothese aufstellte, dass die differentielle miRNA-Expression auch für die forensisch-genetische Identifikation von Spurenarten herangezogen werden kann [82]. Für die Beantwortung dieser Fragestellung ist die eigentliche Funktion der einzelnen miRNAs weniger relevant; ausschlaggebend ist vielmehr ein möglichst deutlich erhöhtes Expressionsniveau der miRNA in ihrer Zielspurenart im Vergleich zu den jeweils verbleibenden Spurenarten.

1.3.3 Experimenteller Nachweis

In den hier vorgestellten Arbeiten wurden zwei distinkte Methoden zur Bestimmung von miRNA-Expressionsniveaus angewendet: miRNA Microarrays und reverse Transkription mit anschließender quantitativer Echtzeit-PCR (RT-qPCR).

Microarrays basieren auf der simultanen Detektion tausender fluoreszenzmarkierter RNA-Moleküle, die mit spezifischen komplementären DNA-Fängersonden, welche auf einem inerten Trägerchip befestigt und in Clustern angeordnet sind, hybridisieren. Auf diese Weise lässt sich das relative Expressionsmuster eines Großteils oder sogar der Gesamtheit des bekannten miRNoms in einer gegebenen Probe darstellen, woraus Rückschlüsse auf Zellart und -zustand gezogen werden können. Diese Herangehensweise benötigt jedoch relativ viel RNA-Ausgangsmaterial, was in forensischen Proben häufig nicht enthalten ist, und ist zudem aufwendig in Bezug auf Kosten und Arbeitszeit. Daher ist sie für die routinemäßige Anwendung im Labor ungeeignet und wurde hier ausschließlich für die Aus-

wahl potentiell körperflüssigkeits- oder organgewebespezifischer miRNA-Kandidaten verwendet. Durch bioinformatische Algorithmen wie der hierarchischen Clusteranalyse können aus der Vielzahl mittels Microarray analysierter miRNAs diejenigen mit den besten Trenneigenschaften, also der am deutlichsten ausgeprägten differentiellen Expression zwischen den untersuchten Zuständen oder, wie im Fall der Spurenartidentifikation, Zelltypen von Interesse ausgewählt werden. Aufgrund des eingeschränkten dynamischen Messbereichs und der lediglich semiquantitativen Natur der Ergebnisse der Microarrayanalysen ist eine anschließende Bestätigung der Tauglichkeit der selektierten Marker mit der für die routinemäßige Anwendung ausgewählten Methode, in diesen Arbeiten RT-qPCR, erforderlich.

Die hier eingesetzte TaqMan®-Assay-basierte RT-qPCR-Strategie verwendet für die reverse Transkription der miRNA-Moleküle sogenannte Stem-Loop-Primer, die eine kurze einzelsträngige, zum 3‘-Ende einer bestimmten reifen miRNA komplementäre Sequenz und eine allen Primern gemeinsame Sequenz, die eine Haarnadelstruktur (englisch: stem loop) bildet, besitzen [253]. Bei der darauffolgenden PCR werden dann ein miRNA-spezifischer Vorwärtsprimer und ein universeller, gegen die sich in der Haarnadelstruktur der RT-Primer befindende Sequenz komplementärer Rückwärtsprimer eingesetzt. Zudem bindet eine spezifische Sonde an das RT-Produkt, welche jeweils zum Teil zu der miRNA-Sequenz, die der Vorwärtsprimer nicht abdeckt, und zu dem in der RT eingesetzten Stem-Loop-Primer komplementär ist. Diese Sonde trägt an einem Ende einen Fluoreszenzfarbstoff und am anderen Ende einen sogenannten Quencher. Bei intakter Sonde befinden sich Fluorophor und Quencher in unmittelbarer räumlicher Nähe, sodass bei Anregung des Fluorophors eine strahlungsfreie Energieübertragung auf den Quencher stattfindet und keine Fluoreszenz emittiert (FRET, Fluoreszenz-Resonanzenergietransfer). Während der PCR-Amplifikation wird die Sonde durch die 5‘-3‘-Exonukleaseaktivität der Taq-Polymerase hydrolysiert, wodurch die molekulare Verbindung zwischen Fluorophor und Quencher zerstört wird und sie sich voneinander entfernen, sodass die emittierte Fluoreszenz messbar wird. Die Fluoreszenz wird in jedem PCR-Zyklus gemessen und ist proportional zur kumulativen Menge des Amplifikationsprodukts in der Reaktion. Zu Beginn ist der Anstieg sehr langsam, da nur wenige Kopien der Zielsequenz vorliegen. Dann erfolgt eine Phase exponentieller Vermehrung des Amplifikationsprodukts und damit Anstiegs der Fluoreszenzintensität, der sich wieder verlangsamt sobald die noch vorhandene Menge mindestens einer der essentiellen Reagenzien reaktionslimitierend wird. Für jede Probe-Assay-Kombi-

nation wird gemessen, in welchem PCR-Zyklus die Fluoreszenzintensität einen definierten Intensitätsschwellenwert erreicht. Dieser Zyklus wird als C_q-Wert (englisch: cycle of quantification) bezeichnet. Je niedriger der C_q-Wert ist, je früher also der Schwellenwert überschritten wird, desto mehr Kopien des Moleküls von Interesse lagen ursprünglich im Reaktionsansatz vor. Dieser Zusammenhang ist sehr genau quantifizierbar, sodass von einem C_q-Wert auf die Ausgangsmenge der Kopien der Zielsequenz rückgeschlossen werden kann.

Für aussagekräftige RT-qPCR-Ergebnisse, die den Vergleich der relativen Expression einer Ziel-miRNA in mehreren Proben erlauben und durch den Einfluss nicht-biologischer Störvariablen nicht beeinträchtigt wird, bedarf es allerdings einer strengen Standardisierung der Proben und des Versuchsablaufs sowie der Normalisierung der erhaltenen Expressionsdaten [225,254]. Hierzu wurden 2009 von Bustin et al. in den „minimum information for publication of qualitative real-time PCR experiments" (MIQE)-Richtlinien [255–257] eine Reihe von Mindestanforderungen für die Dokumentation von qPCR-Experimenten formuliert, die in den hier aufgeführten Studien erfüllt wurden (siehe Kapitel 3).

1.3.4 Vorteile miRNA-basierter Spurenartidentifikation

Neben der für die Spurenartidentifikation notwendigen Voraussetzung der differentiellen Expression in unterschiedlichen Körperflüssigkeiten und Organgeweben besitzen miRNAs eine Reihe vorteilhafter Charakteristika, die sie in besonderem Maße für die hohen Anforderungen im forensischen Kontext geeignet erscheinen lassen: erstens ermöglicht eine simultane Extraktion von miRNA und DNA [139,140,143], wie bei den bereits routinemäßig eingesetzten mRNA-basierten Methoden, die simultane Untersuchung der Fragestellungen auf Quellen- und Handlungsebene ohne zusätzlichen Verbrauch von Probenmaterial, zweitens sind miRNAs aufgrund ihrer geringen Größe und der engen Assoziation mit Argonaut-Proteinen *in vivo* relativ widerstandsfähig gegenüber Degradation durch physikalischen oder chemischen Stress und drittens werden miRNAs in ihrer reifen, biologisch aktiven Form detektiert und quantifiziert, sodass im Gegensatz zur mRNA-Analyse keine Spleißvarianten berücksichtigt werden müssen.

2 Ziele der Arbeit

An Tatorten, Personen oder Gegenständen sichergestellte biologische Spuren können nicht nur der Individualisierung spurenlegender Personen mittels der Erstellung von DNA-Profilen dienen, sondern zusätzlich durch die Analyse der körperlichen Herkunft potentiell entscheidende Informationen zur Kontextualisierung einer Spur für die Tathergangsrekonstruktion liefern. Da die ‚klassischen' Methoden zur Spurenartidentifikation Probleme hinsichtlich ihrer Sensitivität, Spezifität und insbesondere angesichts ihres großen Probenmaterialverbrauches aufweisen, werden seit einigen Jahren alternative, nukleinsäurebasierte Herangehensweisen erforscht. MiRNAs haben sich dabei aufgrund ihrer zelltypspezifischen Expression, der Möglichkeit, sie parallel zur DNA zu extrahieren und ihrer intrinsisch geringen Anfälligkeit für Degradation als in besonderem Maße geeignet für die Untersuchung von typischerweise nur in geringen Mengen vorhandenem und beeinträchtigtem forensischen Spurenmaterial erwiesen.

Ziel der in dieser Dissertation zusammengefassten Studien war daher, die Identifikation robuster miRNA-Marker für die Identifizierung der forensisch relevanten Körperflüssigkeiten Blut, Speichel, Sperma, Vaginalsekret und Menstruationsblut sowie der Organgewebe Gehirn, Lunge, Leber, Niere, Herzmuskel, Skelettmuskel und Haut. Um biologisch aussagekräftige Ergebnisse zu gewährleisten, sollte zunächst jeweils eine evidenzbasierte, auf die Versuchsbedingungen zugeschnittene Strategie zur Datennormalisierung für die qPCR-basierte Analyse von Körperflüssigkeiten respektive Organgeweben im forensischen Kontext erarbeitet und validiert werden. Anschließend sollte eine unvoreingenommene Auswahl potentiell körperflüssigkeits- oder organgewebespezifischer Markerkandidaten anhand von Microarray-Experimenten getroffen werden. Die Tauglichkeit der selektierten miRNAs sollte in forensisch realistischen Proben unter verschiedenen Bedingungen, darunter gealterte Proben, Gemische mehrerer Spurenarten und simulierte Fallproben, evaluiert werden. Neben standardisierten Arbeitsabläufen und einer stringenten Datennormalisierung sollte dabei besonderes Augenmerk auf eine belastbare, non-arbiträre statistische Auswertung, die eine Einschätzung der Zuverlässigkeit der Methode erlaubt, gelegt werden.

3 Etablierung empirisch begründeter Strategien zur Normalisierung quantitativer miRNA-Expressionsdaten aus forensischem Probenmaterial

3.1 Einleitung

Die RT-qPCR ist das meistverwendete und aufgrund seines großen Dynamikumfangs und seiner hohen Reproduzierbarkeit und Genauigkeit allgemein als ‚Goldstandard' angesehene Verfahren zur Detektion und Quantifizierung von RNA [258,259]. Um mit dieser Methode reliable Aussagen zur Ausgangsmenge einer untersuchten RNA zu erhalten und verlässlich sich nur wenig unterscheidende Expressionsniveaus erfassen zu können, bedarf es jedoch einer an die jeweiligen Versuchsbedingungen angepassten Datennormalisierung zur Eliminierung externer und experimentell bedingter nicht-biologischer Varianzen [260–263]. Diese Varianzen umfassen unter anderem Unterschiede in Menge und Qualität des Ausgangsmaterials sowie der Effizienzen der RT- und qPCR-Reaktionen [264–266]. Bei forensischem Spurenmaterial, das charakteristischerweise von unterschiedlichen Individuen und aus unterschiedlichen Geweben stammt und welches häufig für einen unbekannten Zeitraum verschiedenen und variierenden Umwelteinflüssen ausgesetzt war, ist es dabei von besonderer Bedeutung hochgradig standardisiert zu arbeiten. Zudem ist für eine robuste Normalisierungsstrategie eine empirische Etablierung und Validierung einer Gruppe von Referenzgenen essentiell [267–271]. Der Einsatz arbiträr ausgewählter, nicht validierter Referenzgene kann die Interpretation eines Datensatzes quantitativer Expressionsdaten unter Umständen erheblich verzerren [272,273].

Während für mRNA-basierte Methoden bereits zahlreiche Referenzgene beschrieben sind – auch bezogen auf forensische Fragestellungen [274–278] – war das Ziel der im Folgenden aufgeführten Validierungsstudien, erstmals jeweils eine Gruppe geeigneter Referenzgene

zur Normalisierung von miRNA-Expressionsdaten der forensisch relevanten Körperflüssigkeiten venöses Blut, Speichel, Sperma, Vaginalsekret und Menstruationsblut (Kapitel 3.2) beziehungsweise der Organgewebe Gehirn, Niere, Leber, Lunge, Herzmuskel, Skelettmuskel und Haut (Kapitel 3.3) empirisch zu ermittelten.

Um eine verlässliche, unzweideutige Interpretation der qPCR-Ergebnisse und Reproduzierbarkeit der Experimente zu gewährleisten, wurde neben der Anwendung hochgradig standardisierter Protokolle den in den MIQE-Richtlinien geforderten essentiellen Dokumentationskriterien entsprochen.

Forensic Science International: Genetics 11 (2014) 174–181

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 forensically relevant body fluids

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

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

ARTICLE INFO

Article history:
Received 30 August 2013
Received in revised form 27 January 2014
Accepted 20 March 2014

Keywords:
Micro-RNA
Normalization
RT-qPCR
Body fluid identification
Endogenous references

ABSTRACT

Micro-RNA (miRNA) based analysis of body fluids and composition of complex crime stains has recently been introduced as a potential and powerful tool to forensic genetics. Analysis of miRNA has several advantages over mRNA but reliable miRNA detection and quantification using quantitative PCR requires a solid and forensically relevant normalization strategy.

In our study we evaluated a panel of 13 carefully selected reference genes for their suitability as endogenous controls in miRNA qPCR normalization in forensically relevant settings. We analyzed assay performances and variances in venous blood, saliva, semen, menstrual blood, and vaginal secretion and mixtures thereof integrating highly standardized protocols with contemporary methodologies and included several well established computational algorithms.

Based on these empirical results, we recommend normalization to the group of *SNORD24*, *SNORD38B*, and *SNORD43* as this signature exhibits the most stable expression levels and the least expected variation among the evaluated candidate reference genes in the given set of forensically relevant body fluids.

To account for the lack of consensus on how best to perform and interpret quantitative PCR experiments, our study's documentation is compliant to MIQE guidelines, defining the "minimum information for publication of quantitative real-time PCR experiments".

1. Introduction

RNA based analytical methods are on the rise in forensic molecular biology [1] and early international trial exercises for forensic RNA analysis have already been conducted [2–4]. The analysis of differential expression of mRNA may be used in forensic settings to identify body fluid components of mixed stains [5], to estimate wound or stain age [6,7], detect pregnancy [8], and to help discern the cause of death [9].

There are, however, drawbacks associated with the analysis of mRNA, e.g. its susceptibility to degradation and lack of specificity in the identification or discrimination of particular body fluids especially vaginal secretions [10,11]. Therefore, in addition, feasibility and practicability of forensic miRNA analysis [12] based on quantitative PCR (qPCR) is being assayed since recently by several groups [13–15].

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 variances is essential [16–19]. 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 [20–22]. Therefore, in the present study, we present a group of endogenous reference genes selected on the base of empirical evidence for the normalization of qPCR data from expression analysis of 13 preselected miRNAs in forensically relevant body fluids.

2. Material and methods

2.1. Adherence to the MIQE guidelines

To facilitate reliable and unequivocal interpretation of the qPCR results reported herein, all information that is rated 'essential' according to the MIQE guidelines [23] is reported, where applicable.

* 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.2014.03.011
1872-4973/

2.2. Samples

Samples for each tested body fluid, i.e. venous blood, saliva, vaginal fluid, menstrual blood, and semen, were collected from healthy volunteers, after obtaining informed consent.

Venous blood was collected by venipuncture using dry vacutainer tubes and spotted onto sterile cotton swabs. For collection of saliva via buccal swab, donors were asked to abstain from eating, smoking, drinking and oral hygiene at least 30 min prior to sampling. Samples of semen-free vaginal secretion were collected by the female donors themselves using sterile stemmed cotton swabs. Menstrual blood samples were obtained by the female donors using tampons. Freshly ejaculated semen was provided in sealed Falcon tubes by male donors and dried onto sterile stemmed cotton swabs by the researcher immediately after receipt. All samples were dried at room temperature and processed for RNA extraction after 24 h.

2.3. RNA extraction and quantification

All surfaces, devices, and machines utilized in the extraction procedure were thoroughly cleaned using RNase-Zap® (Ambion, Austin, TX, USA) to remove ambient RNases and only RNase-free reagents and plastic consumables were used.

Total RNA was extracted using the mirVana™ miRNA Isolation Kit (Ambion) according to the manufacturer's protocol. Prior to extraction, samples (whole cotton tip or approximately 2 cm^2 of the tampon or blood stain) were cut into pieces and incubated with 350 μl Lysis/Binding Buffer at 56 °C for 1 h. Venous blood samples were additionally treated with RBC Lysis Solution (Qiagen, Hilden, Germany) to selectively lyse red blood cells prior to incubation. Spin baskets (Promega, Mannheim, Germany) were used to separate lysate and substrate by centrifugation at 13.000 × *g* for 1 min. Total RNA eluates were stored at −80 °C until further processing.

For removal of potential traces of genomic DNA, subsequent DNase I digestion was performed with the Turbo DNA-free™ Kit (Ambion), following the manufacturer's protocol. Total RNA concentration and quality, represented by the RNA integrity number (RIN) [24], were determined using the Quant-iT™ RNA Assay Kit on a Qubit fluorometer (both Invitrogen, Darmstadt, Germany) and the RNA 6000 Nano Kit on a Agilent 2100 Bioanalyzer (both Agilent, Böblingen, Germany), respectively.

2.4. Preparation of samples

Individual samples were diluted to 2 ng/μl based on quantification results and were used as single samples only or additionally for the preparation of a pooled sample per body fluid by combining identical volumes of diluted sample. Five individual samples per body fluid were examined. Pooled samples for blood and saliva consisted of 10 donor samples, while those for vaginal secretion, menstrual blood and semen consisted of five donor samples. Further, a mixture of all five body fluids was prepared for efficiency determination experiments containing identical volumes of the above mentioned pooled samples.

2.5. Selection of candidate reference genes

A panel of 13 potential reference genes was selected based on a literature survey, mainly focusing on reference genes previously used in forensic miRNA analyses and the manufacturer's recommended control panel [13–15,25–29]. The selected panel encompassed *hsa-miR-93-5p*, *hsa-miR-191-5p*, *RNU6-1*, *RNU6-2*, *SNORA66*, *SNORA74A*, *SNORD7*, *SNORD24*, *SNORD38B*, *SNORD43*, *SNORD44*, *SNORD48* and *SNORD49A* (Supplementary Table 1).

2.6. Reverse transcription qPCR (RT-qPCR)

Complementary DNA (cDNA) was synthesized using target-specific stem-loop primers (Supplementary Table 1) and the TaqMan® MicroRNA Reverse Transcription Kit (Life Technologies, Weiterstadt, Germany), as per manufacturer's protocol. Each 15 μl reaction volume contained 10 ng total RNA, 1X RT primers, 50 U MultiScribe™ reverse transcriptase, 1 mM dNTPs, 3.8 U RNase inhibitor, and 1X reverse transcription buffer. Reactions were performed on a T3 Thermocycler (Biometra, Göttingen, Germany) with the following cycling conditions: 16 °C for 30 min, 42 °C for 30 min, and 85 °C for 5 min. Besides extraction negative and H_2O controls, we employed RT(−)-controls to control for potential contamination with genomic DNA. For efficiency determination experiments, reverse transcriptions of the mixture containing all body fluids were conducted twice. RT reaction products were stored at −20 °C.

QPCR reactions were performed using target-specific TaqMan® Assays (Supplementary Table 1) and the TaqMan® Universal PCR Master Mix, No AmpErase® UNG (Life Technologies) as per manufacturer's protocol: 1.3 μl of the appropriate RT reaction product were added into a 20 μl reaction volume, containing 1X TaqMan Universal PCR Master Mix and 1X specific TaqMan® Assay. All sample-assay combinations were run in triplicates for the pooled samples and in duplicates for the individual samples, respectively. The internal PCR control from the Quantifiler® Human DNA Quantification Kit (Life Technologies) was used as an inter plate calibrator. PCR cycling conditions consisted of 95 °C for 10 min and 40 cycles of 95 °C for 15 s and 60 °C for 1 min, and were performed on an ABI Prism 7500 (Life Technologies). Data collection was performed during the 60 °C step by the SDS software version 1.2.3 (Life Technologies). 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.

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

The LinRegPCR program version 2012.3 [30] 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 $-0.7 \log_{10}$(fluorescence). C_q-values deviating more than one cycle from the triplicate median were excluded from subsequent pre-processing. For comparison, amplification efficiencies were computed analogously using the Real-time PCR Miner algorithm [31].

Analysis of qPCR data including pre-processing was then performed using the GenEx software version 5.3 (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 [32], geNorm [22], both implemented in the GenEx software, and the Excel-based BestKeeper [33]. 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 body fluid and for individual body fluid samples.

3. Results

Quantity and integrity of total RNA varied notably among samples of the same body fluid as well as between groups with RIN values generally ≤4 (Table 1). Overall, saliva samples exhibited the lowest (total RNA concentration: 3.6–25.3 ng/μl; RIN: n.d.–1.8) and vaginal secretion samples the highest overall values (total RNA concentration: 41.9–257 ng/μl; RIN: 2.9–4).

RT(−)-controls for *RNU6-1* and *SNORD48* showed scarce unspecific amplification and these markers were excluded from further analyses (data not shown). The negative controls were negative for all candidate genes when extracted from stemmed cotton swabs, while extracts from tampons and sterile cotton swabs produced a weak unspecific signal for *miR-93* and *miR-191* (difference to C_q-values of mixture >15 C_q, data not shown).

Table 1
Total RNA quantity and integrity per sample.

Body fluid	Sample number	Gender	Total RNA concentration (ng/μl)	RIN	Additional information
Blood	1	F	8.7	2	
	2	F	12.6	2.4	
	3	M	14.3	1.1	
	4	M	15.2	1.6	
	5	M	15.6	1.2	
	6*	F	18.1	1.8	
	7*	M	19.8	1.3	
	8*	M	19.9	1	
	9*	F	20.8	1.3	
	10*	F	23.8	2.4	
Menstrual blood	1*	F	24.8	2.6	Day 3 of menstruation
	2*	F	27.5	1.8	Not specified
	3*	F	37.2	2.4	Day 4 of menstruation
	4*	F	52.0	2.7	Not specified
	5*	F	150.0	2.6	Day 3 of menstruation
Saliva	1	F	3.6	1	
	2*	F	6.5	1.1	
	3	M	8.2	1	
	4*	F	8.4	1	
	5	M	9.0	1	
	6	F	11.2	1	
	7*	M	12.5	1.2	
	8*	F	16.1	n.d.	
	9*	M	17.8	1.1	
	10	M	25.3	1.8	
Semen	1*	M	11.4	1	
	2*	M	16.0	1	
	3*	M	16.4	2.1	
	4*	M	18.3	1.2	
	5*	M	38.4	2.2	
Vaginal secretion	1*	F	41.9	2.9	Day 10 after menstruation
	2*	F	51.5	3.4	Day 20 after menstruation
	3*	F	85.5	2.9	Not specified
	4*	F	110.0	3.3	Not specified
	5*	F	257.0	4	Day 17 after menstruation

RIN RNA integrity number; *F* female, *M* male; *n.d.* not detectable; *sample used for individual analyses.

Table 2
Amplification efficiencies of candidate reference genes calculated by LinRegPCR software and Real-time PCR Miner algorithm, respectively.

Gene symbol	Amplification efficiency of mixture			
	LinRegPCR		Real-time PCR Miner	
	Mean[a]	SD	Mean[b]	SD
miR-191	1.78	0.016	0.82	0.005
miR-93	1.82	0.017	0.87	0.018
RNU6-2	1.84	0.005	0.87	0.004
SNORA66	1.82	0.010	0.93	0.011
SNORA74A	1.88	0.015	0.96	0.010
SNORD24	1.84	0.028	0.92	0.013
SNORD38B	1.89	0.030	0.98	0.015
SNORD43	1.84	0.035	0.92	0.031
SNORD44	1.61[c]	0.006	–	–
SNORD49A	1.83	0.024	0.92	0.026
SNORD7	1.78[c]	0.028	–	–

SD, standard deviation; –, not computed.
[a] Efficiencies are given as values between 1 and 2, with 2 representing an amplification efficiency of 100%.
[b] Efficiencies are given as values between 0 and 1, with 1 representing an amplification efficiency of 100%.
[c] Excluded from study.

3.1. Amplification efficiency

Amplification efficiency per amplicon was derived from the mixture containing all five body fluids, including two distinct RT-reactions and qPCR triplicates into the computation. Mean efficiencies per amplicon computed with LinRegPCR ranged from 89% (*SNORD38B*) to 61% (*SNORD44*) (Table 2). Due to its grossly outlying amplification efficiency, *SNORD44* was excluded from further analyses. *SNORD7* was excluded since an additional examination of the amplification efficiencies in the pooled samples revealed considerable variation between body fluids ranging from 88% in menstrual blood to 59% in semen (data not shown).

Mean efficiencies per amplicon computed with Real-time PCR Miner ranged from 98% (*SNORD38B*) to 82% (*miR-191*) (Table 2).

3.2. Determination of most suitable reference genes

3.2.1. Pooled samples

A first examination of the most suitable reference genes using LinRegPCR spread sheet exported data for the pooled samples of each body fluid type was performed including the remaining nine candidate genes (Supplementary Table 2).

According to NormFinder *SNORD38B* was the most stable gene with a stability value of 0.3049 standard deviations, followed by *SNORA66, SNORD24* and *SNORD43* (Supplementary Fig. 1, upper panel). The least stable gene was *miR-93* with a stability value of 2.4782. The simultaneously calculated accumulated standard deviation was lowest (0.2731) when the use of two reference genes was assumed (Supplementary Fig. 1, lower panel).

Analyzed with geNorm, the most stable pair of genes was *SNORD49A & SNORD24*, with an M-value of 0.3066, followed by *SNORD43* and *SNORA66*, while *miR-93* was the least stable gene with an M-value of 1.6026 (Supplementary Fig. 2).

The only candidate gene considered as stable (standard deviation of C_q-values <1.0) by BestKeeper was *RNU6-2* with a standard deviation of 0.83. The remaining values ranged between 1.17 (*SNORD43*) and 3.25 (*miR-93*) (Supplementary Fig. 3).

Subsequently, candidate reference gene data put out by the three algorithms, respectively, were transformed into consecutively numbered ranks with 1 representing the most and 9 the least stable gene (Supplementary Table 3A) and a comprehensive gene stability ranking was attained by calculation of the arithmetic mean ranking value per gene. In this comprehensive ranking *SNORA66* and

Table 3
Mean C_q-values and standard deviation per candidate reference gene of five individual body fluid samples after pre-processing (amplification efficiency and C_q-values as per LinRegPCR software).

Body fluid	*miR-191*	*miR-93*	*RNU6-2*	*SNORA66*	*SNORA74A*	*SNORD24*	*SNORD38B*	*SNORD43*	*SNORD49A*
Blood	15.92 ± 0.22	16.39 ± 0.49	24.90 ± 0.73	24.76 ± 0.32	26.21 ± 0.44	22.53 ± 0.57	25.17 ± 0.53	23.43 ± 0.56	21.97 ± 0.76
Menstrual blood	18.28 ± 2.97	19.08 ± 3.94	26.35 ± 0.87	26.91 ± 3.09	27.34 ± 1.94	23.61 ± 1.50	27.24 ± 1.36	23.75 ± 0.84	23.18 ± 1.07
Saliva	21.38 ± 0.38	22.31 ± 0.51	27.47 ± 0.89	26.33 ± 1.48	30.31 ± 1.61	24.79 ± 0.85	28.55 ± 0.76	24.23 ± 0.29	24.06 ± 1.03
Semen	20.57 ± 0.28	22.69 ± 0.44	28.00 ± 1.32	29.82 ± 1.83	31.18 ± 1.12	28.76 ± 1.59	31.34 ± 1.70	28.10 ± 1.30	28.53 ± 1.70
Vaginal secretion	21.95 ± 0.65	24.23 ± 1.55	26.16 ± 0.62	29.37 ± 2.75	28.46 ± 1.67	24.90 ± 0.96	29.14 ± 1.02	24.89 ± 0.40	23.93 ± 0.38

SNORD43 were top ranked, closely followed by *SNORD24* and *SNORD38B*. The least stable genes were *miR-93* and *miR-191*.

Analyzing the SDS software spread sheet exported data corrected with the amplification efficiencies as per Real-time PCR Miner (Supplementary Table 4) resulted in comparable ranking orders (Supplementary Table 3B and Supplementary Figs. 4–6).

3.2.2. Individual samples

Analyses for determination of the most suitable reference genes were simultaneously computed for a set of five individual samples per body fluid (Table 3 and Supplementary Tables 5 and 6).

LinRegPCR spread sheet exported data analyzed with NormFinder designated *SNORD38B* as the most stable gene with a stability value of 0.5730 standard deviations, followed by *SNORD24* and *SNORA74A* (Fig. 1, upper panel). The least stable gene was *miR-93* with a stability value of 2.0461. The simultaneously calculated accumulated standard deviation was lowest (0.4117) when the use of eight reference gene was assumed (Fig. 1, lower panel).

According to geNorm *SNORD24* & *SNORD49A* was the most stable pair with an M-value of 0.7183, followed by *SNORD43* and *SNORD38B*, while *miR-93* was the least stable gene with an M-value of 1.7195 (Fig. 2).

None of the candidate genes was considered stable by the BestKeeper algorithm, with the value of *RNU6-2* being close to the cut-off 1.08, however. The remaining values ranged between 1.37 for *SNORD43* and 2.93 for *miR-93* (Fig. 3).

Analyzing the SDS software spread sheet exported data corrected with the amplification efficiencies as per Real-time PCR Miner resulted in comparable results and the same ranking orders (Supplementary Figs. 7–9).

The comprehensive ranking following both computations designated *SNORD24* followed by *SNORD38B* and *SNORD43* as the best suitable reference genes in the given single sample set (Table 4).

4. Discussion

For qPCR to deliver reliable and biologically meaningful results an accurate and relevant normalization of non biological variances is essential [16–19]. 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 [34–36]. This has to be accounted for especially when, as in most forensic settings, samples have been obtained from different individuals, different body fluids and different time courses. In this study we applied highly standardized protocols starting with the handling, storage and extraction of samples to minimize the external variances. To compensate for internal non-biological variances, the use of endogenous reference genes is essential.

As the so called "housekeeping genes", like *ACTB* and *GAPDH*, that were commonly used as reference genes for normalization in numerous studies have long been shown to be differentially expressed under many different experimental conditions [37–43], which has only recently been confirmed in a forensic setting [44], they should not any longer be used uncritically for normalization purposes. Consequently, reference genes intended for qPCR normalization in a given experimental setting have to be selected beforehand and then based on their empirically proven suitability. This is particularly important for miRNA analysis as a general agreement on methodological standardization of qPCR in miRNA quantification has not been achieved yet. The aim of our study was therefore to present a reliable and empirically derived reference framework for normalization of qPCR data in the analysis of miRNA expression in realistic representations of five forensically relevant body fluids. The selection of 13 candidates for the starting panel of reference genes was based upon a literature survey [13–15,25–27] and several criteria such as a relatively constant and highly abundant expression across a large number of tissues and cell lines [28].

Another important aspect that has to be accounted for in qPCR data analysis is PCR efficiency. Samples from different tissues are known to exhibit different PCR efficiencies caused by variations in RT and PCR due to inhibitors and by variations on the total RNA fraction pattern extracted. It has been shown that omission of correction for differential PCR efficiencies [45] introduces bias in the expression results [21,46–48] and thus, a separate determination of qPCR efficiency for each performed transcript is necessary [16,18,49]. There is as yet no consensus as to which of several algorithms presented so far for determining C_q-values and PCR efficiencies from raw fluorescent data is best suited. However, Ruijter et al. recently published a first comprehensive benchmark study of the evaluation of nine qPCR analysis methods [50] and the LinRegPCR method [30] was top ranked for precision and resolution and also showed high linearity without introducing excessive bias [50]. This software was shown to underestimate efficiencies compared to those determined by standard curve analysis, though. To account for this we did not apply the commonly used standard 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 [31], which was highly ranked in the benchmark study by Ruijter et al. as well but was shown to rather overestimate efficiencies compared to those determined by standard curve analysis [30]. This algorithm determines the cycle threshold for each sample dynamically based on its reaction kinetics. As the composition of analyzed stains is usually unknown in a forensic setting, we decided to take a more robust approach by employing the C_q-values determined with the SDS software.

The results of our efficiency calculations employing these two distinct algorithms indeed reflect their differences in efficiency estimation as demonstrated by Ruijter et al. [30] with Real-time PCR Miner reporting higher and LinRegPCR reporting lower efficiency values, respectively. To investigate whether these divergent estimates of efficiency influence the algorithmic identification of most stable reference genes we performed two separate computations, based on values given by LinRegPCR and Real-time PCR Miner. The reference gene rankings extracted from LinRegPCR data or Real-time PCR Miner data are very similar thus

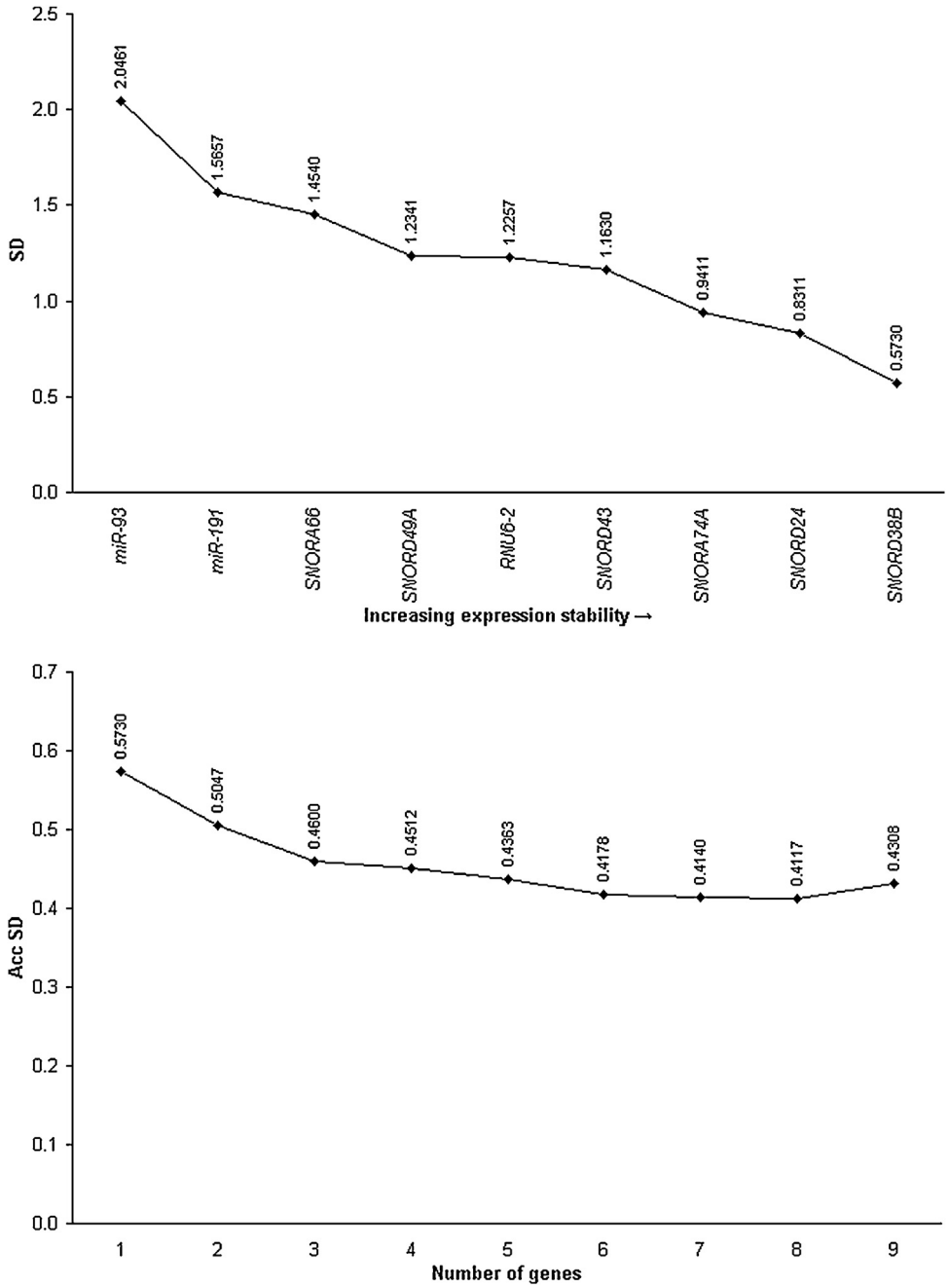


Fig. 1. NormFinder data analysis of the nine candidate genes in individual body fluid samples.
(upper panel) Gene expression stability values of genes – from least (*left*) to most stable (*right*). (lower panel) Determination of the optimal number of reference genes by computation of accumulated standard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

suggesting that the influence of the difference between these algorithms on reference gene selection is negligible.

We used the amplification efficiencies as computed from a mixture of all five body fluids for all subsequent calculations. This is probably the most conservative approach since it is usually unknown which and how many different types of body fluids and of how many individuals are present in a given casework stain (e.g. sexual assault crimes). In addition, a computation of amplification efficiencies in pooled samples per body fluid was performed, however, to assess the variation between the distinct body fluids and resulted in the exclusion of *SNORD7* from further analyses.

Analogous to algorithms for the calculation of PCR efficiency, there is as yet no consensus as to which of several present algorithms performs best in identifying the most suitable endogenous reference out of a set of candidate genes. We therefore employed three well established and commonly used methods – NormFinder, geNorm, and BestKeeper – and reported both the results for each algorithm and in combination as described by Wang et al. [51]. The notably different ranking resulting from BestKeeper might be due to the fact, that this algorithm uses C_q-values directly, while geNorm and NormFinder transform imported C_q-values to relative quantities for stability calculations.

A major challenge for the establishment of a solid normalization strategy in forensic settings is posed by the nature of forensically relevant samples e.g. body fluids containing multiple different types of cells, as well as by the potentially complex composition of typical forensic stain evidence that may include several body fluids in vastly different proportions. Also, sampling and storage

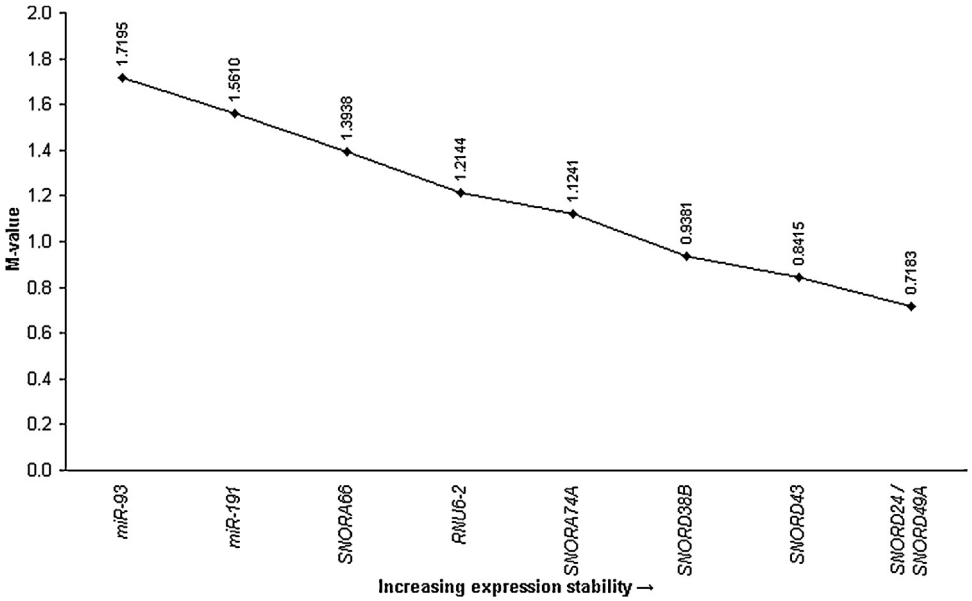


Fig. 2. GeNorm data analysis of the nine candidate genes in individual body fluid samples.
GeNorm proceeds by calculation of the gene stability measure (M-value) per gene – from least (*left*) to most stable (*right*); determination of the optimal pair of reference genes by stepwise exclusion of the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

conditions of evidential material will often not be optimized for potential RNA analysis and thus result in low total RNA concentration and integrity. Therefore, to mimic realistic forensic casework we used dried body fluid stains in our study instead of fresh, less compromised samples. The weak signal detected in negative controls of *miR-93* and *miR-191* can most probably be explained by the type of sample, too, since tampons and cotton swabs were not declared DNA-/RNA-free as was the case for stemmed cotton swabs. However, with C_q-values for these extraction controls being so high and far off those from the actual samples (>15 C_q), the respective candidates had not to be excluded from analyses.

Another important difference to previously published studies on miRNA normalization [26] is, again owing to the forensic scope of this study, that the evaluated samples do not represent two conditions of the same tissue or cell type (e.g. healthy/cancerous or treated/untreated) but up to five distinct body fluids. The more types of body fluids are included in a mixture, the higher the variances of expression values for any one reference gene are expected to be. It was therefore 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 further took into account that the composition of a stain encountered in forensic casework is usually unknown in terms of

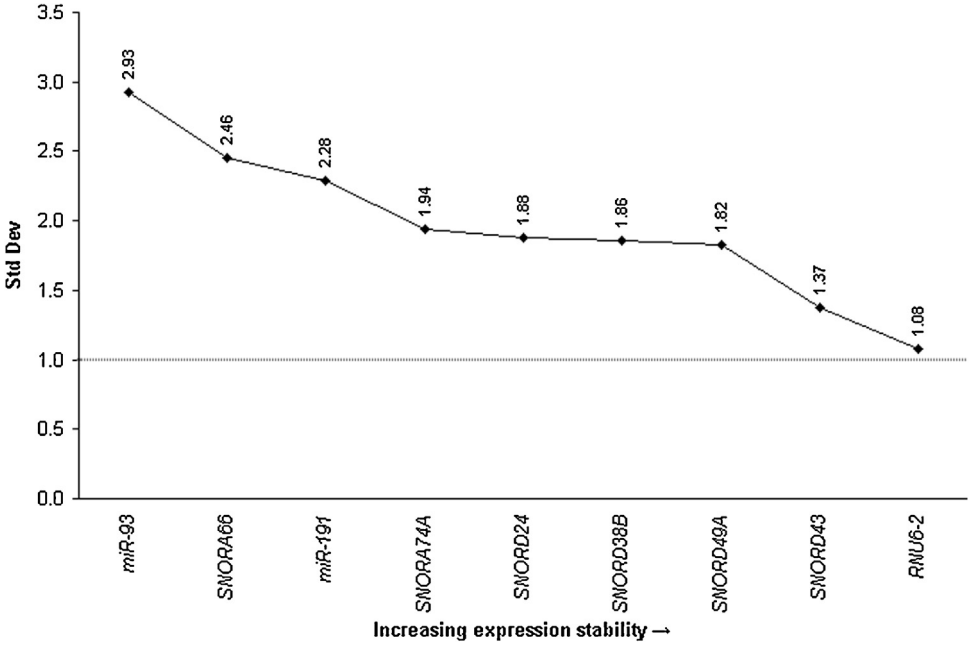


Fig. 3. BestKeeper data analysis of the nine candidate genes in individual body fluid samples.
BestKeeper proceeds by pair wise correlation analysis of the C_q-values of all pairs of candidate reference genes – from least (*left*) to most stable (*right*). Amplification efficiency and C_q-values as per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

Table 4
Comprehensive ranking order of the candidate reference genes for individual body fluid samples, derived by integrating rankings of NormFinder, geNorm, and BestKeeper. Computations with the LinRegPCR software derived values and computations with amplification efficiency as per Real-time PCR Miner algorithm and C_q-values as per SDS software resulted in identical ranking orders.

Ranking order	NormFinder	geNorm	BestKeeper	Comprehensive ranking (mean rank value)
1	*SNORD38B*	*SNORD24 & SNORD49A*	*RNU6-2*	*SNORD24* (2.67)
2	*SNORD24*		*SNORD43*	*SNORD38B/SNORD43* (3.00)
3	*SNORA74A*	*SNORD43*	*SNORD49A*	
4	*SNORD43*	*SNORD38B*	*SNORD38B*	*SNORD49A* (3.33)
5	*RNU6-2*	*SNORA74A*	*SNORD24*	*RNU6-2* (4.00)
6	*SNORD49A*	*RNU6-2*	*SNORA74A*	*SNORA74A* (4.67)
7	*SNORA66*	*SNORA66*	*miR-191*	*SNORA66* (7.33)
8	*miR-191*	*miR-191*	*SNORA66*	*miR-191* (7.67)
9	*miR-93*	*miR-93*	*miR-93*	*miR-93* (9.00)

the types of body fluids present and the number of individual contributors, respectively.

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 body fluid may be present in a given sample. It is possible though, to apply a more specific normalization strategy if, for whatever reason, the composition of a stain is less unknown, if e.g. the presence in the stain of one body fluid can reliably be excluded. This would have to be validated in the given setting.

For an initial screening of suitable reference genes, we used pooled samples in which the expected inter-individual differences are counterbalanced. This analysis indicated *SNORA66* and *SNORD43* to be stable reference genes followed by *SNORD24* and *SNORD38B*.

Subsequent comprehensive computations were performed with five individual samples per body fluid, hence taking into account biological variation. As expected, blood samples showed low inter-individual differences in all markers. Higher standard deviations between individual samples were present in semen, vaginal secretion and menstrual blood, whereas the high standard deviation values in the latter can be attributed to outlying values in a single sample.

As in the initial screening, *SNORD24*, *SNORD38B* and *SNORD43* were again the top ranked markers. *SNORA66*, however, appears to show large inter-individual differences resulting in a high mean rank value, and was therefore excluded from the recommended set of reference genes.

We are aware of the relatively small sample size and aim to further assess the presented set of reference genes in terms of its value for data normalization in future studies in which miRNA candidates for the identification of body fluids will have to be validated with the presented normalization strategy.

5. Conclusion

Herein we analyzed 13 potential reference genes and empirically determined *SNORD24*, *SNORD38B* and *SNORD43* to be the most stable endogenous reference genes for a reliable normalization of qPCR data from forensic miRNA expression analysis of body fluids in a set of body fluid samples.

Acknowledgements

We would like to thank all volunteers for their kind donation of sample material. We also thank the Deutsche Forschungsgemeinschaft (DFG) CO 992/3-1 for funding this project.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2014.03.011.

References

[1] M. Bauer, RNA in forensic science, Forensic Sci. Int. Genet. 1 (2007) 69–74.
[2] 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. (2010).
[3] 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.
[4] 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.
[5] C. Haas, B. Klesser, C. Maake, W. Bar, A. Kratzer, mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR, Forensic Sci. Int. 3 (2009) 80–88.
[6] S. Anderson, B. Howard, G.R. Hobbs, C.P. Bishop, A method for determining the age of a bloodstain, Forensic Sci. Int. 148 (2005) 37–45.
[7] M. Takamiya, K. Saigusa, N. Nakayashiki, Y. Aoki, Studies on mRNA expression of basic fibroblast growth factor in wound healing for wound age determination, Int. J. Legal Med. 117 (2003) 46–50.
[8] J. Gauvin, D. Zubakov, J. van Rhee-Binkhorst, A. Kloosterman, E. Steegers, M. Kayser, Forensic pregnancy diagnostics with placental mRNA markers, Int. J. Legal Med. 124 (2010) 13–17.
[9] D. Zhao, B.L. Zhu, T. Ishikawa, D.R. Li, T. Michiue, H. Maeda, Quantitative RT-PCR assays of hypoxia-inducible factor-1alpha, erythropoietin and vascular endothelial growth factor mRNA transcripts in the kidneys with regard to the cause of death in medicolegal autopsy, Leg. Med. (Tokyo) 8 (2006) 258–263.
[10] M.L. Richard, K.A. Harper, R.L. Craig, A.J. Onorato, J.M. Robertson, J. Donfack, Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis, Forensic Sci. Int. Genet. 6 (2012) 452–460.
[11] K. Sakurada, T. Akutsu, K. Watanabe, Y. Fujinami, M. Yoshino, Expression of statherin mRNA and protein in nasal and vaginal secretions, Leg. Med. (Tokyo) 13 (2011) 309–313.
[12] C. Courts, B. Madea, Micro-RNA: a potential for forensic science? Forensic Sci. Int. 203 (2010) 106–111.
[13] 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.
[14] 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.
[15] 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.
[16] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Res. 29 (2001) e45.
[17] S.A. Bustin, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems, J. Mol. Endocrinol. 29 (2002) 23–39.
[18] 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.
[19] S. Bustin, T. Nolan, Data analysis and interpretation, in: S. Bustin (Ed.), A-Z of Quantitative PCR, 1st ed., International University Line, La Jolla, CA, 2004, pp. 439–492.
[20] 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.
[21] 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.
[22] 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).
[23] 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.

[24] 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.
[25] 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.
[26] 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.
[27] A. Masotti, V. Caputo, L. Da Sacco, A. Pizzuti, B. Dallapiccola, G.F. Bottazzo, Quantification of small non-coding RNAs allows an accurate comparison of miRNA expression profiles, J. Biomed. Biotechnol. 2009 (2009) 659028.
[28] L. Wong, K. Lee, I. Russell, C. Chen, Endogenous controls for real-time quantitation of miRNA using TaqMan microRNA assays, RNA (2010) 1–8.
[29] S. Griffiths-Jones, H.K. Saini, D.S. van, A.J. Enright, miRBase: tools for microRNA genomics, Nucleic Acids Res. 36 (2008) D154–D158.
[30] 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.
[31] S. Zhao, R.D. Fernald, Comprehensive algorithm for quantitative real-time polymerase chain reaction, J. Comput. Biol. 12 (2005) 1047–1064.
[32] 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.
[33] 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.
[34] 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.
[35] 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.
[36] 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.
[37] D. Goidin, A. Mamessier, M.J. Staquet, D. Schmitt, O. Berthier-Vergnes, Ribosomal 18S RNA prevails over glyceraldehyde-3-phosphate dehydrogenase and beta-actin genes as internal standard for quantitative comparison of mRNA levels in invasive and noninvasive human melanoma cell subpopulations, Anal. Biochem. 295 (2001) 17–21.
[38] T.D. Schmittgen, B.A. Zakrajsek, Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR, J. Biochem. Biophys. Methods 46 (2000) 69–81.
[39] P. Bhatia, W.R. Taylor, A.H. Greenberg, J.A. Wright, Comparison of glyceraldehyde-3-phosphate dehydrogenase and 28S-ribosomal RNA gene expression as RNA loading controls for northern blot analysis of cell lines of varying malignant potential, Anal. Biochem. 216 (1994) 223–226.
[40] J. Bereta, M. Bereta, Stimulation of glyceraldehyde-3-phosphate dehydrogenase mRNA levels by endogenous nitric oxide in cytokine-activated endothelium, Biochem. Biophys. Res. Commun. 217 (1995) 363–369.
[41] T.J. Chang, C.C. Juan, P.H. Yin, C.W. Chi, H.J. Tsay, Up-regulation of beta-actin, cyclophilin and GAPDH in N1S1 rat hepatoma, Oncol. Rep. 5 (1998) 469–471.
[42] G. Zhu, Y. Chang, J. Zuo, X. Dong, M. Zhang, G. Hu, et al., Fudenine, a C-terminal truncated rat homologue of mouse prominin, is blood glucose-regulated and can up-regulate the expression of GAPDH, Biochem, Biophys. Res. Commun. 281 (2001) 951–956.
[43] J.B. de Kok, R.W. Roelofs, B.A. Giesendorf, J.L. Pennings, E.T. Waas, T. Feuth, et al., Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes, Lab. Invest. 85 (2005) 154–159.
[44] A. Huth, B. Vennemann, T. Fracasso, S. Lutz-Bonengel, M. Vennemann, Apparent versus true gene expression changes of three hypoxia-related genes in autopsy derived tissue and the importance of normalisation, Int. J. Legal Med. 127 (2013) 335–344.
[45] 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.
[46] 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.
[47] Y. Karlen, A. McNair, S. Perseguers, C. Mazza, N. Mermod, Statistical significance of quantitative PCR, BMC Bioinformatics 8 (2007) 131.
[48] 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.
[49] 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.
[50] 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.
[51] 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.

SUPPLEMENTARY TABLE 1: Specifications of the 13 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-93-5p	-	MIMAT0000093	001090	CAAAGUGCUGUUCGUGCAGGUAG (23)	[26]
hsa-miR-191-5p	-	MIMAT0000440	002299	CAACGGAAUCCCAAAAGCAGCUG (23)	[26]
RNU6-1	*U6, RNU6, U6-1, RNU6A*	NR_004394	001973	GUGCUCGCUUCGGCAGCACAUAUACUAAAAUUGGAACGAUACAG AGAAGAUUAGCAUGGCCCCUGCGCAAGGAUGACACGCAAAUUCG UGAAGCGUUCCAUAUUUU (106)	[25]
RNU6-2	*U6, RNU6B*	NR_002752	001093	CGCAAGGAUGACACGCAAAUUCGUGAAGCGUUCCAUAUUUUU (42)	[13,15,25,28]
SNORA66	*U66, RNU66*	NR_002444	001002	GUAACUGUGGUGAUGGAAAUGUGUUAGCCUCAGACACUACUGAG GUGGUUCUUUCUAUCCUAGUACAGUC (70)	[28]
SNORA74A	*U19, RNU19*	X94290	001003	UUGCACCUCUGAGAGUGGAAUGACUCCUGUGGAGUUGAUCCUAG UCUGGGUGCAAACAAUU (61)	[28]
SNORD7	*Z30, mgU6-47*	AJ007733	001092	UGGUAUUGCCAUUGCUUCACUGUUGGCUUUGACCAGGGUAUGAU CUCUUAAUCUUCUCUCUGAGCUG (67)	[28]
SNORD24	*U24, RNU24*	NR_002447	001001	AUUUGCUAUCUGAGAGAUGGUGAUGACAUUUUAAACCACCAAGA UCGCUGAUGCA (55)	[14,28]
SNORD38B	*U38B, RNU38B*	NR_001457	001004	CCAGUUCUGCUACUGACAGUAAGUGAAGAUAAAGUGUGUCUGAG GAGA (48)	[28]
SNORD43	*U43, RNU43*	NR_002439	001095	GAACUUAUUGACGGGCGGACAGAAACUGUGUGCUGAUUGUCACG UUCUGAUU (52)	[28]
SNORD44	*U44; RNU44*	NR_002750	001094	CCUGGAUGAUGAUAGCAAAUGCUGACUGAACAUGAAGGUCUUAA UUAGCUCUAACUGACU (60)	[14,25,28]
SNORD48	*U48, RNU48*	NR_002745	001006	GAUGACCCCAGGUAACUCUGAGUGUGUCGCUGAUGCCAUCACCG CAGCGCUCUGACC (57)	[14,25,28]
SNORD49A	*U49, U49A, RNU49*	NR_002744	001005	CACUAAUAGGAAGUGCCGUCAGAAGCGAUAACUGACGAAGACUA CUCCUGUCUGAUU (57)	[28]

NCBI National Center for Biotechnology Information, *miRBase* microRNA database [29]

SUPPLEMENTARY TABLE 2: Mean C_q-values of technical replicates per candidate reference gene and pooled body fluid sample after pre-processing (amplification efficiency and C_q-values as per LinRegPCR software).

Body fluid	*miR-191*	*miR-93*	*RNU6-2*	*SNORA66*	*SNORA74A*	*SNORD24*	*SNORD38B*	*SNORD43*	*SNORD49A*
Blood	16.06	15.39	24.81	24.23	27.14	22.72	25.16	23.29	22.23
Menstrual blood	16.42	16.98	26.26	26.14	27.44	23.55	27.55	23.81	23.63
Saliva	22.00	22.27	27.44	26.76	31.08	25.05	28.56	24.68	24.43
Semen	21.04	22.65	27.45	29.07	32.14	28.43	30.63	27.61	28.09
Vaginal secretion	22.32	23.96	26.08	27.37	27.14	24.86	28.44	25.11	24.18

SUPPLEMENTARY TABLE 3: Comprehensive ranking order of the candidate reference genes for pooled body fluid samples, derived from integrating rankings by NormFinder, geNorm, and BestKeeper: (A) computation with amplification efficiency and C_q-values as per LinRegPCR software, (B) computation with amplification efficiency as per Real-time PCR Miner algorithm and C_q-values as per SDS software.

(A)

Ranking Order	NormFinder	geNorm	BestKeeper	Comprehensive ranking (mean rank value)
1	*SNORD38B*	*SNORD24 & SNORD49A*	*RNU6-2*	*SNORA66 / SNORD43* (3.00)
2	*SNORA66*		*SNORD43*	
3	*SNORD24*	*SNORD43*	*SNORA66*	*SNORD24 / SNORD38B* (3.33)
4	*SNORD43*	*SNORA66*	*SNORD38B*	
5	*SNORD49A*	*SNORD38B*	*SNORD49A*	*SNORD49A* (3.67)
6	*RNU6-2*	*RNU6-2*	*SNORD24*	*RNU6-2* (4.33)
7	*SNORA74A*	*SNORA74A*	*SNORA74A*	*SNORA74A* (7.00)
8	*miR-191*	*miR-191*	*miR-191*	*miR-191* (8.00)
9	*miR-93*	*miR-93*	*miR-93*	*miR-93* (9.00)

(B)

Ranking Order	NormFinder	geNorm	BestKeeper	Comprehensive ranking (mean rank value)
1	*SNORA66*	*SNORD24 & SNORD49A*	*RNU6-2*	*SNORA66* (2.67)
2	*SNORD38B*		*SNORD43*	*SNORD24* / *SNORD43* (3.00)
3	*SNORD24*	*SNORD43*	*SNORA66*	
4	*SNORD43*	*SNORA66*	*SNORD38B*	*SNORD38B* (3.67)
5	*SNORD49A*	*SNORD38B*	*SNORD24*	*SNORD49A* (4.00)
6	*RNU6-2*	*RNU6-2*	*SNORD49A*	*RNU6-2* (4.33)
7	*SNORA74A*	*SNORA74A*	*SNORA74A*	*SNORA74A* (7.00)
8	*miR-191*	*miR-191*	*miR-191*	*miR-191* (8.00)
9	*miR-93*	*miR-93*	*miR-93*	*miR-93* (9.00)

SUPPLEMENTARY TABLE 4: Mean C_q-values of technical replicates per candidate reference gene and pooled body fluid samples after pre-processing (amplification efficiency as per Real-time PCR Miner algorithm and C_q-values as per SDS software).

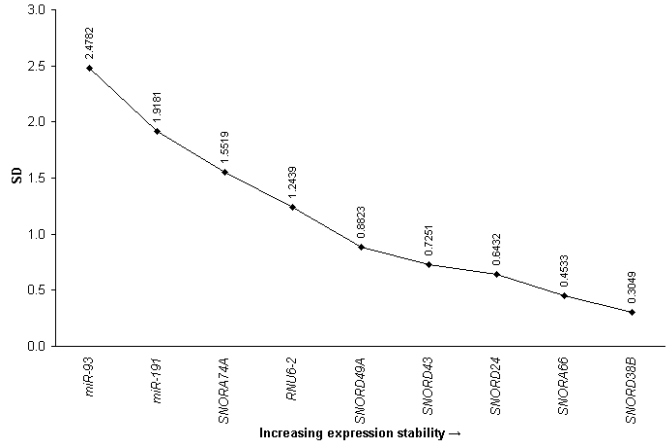
Body fluid	*miR-191*	*miR-93*	*RNU6-2*	*SNORA66*	*SNORA74A*	*SNORD24*	*SNORD38B*	*SNORD43*	*SNORD49A*
Blood	16.93	16.27	25.50	27.14	29.08	24.37	27.12	25.01	24.11
Menstrual blood	17.19	18.07	27.05	28.59	29.38	25.24	29.71	25.58	25.51
Saliva	22.90	23.41	28.27	29.64	33.33	26.87	30.82	26.53	26.52
Semen	21.84	23.80	28.30	32.07	34.35	30.52	33.00	29.77	30.55
Vaginal secretion	23.65	25.18	26.86	30.45	29.11	26.67	30.64	27.00	26.27

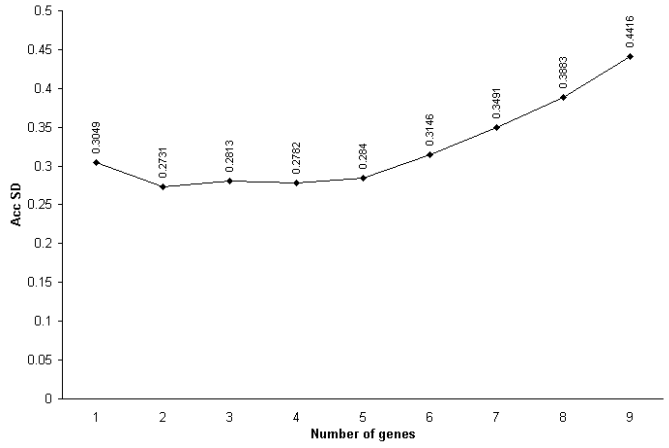
SUPPLEMENTARY TABLE 5: Mean C_q-values of technical replicates per candidate reference gene and individual body fluid samples after pre-processing (amplification efficiency and C_q-values as per LinRegPCR software).

Body fluid	*miR-191*	*miR-93*	*RNU6-2*	*SNORA66*	*SNORA74A*	*SNORD24*	*SNORD38B*	*SNORD43*	*SNORD49A*
Blood 6	15.76	15.94	24.39	24.62	25.76	21.66	24.82	23.08	21.76
Blood 7	15.89	16.27	24.54	24.64	26.02	22.51	24.71	23.10	21.35
Blood 8	15.67	16.15	25.67	24.71	25.91	22.38	25.73	23.86	22.65
Blood 9	16.20	17.22	25.72	25.31	26.61	23.14	25.76	24.20	22.88
Blood 10	16.10	16.35	24.20	24.51	26.73	22.94	24.81	22.92	21.21
Menstrual blood 1	15.30	15.36	25.61	24.42	25.66	22.33	26.59	23.04	21.92
Menstrual blood 2	16.59	16.61	25.50	24.80	26.08	22.75	26.43	23.47	22.75
Menstrual blood 3	19.14	19.84	26.48	27.84	28.11	23.70	26.92	23.94	23.08
Menstrual blood 4	17.40	18.15	26.50	25.58	26.45	23.13	26.59	23.18	23.33
Menstrual blood 5	22.97	25.45	27.66	31.89	30.39	26.14	29.65	25.13	24.84
Saliva 2	20.98	21.96	27.54	25.99	30.59	25.17	28.41	24.44	24.12
Saliva 4	21.45	22.43	26.70	25.35	28.49	24.15	27.53	24.60	23.31
Saliva 7	21.15	21.71	26.81	25.49	29.44	23.77	28.24	24.19	22.91
Saliva 8	21.97	23.04	28.93	28.94	32.80	25.93	29.50	24.00	25.55
Saliva 9	21.37	22.42	27.35	25.89	30.21	24.92	29.06	23.93	24.42
Semen 1	20.38	22.50	25.98	27.22	29.74	27.10	28.62	26.18	26.16
Semen 2	20.41	22.29	27.80	29.81	30.97	28.40	31.94	28.39	28.29
Semen 3	20.40	22.49	29.14	30.93	31.22	29.42	32.00	27.86	28.38
Semen 4	21.05	23.43	29.24	32.04	32.88	31.15	33.15	29.82	30.92
Semen 5	20.60	22.72	27.85	29.11	31.07	27.73	31.00	28.26	28.92
Vaginal secretion 1	20.91	22.28	25.19	24.90	25.59	23.48	28.98	24.31	23.90
Vaginal secretion 2	22.15	23.44	26.51	30.47	28.99	24.50	28.08	24.83	23.66
Vaginal secretion 3	21.99	26.49	26.77	30.62	28.79	26.01	30.52	25.25	24.31
Vaginal secretion 4	22.69	24.42	26.39	32.07	29.98	25.36	29.80	25.29	24.32
Vaginal secretion 5	22.02	24.54	25.97	28.81	28.94	25.15	28.32	24.78	23.49

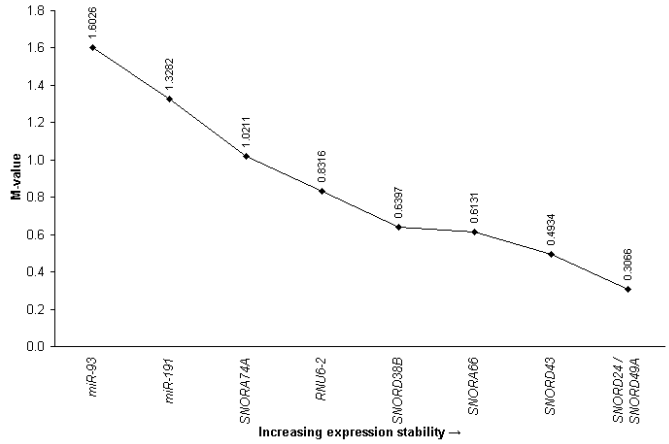
SUPPLEMENTARY TABLE 6: Mean C_q-values of technical replicates per candidate reference gene and individual body fluid samples after pre-processing (amplification efficiency as per Real-time PCR Miner algorithm and C_q-values as per SDS software).

Body fluid	*miR-191*	*miR-93*	*RNU6-2*	*SNORA66*	*SNORA74A*	*SNORD24*	*SNORD38B*	*SNORD43*	*SNORD49A*
Blood 6	16.44	16.76	25.14	27.14	27.56	23.14	26.77	24.75	23.65
Blood 7	16.58	17.11	25.28	27.15	27.79	24.04	26.62	24.78	23.15
Blood 8	16.43	17.03	26.45	27.33	27.85	24.09	27.74	25.68	24.64
Blood 9	16.96	18.11	26.52	28.04	28.67	24.90	27.80	26.06	24.88
Blood 10	16.84	17.22	24.96	27.16	28.71	24.71	26.75	24.68	23.07
Menstrual blood 1	15.96	16.19	26.40	26.96	27.45	23.89	28.70	24.76	23.82
Menstrual blood 2	17.24	17.45	26.08	27.34	27.92	24.34	28.53	25.14	24.72
Menstrual blood 3	19.98	20.85	27.08	30.80	30.19	25.47	29.07	25.77	25.06
Menstrual blood 4	18.20	19.11	27.17	28.30	28.42	24.90	28.70	24.94	25.34
Menstrual blood 5	24.02	26.74	28.40	35.39	32.64	28.13	32.01	27.04	26.98
Saliva 2	21.84	22.97	28.40	28.66	32.68	26.91	30.67	26.19	26.20
Saliva 4	22.28	23.47	27.53	27.97	30.49	25.86	29.74	26.37	25.30
Saliva 7	21.99	22.71	27.65	28.12	31.51	25.42	30.49	25.93	24.90
Saliva 8	22.91	24.18	29.86	32.04	35.23	27.89	31.82	25.84	27.73
Saliva 9	22.32	23.56	28.22	28.67	32.50	26.79	31.33	25.75	26.52
Semen 1	21.24	23.54	26.76	29.94	31.79	29.00	30.93	28.09	28.40
Semen 2	21.13	23.34	28.67	32.75	33.14	30.38	34.73	30.58	30.74
Semen 3	21.08	23.59	30.10	34.09	33.54	31.65	34.52	30.09	30.92
Semen 4	22.13	24.62	30.16	35.43	35.30	33.59	35.84	32.42	33.85
Semen 5	21.59	23.87	28.75	32.19	33.31	29.89	33.52	30.75	31.56
Vaginal secretion 1	21.77	23.30	25.94	27.45	27.38	25.11	31.24	26.06	25.93
Vaginal secretion 2	23.02	24.50	27.34	33.68	31.06	26.20	30.28	26.73	25.71
Vaginal secretion 3	22.93	27.80	27.63	33.92	30.94	27.96	32.81	27.22	26.40
Vaginal secretion 4	23.71	25.66	27.24	35.52	32.31	27.28	32.12	27.40	26.39
Vaginal secretion 5	23.03	25.81	26.81	31.95	31.13	27.06	30.42	26.67	25.54

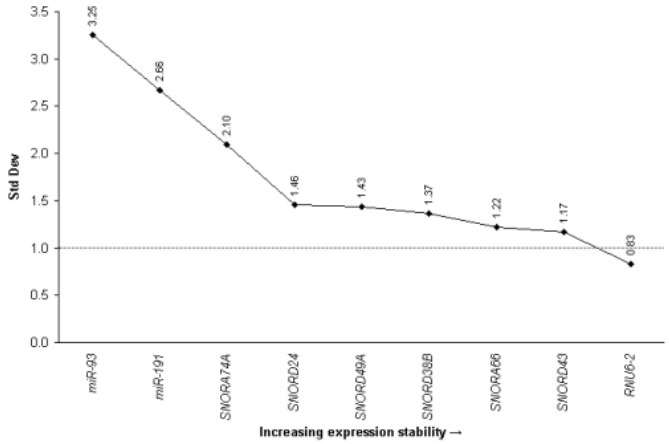




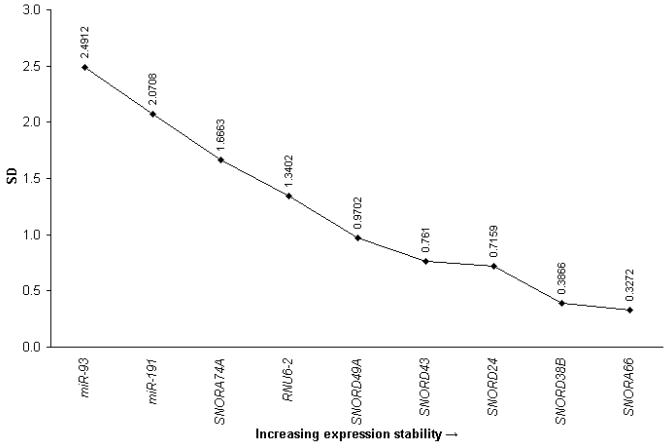
SUPPLEMENTARY FIGURE 1:
NormFinder data analysis of the nine candidate genes in pooled body fluid samples
(Upper panel) Gene expression stability values of genes - from least (*left*) to most stable (*right*).
(Lower panel) Determination of the optimal number of reference genes by computation of accumulated standard deviation values. Amplification efficiency and C_q-values as per LinRegPCR software.

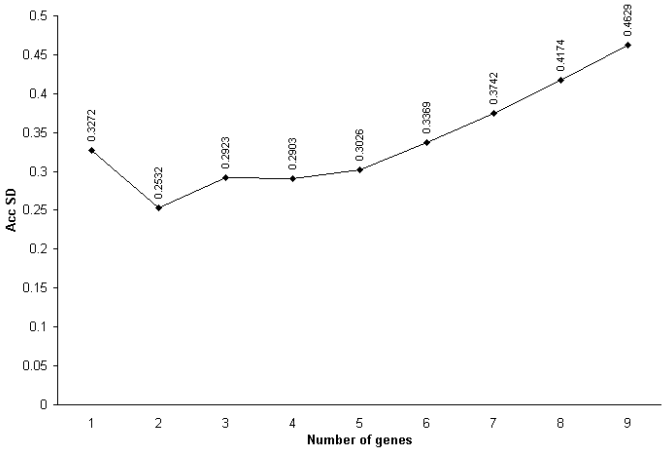


SUPPLEMENTARY FIGURE 2: **GeNorm data analysis of the nine candidate genes in pooled body fluid samples**
GeNorm proceeds by calculation of the gene stability measure (M-value) per gene - from least (*left*) to most stable (*right*); determination of the optimal pair of reference genes by stepwise exclusion of the gene with the highest M-value. Amplification efficiency and C_q-values as per LinRegPCR software.



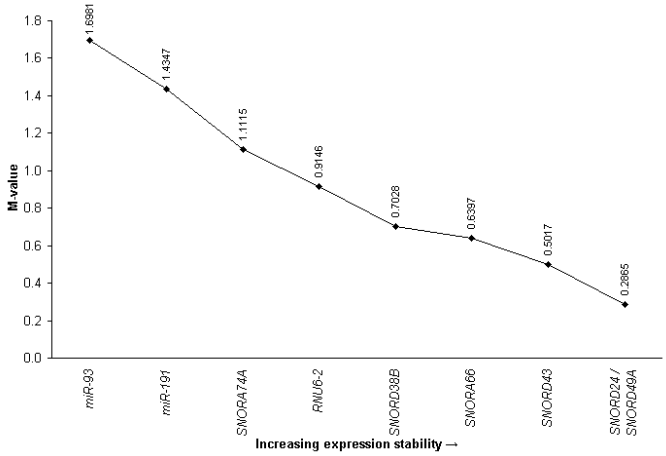
SUPPLEMENTARY FIGURE 3:
BestKeeper data analysis of the nine candidate genes in pooled body fluid samples
BestKeeper proceeds by pair wise correlation analysis of the C_q-values of all pairs of candidate reference genes - from least (*left*) to most stable (*right*). Amplification efficiency and C_q-values as per LinRegPCR software.





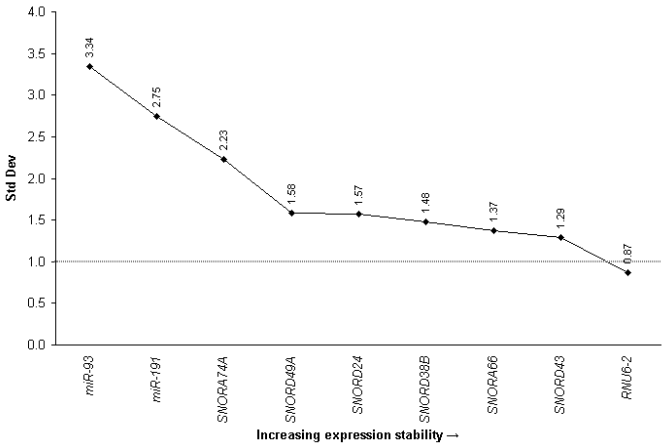
SUPPLEMENTARY FIGURE 4: **NormFinder data analysis of the nine candidate genes in pooled body fluid samples**

(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 as per PCR Miner algorithm and C_q-values as SDS software.



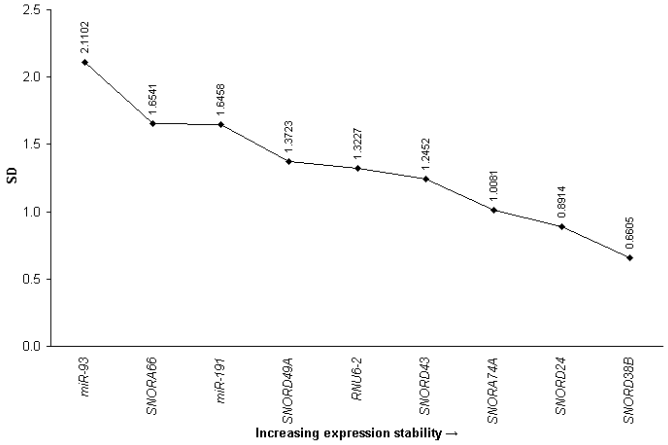
SUPPLEMENTARY FIGURE 5: **GeNorm data analysis of the nine candidate genes in pooled body fluid samples**

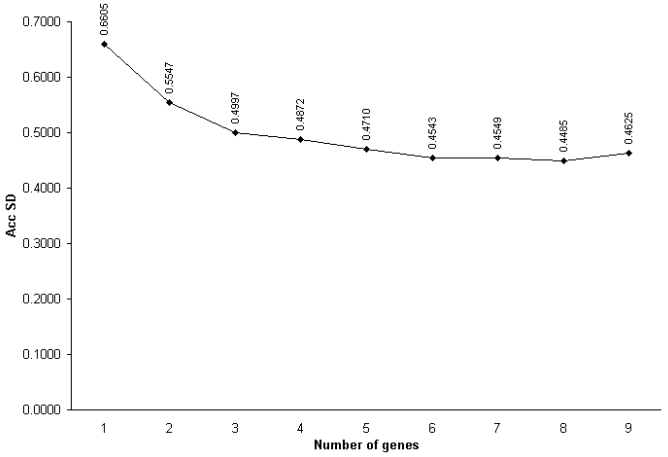
GeNorm proceeds by calculation of the gene stability measure (M-value) per gene - from least (*left*) to most stable (*right*); determination of the optimal pair of reference genes by stepwise exclusion of the gene with the highest M-value. Amplification efficiency as per PCR Miner algorithm and C_q-values as SDS software.



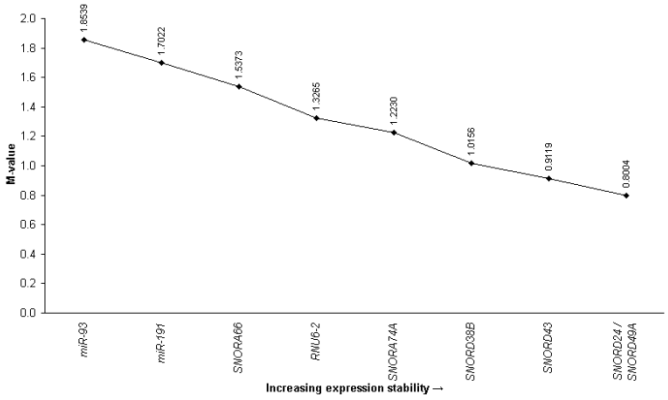
SUPPLEMENTARY FIGURE 6: **BestKeeper data analysis of the nine candidate genes in pooled body fluid samples**

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 as per PCR Miner algorithm and C_q-values as SDS software.

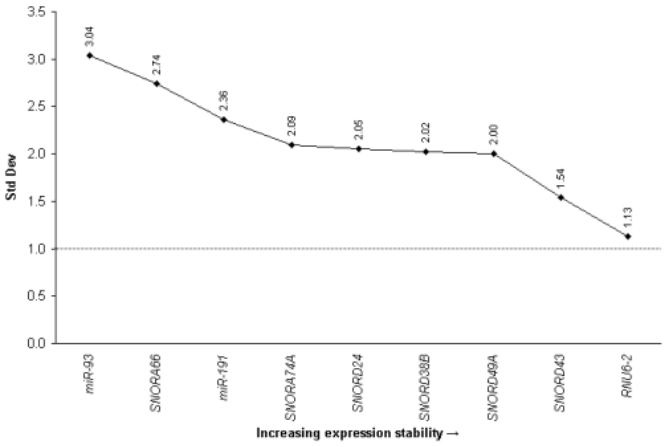




SUPPLEMENTARY FIGURE 7:
NormFinder data analysis of the nine candidate genes in individual body fluid samples
(Upper panel) Gene expression stability values of genes - from least (*left*) to most stable (*right*). (Lower panel) Determination of the optimal number of reference genes by computation of accumulated standard deviation values. Amplification efficiency as per PCR Miner algorithm and C_q-values as SDS software.



SUPPLEMENTARY FIGURE 8: **GeNorm data analysis of the nine candidate genes in individual body fluid 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 as per PCR Miner algorithm and C_q-values as SDS software.



SUPPLEMENTARY FIGURE 9:
BestKeeper data analysis of the nine candidate genes in individual body fluid 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 as per PCR Miner algorithm and C_q-values as SDS software.

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

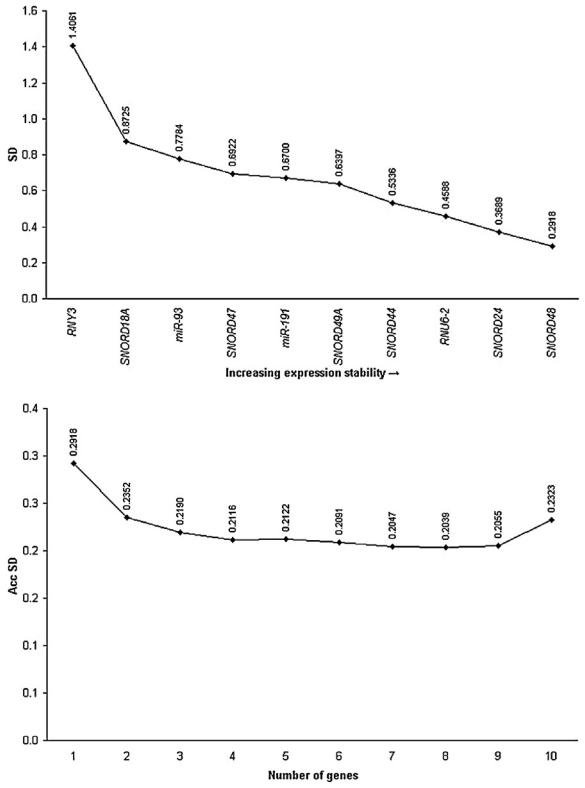


Fig. 1. NormFinder data analysis of 10 candidate genes in individual organ tissue samples. (Upper panel) Gene expression stability values of genes – from least (left) to most stable (right). (Lower panel) Determination of the optimal number of reference genes by computation of accumulated standard deviation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

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.

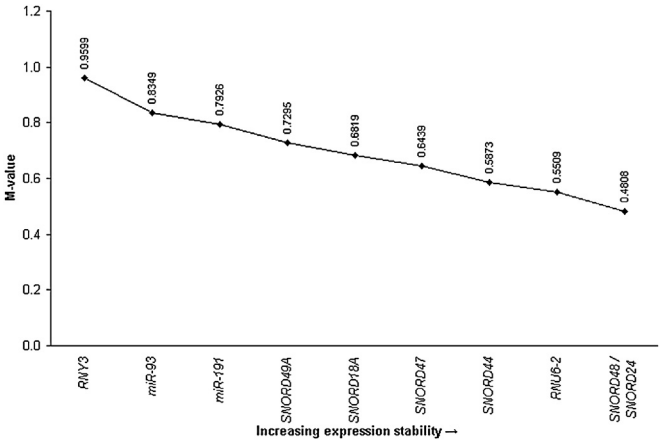


Fig. 2. GeNorm data analysis of 10 candidate genes in individual organ tissue samples. GeNorm proceeds by calculation of the gene stability measure (*M*-value) per gene – from least (left) to most stable (right); determination of the optimal pair of reference genes by stepwise exclusion of the gene with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

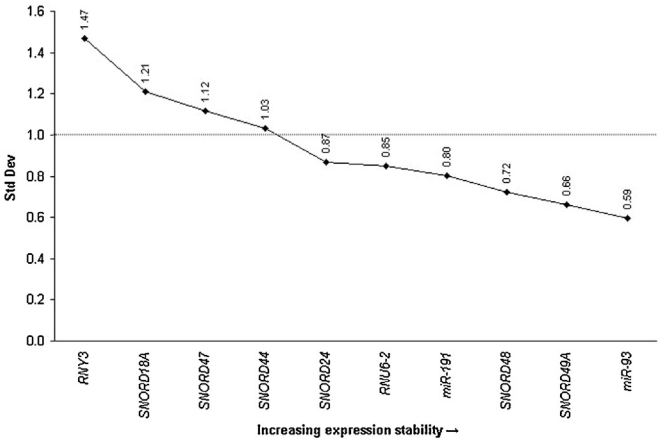


Fig. 3. BestKeeper data analysis of 10 candidate genes in individual organ tissue samples. BestKeeper proceeds by pair-wise correlation analysis of the C_q-values of all pairs of candidate reference genes – from least (left) to most stable (right). Amplification efficiency and C_q-values as per LinRegPCR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

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.

SUPPLEMENTARY TABLE 1: Mean C_q-values of technical replicates per candidate reference gene and pooled organ tissue sample after pre-processing (amplification efficiency and C_q-values as per LinRegPCR software).

Organ	*miR-26b*	*miR-92a*	*miR-93*	*miR-191*	*RNU6-2*	*RNY3*	*SNORD18A*	*SNORD24*	*SNORD44*	*SNORD47*	*SNORD48*	*SNORD49A*	*SNORD58B*	*SNORD75*
Brain	22.23	23.24	26.16	20.59	24.61	21.56	22.02	23.91	23.44	22.55	19.92	22.91	24.87	23.53
Heart	20.15	21.69	26.22	20.88	23.32	21.70	21.01	23.89	22.55	21.76	20.07	22.76	23.01	23.26
Kidney	20.88	20.27	26.38	21.18	24.34	23.95	21.98	23.91	23.34	22.85	20.73	22.69	22.87	22.93
Liver	20.86	21.05	26.80	21.56	23.76	25.31	21.71	24.05	22.39	22.40	20.04	22.15	23.80	22.83
Lung	19.83	19.83	25.07	20.05	23.93	22.83	20.76	22.36	22.11	21.39	20.18	22.15	22.96	22.45
SM	19.29	21.28	26.55	20.31	24.09	20.92	20.83	23.43	22.77	21.78	19.71	22.66	22.38	22.98
Skin	21.43	21.83	26.72	21.71	22.99	24.52	20.62	23.45	22.71	21.12	20.70	22.58	23.96	20.90

SM skeletal muscle

SUPPLEMENTARY TABLE 2: Comprehensive ranking order of the candidate reference genes for pooled 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	*SNORD24*	*SNORD47* & *SNORD18A*	*SNORD49A*	*SNORD44* (3.00)
2	*SNORD44*		*SNORD48*	*SNORD49A* (3.33)
3	*SNORD18A*	*RNU6-2*	*SNORD44*	*SNORD24* (4.00)
4	*SNORD49A*	*SNORD44*	*miR-93*	*SNORD18A* (4.33)
5	*SNORD47*	*SNORD49A*	*SNORD24*	*SNORD47* (4.67)
6	*miR-93*	*SNORD24*	*RNU6-2*	*SNORD48*, *miR-93* (5.67)
7	*SNORD48*	*miR-93*	*miR-191*	
8	*miR-191*	*SNORD48*	*SNORD47*	*RNU6-2* (6.00)
9	*RNU6-2*	*miR-191*	*SNORD18A*	*miR-191* (8.00)
10	*SNORD58B*	*SNORD58B*	*SNORD75*	*SNORD58B* (10.33)
11	*miR-26b*	*miR-26b*	*SNORD58B*	*SNORD75*, *miR-26b* (11.33)
12	*SNORD75*	*SNORD75*	*miR-26b*	
13	*miR-92a*	*miR-92a*	*miR-92a*	*miR-92a* (13.00)
14	*RNY3*	*RNY3*	*RNY3*	*RNY3* (14.00)

SUPPLEMENTARY TABLE 3: Mean C_q-values of technical replicates per candidate reference gene and 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 1	26.09	19.71	24.66	22.13	21.40	23.54	23.33	21.57	20.02	22.80
Brain 2	26.54	19.95	24.69	22.31	21.74	24.70	23.86	22.36	20.51	23.27
Brain 3	26.10	19.65	24.38	21.97	21.46	24.05	24.09	22.28	20.57	22.65
Brain 4	26.22	19.91	24.90	21.83	22.18	24.31	24.05	22.54	20.55	22.68
Brain 5	25.31	19.38	24.83	20.85	22.39	23.97	23.42	22.96	20.60	23.16
Heart muscle 1	26.15	20.79	24.29	21.07	20.44	23.23	22.38	21.25	19.93	23.43
Heart muscle 2	27.66	21.99	24.53	23.77	22.80	24.03	23.77	23.21	20.79	23.90
Heart muscle 3	25.56	20.63	23.84	22.82	20.74	22.54	22.93	21.18	20.03	22.66
Heart muscle 4	26.84	21.39	24.32	22.42	22.35	23.91	23.85	22.53	20.62	24.00
Heart muscle 5	24.90	19.81	22.71	19.57	20.15	22.25	21.23	20.70	19.05	22.34
Kidney 1	25.50	20.12	22.46	22.31	20.27	22.55	21.70	20.92	19.70	21.50
Kidney 2	26.36	22.11	25.92	24.60	24.22	25.64	25.97	25.02	22.07	24.78
Kidney 3	27.07	21.84	26.13	24.63	24.81	25.67	25.30	24.49	22.10	24.39
Kidney 4	27.20	23.44	26.54	24.99	26.48	26.64	27.26	26.28	23.31	25.72
Kidney 5	26.00	20.59	22.48	22.51	20.45	23.01	21.92	21.51	19.68	21.52
Liver 1	27.39	21.82	23.82	25.39	20.35	24.15	22.12	21.45	19.97	22.93
Liver 2	26.03	21.47	26.00	25.10	22.78	25.02	24.75	23.58	21.12	22.98
Liver 3	26.90	21.70	24.95	26.98	21.98	24.13	23.62	22.95	20.24	22.61
Liver 4	28.34	23.00	26.87	27.96	25.51	26.82	26.33	26.87	24.71	24.71
Liver 5	26.91	21.82	24.85	24.47	22.12	24.76	22.94	22.25	21.17	23.29
Lung 1	25.43	20.04	23.83	23.08	20.15	23.09	22.45	21.09	20.36	23.09
Lung 2	25.61	20.01	23.91	23.61	21.35	23.68	22.82	21.94	20.34	23.36
Lung 3	26.16	20.64	24.89	24.49	22.58	24.44	24.34	23.85	20.99	22.86
Lung 4	25.90	20.89	23.46	23.55	21.99	24.15	23.53	23.12	20.76	22.74
Lung 5	25.74	20.00	23.88	22.90	21.24	24.32	22.77	21.50	20.57	22.26
Skeletal Muscle 1	25.58	20.40	22.62	21.39	20.76	22.50	22.01	21.52	19.56	23.37
Skeletal Muscle 2	25.59	20.14	23.01	20.77	21.06	22.78	22.04	21.75	19.43	22.06
Skeletal Muscle 3	25.94	20.36	23.28	21.93	21.08	22.62	23.17	21.86	19.67	22.89
Skeletal Muscle 4	26.22	20.58	24.63	22.99	22.14	23.11	22.73	22.22	20.16	22.72
Skeletal Muscle 5	25.94	20.73	23.89	23.13	20.68	23.02	22.19	21.16	19.65	23.36
Skin1	27.42	21.49	24.44	24.95	20.30	23.48	23.11	21.09	20.70	22.43
Skin2	26.81	21.04	23.41	24.04	19.22	22.74	22.09	20.74	19.71	22.22
Skin3	26.92	22.28	25.29	26.73	20.68	23.49	24.24	21.85	21.22	23.53
Skin4	26.46	21.20	24.55	24.35	20.04	23.52	22.36	20.90	20.10	22.39
Skin5	26.26	20.51	23.69	22.34	20.19	23.00	22.75	20.82	20.37	22.22

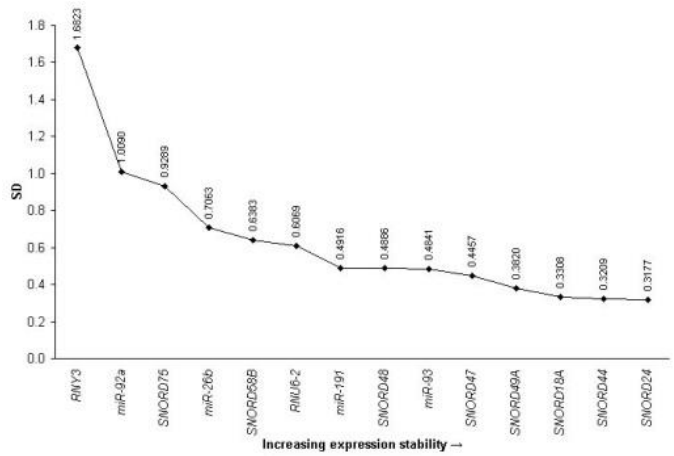
SUPPLEMENTARY TABLE 4: Comprehensive ranking order of the candidate reference genes for (A) pooled and (B) individual organ tissue samples, derived by integrating rankings of NormFinder, geNorm, and BestKeeper. Computation with amplification efficiency as per Real-time PCR Miner algorithm and C_q-values as per SDS software.

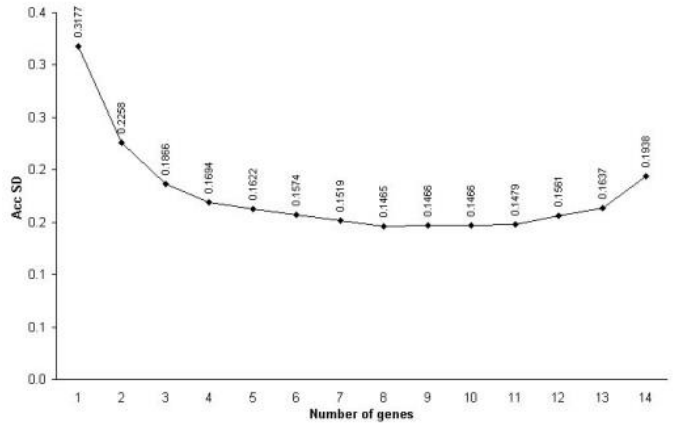
(A)

Ranking Order	NormFinder	geNorm	BestKeeper	Comprehensive ranking (mean rank value)
1	*SNORD44*	*SNORD18A & SNORD47*	*SNORD44*	*SNORD44* (2.00)
2	*SNORD24*		*SNORD49A*	*SNORD24* (3.33)
3	*SNORD48*	*SNORD24*	*SNORD48*	*SNORD47, SNORD48, SNORD49A* (4.33)
4	*SNORD47*	*SNORD44*	*miR-93*	
5	*miR-93*	*SNORD49A*	*SNORD24*	
6	*SNORD49A*	*miR-93*	*RNU6-2*	*miR-93* (5.00)
7	*SNORD18A*	*SNORD48*	*miR-191*	*SNORD18A* (5.67)
8	*miR-191*	*miR-191*	*SNORD47*	*miR-191* (7.67)
9	*RNU6-2*	*RNU6-2*	*SNORD18A*	*RNU6-2* (8.00)
10	*SNORD58B*	*SNORD58B*	*SNORD75*	*SNORD58B* (10.67)
11	*miR-26b*	*miR-26b*	*miR-92a*	*miR-26b, miR-92a* (11.67)
12	*miR-92a*	*miR-92a*	*SNORD58B*	
13	*SNORD75*	*SNORD75*	*miR-26b*	*SNORD75* (12.00)
14	*RNY3*	*RNY3*	*RNY3*	*RNY3* (14.00)

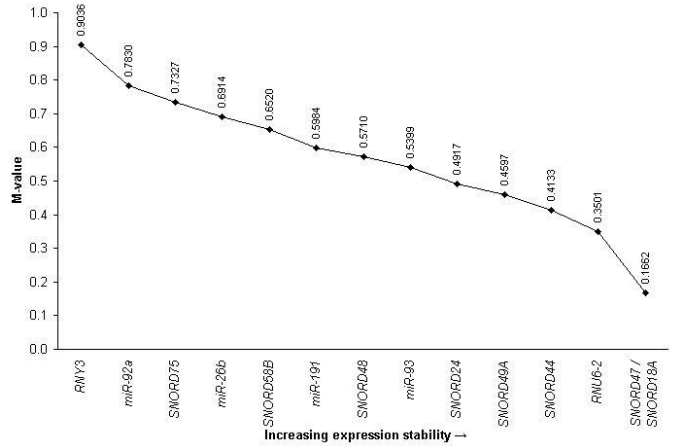
(B)

Ranking Order	NormFinder	geNorm	BestKeeper	Comprehensive ranking (mean rank value)
1	*SNORD48*	*SNORD24 & SNORD48*	*miR-93*	*SNORD48* (1.67)
2	*SNORD24*		*SNORD49A*	*SNORD24* (3.00)
3	*RNU6-2*	*SNORD44*	*SNORD48*	*RNU6-2, SNORD44* (4.67)
4	*SNORD44*	*SNORD47*	*miR-191*	
5	*miR-191*	*SNORD18A*	*RNU6-2*	*SNORD49A* (5.00)
6	*SNORD49A*	*RNU6-2*	*SNORD24*	*miR-191* (5.67)
7	*SNORD47*	*SNORD49A*	*SNORD44*	*miR-93* (6.00)
8	*miR-93*	*miR-191*	*SNORD47*	*SNORD47* (6.33)
9	*SNORD18A*	*miR-93*	*SNORD18A*	*SNORD18A* (7.67)
10	*RNY3*	*RNY3*	*RNY3*	*RNY3 (10.00)*

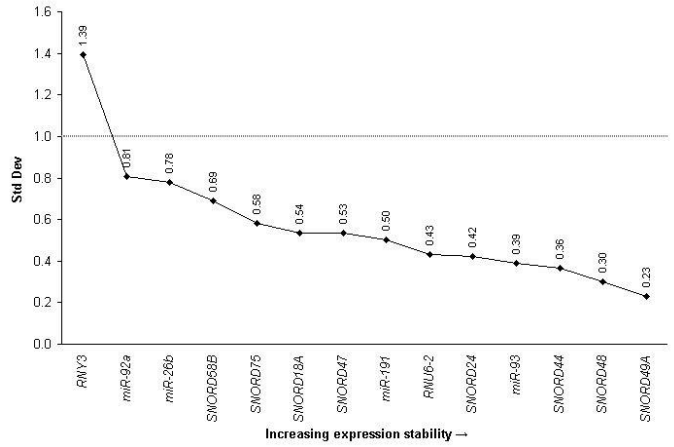




SUPPLEMENTARY FIGURE 1:
NormFinder data analysis of 14 candidate genes in pooled organ tissue samples
(Upper panel) Gene expression stability values of genes - from least (left) to most stable (right). (Lower panel) Determination of the optimal number of reference genes by computation of accumulated standard deviation values. Amplification efficiency and C_q-values as per LinRegPCR software.



SUPPLEMENTARY FIGURE 2: **GeNorm data analysis of 14 candidate genes in pooled 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 Cq-values as per LinRegPCR software.



SUPPLEMENTARY FIGURE 3:
BestKeeper data analysis of 14 candidate genes in pooled organ tissue samples
BestKeeper proceeds by pair wise correlation analysis of the Cq-values of all pairs of candidate reference genes - from least (left) to most stable (right). Amplification efficiency and Cq-values as per LinRegPCR software.

3.4 Zusammenfassung und Diskussion

Bei der Anwendung der RT-qPCR ist der Einsatz von Referenzgenen, deren Eignung empirisch für das jeweilige Experiment belegt wurde, essentiell für die Eliminierung nichtbiologischer Varianzen. Dies ist besonders wichtig für die Analyse von miRNA-Expressionsdaten, da hierzu bisher kein allgemeiner Konsens über die bestgeeignete Methode der Datennormalisierung besteht. In den hier vorgestellten Experimenten sollten daher robuste Normalisierungsstrategien etabliert und so eine methodische Grundlage für die Folgestudien zur Etablierung und Validierung körperflüssigkeits- und organgewebespezifischer miRNA-Kandidaten geschaffen werden.

Zunächst wurde jeweils eine Vorauswahl an miRNAs, small nuclear RNAs (snRNAs) und small nucleolar RNAs (snoRNAs) als potentielle Referenzgene getroffen, für die eine möglichst gleichmäßige und hohe Expression über verschiedene Gewebearten und Zelllinien beschrieben war oder die zuvor bereits im forensischen Kontext eingesetzt wurden [82,135,137,141,279–281]. Die selektierten Kandidaten wurden dann einem Auswahlprozess unterzogen, in dem zunächst die technische Performanz der verwendeten TaqMan®-Assays (Amplifikationserfolg und -effizienz) und anschließend ihre Expressionsstabilität in den untersuchten Körperflüssigkeiten und Organgeweben betrachtet wurde.

Die Berücksichtigung der Amplifikationseffizienz pro Amplikon ist notwendig für die korrekte Analyse von qPCR-Daten, da jene je nach Versuchsbedingungen variieren und ein Unterlassen zu verzerrten Quantifizierungsergebnissen führen kann [260,262,268,282,283]. Da eine Vielzahl an Algorithmen zur Berechnung der Effizienzen aus qPCR-Rohdaten existiert, aber keine Einigkeit darüber besteht welcher zu bevorzugen ist, wurden die Analysen mit zwei verschiedenen soliden [284] und MIQE-konformen [285] Programmen – ‚LinReg-PCR' [286–288] und ‚Real-time PCR Miner' [289] – durchgeführt. Die Amplifikationseffizienzen wurden dabei jeweils in Gemischen aller Körperflüssigkeiten respektive Organgewebe bestimmt. Dies ist plausiblermaßen die konservativste Herangehensweise, also die, die die wenigsten Annahmen erfordert, da im Falle realen Spurenmaterials, dessen Herkunft und Beschaffenheit unbekannt ist, provisorisch von einer maximal komplexen Zusammensetzung ausgegangen werden muss.

Auch hinsichtlich der derzeit verfügbaren Algorithmen zur Darstellung der Genexpressionsstabilität in einer Vorauswahl von Kandidaten-Referenzgenen besteht derzeit kein

Konsens über eine zu bevorzugende Methode. Daher wurden drei etablierte und vielfach eingesetzte Algorithmen –geNorm [269], NormFinder [290] und BestKeeper [291] – angewendet und die jeweils ermittelten Rangfolgen der Kandidaten in einem kombinierten Ranking zusammengefasst [277]. Eine erste Überprüfung erfolgte jeweils in vereinigten Proben mehrerer Individuen, um die durchschnittlichen Expressionsniveaus der Kandidaten-Referenzgene pro Körperflüssigkeit beziehungsweise Organgewebe abzuschätzen. Durch die anschließende Analyse in Einzelproben wurde das Bild um die interindividuellen Expressionsunterscheide erweitert und somit vervollständigt. Diese Analysen wurden parallel für die mit den in ‚LinRegPCR' und ‚Real-time PCR Miner' ermittelten Effizienzen korrigierten C_q-Werten durchgeführt und ergaben vergleichbare bis identische Rangreihenfolgen.

Die Art der Fragestellung dieses Projektes erforderte die Einbeziehung einer Variabilitätsebene, die bei der Etablierung bestgeeigneter, also möglichst varianzarm exprimierter Referenzgene eigentlich vermieden werden sollte [271]: die Analyse der Expressionsstabilität in fünf respektive sieben distinkten und durchaus heterogenen Körperflüssigkeiten beziehungsweise Organgeweben. Weil sich jedoch dadurch ein Zirkelschluss ergäbe, bei unbekanntem Spurenmaterial, dessen gewebliche Herkunft und Zusammensetzung bestimmt werden soll, eine Annahme bezüglich dieser Quelle zu treffen, um davon ausgehend die bestgeeigneten Referenzgene auszuwählen, ist der gewählte der konservativste Ansatz. Da *a priori* kein Grund besteht anzunehmen, dass ein Referenzgen in unterschiedlichen Zelltypen gleich stark exprimiert ist [271], wurden daher von vornherein vergleichsweise höhere Varianzen erwartet und in Kauf genommen.

Schlussendlich ergaben sich unter den untersuchten experimentellen Bedingungen die snoRNAs SNORD24, SNORD38B und SNORD43 als die am stabilsten exprimierten Gene in den Körperflüssigkeitsproben sowie die snoRNAs SNORD48 und SNORD24 und die snRNA RNU6-2 als am besten geeignete Referenzgene in Organgeweben und Haut. In den Folgearbeiten werden die gemittelten Expressionsdaten dieser Referenzgenkombinationen für die Normalisierung der qPCR-Daten körperflüssigkeits- und organgewebespezifischer miRNA-Kandidaten herangezogen und ihre Performanz und Eignung in realistischem forensischen Probenmaterial weiter evaluiert.

4 miRNA-basierte Identifizierung forensisch relevanter Körperflüssigkeiten

4.1 Einleitung

Die am häufigsten in der forensischen Fallarbeit auftretenden und damit relevantesten biologischen Spurenarten sind die Körperflüssigkeiten Blut, Speichel, Sperma, Vaginalsekret und Menstruationsblut. Diese nicht nur dem/der Spurenverursacher/in, sondern auch der körperlichen Herkunft zuzuordnen, kann, insbesondere bei Sexualdelikten, entscheidend dazu beitragen, einen Handlungsablauf zu rekonstruieren oder die Aussagen der Tatbeteiligten auf Plausibilität zu prüfen.

Da die konventionellen Methoden zur Identifikation von Körperflüssigkeiten unstreitige Nachteile aufweisen (siehe Kapitel 1.2.1), werden seit einigen Jahren alternative molekulargenetische Herangehensweisen erforscht, darunter die Analyse differentieller miRNA-Expressionsmuster. Die hierzu bisher veröffentlichten Studien befassten sich dabei entweder mit allen fünf Körperflüssigkeiten von Interesse [82,135,141,142], fokussierten sich auf eine Teilmenge [137,139,140,143,144] oder behandelten nur eine bestimmte Spurenart [134,145]. Die verwendete Methode zur miRNA-Detektion war jeweils entweder Endpoint-PCR mit anschließender kapillarelektrophoretischer Auftrennung der PCR-Produkte [140,143], SYBR®Green-basierte qPCR [82,134,137,142,144] oder TaqMan®-basierte qPCR [135,139,141,145].

Von den insgesamt 40 beschriebenen miRNA-Kandidaten wurden jedoch lediglich sechs von mehr als einer Arbeitsgruppe für die Inferenz derselben Körperflüssigkeit identifiziert: hsa-miR-16-5p [82,141,142] und hsa-miR-451a [82,137,139,140,142,143] für die Identifikation von venösem Blut oder Blut im allgemeinen Sinne (venöses Blut und Menstruationsblut), hsa-miR-200c-3p [137,145], hsa-miR-203a-3p [137,145] und hsa-miR-205-5p [82,137,139,140,142,145] für die Identifikation von Speichel sowie

hsa-miR-891a-5p für die Identifikation von Sperma [135,141,143,145]. Für Vaginalsekret und Menstruationsblut fehlen bisher bestätigte Kandidaten.

Eine Erklärung für die fehlende Übereinstimmung der Ergebnisse ist zum einen in der Anwendung unterschiedlicher Methoden zu suchen. Zum anderen ist die Ergebnisqualität vorheriger Studien dadurch eingeschränkt, dass meist nur eine unzureichende Normalisierung der Quantifizierungsdaten durchgeführt wurde, etwa durch die Verwendung nicht ausreichend validierter Referenzgene. Zudem wurden häufig mehr oder weniger willkürliche Kriterien für die Entscheidung, ob eine Körperflüssigkeit vorliegt oder nicht, festgelegt.

Die in der nachfolgenden Originalpublikation beschriebene Arbeit ist die erste Studie im forensischen Kontext zur miRNA-basierten Identifikation der fünf relevanten Körperflüssigkeiten Blut, Speichel, Sperma, Vaginalsekret und Menstruationsblut auf der Grundlage nicht nur hochgradig standardisierter Protokolle und deren Dokumentation gemäß den Angaben der MIQE-Richtlinien, sondern auch einer zuvor umfassend validierten, empirisch für die gegebenen Versuchsbedingungen erarbeiteten Normalisierungsstrategie (siehe Kapitel 3.2). Unter Einbeziehung der zwischen den Körperflüssigkeiten differentiell exprimierten miRNAs sollte zudem eine möglichst intuitive, einfach anzuwendende Entscheidungshilfe entwickelt werden, die es gestattet, die körperliche Herkunft unbekannter forensischer Proben auf objektiver Grundlage zu identifizieren.

Forensic Science International: Genetics 22 (2016) 89–99

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR

Eva Sauer, Ann-Kathrin Reinke, Cornelius Courts*

Institute of Legal Medicine, University of Bonn, Bonn, Germany

ARTICLE INFO

Article history:
Received 3 November 2015
Accepted 26 January 2016
Available online 12 February 2016

Keywords:
Forensic genetics
MicroRNA
qPCR
Body fluid identification

ABSTRACT

Applying molecular genetic approaches for the identification of forensically relevant body fluids, which often yield crucial information for the reconstruction of a potential crime, is a current topic of forensic research. Due to their body fluid specific expression patterns and stability against degradation, microRNAs (miRNA) emerged as a promising molecular species, with a range of candidate markers published. The analysis of miRNA via quantitative Real-Time PCR, however, should be based on a relevant strategy of normalization of non-biological variances to deliver reliable and biologically meaningful results. The herein presented work is the as yet most comprehensive study of forensic body fluid identification via miRNA expression analysis based on a thoroughly validated qPCR procedure and unbiased statistical decision making to identify single source samples.

1. Introduction

Regarding biological evidence found at the scene of a crime, it can be of crucial importance for the reconstruction of the events leading to its deposition to infer not only the DNA source, i.e. the person who deposited the stain, but also the bodily origin of the DNA, i.e. the body fluid(s) constituting the evidence material. Conventional methods of body fluid identification include immunological, chemical and enzymatic tests. These, however, vary greatly in terms of sensitivity and specificity and no reliable tests for e.g. vaginal secretion and menstrual blood are available. The comparably large amount of sample material required to perform these tests is another problematic aspect since the amount of evidential biological material is usually limited in forensic casework. In order to address these issues, a number of different molecular genetic approaches have been or are currently being explored by several groups (reviewed in Refs. [1–3]), including microRNA (miRNA) based body fluid identification.

Besides the essential prerequisite of cell type specific expression [4–6], these non-coding small RNAs exhibit certain characteristics that render them well suited to the challenging demands of the forensic setting: Firstly, due to their intrinsically small size of 18–25 nt, miRNAs are less prone to degradation caused by chemical and/or physical strains, secondly, miRNA is detected and quantified in its biologically active form so, in contrast to mRNA, no potential splice variants have to be differentiated, and thirdly, there is no gratuitous sample consumption as miRNA and DNA can be extracted simultaneously from the same specimen [7–9]. Previous studies have reported several body fluid specific miRNA markers that were identified using different methods [7–18] but only some of those markers were confirmed in more than one study. Also, most of these studies applied more or less arbitrary criteria to decide whether in a given sample a particular body fluid is present or not.

Quantitative PCR (qPCR) 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 variances is essential [19–22]. A robust normalization strategy that is appropriate for a particular experimental setup should be based on an individual and evidence based selection of one or a group of reference genes [19,23,24]. We therefore applied a normalization strategy specifically designed for body fluid identification in forensic type samples [25].

The aim of the herein presented study was to identify reliable miRNA markers for the statistically robust inference of the

* Corresponding author. Current address: Institute of Forensic Medicine, University Medical Center Schleswig-Holstein, Arnold-Heller-Straße 3, 24105 Kiel, Germany.
E-mail addresses: cornelius.courts@uksh.de, forensische.genetik@gmail.com (C. Courts).

http://dx.doi.org/10.1016/j.fsigen.2016.01.018
1872-4973/

forensically relevant body fluids venous blood, saliva, semen, vaginal secretion and menstrual blood, and to develop an intuitive and easily conveyable approach for the prediction of the origin of unknown forensic samples.

2. Material and methods

2.1. Samples

Samples for each tested body fluid, i.e. venous blood, saliva, menstrual blood, vaginal secretion and semen, were collected from healthy volunteers. All volunteers provided informed consent and the study protocol was reviewed and approved of by the ethics committee of the Hospital of the University of Bonn.

Venous blood was collected by venipuncture using dry vacutainer tubes and spotted onto sterile cotton swabs. For collection of saliva via buccal swab, donors were asked to abstain from eating, smoking, drinking and oral hygiene at least 30 min prior to sampling. Samples of semen-free vaginal secretion were collected by the female donors themselves using sterile stemmed cotton swabs. Menstrual blood samples were obtained by the female donors using tampons. Male donors provided freshly ejaculated semen in sealed Falcon tubes that was then transferred onto sterile stemmed cotton swabs and dried by the researcher immediately after receipt. All samples used for the selection of body fluid specific miRNAs and the blinded specimens were dried at room temperature and processed for RNA extraction after 24 h. The samples used in mixtures (combinations of different body fluids) were stored frozen (−80 °C) after initial drying at room temperature for 24 h, then thawed and dried for another 24 h prior to combining and extraction. Aged samples were stored at room temperature for the period of time indicated in Table 1. A subset of the aged venous blood stains was made available from the Institute of Legal Medicine, Halle (Saale), Germany. These blood stains were either spotted on cotton fabric or presented as dried blood without a carrier substrate.

2.2. RNA extraction and quantification

All surfaces, devices and instruments utilized in the extraction procedure were thoroughly cleaned using RNase-Zap® (Ambion™, Austin, TX, USA) to remove ambient RNases and only RNase-free reagents and plastic consumables were used.

Total RNA was extracted using the mirVana™ miRNA Isolation Kit (Ambion) according to the manufacturer's protocol. Prior to extraction, samples (whole swab tip or approximately 2 cm^2 of the tampon or blood stain) were cut into pieces and incubated with 350 µl Lysis/Binding Buffer at 56 °C for 1 h. Forensic Filters (Macherey-Nagel, Düren, Germany) were used to separate lysate and substrate by centrifugation at 13,000 × *g* for 1 min. Total RNA eluates were stored at −80 °C until further processing.

Aged and blinded samples as well as mixture of different body fluids were treated likewise with the exception of an extended incubation time of 3 h in Lysis/Binding Buffer.

To remove potential traces of genomic DNA, subsequent DNase I digestion was performed with the Turbo DNA-free™ Kit (Ambion), following the manufacturer's protocol. Total RNA concentration and quality, represented by the RNA integrity number (RIN) [26], were determined using the Quant-iT™ RNA Assay Kit on a Qubit fluorometer (both Invitrogen™, Darmstadt,

Table 1
Specification of storage period and total RNA concentration per aged sample.

Body fluid	Sample name	Gender	Date of sampling	Date of extraction	Age		Total RNA concentration (ng/µl)
					years	days	
Venous blood	oldB1	M	13.01.2015	10.06.2015		148	3.8
	oldB2	F	25.07.2013	10.06.2015	1	320	2.6
	oldB3	F	09.05.2012	10.06.2015	3	32	6.4
	oldB4	M	14.09.2011	10.06.2015	3	269	3.6
	oldB5	F	14.09.2011	10.06.2015	3	269	2.7
	bloodHalle1	n.a.	09.01.1989	15.06.2015	26	157	3.7
	bloodHalle2	n.a.	09.01.1989	15.06.2015	26	157	4.3
	bloodHalle3	n.a.	22.09.1980	15.06.2015	34	266	1.0
	bloodHalle4	n.a.	08.09.1980	15.06.2015	34	280	1.5
	bloodHalle5	n.a.	1979[a]	15.06.2015	35	min.166	2.2
	bloodHalle6	n.a.	1979[a]	15.06.2015	35	min.166	2.2
	bloodHalle7	n.a.	01.10.1979	15.06.2015	35	257	2.1
Menstrual blood	oldMB1	F	30.03.2014	17.08.2015	1	140	30.5
	oldMB2	F	18.07.2012	17.08.2015	3	30	26.2
	oldMB3	F	12.10.2011	17.08.2015	3	309	22.8
Saliva	oldSA1	M	03.01.2015	17.08.2015		226	5.2
	oldSA2	F	23.05.2012	17.08.2015	3	86	6.7
	oldSA3	M	14.09.2011	17.08.2015	3	337	5.1
Semen	oldSE1	M	05.05.2014	17.08.2015	1	104	15.1
	oldSE2	M	05.05.2014	17.08.2015	1	104	5.7
	oldSE3	M	16.09.2011	17.08.2015	3	335	10.0
	oldSE4	M	15.09.2011	17.08.2015	3	336	0.9
Vaginal secretion	oldVS1	F	23.10.2014	17.08.2015		298	67.3
	oldVS2	F	11.06.2012	17.08.2015	3	67	56.6
	oldVS3	F	14.10.2011	17.08.2015	3	307	259.0

F—female; M—male; n.a.—gender information not available.
[a] Not otherwise specified.

Germany) and the RNA 6000 Nano Kit on an Agilent 2100 Bioanalyzer (both Agilent, Böblingen, Germany), respectively.

2.3. Selection of candidate miRNAs and reference genes

For the selection of body fluid specific candidate miRNAs, miRNA microarray profiling was conducted by Exiqon Services (Vedbaek, Denmark). Briefly, after quality control of the total RNA samples by an Agilent 2100 Bioanalyzer profile, 400 ng total RNA from the samples and a reference RNA sample was labeled with Hy3™ and Hy5™ fluorescent label, respectively, using the miRCURY LNA™ microRNA Hi-Power Labeling Kit, Hy3™/Hy5™ (Exiqon). The Hy3™-labeled samples and the Hy5™-labeled reference RNA sample were mixed pair-wise and hybridized to the miRCURY LNA™ microRNA Array 7th (Exiqon, Denmark), which contains capture probes targeting all human miRNAs registered in the miRBASE v.18.0 [27]. The hybridization was performed according to the miRCURY LNA™ microRNA Array Instruction manual using a Tecan HS4800™ hybridization station (Tecan, Austria). The miRCURY LNA™ microRNA Array slides were scanned using the Agilent G2565BA Microarray Scanner System (Agilent Technologies, Inc., USA) and the image analysis was carried out using the ImaGene 9.0 software (BioDiscovery, Inc., USA). The quantified signals were background corrected (Normexp with offset value 10 [28]) and normalized using the quantile normalization method. Subsequently, unsupervised as well as supervised data analyses were performed. In addition, quantile normalization was performed for each subgroup (i.e. body fluid) separately. The miRNAs were then ranked according to their expression levels and candidate miRNAs were selected by comparing the mean signal intensities between body fluids manually.

In addition, the available literature on miRNA based forensic body fluid identification (BFI) [7–17] was perused for promising markers.

The selection of reference genes for this study was described elsewhere [25].

2.4. Sample preparation

Individual samples were diluted to 4 ng/μl based on quantification results and were used as single samples only or additionally in preparing a pooled sample per body fluid by combining identical volumes of diluted samples.

The pooled samples used in the first step of marker selection consisted of ten donor samples for venous blood and saliva, while those for vaginal secretion, menstrual blood and semen contained five donor samples. Further, a mixture of all five body fluids containing identical volumes of the above mentioned pooled samples was prepared for the assessment of PCR efficiency.

In the second step of marker selection, we examined 15 individual samples of the respective body fluid of interest and at least five per every non-target body fluid per candidate miRNA (e.g. for hsa-miR-144-3p, a candidate miRNA for the identification of venous blood, 15 venous blood samples and at least five samples each for saliva, semen, menstrual blood and vaginal secretion were examined).

Aged and blinded samples as well as the mixture of different body fluids tested with the selected miRNAs were also diluted to 4 ng/μl, with the exception of the aged blood samples and one aged semen sample that were either diluted to 2 ng/μl or used undiluted.

2.5. Reverse transcription qPCR (RT-qPCR)

Complementary DNA (cDNA) was synthesized using target-specific stem-loop primers (Supplementary Table 1) and the TaqMan® MicroRNA Reverse Transcription Kit (Applied Biosystems™, Weiterstadt, Germany) as per manufacturer's protocol. For the body fluid specific markers, each 15 μl reaction volume contained 10 ng total RNA, 1X RT primers, 50 U MultiScribe™ reverse transcriptase, 1 mM dNTPs, 3.8 U RNase inhibitor, and 1X reverse transcription buffer. The reference genes were transcribed in pooled reactions containing 20 ng total RNA, 1X RT primers of SNORD24, SNORD38B and SNORD43, 100 U MultiScribe™ reverse transcriptase, 1 mM dNTPs, 7.6 U RNase inhibitor, and 1X reverse transcription buffer in a total reaction volume of 30 μl. For a subset of aged blood stains the RNA concentrations were less than 2 ng/μl, thus the maximum volume of 5 μl was used. For samples exhibiting less than 4 ng/μl, the reference genes were transcribed in individual reactions. Reactions were performed on a T3 Thermocycler (Biometra, Göttingen, Germany) with the following cycling conditions: 16 °C for 30 min, 42 °C for 30 min, and 85 °C for 5 min. Extraction negative and H_2O controls were performed on a sample basis and RT(−)-controls were set up to control for potential contamination with genomic DNA. All reverse transcription reactions that showed specific amplification were technically replicated in an independent second reaction. RT reaction products were stored at −20 °C.

QPCR reactions were performed using target-specific TaqMan® Assays (Supplementary Table 1) and the TaqMan® Universal PCR Master Mix II, No AmpErase® UNG (Applied Biosystems) as per manufacturer's protocol: 1.3 μl of the appropriate RT reaction product were added to a 20 μl reaction volume, containing 1X TaqMan Universal PCR Master Mix and 1X specific TaqMan® Assay. QPCR reactions were conducted in duplicate for each RT reaction product, resulting in four technical replicates per sample. As an exception, for the mixture containing all body fluids for efficiency determination qPCR triplicates were performed amounting to six technical replicates. The internal PCR control (IPC) from the Quantifiler® Human DNA Quantification Kit (Applied Biosystems), devised to produce a reference C_q-value, was used as an inter-plate calibrator. PCR cycling conditions consisted of 95 °C for 10 min and 40 cycles of 95 °C for 15 s and 60 °C for 1 min, and were performed on a 7500 Real-Time PCR System (Applied Biosystems). Data collection was performed during the 60 °C step by the SDS software version 2.0.6 (Life Technologies™, Darmstadt, Germany) and R_n-values were exported for further analyses.

2.6. Data analysis

The LinRegPCR program version 2014.8 [29] was employed to compute C_q-values and amplification efficiencies from R_n-values in a MIQE compliant manner [30] (see below). The arithmetic mean values of amplification efficiencies per six technical repeats were used in further analyses, 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 $\log_{10}$ (fluorescence). C_q-values deviating more than one cycle from the median value of the respective technical repeat were excluded from subsequent analyses. C_q-values ≥35 were regarded to result from non-specific amplification [31].

Analysis of qPCR data, including pre-processing, was then performed using the GenEx software version 5.3 (multiD Analyses, Göteborg, Sweden) into which LinRegPCR spread sheet exported data was imported. Pre-processing of qPCR encompassed the following steps in the given order: efficiency correction, averaging of technical qPCR and RT replicates and normalization with reference genes resulting in ΔC_q-values. For relative quantification, $\Delta\Delta C_q$-values were calculated by using the respective target body fluid as a calibrator sample $\Delta\Delta C_q = \Delta C_q$(non-target body fluid) - ΔC_q(target body fluid). For all samples not generating specific

amplification, ΔC_q- and $\Delta\Delta C_q$- values were set to the arbitrary values of 15 and 20, respectively. Inter-plate calibration was omitted, since evaluation of the C_q-values of the IPC as well as concordant technical repeats between plates indicated negligible differences between plate performances (data not shown).

Further statistical analyses, including discriminant function analysis, were performed with IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp., Armonk, NY).

2.7. Compliance to the MIQE guidelines

To facilitate reliable and unequivocal interpretation of the qPCR results reported herein, information that is rated 'essential' according to the MIQE guidelines [32] is reported, where applicable.

3. Results and discussion

3.1. Sample quantity and integrity

Quantity and integrity of total RNA varied notably among samples of the same body fluid as well as between groups (Supplementary Table 1), with venous blood samples exhibiting the lowest (total RNA concentration: 4.7–20.3 ng/μl; RIN: n.d.–2.4) and vaginal secretion samples the highest overall values (total RNA concentration: 39.4–348.0 ng/μl; RIN: 2.5–6.9). With the exception of one saliva and two vaginal secretion samples, all RIN values were <4. However, since studies suggest that low RNA integrity values do not adversely affect RT-qPCR profiling results [33,34] and no difficulties were encountered using the same type of specimens mimicking forensic evidence in identifying small nucleolar RNAs (snoRNAs) as reference genes [25], the recommendation of a minimum RIN of 5 [35,36] for samples to be proceeded was disregarded.

The quantity of the blinded specimens and also the aged samples fit well into the range of the samples used for the marker selection (Supplementary Table 2 and Table 1, respectively), with the exception of the aged blood samples that showed concentrations between 1.0 and 6.4 ng\μl and one aged semen sample for which the concentration was 0.9 ng/μl. Quantification results for the mixture samples are given in Table 2 .

All extraction negatives, RT(−)- and H_2O-controls were free of specific amplification (data not shown), evincing that the observed specific amplification truly depicts the miRNA expression status of a given sample.

3.2. Micro array data analysis and selection of candidate miRNAs for body fluid specific signatures

After the removal of miRNAs exhibiting signal intensities above background in less than 20% of samples, a total of 743 miRNAs remained for further analyses. Unsupervised hierarchical clustering including the 50 miRNAs with highest standard deviation collocated almost all samples according to their biological origin indicating that differences between the body fluids explain the largest percentage of observed expression variation (Fig. 1). As an exception to this general observation, two menstrual blood samples clustered within the vaginal secretion samples and in between samples of venous blood and saliva, respectively reflecting the complexity of this body fluid as a composite of venous blood, endometrial mucosal tissue and vaginal secretions.

Candidate miRNA selection encompassed expression analysis with the entire data set being quantile normalized followed by pairwise comparison of two body fluids at a time as well as quantile normalization for each body fluid subgroup separately with subsequent ranking of the miRNAs according to their expression levels. The latter was conducted in consultation with the service provider, as the heterogeneity of the five body fluids examined hindered accurate common normalization of the data. This approach is scientifically sound and generated a correct dataset, but did not allow any significance levels to be applied to the results.

Thus, for candidate selection, we focused on miRNAs that were either exclusively expressed in one body fluid, showed high signal intensities in one body fluid as compared to the remaining four, or displayed differential expression patterns between body fluids of which the distinction is of particular importance, e.g. saliva versus vaginal secretion. While a large number of promising venous blood specific miRNA candidates could be derived from the microarray expression data, this proved to be more difficult for the remaining body fluids. For example, the majority of miRNAs with high expression in menstrual blood were expressed, as expected, at similar or even higher levels in venous blood (data not shown). Hence, good candidates for identification of venous blood were excluded in favor of more promising ones, while for the other body fluids prima facie less suitable candidates were included. In any case, it has to be noted, that candidate marker selection based on microarray expression data as a matter-of-fact bears some degree of arbitrariness, given an array's limitation of target representation

Table 2
Specification of the composition of mixture samples containing two or three body fluids and the respective total RNA concentration.

Sample name	Composition	Total RNA concentration (ng \μl)
M01	Blood (ca. 2 cm^2) and vaginal secretion (whole swab) from the same female donor	56.0
M02	Blood (ca. 2 cm^2) from a female donor and semen (whole swab) from a male donor	4.7
M03	Blood (ca. 2 cm^2) from a female donor and semen (ca. 0.2 × 0.2 cm) from a male donor (same donors as in M02)	5.7
M04	Blood (ca. 2 cm^2) from a female donor and saliva (whole swab) from a male donor	8.4
M05	Blood (ca. 2 cm^2) and menstrual blood (ca. 2 cm^2) from the same female donor	32.1
M06	Semen (whole swab) from a male donor and vaginal secretion (whole swab) from a female donor	9.6
M07	Semen (whole swab) from a male donor and vaginal secretion (whole swab) from a female donor	51.5
M08	Semen (ca. 0.2 × 0.2 cm) from a male donor and vaginal secretion (whole swab) from a female donor (same donors as in M07)	7.1
M09	Semen (whole swab) from a male donor and menstrual blood (ca. 2 cm^2) from a female donor	53.0
M10	Semen (whole swab) from a male donor and menstrual blood (ca. 2 cm^2) from a female donor	21.6
M11	Semen (whole swab) and saliva (whole swab) from the same male donor	21.4
M12	Saliva (whole swab) from a male donor and vaginal secretion (whole swab) from a female donor	10.1
M13	Semen (ca. 0.2 × 0.2 cm) from a male donor and blood (ca. 2 cm^2) and vaginal secretion (whole swab) from the same female donor	54.5
M14	Semen (ca. 0.2 × 0.2 cm) from a male donor and blood (ca. 2 cm^2) and vaginal secretion (whole swab) from the same female donor	10.7

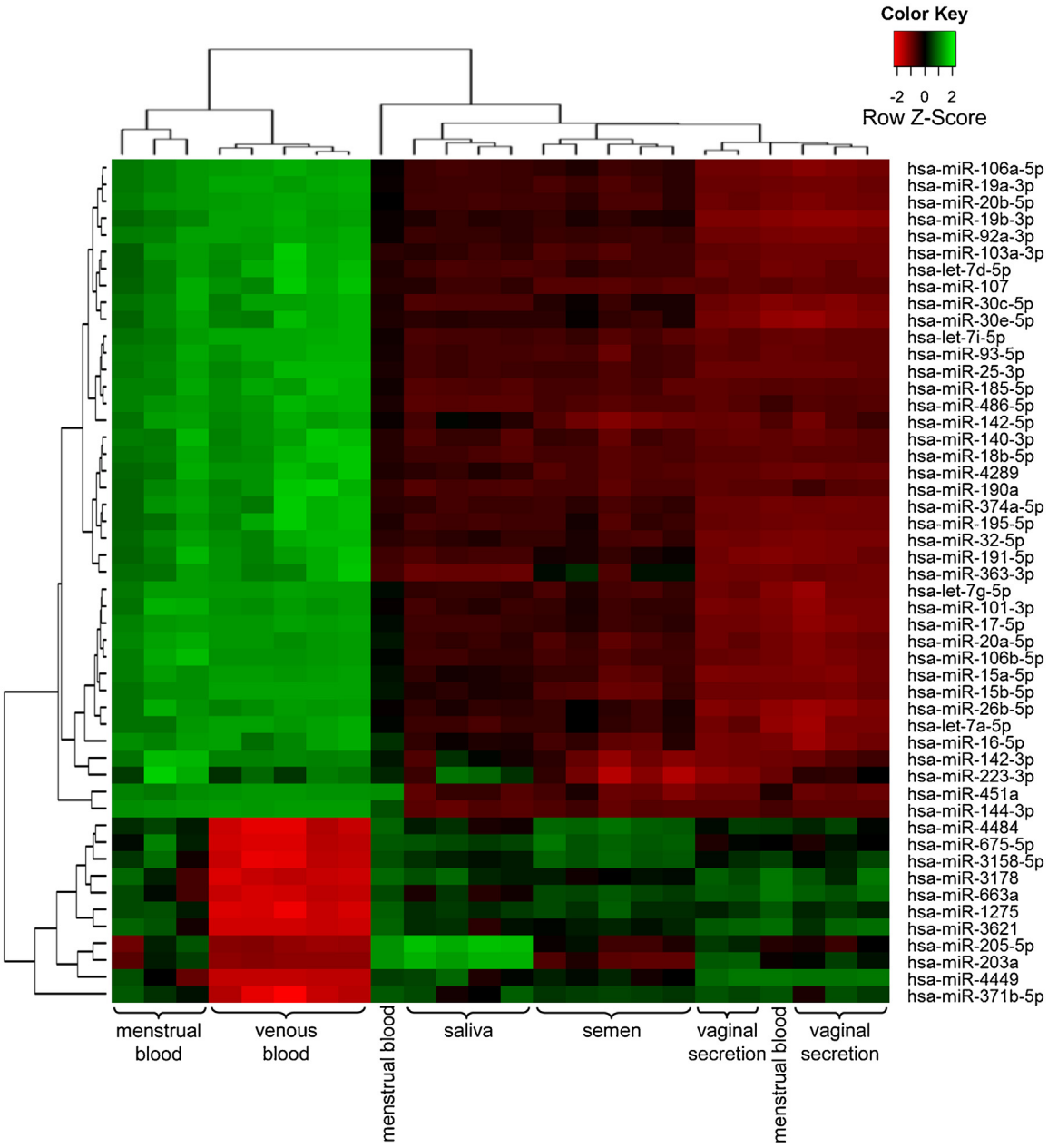


Fig. 1. Unsupervised hierarchical clustering of 24 samples from five forensically relevant body fluids (five samples per venous blood, menstrual blood, semen and vaginal secretion, and four saliva samples—one saliva sample did not meet the quantity requirements for the microarray experiment (data not shown)) using the 50 miRNAs with highest standard deviation when quantile normalization was applied to the whole dataset.

and dynamic range as well as the lack of fixed rules as to which criteria have to be met.

Therefore, for an additional selection of markers for evaluation in RT-qPCR experiments, the pertinent forensic literature was perused and the results integrated with those of the microarray based selection applying differential weighting, as illustrated in the following examples: The selection of hsa-miR-451a as a marker for blood in general was straight forward as it showed the highest signal in both venous blood and menstrual blood in the microarray and exhibited no expression in saliva, semen and vaginal secretion. Furthermore, this marker was described to be blood specific in six previous studies applying different approaches (Supplementary Table 3). According to the microarray, hsa-miR-891a-5p, a candidate miRNA for semen, exhibited only weak expression in vaginal secretion and none in the remaining body fluids, but since Zubakov et al. described a discrepancy between microarray and RT-qPCR results for this marker [11] and a semen specific expression was found in three independent studies (Supplementary Table 3) it was included in the candidate panel. The selection of markers for vaginal secretion is notoriously challenging, displayed inter alia by the comparably low number of studies presenting reliable vaginal secretion markers. Therefore, hsa-miR-155-5p and hsa-miR-196a-3p were selected for evaluation in RT-qPCR experiments although they had not been described before, because they displayed a higher expression level in vaginal secretion samples as compared to saliva and semen specimens. A comprehensive overview of the selected as well as the previously published miRNA candidates is given in Supplementary Table 3.

The 36 candidate miRNAs selected for evaluation in quantitative PCR experiments comprised seven markers for venous blood or

blood in general – from now on referred to as 'blood' – (hsa-miR-16-5p, hsa-miR-20a-5p, hsa-miR-22-5p, hsa-miR-106a-5p, hsa-miR-126-3p, hsa-miR-144-3p and hsa-miR-451a), menstrual blood (hsa-miR-144-5p, hsa-miR-185-5p, hsa-miR-214-3p, hsa-miR-371b-5p, hsa-miR-675-5p, hsa-miR-2116-5p and hsa-miR-4444) and saliva (hsa-miR-23a-3p, hsa-miR-34a-5p, hsa-miR-203a-3p, hsa-miR-205-5p, hsa-miR-658, hsa-miR-1290 and hsa-miR-4732-5p), respectively, six markers for semen (hsa-miR-10a-5p, hsa-miR-10b-5p, hsa-miR-135a-5p, hsa-miR-585-3p, hsa-miR-891a-5p and hsa-miR-1260a), as well as nine markers for vaginal secretion (hsa-miR-124-3p, hsa-miR-155-5p, hsa-miR-196a-3p, hsa-miR-654-5p, hsa-miR-1260b, hsa-miR-3685, hsa-miR-4286, hsa-miR-4653-3p and hsa-miR-4795-3p).

3.3. Amplification efficiency and amplification success

Amplification efficiency per amplicon was derived from the mixture containing all body fluids, including duplicated RT reactions and qPCR triplicates into the computation, while amplification success was evaluated in the pooled samples per body fluid.

No specific amplification in the pooled body fluid samples was observed for hsa-miR-658 ('saliva'), hsa-miR-585-3p ('semen'), as well as hsa-miR-196a-3p, hsa-miR-3685 and hsa-miR-4795-3p (all 'vaginal secretion'). Hsa-miR-658 was initially selected based on results published by Hanson et al. [10,15], where it showed good separation properties between saliva and other body fluids. The lack of specific amplification in our pooled samples is, however, in concordance with a study by Wang et al. [18] and is most likely due to methodological differences since Hanson et al. used SYBR Green, while both Wang et al. and our group employed TaqMan assays. The other four candidates were selected from the microarray and the non-compliance of results is most likely explained by either the distinct methods or the contingency that due to the high number of genes being tested in parallel, microarrays are prone to false positive results.

Two candidate markers for menstrual blood were excluded from further analyses due to their grossly outlying amplification efficiency in the mixture (hsa-miR-371b-5p: 32%; hsa-miR-2116-5p: 53%) and across all pooled body fluid samples (data not shown). Further, hsa-miR-654-5p was removed due to its low amplification efficiency in the target body fluid vaginal secretion (data not shown). For the remaining candidates, mean amplification efficiencies ranged between 102% for hsa-miR-4653-3p and 79% for hsa-miR-124-3p (Supplementary Table 4). Amplification efficiencies of the reference genes were taken from Sauer et al. [25].

Moreover, hsa-miR-675-5p, hsa-miR-4444 (both 'menstrual blood'), hsa-miR-4732-5p ('saliva'), as well as hsa-miR-155-5p and hsa-miR-4653-3p (both 'vaginal secretion') were not further investigated as their expression levels lay below those of the reference genes in the respective target body fluid (i.e. $\Delta C_q > 0$, data not shown).

3.4. Evaluation of candidate miRNAs for body fluid specific signatures with RT-qPCR

An initial screening of the selected candidate miRNAs was performed in pooled samples to evaluate their general expression levels in the five examined forensically relevant body fluids (results summarized in Supplementary Fig. 1). Based on these results the most promising markers were selected for testing in individual samples, thereby adding inter-individual differences to the picture (summarized in Supplementary Fig. 2).

Of the seven miRNAs selected as potential blood specific markers, five exhibited expression patterns in the pooled samples suggestive of sufficient specificity for the distinction of blood in general from the remaining forensically relevant body fluids: while the expression differences between the venous blood and menstrual blood pooled samples were comparable ($2.0 < \Delta\Delta C_q < 2.5$) for all tested markers, hsa-miR-16-5p, hsa-miR-20a-5p, hsa-miR-126-3p, hsa-miR-144-3p and hsa-miR-451a exhibited considerably lower expression levels in saliva, semen and vaginal secretion, with hsa-miR-144-3p devoid of any specific amplification in those three pooled samples. Hsa-miR-22-5p, selected from the microarray as a potential marker for the differentiation of venous blood and menstrual blood not only showed no larger difference between these two body fluids than the other miRNAs, but was expressed at similar levels in all body fluids.

For further evaluation in the individual sample set, we focused on hsa-miR-126-3p, hsa-miR-144-3p and hsa-miR-451a, bearing the most pronounced separation properties. The general expression patterns of these markers could be confirmed, however, as expected due to the composition of menstrual blood, a considerable overlap in expression levels between venous blood and menstrual blood samples was observed for all three miRNAs.

But while the expression values for hsa-miR-126-3p and hsa-miR-451a in menstrual blood overlapped with those in semen (both) and saliva (hsa-miR-451a), hsa-miR-144-3p lacked specific amplification in all non-blood samples (i.e. saliva, semen and vaginal secretion), rendering it a solid marker for identification of blood in general.

The three candidate markers for menstrual blood hsa-miR-144-5p, hsa-miR-185-5p and hsa-miR-214-3p displayed similar expression patterns than those for blood, with slightly higher expression in the venous blood sample compared to menstrual blood. Nonetheless, they were tested in individual samples, but did not add to the information obtainable by the blood markers described above. (In the course of testing hsa-mir-214-3p in individual samples, the TaqMan assay's amplification efficiencies varied greatly between batches, and even within the same sample, which remained unexplained (data not shown). Therefore, in consultation with the manufacturer's technical support, the assay was excluded from analyses.)

All possible saliva specific markers showed only minor differences in expression between the pooled saliva and vaginal secretion samples, with slightly higher expression in saliva ($\Delta\Delta C_q \leq 1.2$). The expression patterns of hsa-miR-203a-3p and hsa-miR-205-5p, however, indicated sufficient separation properties for saliva and vaginal secretion from the remaining three body fluids with $\Delta\Delta C_q$-values >4 between saliva and semen, venous blood and menstrual blood. Consequently, both were tested in the set of individual samples.

While the general expression patterns were concordant with those of the pooled samples, blending the inter-individual differences into the picture revealed an overlap in expression levels between saliva and vaginal secretion and menstrual blood samples for both markers. This suggests that hsa-miR-203a-3p and hsa-miR-205-5p might rather be markers for the presence of epithelial (or mucosal cells) in general, than for saliva. Another result of interest was the considerable difference in expression levels between venous blood and menstrual blood samples for both markers.

Hsa-miR-891a-5p, a candidate marker for the identification of semen, was exclusively found in the target body fluid, both in pooled and individual samples, thereby confirming previous findings indicating semen specificity [11,14]. Also, hsa-miR-10a-5p, hsa-miR-10b-5p and hsa-miR-135a-5p displayed distinctly varying expression levels between semen and the other body fluids investigated, as indicated by $\Delta\Delta C_q$-values >3 in the pooled samples and confirmed by non-overlapping expression values in the individual sample set.

Of the three remaining candidate markers for vaginal secretion, hsa-miR-1260b and hsa-miR-4286 were expressed slightly higher in semen than in vaginal secretion and displayed only minor differences in expression between the pooled vaginal secretion and saliva samples. Hsa-miR-124a-3p, however, exhibited its highest expression in vaginal secretion and $\Delta\Delta C_q$-values of approximately 2 between vaginal secretion and semen and saliva, respectively. The evaluation in individual samples revealed, that expression levels of hsa-miR-124a-3p in vaginal secretion overlapped with those in semen, menstrual blood and even an outlier venous blood sample, but notably, almost but not yet overlapping expression levels between vaginal secretion and saliva samples.

While hsa-miR-891a-5p presented as a truly semen specific marker within the five body fluids investigated and hsa-miR-144-3p allowed the unambiguous differentiation between blood (i.e. venous blood and menstrual blood) and non-blood (i.e. saliva, semen and vaginal secretion) samples, the distinction between venous blood and menstrual blood as well as saliva and vaginal secretion proved to be more challenging.

We hypothesized that, for single source samples, detection of each of the five body fluids could be achieved by an algorithm of decisions, explained hereinafter, that for simplicity's sake includes as few markers as possible: In a first step, semen samples are differentiated from the remaining body fluid specimens based on specific amplification results for hsa-miR-891a-5p, that occur exclusively in semen. Separation of blood samples from saliva and vaginal secretion specimens is then achieved on the basis of specific amplification results for hsa-miR-144-3p occurring exclusively in samples containing blood in general. For the distinction between venous blood and menstrual blood, combining the results for hsa-miR-144-3p and hsa-miR-203a-3p appeared to be promising and likewise the combination of hsa-miR-203a-3p and hsa-miR-124a-3p to differentiate between saliva and vaginal secretion.

For this purpose we partially extended the final set of results to encompass nine samples per body fluid type with expression results for all four miRNAs and 15 samples per body fluid to be distinguished by the respective miRNA (i.e. for hsa-miR-203a-3p 15 samples each were analyzed for venous blood, menstrual blood, saliva and vaginal secretion, since this marker was applied to both differentiations venous/menstrual blood and saliva/vaginal secretion) (Fig. 2).

The exclusivity of expression in their respective target body fluids rendered statistical analyses for the markers hsa-miR-891a-5p and hsa-miR-144-3p unnecessary. Discriminant function analysis as a means of predicting group affiliations of unknown samples was applied to the data subsets with the ΔC_q-values of

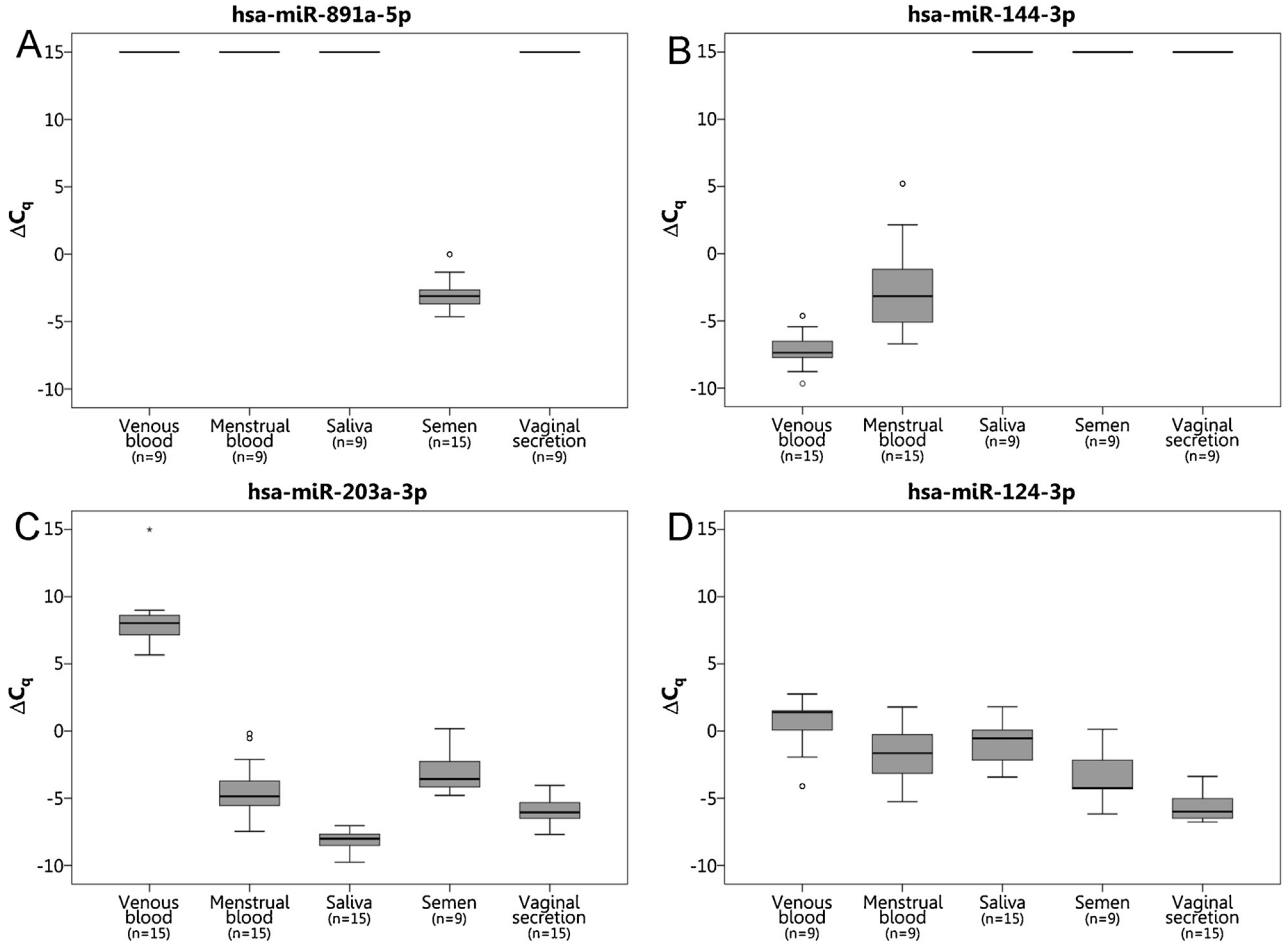


Fig. 2. Expression patterns of the four miRNA markers comprised in the proposed decision process evaluated in individual samples of forensically relevant body fluids. Box blots represent the lower quartile, median and upper quartile of a respective samples set, with the whiskers indicating the (non-extreme) maximum and minimum values, and outliers (1.5× interquartile range) and extreme outliers (3× interquartile range) are depicted as circles and asterisks, respectively. For all samples without specific amplification ΔC_q-values were arbitrarily set to the value 15. Selected from the miRNAs candidates for semen (A), venous blood (B), saliva (C) and vaginal secretion (D).

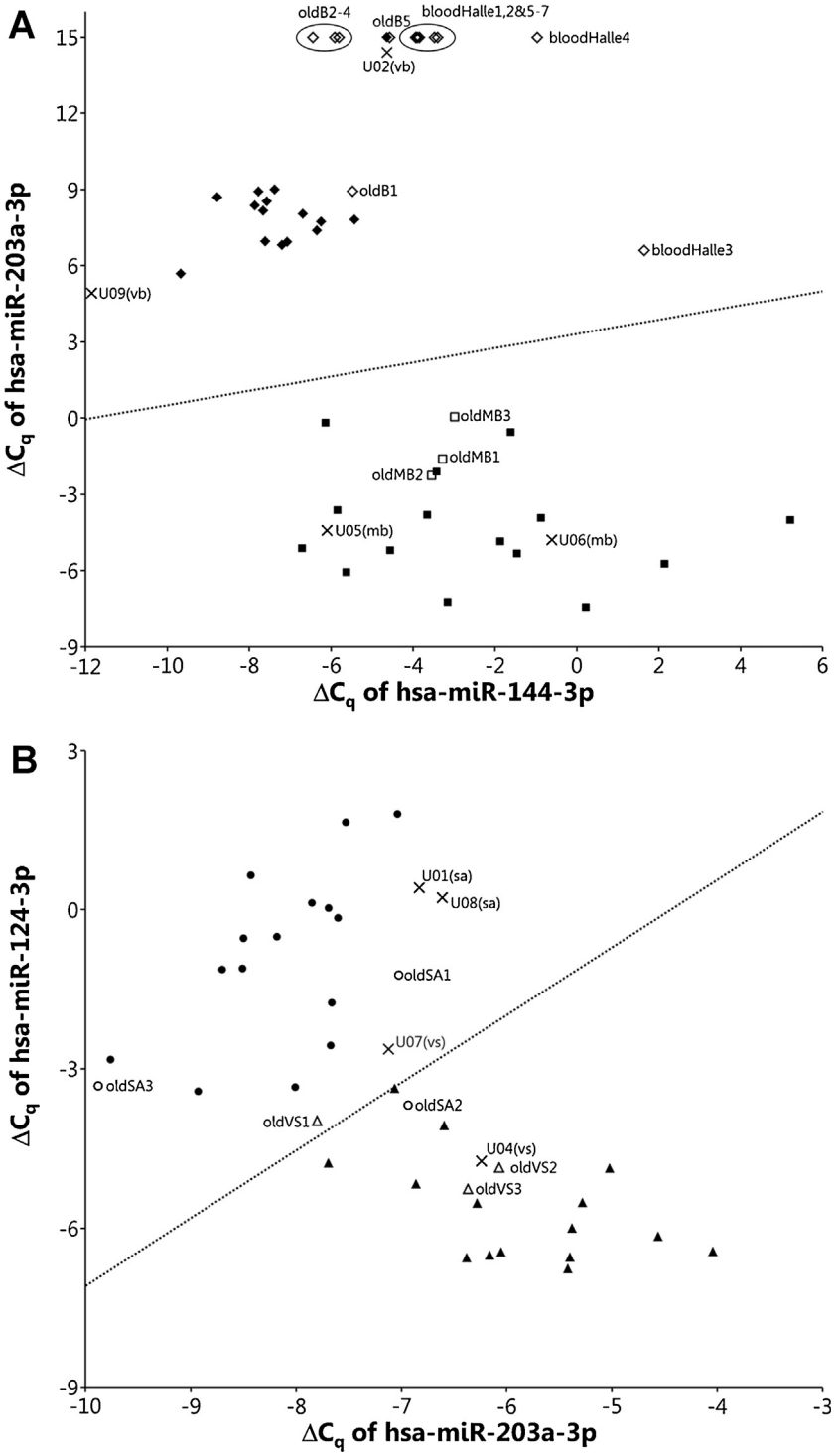


Fig. 3. (A) Scatter plot of the ΔC_q-values of hsa-miR-144-3p (*x*-axis) and hsa-miR-203a-3p (*y*-axis) to distinguish between venous blood samples (closed diamonds, $n = 15$) and menstrual blood samples (closed squares, $n = 15$). Additionally, open diamonds and squares represent aged venous and menstrual blood, respectively, while the *x*' depict blinded venous and menstrual blood specimens. (B) Scatter plot of the ΔC_q-values of hsa-miR-203a-3p (*x*-axis) and hsa-miR-124-3p (*y*-axis) to distinguish between saliva samples (closed circles, $n = 15$) and vaginal secretion specimens (closed triangles, $n = 15$). Additionally, open circles and triangles represent aged saliva and vaginal secretion samples, respectively, while the *x*' depict blinded saliva and vaginal secretion specimens. The dotted lines correspond to the respective decision boundary which is orthogonal to the axis described by the discriminant function.

hsa-miR-203a-3p and hsa-miR-124-3p as independent variables and the groups 'saliva' and 'vaginal secretion' as dependent variables and the ΔC_q-values of hsa-miR-144-3p and hsa-miR-203a-3p as independent variables and the groups 'venous blood' and 'menstrual blood' as dependent variables, respectively.

This resulted in the following term for the separation of saliva and vaginal secretion samples:

$$D = -3.305 - 0.743 \times \Delta C_q(\text{hsa} - \text{miR} - 203\text{a} - 3\text{p}) + 0.582 \times \Delta C_q(\text{hsa} - \text{miR} - 124 - 3\text{p}).$$

A sample is classified as saliva if $D > 0$ and as vaginal secretion if $D < 0$.

The corresponding term for the distinction between venous blood and menstrual blood specimens was:

$$D = -1.504 - 0.127 \times \Delta C_q(\text{hsa} - \text{miR} - 144 - 3\text{p}) + 0.454 \times \Delta C_q(\text{hsa} - \text{miR} - 203\text{a} - 3\text{p}).$$

A sample is classified as venous blood if $D > 0$ and as menstrual blood if $D < 0$.

Both functions classified 100% of the respective data subset correctly. (For detailed information on the discriminant function analysis, please refer to the Supplementary material.)

3.5. Testing of the hypothesized decision algorithm with specimens blinded to the researcher

To test the proposed series of decisions to identify forensically relevant body fluids, a set of total RNA extracts (two samples per body fluid type) was compiled, blinded, then committed to the analyst performing the RT-qPCR experiments and finally the R_n-values were passed to the researcher conducting the data analyses (Supplementary Tables 2 and 5).

Specific amplification of hsa-miR-891a-5p was found in two specimens. Consequently, these were classified as 'semen'. Four samples exhibited specific amplification of hsa-miR-144-3p and were therefore identified as 'blood in general'. To distinguish between venous blood and menstrual blood, the respective discriminant function was applied and yielded positive discriminant scores for samples U02 and U09 ($D = 5.627$ and $D = 2.233$, respectively) and negative values for samples U05 and U06 ($D = -2.741$ and $D = -3.606$, respectively). Thus, the previous two were classified as 'venous blood' and the latter two as 'menstrual blood' (Fig. 3A). Due to the lack of specific amplification of both hsa-miR-891a-5p and hsa-miR-144-3p, the remaining samples were presumed to be either saliva or vaginal secretion specimens. Implementation of the obtained ΔC_q-values into the corresponding term resulted in the determination of samples U01, U07 and U08 as 'saliva' ($D = 2.014$,$D = 0.455$ and $D = 1.740$, respectively) while sample U04 ($D = -1.427$) was identified as 'vaginal secretion' (Fig. 3B).

Disclosing the actual body fluid type of the blinded specimens, it became apparent that nine out of ten samples were identified correctly. Sole exception was U07, a vaginal secretion sample erroneously classified as 'saliva'.

3.6. Evaluation of expression stability in aged samples

All analyzed stains stored for up to 36 years not only yielded analyzable results (Table 1 and Supplementary Table 5), thereby confirming the suitability of miRNA markers for the analysis of aged forensic material, but further corroborated the utility of the devised decision algorithm. These are to our knowledge the oldest samples investigated so far in terms of miRNA expression.

Hsa-miR-891a-5p was detected in all four aged semen samples employed in this study, even in one sample with a very low RNA concentration (oldSE4). However, in the two samples aged for over three years, the reference genes dropped out.

All blood samples exhibited expression of hsa-miR-144-3p, while all non-blood samples were devoid of any specific amplification for this marker. Furthermore, all blood samples were correctly classified as 'venous blood' or 'menstrual blood' (Fig. 3A). While the expression levels of hsa-miR-144-3p were only slightly lower in venous blood samples dating back as far as 1979 compared to those extracted after 24 h, hsa-miR-203a-3p showed no specific amplification in ten out of eleven venous blood samples stored for more than one year. This loss of amplification of hsa-miR-203a-3p appears to occur exclusively in venous blood samples as menstrual blood samples aged for comparable periods of time (oldB2–5 and oldMB1–3) clustered well with the samples used for marker evaluation. Markedly deviant results were obtained for venous blood samples collected in 1980. We speculate, that the substrate on which these stains were stored (cotton fabric dyed green) acted as an inhibitor or promoted degradation. However, even these samples were classified correctly by the discriminant function analysis.

In all, saliva and vaginal secretion samples lacked specific amplification of both hsa-miR-891a-5p and hsa-miR-144-3p which is in line with previous results. Application of discriminant function analysis, however, misclassified one samples for both body fluids (Fig. 3B).

Combining the results of the blinded and aged samples and the sample set used in the initial evaluation, an unambiguous identification of the body fluid types 'semen', 'venous blood' and 'menstrual blood' can be ascertained for single source sample by analyzing the miRNAs hsa-miR-891a-5p, hsa-miR-144-3p and hsa-miR-203a-3p as described above. The lack of specific amplification of hsa-miR-891a-5p and hsa-miR-144-3p further determines a sample to be either saliva or vaginal secretion. Yet, the clear separation between these two body fluids remains challenging, with two vaginal secretion samples being erroneously classified as 'saliva' and one saliva sample being wrongly identified as 'vaginal secretion'. An extension of the sample cohort and subsequent adjustment of the discriminant function analysis might enhance the separation properties of this marker combination, while a certain overlap will most likely remain as the two body fluids both consist of mucosal cells and are consequently quite similar. An additional option would be to include markers in the signature that stem from vaginal secretion specific microbiota [37].

3.7. Mixtures

Due to the lack of truly evidential markers for the identification of menstrual blood, saliva and vaginal secretion, and the overlapping expression levels of hsa-miR-203a-3p and hsa-miR-124-3p in all five body fluids, the decision algorithm described above is not fully applicable to mixtures comprising different body fluids. Thus, our analyses of mixture samples aimed primarily at the designation of the presence or absence of semen and blood in general in a number of mixtures mimicking evidential stains that are characteristic for sexual assaults (Table 2 and Supplementary Table 5).

Again, hsa-miR-891a-5p proved to be a well suited and robust marker for the detection of semen. Specific amplification for this marker was not only exclusively found in mixtures containing semen, thus not giving false positive results, but was also found in mixtures with only small amounts of semen present relative to the other body fluids, thereby attesting a reasonable sensitivity.

The sample comprising saliva and vaginal secretion was devoid of amplification for hsa-miR-891a-5p and hsa-miR-144-3p, thus the absence of semen and blood was determined correctly and the sample could be assumed to be of mucosal origin.

While, with all blood samples displaying strong expression of hsa-miR-144-3p, no false negatives were observed, two non-blood samples also exhibited slight expression of this marker. These false positive findings may result from slight contamination of a sample with blood during the sampling procedure (e.g. high pressure while performing the buccal swab). It is however more likely, that hsa-miR-144-3p is expressed at low levels in non-blood samples adding up to an expression slightly above the applied threshold.

4. Conclusion

Herein, we present the as yet most comprehensive study of forensic body fluid identification via miRNA expression analysis based on a thoroughly validated, state-of-the-art qPCR procedure and unbiased statistical decision making.

With hsa-miR-891a-5p we confirmed a truly semen specific marker among the five forensically relevant body fluids investigated. Further, hsa-miR-10a-5p, hsa-miR-10b-5p and hsa-miR-135-5p proved to be overexpressed in semen samples compared to venous blood, menstrual blood, saliva and vaginal secretion specimens, rendering them additional suitable markers. Hsa-miR-144-3p as well as hsa-miR-144-5p and hsa-miR-451a exhibited promising expression patterns for the separation of blood samples from non-blood samples. Truly unambiguous miRNA markers for either venous blood or menstrual blood could, however, not be confirmed. Also, the determination of the body fluids saliva and vaginal secretion proved difficult, as no definite markers were identified.

With the given set of markers, a decision algorithm was proposed to detect each of the five body fluids employing as few markers as possible to simplify the analysis procedure: hsa-miR-891a-5p was used for the identification of semen, hsa-miR-144-3p for separating blood from non-blood samples and then a combination of hsa-miR-144-3p and hsa-miR-203a-3p to distinguish between venous blood and menstrual blood samples and likewise the combination of hsa-miR-203a-3p and hsa-miR-124-3p to differentiate between saliva and vaginal secretion specimens.

While semen and blood in general could be determined with high certainty, differentiation of venous blood and menstrual blood as well as saliva and vaginal secretion remained most challenging, especially in mixtures. Further research is needed to identify a set of unambiguous markers or marker signatures for all forensically relevant body fluids. While a few markers have been reported by several groups, there are still inconsistencies in the findings that may be due to different approaches and methodologies. In this regard, miRNA based body fluid identification would, in our opinion, benefit greatly from inter-laboratory evaluation trials (similar to the EDNAP exercise). Also, the analysis of miRNA expression patterns in body fluid mixtures requires further comprehensive and detailed investigation, as miRNAs are, in most cases, not restricted to one body fluid.

Great advances in forensic body fluid identification are to be expected of the implication of the massive parallel sequencing technology in routine casework, for example by combining miRNAs, mRNAs, bacterial markers and methylation patterns in highly specific signatures.

Acknowledgements

We would like to thank all volunteers for their kind donation of sample material and Dr. Uta Immel from the Institute of Legal Medicine, Halle (Saale), Germany for committing a collective of aged blood samples to our disposal. We further thank Dr. Rolf Fimmers for the statistical consultations.

We thank the Deutsche Forschungsgemeinschaft (DFG) for funding this project (CO 992/3-1).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2016.01.018.

References

[1] K. Virkler, I.K. Lednev, Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene, Forensic Sci. Int. 188 (1–3) (2009) 1–17.
[2] J.H. An, K.J. Shin, W.I. Yang, H.Y. Lee, Body fluid identification in forensics, BMB Rep. 45 (10) (2012) 545–553.
[3] T. Sijen, Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers, Forensic Sci. Int. Genet. 18 (2015) 21–32.
[4] P. Sood, A. Krek, M. Zavolan, G. Macino, N. Rajewsky, Cell-type-specific signatures of microRNAs on target mRNA expression, Proc. Natl. Acad. Sci. U. S. A. 103 (8) (2006) 2746–2751.
[5] P. Landgraf, M. Rusu, R. Sheridan, A. Sewer, N. Iovino, A. Aravin, S. Pfeffer, A. Rice, A.O. Kamphorst, M. Landthaler, C. Lin, N.D. Socci, L. Hermida, V. Fulci, S. Chiaretti, R. Foa, J. Schliwka, U. Fuchs, A. Novosel, R.U. Muller, B. Schermer, U. Bissels, J. Inman, Q. Phan, M. Chien, D.B. Weir, R. Choksi, V.G. De, D. Frezzetti, H. I. Trompeter, V. Hornung, G. Teng, G. Hartmann, M. Palkovits, L.R. Di, P. Wernet, G. Macino, C.E. Rogler, J.W. Nagle, J. Ju, F.N. Papavasiliou, T. Benzing, P. Lichter, W. Tam, M.J. Brownstein, A. Bosio, A. Borkhardt, J.J. Russo, C. Sander, M. Zavolan, T. Tuschl, A mammalian microRNA expression atlas based on small RNA library sequencing, Cell 129 (7) (2007) 1401–1414.
[6] Y. Liang, D. Ridzon, L. Wong, C. Chen, Characterization of microRNA expression profiles in normal human tissues, BMC Genomics 8 (2007) 166.
[7] 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 (4) (2013) 967–971.
[8] E.J. Omelia, M.L. Uchimoto, G. Williams, Quantitative PCR analysis of blood- and saliva-specific microRNA markers following solid-phase DNA extraction, Anal. Biochem. 435 (2) (2013) 120–122.
[9] 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.
[10] E.K. Hanson, H. Lubenow, J. Ballantyne, Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs, Anal. Biochem. 387 (2) (2009) 303–314.
[11] 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 (3) (2010) 217–226.
[12] C. Courts, B. Madea, Specific micro-RNA signatures for the detection of saliva and blood in forensic body-fluid identification, J. Forensic Sci. 56 (6) (2011) 1464–1470.
[13] Z. Wang, H. Luo, X. Pan, M. Liao, Y. Hou, A model for data analysis of microRNA expression in forensic body fluid identification, Forensic Sci. Int. Genet. 6 (3) (2012) 419–423.
[14] 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.
[15] E.K. Hanson, K. Rekab, J. Ballantyne, Binary logistic regression models enable miRNA profiling to provide accurate identification of forensically relevant body fluids and tissues, Forensic Sci Int. Genet. Suppl. Ser. 4 (1) (2013) e127–e128.
[16] E.K. Hanson, M. Mirza, K. Rekab, J. Ballantyne, The identification of menstrual blood in forensic samples by logistic regression modeling of miRNA expression, Electrophoresis 35 (21–22) (2014) 3087–3095.
[17] J.L. Park, S.M. Park, O.H. Kwon, H.C. Lee, J.Y. Kim, H.H. Seok, W.S. Lee, S.H. Lee, Y. S. Kim, K.M. Woo, S.Y. Kim, Microarray screening and qRT-PCR evaluation of microRNA markers for forensic body fluid identification, Electrophoresis 35 (21–22) (2014) 3062–3068.
[18] Z. Wang, J. Zhang, W. Wei, D. Zhou, H. Luo, X. Chen, Y. Hou, Identification of saliva using microRNA biomarkers for forensic purpose, J. Forensic Sci. 3 (2015) 702–706.
[19] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Res. 29 (9) (2001) e45.
[20] 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 (9) (2002) e36.
[21] S.A. Bustin, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems, J. Mol. Endocrinol. 29 (1) (2002) 23–39.
[22] S. Bustin, T. Nolan, Data analysis and interpretation, in: S. Bustin (Ed.), A–Z of Quantitative PCR, 1st (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 (4) (2001) 402–408.
[24] J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. van Roy, A. De Paepe, F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, Genome Biol. 3 (7) (2002) RESEARCH0034.
[25] 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.
[26] A. Schroeder, O. Mueller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, S. Lightfoot, W. Menzel, M. Granzow, T. Ragg, The RIN: an RNA integrity number for assigning integrity values to RNA measurements, BMC Mol. Biol. 7 (2006) 3.
[27] S. Griffiths-Jones, The microRNA registry, Nucleic Acids Res. 32 (Database issue) (2004) D109–D111.
[28] M.E. Ritchie, J. Silver, A. Oshlack, M. Holmes, D. Diyagama, A. Holloway, G.K. Smyth, A comparison of background correction methods for two-colour microarrays, Bioinformatics 23 (20) (2007) 2700–2707.
[29] J.M. Ruijter, C. Ramakers, W.M. Hoogaars, Y. Karlen, O. Bakker, M.J. van den Hoff, A.F. Moorman, Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data, Nucleic Acids Res. 37 (6) (2009) e45.
[30] S. Pabinger, S. Rödiger, A. Kriegner, K. Vierlinger, A. Weinhäusel, A survey of tools for the analysis of quantitative PCR (qPCR) data, Biomol. Detect. Quantif. 1 (1) (2014) 22–33.
[31] J.L. Guthrie, C. Seah, S. Brown, P. Tang, F. Jamieson, S.J. Drews, Use of Bordetella pertussis BP3385 to establish a cutoff value for an IS481-targeted real-time PCR assay, J. Clin. Microbiol. 46 (11) (2008) 3798–3799.
[32] S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, C.T. Wittwer, The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments, Clin. Chem. 55 (4) (2009) 611–622.
[33] M. Jung, A. Schaefer, I. Steiner, C. Kempkensteffen, C. Stephan, A. Erbersdobler, K. Jung, Robust microRNA stability in degraded RNA preparations from human tissue and cell samples, Clin. Chem. 56 (6) (2010) 998–1006.
[34] J.S. Hall, J. Taylor, H.R. Valentine, J.J. Irlam, A. Eustace, P.J. Hoskin, C.J. Miller, C. M. West, Enhanced stability of microRNA expression facilitates classification of FFPE tumour samples exhibiting near total mRNA degradation, Br. J. Cancer 107 (4) (2012) 684–694.
[35] C. Becker, A. Hammerle-Fickinger, I. Riedmaier, M.W. Pfaffl, mRNA and microRNA quality control for RT-qPCR analysis, Methods 50 (4) (2010) 237–243.
[36] 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 (19) (2006) 1601–1613.
[37] S. Giampaoli, A. Berti, F. Valeriani, G. Gianfranceschi, A. Piccolella, L. Buggiotti, C. Rapone, A. Valentini, L. Ripani, S.V. Romano, Molecular identification of vaginal fluid by microbial signature, Forensic Sci. Int. Genet. 6 (5) (2012) 559–564.

SUPPLEMENTARY TABLE 1: Total RNA concentration and integrity per sample used in the validation of candidate miRNAs.

Body fluid	Sample name	Gender	Total RNA concentration	RIN	Additional information
Venous blood	B1	M	4.7	1	
	B2[P]	F	4.9	n.d.	
	B3	F	4.9	1	
	B4[P]	M	4.9	1.1	
	B5[P]	F	5.8	n.d.	
	B6	F	6.3	1	
	B7[P]	F	7.1	n.d.	
	B8	F	7.7	n.d.	
	B9[P]	M	7.9	n.d.	
	B10	F	8.1	1	
	B11[P]	F	8.6	1	
	B1[P]	M	8.9	1.1	
	B12[MA]	M	11.8	1.1	
	B13[MA]	M	12.0	n.d.	
	B14[P]	M	12.3	1	
	B15[MA]	M	13.2	1.1	
	B16[MA]	F	15.3	1	
	B17[P]	M	16.3	1	
	B18[P]	F	20.2	1	
	B19[MA]	F	20.3	2.4	
Menstrual blood	MB1	F	4.0	n.d.	Not specified
	MB2	F	8.6	1.4	Day 4 of menstruation
	MB3[P.MA]	F	15.3	1.1	Not specified
	MB4	F	15.6	n.d.	Day 4 of menstruation
	MB5	F	18.6	1.7	Day 1 of menstruation
	MB6[MA]	F	20.4	2.5	Day 3 of menstruation
	MB7	F	20.5	2.5	Not specified
	MB8[MA]	F	32.9	2.4	Day 4 of menstruation
	MB9[P]	F	33.8	2.2	Day 2 of menstruation
	MB10[P]	F	50.6	2.6	Not specified
	MB11[P.MA]	F	56.0	2.5	Not specified
	MB12	F	68.5	2.5	Day 2 of menstruation
	MB13	F	100.0	2.5	Day 3 of menstruation
	MB14[MA]	F	141.0	2.6	Day 3 of menstruation
	MB14.2	F	150.0	2.6	"
	MB15	F	150.0	2.5	Day 1 of menstruation
	MB16[P]	F	201.0	1.2	Day 3 of menstruation
Saliva	SA1[MA]	F	4.9	n.d.	
	SA2[P]	F	5.5	1	
	SA3	F	6.5	1	
	SA4[P]	F	7.5	n.d.	
	SA5	M	8.2	1	
	SA6[P]	M	9.8	1	
	SA7[P]	M	9.8	n.d.	
	SA8[MA]	F	10.6	n.d.	
	SA9[P]	F	11.6	1	

(Supplementary Table 1 continued)

Body fluid	Sample name	Gender	Total RNA concentration	RIN	Additional information
	SA10	M	12.5	1.2	
	SA5.2^{MA}	M	13.7	1.1	
	SA11^{M}	M	15.6	1.2	
	SA12^{P}	M	16.9	n.d.	
	SA13^{MA}	M	17.8	1.1	
	SA14	M	18.0	1	
	SA15^{P}	F	18.5	2.3	
	SA16	F	21.5	1	
	SA117^{P}	F	37.5	2.4	
	SA18^{P}	M	42.5	6.6	
	SA19^{P}	F	77.0	2.8	
Semen	SE1	M	4.4	n.d.	
	SE2	M	6.3	1	
	SE3^{MA}	M	6.8	n.d.	
	SE4	M	7.5	1.2	
	SE5	M	9.4	1.4	
	SE6^{P}	M	9.5	1	
	SE7	M	11.0	1	
	SE8^{MA}	M	11.7	1	
	SE9^{MA}	M	12.1	1	
	SE10^{MA}	M	12.4	n.d.	
	S11^{P}	M	12.8	n.d.	
	SE12^{P}	M	15.4	1.1	
	SE13	M	15.9	1	
	SE14$^{P.MA}$	M	22.7	1.4	
	SE15	M	27.5	1.2	
	SE16^{P}	M	39.0	2.3	
Vaginal secretion	VS1^{P}	F	39.4	4.8	In menopause
	VS2	F	41.5	2.5	Day 6 after menstruation
	VS3	F	49.5	2.6	Day 6 after menstruation
	VS4^{MA}	F	56.5	2.7	Day 4 after menstruation
	VS5	F	56.7	3.1	Day 10 after
	VS6^{P}	F	62.0	2.7	Day 15 after
	VS7	F	69.6	3.5	Day 12 after
	VS8^{P}	F	95.2	3.3	Day 10 after
	VS9	F	95.3	3.2	Day 8 after menstruation
	VS10^{MA}	F	130.0	2.8	Not specified
	VS11	F	140.0	2.6	Day 16 after
	VS12^{MA}	F	142.0	3.4	Not specified
	VS13^{MA}	F	160.0	3.1	Day 10 after
	VS14^{P}	F	162.0	6.9	Day 21 after
	VS15	F	218.5	2.5	Day 18 after
	VS16	F	263.0	2.9	Day 9 after menstruation
	VS17	F	315.0	3	Day 4 after menstruation
	VS18^{MA}	F	344.0	3.4	Day 18 after
	VS19^{P}	F	348.0	3.1	Day 4 after menstruation

RIN, RNA integrity number; F, female; M, male; ^{P}, sample used in pooled body fluid samples; ^{MA}, sample used for Micro-Array analysis; n.d., not detectable.

SUPPLEMENTARY TABLE 2: Total RNA concentration and integrity per blinded sample and correct classification (unknown to the researcher performing the analyses).

Sample name	Total RNA concentration (ng\µl)	RIN	Body fluid	Gender
U01	7.2	1.6	Saliva	M
U02	12.7	1	Venous blood	M
U03	6.8	n.d.	Semen	M
U04	110.0	3.3	Vaginal secretion	F
U05	32.3	2.5	Menstrual blood	F
U06	131.0	2.6	Menstrual blood	F
U07	176.0	2.6	Vaginal secretion	F
U08	38.5	5.1	Saliva	F
U09	4.2	n.d.	Venous blood	F
U10	14.2	1	Semen	M

RIN, RNA integrity number; n.d., not detectable; F, female; M, male.

SUPPLEMENTARY TABLE 3: Specifications of the 36 selected candidate miRNAs for body fluid specific signatures and an overview of the miRNA markers previously reported in the forensic literature.
References are classified according to the respective method used for validation of body fluid specific miRNAs. Microarray average signal intensities apply to the dataset in which quantile normalization was performed separately per body fluid group.
Target sequences and TagMan assay IDs are given for the miRNAs selected for validation in qPCR only.

		References				Microarray Average Signal								
Body Fluid	**Official Gene Symbol**	**TaqMan assays**	**SYBR-Green**	**Endpoint-PCR**	**Count**	**VB**	**MB**	**SA**	**SE**	**VS**	**miRBase-/ NCBI-Mature Sequence Accession**	**Previous miRBase-IDs / NCBI-Aliases**	**TaqMan Assay ID**	**Target Sequence (amplicon length)**
Venous blood / blood in general	**hsa-miR-16-5p**	[13], [14]	[10], [15]		4	++	++	o	-	.	MIMAT0000069	hsa-miR-16	000391	UAGCAGCACGUAAAUAUUGGCG (22)
	hsa-miR-20a-5p	[11]			1	+	+	-	.	.	MIMAT0000075	hsa-miR-20, hsa-miR-20a	000580	UAAAGUGCUUAUAGUGCAGGUAG (23)
	hsa-miR-22-5p				-	o	.	.	.	.	MIMAT0004495	hsa-miR-22*	002301	AGUUCUUCAGUGGCAAGCUUUA (22)
	hsa-miR-106a-5p	[11]			1	+	.	.	.	.	MIMAT0000103	hsa-miR-106a	002169	AAAAGUGCUUACAGUGCAGGUAG (23)
	hsa-miR-126-3p		[12]		1	+	.	.	.	.	MIMAT0000445	hsa-miR-126	002228	UCGUACCGUGAGUAAUAAUGCG (22)
	hsa-miR-144-3p	[11]	[15]*, [16]*		1 (3)	++	.	.	.	.	MIMAT0000436	hsa-miR-144	002676	UACAGUAUAGAUGAUGUACU (20)
	hsa-miR-150-5p		[12]		1	-	-	-	.	o	MIMAT0000451	hsa-miR-150		
	hsa-miR-182-5p		[17]		1	o	.	.	.	.	MIMAT0000259	hsa-miR-182		
	hsa-miR-451a	[8]	[10], [12]	[7], [9]	6	++	++	.	.	.	MIMAT0001631	hsa-miR-451	001141	AAACCGUUACCAUUACUGAGUU (22)
	hsa-miR-484		[15]		1	-	-	.	.	.	MIMAT0002174			
	hsa-miR-486-5p	[14], [18]	[17]		2	+	+	.	.	.	MIMAT0002177	hsa-miR-486		
Menstrual blood	hsa-miR-142-3p		[15]		1	++	++	o	.	.	MIMAT0000434			
	hsa-miR-144-5p		[15], [16]		2	+	.	.	.	.	MIMAT0004600	hsa-miR-144*	002148	GGAUAUCAUCAUAUACUGUAAG (22)
	hsa-miR-185-5p	[11]**	[15], [16]		2 (3)	+	+	.	.	.	MIMAT0000455	hsa-miR-185	002271	UGGAGAGAAAGGCAGUUCCUGA (22)
	hsa-miR-214-3p	[14], [18]		[9]	3	.	.	.	.	.	MIMAT0000271	hsa-miR-214	002306	ACAGCAGGCACAGACAGGCAGU (22)
	hsa-miR-371b-5p				-	o	++	++	++	++	MIMAT0019892		463886_mat	ACUCAAAAGAUGGCGGCACUUU (22)
	hsa-miR-412-3p		[10]		1	.	.	.	.	.	MIMAT0002170			
	hsa-miR-675-5p				-	.	+	+	+	o	MIMAT0004284	hsa-miR-675	002005	UGGUGCGGAGAGGGCCCACAGUG (23)
	hsa-miR-2116-5p				-	.	+	-	.	o	MIMAT0011160	hsa-miR-2116	241122_mat	GGUUCUUAGCAUAGGAGGUCU (21)
	hsa-miR-4444				-	.	-	.	.	.	MIMAT0018962		465074_mat	CUCGAGUUGGAAGAGGCG (18)
Saliva	**hsa-miR-23a-3p**				-	-	o	o	-	.	MIMAT0000078	hsa-miR-23a	000399	AUCACAUUGCCAGGGAUUUCC (21)
	hsa-miR-34a-5p				-	.	.	.	-	.	MIMAT0000255	hsa-miR-34a	000426	UGGCAGUGUCUUAGCUGGUUGU (22)
	hsa-miR-124-5p		[15]		1	.	.	.	.	.	MIMAT0004591	hsa-miR-124*		
	hsa-miR-145-5p		[17]		1	.	.	.	.	.	MIMAT0000437	hsa-miR-145		
	hsa-miR-200c-3p	[18]	[12]		2	.	-	o	o	-	MIMAT0000617	hsa-miR-200c		
	hsa-miR-203a-3p	[18]	[12]		2	.	o	+	.	o	MIMAT0000264	hsa-miR-203a	000507	GUGAAAUGUUUAGGACCACUAG (22)
	hsa-miR-205-5p	[8], [18]	[10], [12]	[7]	6	.	o	+	-	o	MIMAT0000266	hsa-miR-205	000509	UCCUUCAUUCCACCGGAGUCUG (22)
	hsa-miR-223-3p		[17]		1	o	+	o	.	-	MIMAT0000280	hsa-miR-223		
	hsa-miR-658		[10], [15]		2	.	.	-	.	.	MIMAT0003336		001513	GGCGGAGGGAAGUAGGUCCGUUGGU (25)
	hsa-miR-1290				-	.	o	+	-	-	MIMAT0005880		002863	UGGAUUUUUGGAUCAGGGA (19)
	hsa-miR-4732-5p				-	.	o	+	o	-	MIMAT0019855		465097_mat	UGUAGAGCAGGGAGCAGGAAGCU (23)
Semen	**hsa-miR-10a-5p**	[11]			1	.	.	.	-	.	MIMAT0000253	hsa-miR-10a	000387	UACCCUGUAGAUCCGAAUUUGUG (23)
	hsa-miR-10b-5p		[10]		1	.	.	.	-	.	MIMAT0000254	hsa-miR-10b	002218	UACCCUGUAGAACCGAAUUUGUG (23)
	hsa-miR-135a-5p	[11]			1	.	.	.	o	.	MIMAT0000428	hsa-miR-135a	000460	UAUGGCUUUUUAUUCCUAUGUGA (23)
	hsa-miR-135b-5p		[10]		1	.	.	.	.	.	MIMAT0000758	hsa-miR-135b		
	hsa-miR-507	[11]			1	.	.	.	.	.	MIMAT0002879			
	hsa-miR-585-3p				-	.	o	-	+	o	MIMAT0003250		001625	UGGGCGUAUCUGUAUGCUA (19)
	hsa-miR-888-5p	[14]		[9]	2	.	.	.	.	.	MIMAT0004916	hsa-miR-888		
	hsa-miR-891a-5p	[11], [14]		[9]	4	.	.	.	.	-	MIMAT0004902		002191	UGCAACGAACCUGAGCCACUGA (22)
	hsa-miR-891b		[15]		1	.	.	.	.	.	MIMAT0004913			
	hsa-miR-892a		[15]		1	.	.	.	.	.	MIMAT0004907			
	hsa-miR-943	[11]			1	.	-	-	o	o	MIMAT0004986			
	hsa-miR-1260a				-	.	-	-	+	o	MIMAT0005911	hsa-miR-1260	002896	AUCCCACCUCUGCCACCA (18)
	hsa-miR-2392		[17]		1	.	.	.	.	.	MIMAT0019043			
	hsa-miR-3197		[17]		1	.	.	.	.	.	MIMAT0015082			

(Supplementary Table 3 continued)

Body Fluid	Official Gene Symbol	References assays	References SYBR-Green PCR	Count	Microarray Average Signal VB	MB	SA	SE	VS	Mature Sequence Accession	Previous miRBase-IDs / NCBI-Aliases	TaqMan Assay ID	Target Sequence (amplicon length)
Vaginal secretion	**hsa-miR-124-3p**		[10], [15]	2	.	-	.	.	.	MIMAT0000422	hsa-miR-124a, hsa-miR-124	001182	UAAGGCACGCGGUGAAUGCC (20)
	hsa-miR-155-5p			-	-	-	.	.	o	MIMAT0000646	hsa-miR-155	002623	UUAAUGCUAAUCGUGAUAGGGGU (23)
	hsa-miR-196a-3p			-	.	o	.	.	o	MIMAT0004562	hsa-miR-196a*	002336	CGGCAACAAGAAACUGCCUGAG (22)
	hsa-miR-372-3p		[10]	1	.	.	.	.	.	MIMAT0000724			
	hsa-miR-654-5p		[17]	1	.	.	.	.	.	MIMAT0003330	hsa-miR-654	001611	UGGUGGGCCGCAGAACAUGUGC (22)
	hsa-miR-1260b		[17]	1	+	+	+	++	++	MIMAT0015041		242525_mat	AUCCCACCACUGCCACCAU (19)
	hsa-miR-1280		[15]	1	.	.	.	.	.	removed***			
	hsa-miR-3685			-	.	o	o	-	+	MIMAT0018113		462769_mat	UUUCCUACCCUACCUGAAGACU (22)
	hsa-miR-4286		[15]	1	-	o	-	+	o	MIMAT0016916		241500_mat	ACCCCACUCCUGGUACC (17)
	hsa-miR-4653-3p			-	.	.	.	.	o	MIMAT0019719		464305_mat	UGGAGUUAAGGGUUGCUUGGAGA (23)
	hsa-miR-4795-3p			-	.	-	.	-	o	MIMAT0019969		462535_mat	AUAUUAUUAGCCACUUCUGGAU (22)
Reference genes	**SNORD24**	[11], [25], [a]								NR_002447	U24, RNU24	001001	AUUUGCUAUCUGAGAGAUGGUGAUGACAUUU
	SNORD38B	[25], [a]								NR_001457	U38B, RNU38B	001004	CCAGUUCUGCUACUGACAGUAAGUGAAGAUA AAGUGUGUCUGAGGAGA (48)
	SNORD43	[25], [a]								NR_002439	U43, RNU43	001095	GAACUUAUUGACGGGCGGACAGAAACUGUGU GCUGAUUGUCACGUUCUGAUU (52)

VB, venous blood; MB, menstrual blood; SA, saliva; SE, semen; VS, vaginal secretion; miRBase, microRNA database [27]; NCBI, National Center for Biotechnology Information; *, as menstrual blood marker; **, as venous blood marker; ++, very high signal intensity; +, high signal intensity; o, medium to low signal intensity; -, very low signal intensity; ., signal not distinguishable from background or marker not detected. ***, hsa-miR-1280 was removed from miRBase since a study by Schopman et al. [b] indicates that the sequence annotated as miR-1280 is likely to be a fragment of a tRNA. The publication by Wang et al. [18] was not available to our group at the time of candidate marker selection.

[a]Wong L, Lee K, Russell I, Chen C. Endogenous Controls for Real-Time Quantitation of miRNA Using TaqMan MicroRNA Assays Application Note (Applied Biosystems). 2010, 1–8 (127AP11-01).

[b] Schopman NC, Heynen S, Haasnoot J, Berkhout B. A miRNA-tRNA mix-up: tRNA origin of proposed miRNA. RNA Biol. 2010;7(5):573-576.

SUPPLEMENTARY TABLE 4: Amplification efficiencies of candidate miRNAs calculated by LinRegPCR software.

		Amplification efficiency of mixture	
Body fluid	Gene Symbol	Mean[a]	SD
Venous blood/blood in general	hsa-miR-16-5p	1.82	0.015
	hsa-miR-20a-5p	1.95	0.011
	hsa-miR-22-5p	1.91	0.016
	hsa-miR-106-5p	2.00	0.018
	hsa-miR-126-3p	1.90	0.020
	hsa-miR-144-3p	1.90	0.028
	hsa-miR-451a	1.85	0.007
Menstrual blood	hsa-miR-144-5p	1.91	0.018
	hsa-miR-185-5p	1.84	0.015
	hsa-miR-214-3p	1.91	0.029
	hsa-miR-371b-5p[b]	1.32	0.045
	hsa-miR-675-5p	1.86	0.032
	hsa-miR-2116-5p[b]	1.53	0.028
	hsa-miR-4444	1.88	0.018
Saliva	hsa-miR-23a-3p	1.93	0.021
	hsa-miR-34a-5p	1.92	0.020
	hsa-miR-203a-3p	1.93	0.028
	hsa-miR-205-5p	1.92	0.017
	hsa-miR-1290	1.96	0.035
	hsa-miR-4732-5p	1.90	0.019
Semen	hsa-miR-10a-5p	1.87	0.017
	hsa-miR-10b-5p	1.92	0.024
	hsa-miR-135a-5p	1.97	0.021
	hsa-miR-891a	1.89	0.022
	hsa-miR-1260a	1.83	0.023
Vaginal secretion	hsa-miR-124a-3p	1.79	0.030
	hsa-miR-155-5p	1.91	0.021
	hsa-miR-654-5p[b]	1.82	0.005
	hsa-miR-1260b	1.84	0.025
	hsa-miR-4286	1.80	0.017
	hsa-miR-4653-3p	2.02	0.025

SD, standard deviation

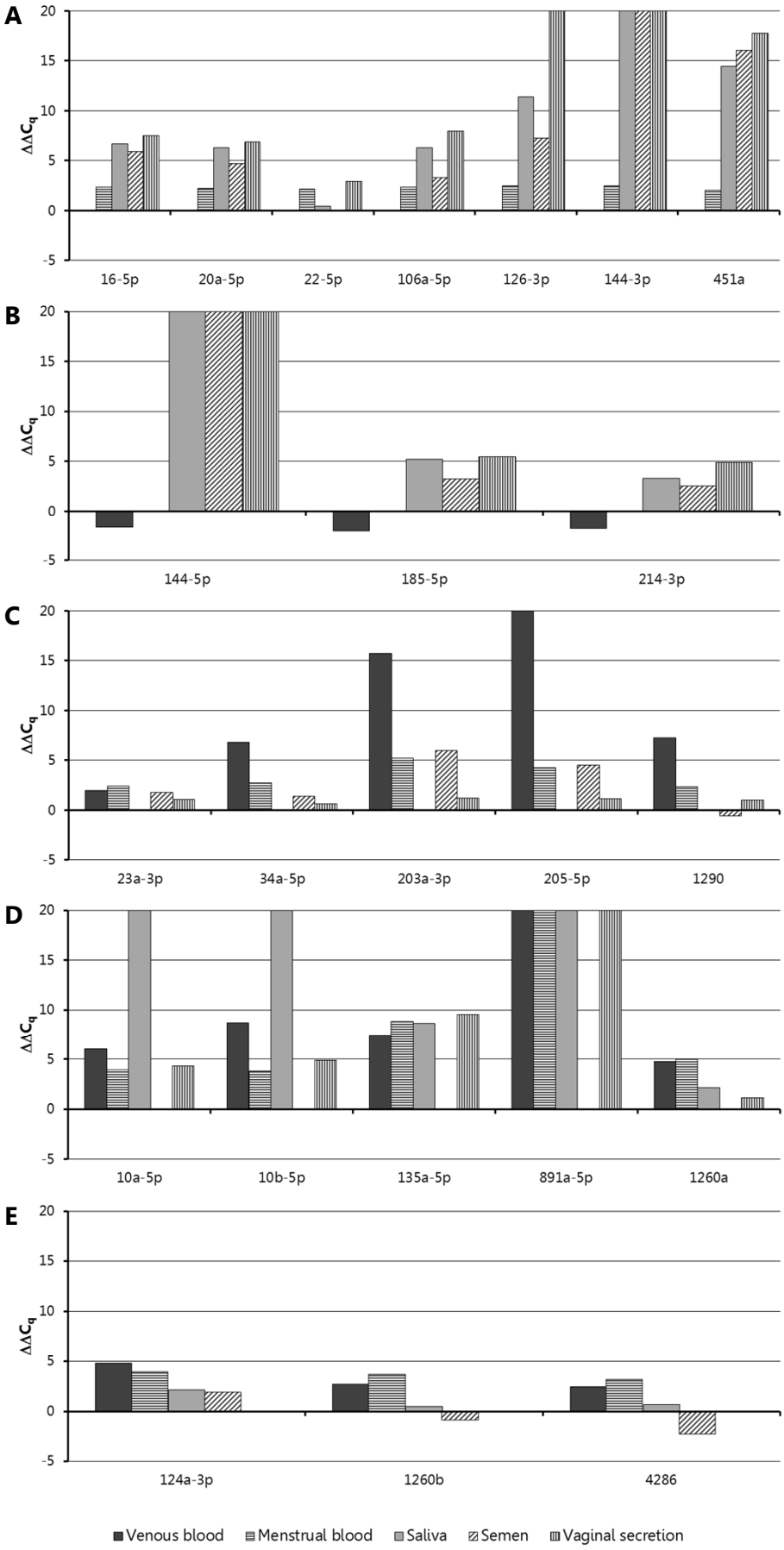
[a] An efficiency value of 1 represents an amplification efficiency of 0 %, while a value of 2 represents an amplification efficiency of 100 %.

[b] Excluded from study.

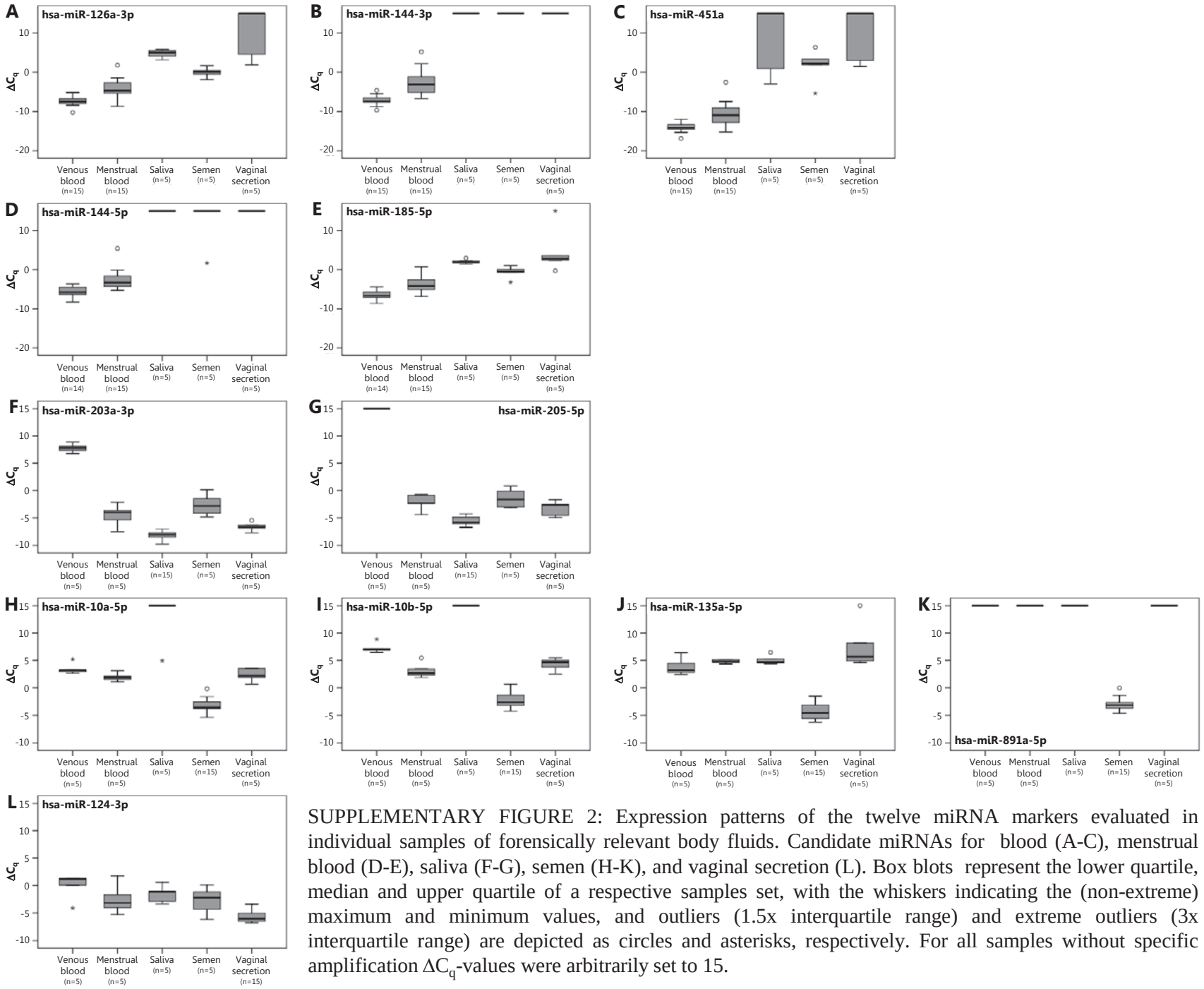
SUPPLEMENTARY TABLE 5: ΔC_q-values of the miRNAs used in the decision series for the blinded specimens (U01-10), aged samples of venous blood, menstrual blood, saliva, semen and vaginal secretion, and in mixed samples (combinations of different body fluids, Mix01-14). The respective body fluid per blinded specimen – not known to the researcher at the time of analysis – is given in parentheses. The included body fluid types per mixture are given in square brackets. For more detailed sample information please refer to Tables 1 & 2 and Supplementary Table 2. For all samples without specific amplification ΔC_q-values were arbitrarily set to 15.

Sample name	hsa-miR891a-5p	hsa-miR-144-3p	hsa-miR-203a-3p	hsa-miR-124-3p
U01 (SA)	15	15	-6.83	0.42
U02 (VB)	15	-4.64	14.41	-0.70
U03 (SE)	-3.53	15	-2.37	-3.40
U04 (VS)	15	15	-6.24	-4.74
U05 (MB)	15	-6.10	-4.43	-0.01
U06 (MB)	15	-0.61	-4.80	-2.36
U07 (VS)	15	15	-7.12	-2.63
U08 (SA)	15	15	-6.61	0.23
U09 (VB)	15	-11.84	4.92	15
U10 (SE)	-4.27	15	-3.38	-2.49
oldB1	15	-5.48	8.94	0.97
oldB2	15	-5.91	15	-0.39
oldB3	15	-6.45	15	0.65
oldB4	15	-5.81	15	0.34
oldB5	15	-4.57	15	0.39
bloodHalle1	15	-3.48	15	-2.86
bloodHalle2	15	-3.84	15	-1.82
bloodHalle3	15	1.65	6.61	-1.50
bloodHalle4	15	-0.97	15	-1.42
bloodHalle5	15	-3.89	15	15
bloodHalle6	15	-3.40	15	15
bloodHalle7	15	-3.94	15	15
oldMB1	15	-3.28	-1.60	-1.51
oldMB2	15	-3.55	-2.24	-1.31
oldMB3	15	-2.99	0.07	1.31
oldSA1	15	15	-7.03	-1.22
oldSA2	15	15	-6.94	-3.68
oldSA3	15	15	-9.88	-3.31
oldSE1	-2.70	15	-3.10	-5.42
oldSE2	-3.98	15	-2.15	-4.94
oldVS1	15	15	-7.80	-3.97
oldVS2	15	15	-6.07	-4.85
oldVS3	15	15	-6.37	-5.26
Mix01 [VB+VS]	15	-7.08	-5.99	-3.80
Mix02 [VB+SE]	-2.58	-10.66	4.38	-3.11
Mix03 [VB+SE]	2.20	-7.87	3.55	-0.60
Mix04 [VB+SA]	15	-7.78	-2.54	0.05
Mix05 [VB+MB]	15	-5.85	-3.34	-1.53
Mix06 [SE+VS]	3.67	15	-6.75	-6.35
Mix07 [SE+VS]	0.31	3.96	-8.80	-4.83
Mix08 [SE+VS]	3.94	15	-7.59	-4.02
Mix09 [SE+MB]	-0.27	-7.05	-4.93	-2.98
Mix10 [SE+MB]	2.74	-6.73	-2.47	-2.29
Mix11[SE+SA]	-0.45	4.81	-6.43	-4.13
Mix12 [SA+VS]	15	15	-6.03	-5.51
Mix13 [SE+VB+VS]	5.81	-6.78	-5.37	-2.90
Mix14 [SE+VB+VS]	4.42	-5.54	-5.54	-6.16

VB, venous blood; MB, menstrual blood; SA, saliva; SE, semen; VS, vaginal secretion. ΔC_q-values of sample/assay combinations without specific amplification were arbitrarily set to 15. No ΔC_q-values are given for oldSE3 and oldSA4, as for these samples the reference genes dropped out.



SUPPLEMENTARY FIGURE 1: Relative expression patterns of candidate miRNAs in pooled samples of forensically relevant body fluids depicted as $\Delta\Delta C_q$-values with the respective targeted body fluid samples used as a calibrator sample. Candidate miRNAs for blood (A), menstrual blood (B), saliva (C), semen (D), and vaginal secretion (E). Pooled samples consisted of ten donor samples for venous blood and saliva, and five donor samples for menstrual blood, semen and vaginal secretion. For all samples without specific amplification $\Delta\Delta C_q$-values were arbitrarily set to 20.



SUPPLEMENTARY FIGURE 2: Expression patterns of the twelve miRNA markers evaluated in individual samples of forensically relevant body fluids. Candidate miRNAs for blood (A-C), menstrual blood (D-E), saliva (F-G), semen (H-K), and vaginal secretion (L). Box blots represent the lower quartile, median and upper quartile of a respective samples set, with the whiskers indicating the (non-extreme) maximum and minimum values, and outliers (1.5x interquartile range) and extreme outliers (3x interquartile range) are depicted as circles and asterisks, respectively. For all samples without specific amplification ΔC_q-values were arbitrarily set to 15.

SUPPLEMENTARY MATERIAL: DISCRIMINANT FUNCTION ANALYSIS

Discriminant function analysis aims to explain a categorical dependent variable (=categorical grouping variable) by the means of the values of one or more independent variables. It not only detects coherences between the variables but also predicts unknown values of the dependent variable, thus classifying the samples into groups.

The general form of a linear discriminant function is as follows: $D = a_0 + a_1 \times X_1 + a_2 \times X_2 + \ldots + a_n \times X_n$, with X_i being the independent variables and a_i the respective coefficients.

The first step of discriminant function analysis is to estimate the coefficients, thereby assuming discrete groups. Coefficients are estimated to maximize the quotient of the sum of squares of the discriminant scores between groups and the sum of squares of the discriminant scores within groups.

Using the example of hsa-miR-203a-3p and hsa-miR-124-3p as independent variables and the groups 'saliva' ('4') and 'vaginal secretion'('5') as dependent variables the linear discriminant function term is as follows:

$$D = -3.305 - 0.743 \times \Delta C_q(\text{hsa- miR- 203a- 3p}) + 0.582 \times \Delta C_q(\text{hsa- miR- 124-3p})$$

Table A: Canonical Discriminant Function Coefficients

	Function
	1
ΔC_q of hsa-miR-203a-3p	-.743
ΔC_q of hsa-miR-124-3p	.582
(Constant)	-3.305

Unstandardized coefficients

Table B gives the discriminant score D as computed for each case in the sample set.

Table B: Casewise Statistics

			Highest Group				Second Highest Group			Discriminant Scores
			P(D>d | G=g)							
Case Nr	Actual Group	Predicted Group	p	df	P(G=g | D=d)	Squared Mahalanobis Distance to Centroid	Group	P(G=g | D=d)	Squared Mahalanobis Distance to Centroid	Function 1
Original 1	4	4	.809	1	1.000	.058	5	.000	22.111	2.472
2	4	4	.790	1	1.000	.071	5	.000	22.347	2.497
3	4	4	.386	1	.998	.751	5	.002	12.917	1.364
4	4	4	.125	1	.957	2.348	5	.043	8.574	.698
5	4	4	.848	1	1.000	.037	5	.000	21.648	2.423
6	4	4	.712	1	1.000	.136	5	.000	23.327	2.600
7	4	4	.944	1	1.000	.005	5	.000	20.524	2.300
8	4	4	.984	1	1.000	.000	5	.000	20.080	2.251
9	4	4	.309	1	1.000	1.034	5	.000	30.000	3.247
10	4	4	.183	1	.982	1.771	5	.018	9.796	.900
11	4	4	.644	1	1.000	.213	5	.000	24.229	2.692
12	4	4	.891	1	1.000	.019	5	.000	21.140	2.368
13	4	4	.456	1	1.000	.557	5	.000	27.108	2.976
14	4	4	.370	1	.997	.802	5	.003	12.707	1.334
15	4	4	.270	1	1.000	1.216	5	.000	30.950	3.333
16	5	5	.713	1	1.000	.136	4	.000	23.318	-2.599
17	5	5	.313	1	.996	1.017	4	.004	11.915	-1.222
18	5	5	.321	1	1.000	.985	4	.000	29.732	-3.222
19	5	5	.735	1	1.000	.114	4	.000	23.027	-2.568
20	5	5	.713	1	1.000	.136	4	.000	16.745	-1.862
21	5	5	.771	1	1.000	.085	4	.000	22.580	-2.522
22	5	5	.379	1	1.000	.773	4	.000	28.510	-3.109
23	5	5	.875	1	1.000	.025	4	.000	21.322	-2.387
24	5	5	.566	1	1.000	.329	4	.000	25.342	-2.804
25	5	5	.064	1	.844	3.430	4	.156	6.804	-.378
26	5	5	.027	1	.528	4.863	4	.472	5.086	-.025
27	5	5	.068	1	1.000	3.330	4	.000	39.505	-4.055
28	5	5	.147	1	.970	2.099	4	.030	9.071	-.782
29	5	5	.202	1	1.000	1.627	4	.000	32.903	-3.506
30	5	5	.855	1	1.000	.033	4	.000	21.562	-2.413

A discriminant score $D > 0$ classifies as sample as 'saliva' while a discriminant score $D < 0$ classifies as sample as 'vaginal secretion'.

Table C: Classification Results[a]

			Predicted Group Membership		
		sample type	Saliva	Vaginal secretion	Total
Original	Count	Saliva	15	0	15
		Vaginal secretion	0	15	15
	%	Saliva	100.0	.0	100.0
		Vaginal secretion	.0	100.0	100.0

a. 100,0% of original grouped cases correctly classified.

The group affiliation of unknown samples can be determined in the same manner.

The goodness of the applied function can not only be evaluated by comparing the predicted group of a sample with its actual group, but additionally using the following parameters:

Eigenvalues are an indicator for the separation properties of a given discriminant function (ratio of the sum of squares of the deviation of the discriminant scores between groups (=explained variation) and the sum of squares of the deviation of the discriminant scores within groups (=unexplained variation)).
The canonical correlation of a discriminant function corresponds to the square root of to the ratio of the sum of squares of the deviation of the discriminant scores between groups (=explained variation) and the sum of squares of the overall deviation of the discriminant scores (=total variation). Its square equals the proportion of the overall variation explained by differences between the groups.

Table D: Eigenvalue

Function	Eigenvalue	Canonical Correlation
1	5.329[a]	.918

a. First 1 canonical discriminant functions were used in the analysis.

The eigenvalue of 5.329 indicates that the variation between groups is approximately 5.3-fold higher than the variation within groups.
About 84 % of the overall variation is explained by differences between the groups ($0.981^2 = 0.8427$).

Accordingly, the Wilks‘ Lambda (quotient of the sum of squares of the deviation of the discriminant scores within groups (=unexplained variation) and the sum of squares of the overall deviation of the discriminant scores (=total variation)) yields a relatively low value.

Table E: Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.158	49.820	2	.000

Approximately 16 % of the overall variation is not explained by differences between the groups
The displayed significance value allows the rejection of the null hypothesis that no differences in the mean discriminant function values are given in the parent population, the distinction of groups the groups ‘saliva’ and ‘vaginal secretion by means of this discriminant function is thus significant.

Analysis of the mean ΔC_q-values per group and in the total data set further confirmed that their differences are significant.

Table F: Group Statistics

sample type		Mean	Std. Deviation	Valid N (listwise) Unweighted	Valid N (listwise) Weighted
Saliva	ΔC_q of hsa-miR-203a-3p	-8.1373	.68362	15	15.000
	ΔC_q of hsa-miR-124-3p	-.8707	1.65986	15	15.000
Vaginal secretion	ΔC_q of hsa-miR-203a-3p	-5.8780	.98452	15	15.000
	ΔC_q of hsa-miR-124-3p	-5.6500	1.02122	15	15.000
Total	ΔC_q of hsa-miR-203a-3p	-7.0077	1.41904	30	30.000
	ΔC_q of hsa-miR-124-3p	-3.2603	2.78226	30	30.000

Table G: Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
ΔC_q of hsa-miR-203a-3p	.344	53.298	1	28	.000
ΔC_q of hsa-miR-124-3p	.237	90.213	1	28	.000

Limiting has to be noted, that the theoretical conditions of equal covariance and variance is not met (Box' M Test). The covariance matrices, however, do not indicate gross violations of the prerequisites for discriminant function analyses.

Table H: Box'M Test Results

Box's M		14.969
F	Approx.	4.604
	df1	3
	df2	141120.000
	Sig.	.003

Tests null hypothesis of equal population covariance matrices.

Table I: Covariance Matrices[a]

sample type		ΔC_q of hsa-miR-203a-3p	ΔC_q of hsa-miR-124-3p
Saliva	ΔC_q of hsa-miR-203a-3p	.467	.633
	ΔC_q of hsa-miR-124-3p	.633	2.755
Vaginal secretion	ΔC_q of hsa-miR-203a-3p	.969	-.541
	ΔC_q of hsa-miR-124-3p	-.541	1.043
Total	ΔC_q of hsa-miR-203a-3p	2.014	-2.748
	ΔC_q of hsa-miR-124-3p	-2.748	7.741

The respective tables for hsa –miR-144-3p and hsa-miR-203a-as independent variables and the groups 'venous blood' ('1') and 'menstrual blood'('2') as dependent variables are presented below. The resulting term is:

$$D = -1.504 - 0.127 \times \Delta C_{\mathrm{q}}(\text{hsa-miR-144-3p}) + 0.454 \times \Delta C_{\mathrm{q}}(\text{hsa-miR-203a-3p}).$$

A discriminant score $D > 0$ classifies as sample as 'venous' while a discriminant score $D < 0$ classifies as sample as 'menstrual blood'.

Table J: Canonical Discriminant Function Coefficients

	Function
	1
ΔC_q of hsa-miR-144-3p	-.127
ΔC_q of hsa-miR-203a-3p	.454
(Constant)	-1.504

Unstandardized coefficients

Table K: Casewise Statistics

			Highest Group					Second Highest Group			Discriminant Scores
				P(D>d | G=g)			Squared Mahalanobis Distance to Centroid			Squared Mahalanobis Distance to Centroid	
	Case Nr	Actual Group	Predicted Group	p	df	P(G=g | D=d)		Group	P(G=g | D=d)		Function 1
Original	1	1	1	.510	1	1.000	.435	2	.000	32.096	2.503
	2	1	1	.587	1	1.000	.296	2	.000	33.421	2.619
	3	1	1	.389	1	1.000	.742	2	.000	29.848	2.301
	4	1	1	.993	1	1.000	.000	2	.000	40.115	3.171
	5	1	1	.668	1	1.000	.184	2	.000	34.756	2.733
	6	1	1	.006	1	1.000	7.473	2	.000	82.055	5.896
	7	1	1	.714	1	1.000	.135	2	.000	35.494	2.795
	8	1	1	.535	1	1.000	.385	2	.000	32.542	2.542
	9	1	1	.608	1	1.000	.263	2	.000	33.773	2.649
	10	1	1	.692	1	1.000	.157	2	.000	45.176	3.559
	11	1	1	.866	1	1.000	.028	2	.000	42.165	3.331
	12	1	1	.893	1	1.000	.018	2	.000	41.718	3.297
	13	1	1	.713	1	1.000	.135	2	.000	44.792	3.530
	14	1	1	.866	1	1.000	.029	2	.000	37.893	2.993
	15	1	1	.724	1	1.000	.125	2	.000	44.597	3.516
	16	2	2	.989	1	1.000	.000	1	.000	40.184	-3.177
	17	2	2	.564	1	1.000	.332	1	.000	47.623	-3.739
	18	2	2	.019	1	.994	5.538	1	.006	15.773	-.809
	19	2	2	.710	1	1.000	.139	1	.000	44.853	-3.535
	20	2	2	.078	1	1.000	3.104	1	.000	65.391	-4.924
	21	2	2	.406	1	1.000	.689	1	.000	51.193	-3.993
	22	2	2	.256	1	1.000	1.289	1	.000	26.930	-2.027
	23	2	2	.451	1	1.000	.569	1	.000	31.030	-2.408
	24	2	2	.756	1	1.000	.096	1	.000	44.028	-3.473
	25	2	2	.224	1	1.000	1.482	1	.000	56.882	-4.380
	26	2	2	.216	1	1.000	1.528	1	.000	57.169	-4.399
	27	2	2	.848	1	1.000	.037	1	.000	37.614	-2.971
	28	2	2	.695	1	1.000	.154	1	.000	35.190	-2.770
	29	2	2	.105	1	1.000	2.622	1	.000	22.143	-1.543
	30	2	2	.899	1	1.000	.016	1	.000	41.628	-3.290

Table L: Classification Results[a]

			Predicted Group Membership		
		sample type	Venous blood	Menstrual blood	Total
Original	Count	Venous blood	15	0	15
		Menstrual blood	0	15	15
	%	Venous blood	100.0	.0	100.0
		Menstrual blood	.0	100.0	100.0

a. 100.0% of original grouped cases correctly classified.

Table M: Eigenvalues

Function	Eigenvalue	Canonical Correlation
1	10.715[a]	.956

a. First 1 canonical discriminant functions were used in the analysis.

Table N: Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.085	66.443	2	.000

Table O: Group Statistics

				Valid N (listwise)	
sample type		Mean	Std. Deviation	Unweighted	Weighted
Venous blood	ΔC_q of hsa-miR-144-3p	-7.1907	1.24104	15	15.000
	ΔC_q of hsa-miR-203a-3p	8.2620	2.07536	15	15.000
Menstrual blood	ΔC_q of hsa-miR-144-3p	-2.4913	3.29699	15	15.000
	ΔC_q of hsa-miR-203a-3p	-4.3500	2.12921	15	15.000
Total	ΔC_q of hsa-miR-144-3p	-4.8410	3.42089	30	30.000
	ΔC_q of hsa-miR-203a-3p	1.9560	6.73831	30	30.000

Table P: Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
ΔC_q of hsa-miR-144-3p	.512	26.692	1	28	.000
ΔC_q of hsa-miR-203a-3p	.094	269.883	1	28	.000

Table Q: Box'M Test Results

Box's M		17.627
F	Approx.	5.422
	df1	3
	df2	141120.000
	Sig.	.001

Tests null hypothesis of equal population covariance matrices.

Table R: Covariance Matrices[a]

sample type		ΔC_q of hsa-miR-144-3p	ΔC_q of hsa-miR-203a-3p
Venous blood	ΔC_q of hsa-miR-144-3p	1.540	1.469
	ΔC_q of hsa-miR-203a-3p	1.469	4.307
Menstrual blood	ΔC_q of hsa-miR-144-3p	10.870	-1.264
	ΔC_q of hsa-miR-203a-3p	-1.264	4.534
Total	ΔC_q of hsa-miR-144-3p	11.703	-15.229
	ΔC_q of hsa-miR-203a-3p	-15.229	45.405

a. The total covariance matrix has 29 degrees of freedom.

4.3 Zusammenfassung und Diskussion

Für eine unvoreingenommene Selektion differentiell exprimierter miRNA-Kandidaten für die untersuchten Körperflüssigkeiten Blut, Speichel, Sperma, Vaginalsekret und Menstruationsblut in forensisch realistischem Probenmaterial wurde zunächst ein miRNA Microarray-Experiment durchgeführt, bei dem ca. 1.900 bekannte humane miRNAs abgefragt wurden. Hierbei bestätigte sich in einer unbeaufsichtigten, hierarchischen Clusteranalyse, in der der Großteil der Proben entsprechend ihrer biologischen Herkunft gruppiert wurde, dass die Unterschiede zwischen den distinkten Körperflüssigkeiten den größten Anteil der beobachteten Expressionsvarianzen erklärten. Lediglich zwei Menstruationsblutproben wurden nicht korrekt, sondern in die Gruppe der Vaginalsekretproben beziehungsweise zwischen die Speichel- und Blutproben eingeordnet, worin sich die komplexe Zusammensetzung dieser Probenart aus Blut, Uterusschleimhaut und Vaginalsekret widerspiegelt.

Bei der Auswertung der Microarray-Ergebnisse hinsichtlich geeigneter Kandidaten wurden miRNAs ausgewählt, die entweder ausschließlich in einer Körperflüssigkeit exprimiert wurden, in einer Körperflüssigkeit ein deutlich höheres Expressionsniveau aufwiesen als in den verbleibenden vier Körperflüssigkeiten oder in zwei Körperflüssigkeiten, deren Unterscheidung von besonderem Interesse ist (etwa Speichel und Vaginalsekret), unterschiedlich stark exprimiert waren. Während für venöses Blut etliche vielversprechende miRNA-Kandidaten gefunden werden konnten, stellte sich dies für die restlichen Körperflüssigkeiten schwieriger dar. Der Grund hierfür ist nicht nur in der Komplexität der Probenarten zu finden, sondern beruht zu einem Teil auch auf den Einschränkungen der Microarray-Technologie, wie etwa dem relativ geringen Dynamikbereich. Für eine erweiterte Auswahl potentiell körperflüssigkeitsspezifischer miRNAs wurden daher zusätzlich die in der einschlägigen forensischen Literatur publizierten Ergebnisse analysiert und in den Auswahlprozess integriert – auf diese Weise wurden insgesamt 36 miRNA-Kandidaten zur Evaluation mittels RT-qPCR ausgewählt.

Analog zu den vorangegangenen Normalisierungsstudien wurde zunächst die technische Performanz (Amplifikationserfolg und -effizienz) der Marker analysiert, gefolgt von einer ersten Überprüfung in vereinigten Proben mehrerer Individuen, um die durchschnittlichen Expressionsniveaus der jeweiligen Kandidaten-miRNA pro Körperflüssigkeit abzuschätzen. Die Marker, die in diesen vereinigten Proben die besten Trenneigenschaften

aufwiesen, also in ihrer Zielkörperflüssigkeit deutlich höher exprimiert waren als in den restlichen Körperflüssigkeiten, wurden anschließend in Einzelproben evaluiert, wodurch die Analyse um die interindividuellen Unterschiede erweitert und somit vervollständigt wurde.

Die miRNA hsa-miR-891a-5p zeigte ausschließlich in Spermaproben spezifische Amplifikation und konnte somit als vollständig spermaspezifischer Marker identifiziert werden. Außerdem lagen die Expressionsniveaus von hsa-miR-10a-5p, hsa-miR-10b-5p und hsa-miR-135a-5p in Spermaproben nicht nur höher, als in den Proben der restlichen Körperflüssigkeiten, sondern zeigten überdies auch keine Überlappungen, sodass diese ebenfalls als gut geeignete Marker für die Identifikation von Sperma angesehen werden können. Der Marker hsa-miR-144-3p erlaubte die eindeutige Unterscheidung zwischen Blutproben im allgemeinen Sinne (venöses Blut und Menstruationsblut) und Nicht-Blutproben (Speichel, Sperma und Vaginalsekret), da lediglich in den Blutproben spezifische Amplifikation auftrat. Hsa-miR-144-5p und hsa-miR-451a zeigten ebenfalls gute Trenneigenschaften zwischen Blut- und Nicht-Blutproben. Für die Differenzierung von venösem Blut und Menstruationsblut konnten dagegen keine eindeutig spezifischen Marker identifiziert werden, da die Expressionsniveaus aller potentiellen Marker in diesen beiden Körperflüssigkeiten deutlich überlappten. Auch die Identifikation von Speichel und Vaginalsekret erwies sich als schwierig, da hierfür keine unzweideutigen Marker gefunden werden konnten: die als potentiell speichelspezifisch ausgewählten miRNAs hsa-miR-203a-3p und hsa-miR-205-5p zeigten in einigen Vaginalsekret- uns Menstruationsblutproben ähnlich hohe Ausprägung wie in den Speichelproben, während für hsa-miR-124-3p die Expressionsniveaus der Zielkörperflüssigkeit Vaginalsekret teilweise mit denen von Menstruationsblut- und Spermaproben überlappten.

Um dennoch eine Unterscheidung aller fünf forensisch relevanten Körperflüssigkeiten in Einzelquellproben mit möglichst geringem experimentellen Aufwand zu erreichen, wurde basierend auf den quantitativen Expressionsdaten der Einzelproben folgender Entscheidungsalgorithmus entwickelt: zunächst werden Spermaproben basierend auf der exklusiven Expression von hsa-miR-891a-5p von den anderen Körperflüssigkeiten abgegrenzt; die verbleibenden Proben werden anschließend anhand des Expressionsniveaus von hsa-miR-144-3p in Blut- oder Nicht-Blutproben differenziert; und für die Unterscheidung von venösen Blut- und Menstruationsblutproben sowie Speichel- und Vaginalsekretproben

wird jeweils die Kombination zweier miRNAs (hsa-miR-144-3p/hsa-miR-203a-3p beziehungsweise hsa-miR-203a-3p/ hsa-miR-124-3p) angewendet. Zur objektiven, statistisch begründbaren Gruppenzuordnung beim Einsatz zweier kombinierter Marker wurde eine Diskriminanzanalyse durchgeführt. Alle Proben dieses Trainingssets konnten so korrekt klassifiziert werden.

Die Tauglichkeit der identifizierten miRNAs sowie die Anwendbarkeit des erarbeiteten Entscheidungsbaumes wurden anschließend anhand von verblindeten und gealterten Proben, sowie Mischungen mehrerer Körperflüssigkeiten getestet. Analysierbare Expressionsergebnisse für alle gealterten Proben, inklusive 36 Jahre alter Blutflecken, bestätigten die sehr gute Eignung von miRNA als Biomarker zur Spurenartidentifikation in typischerweise gealtertem und etwa durch Umwelteinflüsse oder Fäulnisprozesse beeinträchtigtem forensisch-biologischen Spurenmaterial. Außerdem wurden sämtliche verblindeten und gealterten Sperma-, venöse Blut- und Menstruationsblutproben mittels des Entscheidungsalgorithmus den jeweils korrekten Körperflüssigkeiten zugeordnet. Ferner zeigten alle Speichel- und Vaginalsekretproben Amplifikation der Marker hsa-miR-203a-3p und hsa-miR-124-3p sowie in keinem Fall spezifische Amplifikation der miRNAs hsa-miR-891a-5p und hsa-miR-144-3p und konnten so als Schleimhautproben klassifiziert werden. Es ergab sich jedoch eine kleine Teilmenge dieser Schleimhautproben, die nahe an der Entscheidungsgrenze der Diskriminanzanalyse gruppierten und teilweise falsch klassifiziert wurden.

In den Mischungen mehrerer Körperflüssigkeiten erwies sich hsa-miR-891a-5p abermals als gut geeigneter robuster Marker zu Identifizierung von Sperma. Die miRNA zeigte nicht nur eine hohe Sensitivität, da sie auch in Mischungen nachweisbar war, in denen Sperma den weit geringeren Mischungsanteil ausmachte, sondern auch eine gute Spezifität, da sich keine falsch-positiven Befundungen in Mischungen, die kein Sperma enthielten, ergaben. Die Spezifität ist im Falle des Spermanachweises besonders herauszustellen, da diesem bei Sexualdelikten spezielle Bedeutung zukommen kann und falsch-positive Ergebnisse potentiell zu einer ungerechtfertigten Belastung eines Tatverdächtigen führen können. Die Differenzierung des Vorhandenseins von venösem Blut gegenüber Menstruationsblut sowie Vaginalsekret gegenüber Speichel war in den Mischspuren dagegen nur bedingt erfolgreich.

Zusammenfassend wurden in der in diesem Kapitel vorgestellten Studie die bisher umfassendsten Arbeiten zur miRNA-basierten Identifizierung forensisch relevanter Körperflüssigkeiten basierend auf einer umfassend validierten Normalisierungsstrategie vorgelegt und sehr gut geeignete, robuste Marker für die Identifikation von Sperma und Blut im allgemeinen Sinne identifiziert. Für die Unterscheidung aller fünf relevanten Probenarten wurde außerdem ein vielversprechender, praxistauglicher Entscheidungs-algorithmus zur unvoreingenommenen Datenauswertung vorgeschlagen.

5 miRNA-basierte Identifizierung von Organgeweben im forensischen Kontext

5.1 Einleitung

Während die Inferenz von Körperflüssigkeiten eine häufige Fragestellung in der forensisch-genetischen Routinearbeit darstellt, ist die Identifizierung von Organgeweben seltener gefordert – ihr kann dennoch fundamentale Bedeutung bei der Rekonstruktion von Tathergängen, etwa bei der Aufklärung von Verletzungs- und Tötungsdelikten unter Einsatz von Waffengewalt, zukommen.

Da die bislang in der forensischen Routine verwendeten immunhistochemischen oder enzymatischen Methoden zur Organidentifizierung eine Reihe von Unzulänglichkeiten aufweisen und nur ein Teil der Organgewebe sicher identifiziert werden konnte, bieten sich hier ebenfalls alternativ nukleinsäureanalytische Herangehensweisen an und einige Studien haben bereits die generelle Eignung von miRNA als Biomarker für die Unterscheidung verschiedener humaner Organgewebe gezeigt [245–252,292–295]. Die in diesem Kapitel vorgestellte Studie befasste sich erstmals mit der Erforschung und Etablierung von miRNA-Kandidaten für die reliable Identifizierung von Gehirn-, Lungen-, Leber-, Nieren-, Herzmuskel - und Skelettmuskelgewebe sowie Haut im forensischen Kontext.

Analog zur vorangegangenen Studie zur Markeretablierung für die Körperflüssigkeitsidentifikation fand eine zuvor umfassend validierte, empirisch erarbeitete Normalisierungsstrategie Anwendung, um reliable, biologisch aussagekräftige Ergebnisse zu erhalten (siehe Kapitel 3.3). Zudem wurden die Arbeitsabläufe hochgradig standardisiert durchgeführt, die Dokumentation und Berichterstattung erfolgte entsprechend der MIQE-Richtlinien und die Expressionsergebnisse wurden einer biasfreien statistischen Auswertung unterzogen.

Identification of organ tissue types and skin from forensic samples by microRNA expression analysis

Eva Sauer[1], Antje Extra[1], Philipp Cachée[2], Cornelius Courts[1]

[1]Institute of Forensic Medicine, University Medical Center Schleswig-Holstein, Kiel, Germany
[2]Sachverständigenbüro Cachée, Saarmunder Straße 3A, 14552 Michendorf, Germany

ABSTRACT

The identification of organ tissues in traces recovered from scenes and objects with regard to violent crimes involving serious injuries can be of considerable relevance in forensic investigations. Molecular genetic approaches are provably superior to histological and immunological assays in characterizing organ tissues, and micro-RNAs (miRNAs), due to their cell type specific expression patterns and stability against degradation, emerged as a promising molecular species for forensic analyses, with a range of tried and tested indicative markers. Thus, herein we present the first miRNA based approach for the forensic identification of organ tissues. Using quantitative PCR employing an empirically derived strategy for data normalization and unbiased statistical decision making we assessed the differential expression of 15 preselected miRNAs in tissues of brain, kidney, lung, liver, heart muscle, skeletal muscle and skin. We show that not only can miRNA expression profiling be used to reliably differentiate between organ tissues but also that this method, which is compatible with and complementary to forensic DNA analysis, is applicable to realistic forensic samples e.g. mixtures, aged and degraded material as well as traces generated by mock stabbings and experimental shootings at ballistic models.

Key words:

Forensic genetics; microRNA; qPCR; organ tissue identification

1. INTRODUCTION

In the reconstruction of violent crimes, the identification of organ tissues can provide crucial information complementing the usual DNA based source attribution. Proving that a tissue sample recovered, for instance, from the blade of a knife or a bullet secured at a crime scene originates from an internal organ confirms that a serious injury has been inflicted with the weapon, while the mere presence of the victim's DNA established by standard STR profiling may be argued to result from a superficial cut or innocuous graze. Also, in cases involving post-mortal dismemberment to enable covert disposal of a corpse the identification of even mangled and degraded tissue recovered, for instance, from sewers or waste pipes may be required.

Conventional immunological, histological and/or enzymatic techniques for the assignment of tissue origin [1–4] may, however, pose problems in terms of specificity and sensitivity, especially when only trace amounts of material are present, and may interfere with DNA profiling. In recent years, molecular genetic approaches for the identification of cell types have been or are currently being explored by several groups, including messenger RNA- (mRNA) and microRNA- (miRNA) based methods for body fluid identification (BFI) as well as studies on the inference of organ tissue types using mRNA profiling [5,6] (extensively reviewed in [7–9]).

MiRNAs not only exhibit the prerequisite of cell type specific expression [10–18], but these small non-coding RNAs also feature certain characteristics that render them well suited to the challenging demands of the analysis of forensic samples. Firstly, due to their intrinsically small size of 18 - 25 nt, miRNAs are less prone to degradation caused by chemical and/or physical strains, secondly, miRNAs are detected and quantified in their biologically active form so, in contrast to mRNAs, no potential splice variants have to be differentiated, and thirdly, there is no wasteful sample consumption as miRNA and DNA can be extracted simultaneously from the same specimen [19–21].

Quantitative PCR (qPCR) 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 to eliminate non-biological variance is essential [22–25]. A robust normalization strategy that is appropriate for a particular experimental setup should be based on an individual and evidence based selection of one or a group of reference genes [26–28]. We therefore applied a normalization strategy specifically designed for the herein examined set of organ tissues [29].

The present study is, to the best of our knowledge, the first approach to employ miRNA markers for the identification of brain, kidney, liver, lung, skin, heart muscle and skeletal muscle in a forensic context.

2. MATERIALS AND METHODS

2.1 Samples

All samples were anonymized after collection and discarded after use. The study design and

experimental procedures were approved by the ethics committees of the University Hospital of Bonn and the University Medical Center of Schleswig-Holstein.

2.1.1 Native organ tissue samples and dried swabs for marker validation

Specimens of each of the seven examined human organs, i.e. brain, kidney, liver, lung, skin, heart muscle (hereinafter referred to as heart) and skeletal muscle (hereinafter referred to as muscle), were obtained during medico-legal autopsies at the Institutes of Forensic Medicine in Bonn and Kiel. Marked signs of putrefaction or pathological change of the sampled organs were set as exclusion criteria. Within one hour of excision, tissue samples were immersed in RNAlater® solution (Ambion™, Austin, TX, USA) and stored at -80 °C until further processing, when samples were allowed to thaw on ice overnight and approximately 100 mg per sample were subjected to nucleic acid extraction.

Additionally, dried organ tissue samples were prepared by rubbing and pricking into the excised native tissue specimens (except skin) with dry stemmed DNA-free cotton swabs (Sarstedt AG & Co, Nümbrecht, Germany). Swabs were allowed to dry at room temperature for 24 h, then stored at -80 °C and prior to extraction thawed at room temperature and let dry for another 24 h. Dry skin samples were obtained by thoroughly swabbing the forehead of informed and consenting adult volunteers with dry stemmed DNA-free nylon swabs (Copan, Brescia, Italy) that were processed after storage at room temperature for 24 h.

Venous blood from informed and consenting adult volunteers was collected by venipuncture using dry vacutainers and spotted onto sterile cotton swabs. Samples were dried at room temperature and processed after 24 h.

2.1.2 Aged specimens, mixtures and mock case samples

Two dried swabs per organ tissue type were aged for 28 days prior to extraction in a clean environment with between 18 and 25 °C and without direct exposure to sunlight (see 2.1.1 for sampling procedure).

To further assess the effects of putrefaction on miRNA organ profiling success, native organ tissue samples measuring approximately 1 cm^3 were placed in an outside location immediately after autopsy (without prior storage in RNAlater® solution) for 28 days during June/July (n = 2 per organ tissue type; for weather parameters see Supplementary Material 1). Specimens were put into translucent plastic containers with perforated lids, allowing air circulation and exposure to circadian variations in temperature, humidity and insolation while restricting access for birds and insects. Macroscopically visible mold growth was excised prior to total RNA extraction.

Mixtures of different organs of the same individual were prepared by combining portions of two to four tissue samples (previously stored in RNAlater® solution at -80 °C) prior to extraction, either in balanced mixtures or with one tissue type as a minor component (Supplementary Table 4).

In an attempt to simulate human gunshot injuries involving perforated organs, experimental contact

shots at basic ballistic models (modified after [30,31]) were conducted to generate backspatter and forward spatter samples. Briefly, the model consisted of 1 l polyethylene bottles, filled with 0.9 l ballistic gelatine (10 %, type 'Ballistic III' (Gelita Ebersbach, Germany), prepared as per manufacturer's instructions). 3 x 3 cm plastic foil bags containing a mixture of 3 ml blood from informed and consenting adult volunteers and 300 mg ground organ tissue obtained during medico-legal autopsy were fixed to the front and back of each model and covered with a layer of approximately 3 mm of silicon. The model was fixated in a self-made contraption (see Supplementary Material 2) and contact shots were fired in such a manner that both organ tissue containing bags were perforated by the bullet. We performed three series of shots employing three distinct firearms, so that each organ tissue type, i.e. brain, kidney, liver, lung, heart and muscle, was shot at by three distinct weapons and each weapon was used to shoot all six organ types, totaling 18 shots altogether. The weapons used were a revolver Ruger Speed-Six with caliber .357 158 grain semi-jacketed soft point ammunition, a semi-automatic pistol Sig Sauer P 225 and a semi-automatic pistol Star, both with 9 x 19 mm 115 grain full metal jacket ammunition. All shots were fired in a forensic ballistics laboratory by an authorized expert and according to German laws and regulations. Prior to each shot, the firearms were thoroughly cleaned with Roti® Nucleic Acid-free (Carl Roth, Karlsruhe, Germany) to avoid contamination. Backspatter from within the gun barrel was sampled using a modified double-swab technique [32] employing stemmed DNA-free cotton swabs with one half moistened with 30 µL of 70 % ethanol. The swabs were stored at room temperature for 96 h prior to extraction. Forward spatter was caught on a 38 x 46 cm screen of 0.35 mm filter paper (Carl Roth) that was tautened 40 cm behind the ballistic model and was replaced prior to every shot. Filter papers were stored at room temperature for 120 h before sampling cuttings totaling 9 cm^2 presumably containing adherent organ tissue prior to extraction.

To mimic case work samples of abdominal stabbings, identical kitchen knives with 12 x 1.8 cm blades were used to stab stacks of autopsy samples consisting of the following strata: skin on muscle, skin on muscle on kidney, skin on muscle on liver, skin on muscle on lung and skin on muscle on heart. Skin samples were excised from the abdomen and had a thickness of 1 to 2 mm (plus approximately 10 mm of adipose tissue) while the strata of internal organs were approximately 15 mm thick. Experiments were performed using three biological replicates per tissue combination. Smears and tissue material were let dry on the knives at room temperature for 24 h and then collected by swabbing the blade with a stemmed DNA-free cotton swab moistened with 100 µl 70 % ethanol. Swabs were stored at room temperature for 24 h prior to extraction.

2.2 Selection of candidate miRNAs and reference genes

For the selection of organ tissue specific candidates, miRNA microarray profiling was performed by Exiqon Services (Vedbaek, Denmark) comprising five native tissue samples per organ type. Briefly, after extraction with the miRNeasy kit (QIAGEN, Hilden, Germany) and quality control of the total RNA samples based on an Agilent 2100 Bioanalyzer electropherogram, 640 ng of total RNA of each

sample as well as a reference RNA sample was labeled with Hy3™ and Hy5™ fluorescent label, respectively, using the miRCURY LNA™ microRNA Hi-Power Labeling Kit, Hy3™/Hy5™ (Exiqon). The Hy3™-labeled samples and the Hy5™-labeled reference RNA sample were mixed pair-wise and hybridized to the miRCURY LNA™ microRNA Array 7th Generation (Exiqon, Denmark) which contains capture probes that represent all human miRNAs registered in the miRBASE v.18.0 [33]. The hybridization was performed according to the miRCURY LNA™ microRNA Array Instruction manual employing a Tecan HS4800™ hybridization station (Tecan, Austria). The miRCURY LNA™ microRNA Array slides were scanned using the Agilent G2565BA Microarray Scanner System (Agilent Technologies, Inc., USA) and the image analysis was carried out applying the ImaGene 9.0 software (BioDiscovery, Inc., USA). The quantified signals were background corrected (Normexp with offset value 10 [34]) and normalized using the quantile normalization method. Subsequently, unsupervised as well as supervised data analyses were performed. In addition, quantile normalization and ranking of the miRNAs according to their expression level was performed for each subgroup (i.e. organ type) separately and candidate miRNAs were selected by comparing the mean signal intensities between organ tissues manually.
The selection of reference genes for this study was done as described by Sauer et al. [29].

2.3 RNA extraction and quantification

All surfaces, devices and instruments utilized during the extraction procedure and when working with RNA eluates were thoroughly cleaned with Roti® Nucleic Acid-free and RNase-Zap® (Ambion) to remove traces of nucleic acid contaminations and ambient RNases. Also, only RNase-free reagents and plastic consumables were used.

Total RNA was extracted using the mirVana™ miRNA Isolation Kit (Ambion) following the manufacturer's protocol with slight modifications. Prior to organic extraction, organ tissue samples were minced using scissors and incubated immersed in 1 ml Lysis/Binding Buffer at 56 °C for 2 h. Unlysed residual tissue was passed through QIAshredder™ columns (Qiagen, Hilden, Germany) by centrifugation at 13,000 × g for 2 min and added to the lysate. Tissue debris remaining in the columns was discarded. Swabs and filter paper cuttings were cut into pieces ≤ 2 mm^2 and incubated immersed in 350 µl Lysis/Binding Buffer at 56 °C for 1 h and in case of mock case samples for 2 h. Forensic Filters (Macherey-Nagel, Düren, Germany) were used to recover the lysate from the substrate by centrifugation at 13,000 × g for 1 min. For the aged organ tissues as well as for the samples from mock stabbings and shooting experiments 20 µl lysate was diverted for DNA co-extraction (see below) and stored at -20 °C until further processing. Total RNA was eluted in 100 µl elution solution and eluates were stored at -80 °C until further processing. Extraction negative controls underwent the same procedure.

To remove potential traces of genomic DNA, subsequent DNase I digestion was performed using the Turbo DNA-free™ Kit (Ambion), according to the manufacturer's routine protocol for swabs and filter paper cuttings while the rigorous treatment was applied to native tissue samples.

Total RNA concentration was determined fluorometrically using the QuantiFluor® RNA System on a Quantus™ Fluorometer (both Promega, Mannheim, Germany). For samples used in the marker validation and aged specimens, the RNA quality, represented by RNA integrity number (RIN) [35] was determined by applying the RNA 6000 Nano Kit on an Agilent 2100 Bioanalyzer (both Agilent, Böblingen, Germany).

2.4 DNA extraction, quantification and STR profiling

DNA was extracted from the diverted volumes of mirVana™ lysates (as described in section 2.3) using the PrepFiler® Forensic DNA Extraction Kit (Applied Biosystems™, Weiterstadt, Germany) according to the manufacturer's protocol. DNA was eluted in 50 µl elution buffer and eluates were stored at -20 °C until further processing.

The qPCR-based PowerQuant® System (Promega) was applied to determine the DNA concentration and to assess the presence of PCR inhibitors and DNA degradation for each sample. Reactions were performed as per manufacturer's protocol on a 7500 Fast Real-Time PCR System (Applied Biosystems).

STR profiling was conducted for selected samples using the PowerPlex® ESI 17 Kit (Promega) following the manufacturer's instructions. PCR products were detected by capillary electrophoresis on an ABI 3500 Genetic Analyzer (Applied Biosystems) and data analyses were performed with the GeneMapper® ID-X software version 1.5 (Life Technologies™, Darmstadt, Germany).

2.5 Sample preparation

Concentrations of total RNA samples from organ tissues were normalized to 4 ng/µl based on quantification results. For the first step of marker selection, pooled samples per organ tissue type were prepared by combining identical volumes of normalized RNA eluates of native specimens from ten donors. Further, a mixture of all seven organ tissues containing identical volumes of the above mentioned pooled samples was prepared for the determination of amplification efficiency per marker. In the next step, ten biological replicates per respective organ tissue of interest and at least five biological replicates per every non-target organ tissue were examined (e.g. for hsa-miR-219a-5p, a candidate for the identification of brain tissue, ten brain tissue samples and five samples each for kidney, liver, lung, heart, muscle and skin were examined).

The selected markers were further validated by analyzing their expression levels in dried swabs of the respective tissues (five biological replicates per respective organ of interest and at least three biological replicates per every non-target organ tissue) and the aged specimens, mixtures and mock case samples described above. These samples were also normalized to a concentration of 4 ng/µl, with the exception of the backspatter specimens and a number of skin swabs that were either diluted to 2 ng/µl or used undiluted.

2.6 Reverse transcription qPCR (RT-qPCR)

Reverse transcription reactions to synthesize complementary DNA (cDNA) were performed using

target-specific stem-loop primers (Applied Biosystems, Supplementary Table 1) and the TaqMan® MicroRNA Reverse Transcription Kit (Applied Biosystems) as per manufacturer's protocol. For the organ tissue specific markers, 10 ng of total RNA were added to the reaction mix containing 1X RT primers, 50 U MultiScribe™ reverse transcriptase, 1 mM dNTPs, 3.8 U RNase inhibitor and 1X reverse transcription buffer, resulting in a reaction volume of 15 µl. The reference genes were transcribed in pooled reactions containing 20 ng total RNA, 1X RT primers for SNORD24, SNORD48, and RNU6-2, 100 U MultiScribe™ reverse transcriptase, 1 mM dNTPs, 7.6 U RNase inhibitor and 1X reverse transcription buffer in a total reaction volume of 30 µl. For samples concentrated lesser than 4 ng/µl, the reference genes were transcribed in individual reactions. If the RNA concentration was below 2 ng/µl, the maximum input volume of 5 µl was used. Reactions were performed on a Veriti™ Thermal Cycler (Applied Biosystems) with the following cycling conditions: 16 °C for 30 min, 42 °C for 30 min and 85 °C for 5 min. Extraction negative and H_2O controls were performed on a sample basis and RT(-)-controls were set up to control for potential contamination with genomic DNA. For initial marker validation with native organ tissues and dry swabs, RT reactions that resulted in specific amplification were technically replicated in an independent second reaction. RT reactions for the aged specimens, mixtures and mock case samples were not replicated. RT reaction products were stored at -20 °C.

QPCR reactions were performed using target-specific TaqMan® Assays (Applied Biosystems, Supplementary Table 1) and the TaqMan® Universal PCR Master Mix II, No AmpErase® UNG (Applied Biosystems) as per manufacturer's protocol: Each 20 µl reaction volume contained 1.3 µl of the appropriate RT reaction product, 1X TaqMan® Universal PCR Master Mix and 1X specific TaqMan® Assay. QPCR reactions were conducted in duplicate for each RT reaction product, thus resulting in four technical replicates for the samples used in the initial validation and in two technical replicates for the aged specimens, mixtures and mock case samples. For efficiency determination, qPCR triplicates of the mixture containing all organ tissues were performed, amounting to six technical replicates.

The internal PCR control (IPC) from the Quantifiler® Human DNA Quantification Kit (Applied Biosystems), devised to produce a set C_q-value under standard conditions, was used as an inter-plate calibrator to control for inter-run variability. PCR reactions were performed on a 7500 Fast Real-Time PCR System (Applied Biosystems) with 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 2.3 (Life Technologies) and R_n-values were exported for further analyses.

2.7 Data analysis

The LinRegPCR program version 2014.8 [36] was employed to compute C_q-values and amplification efficiencies from R_n-values in a MIQE compliant manner [37] (see below). The arithmetic mean values of amplification efficiencies per six technical repeats (of the mixture

containing the pooled samples of all seven tissue types) were used in further analyses, 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 $\log_{10}$ (fluorescence). C_q-values deviating more than one cycle from the median value of the respective technical repeat were excluded from subsequent analyses. C_q-values ≥35 after efficiency correction were regarded to result from non-specific amplification [38].

Analysis of qPCR data including pre-processing was then performed using the GenEx software version 6 (multiD Analyses, Göteborg, Sweden) into which LinRegPCR spread sheet exported data was imported. Pre-processing of qPCR encompassed the following steps in the given order: efficiency correction, averaging of technical qPCR and RT replicates and normalization with reference genes resulting in ΔC_q-values. To illustrate differences in expression between pooled samples, $\Delta\Delta C_q$-values were calculated by using the respective target organ tissue as a calibrator sample: $\Delta\Delta C_q = \Delta C_{q(\text{non-target organ tissue})} - \Delta C_{q(\text{target organ tissue})}$. For all samples not generating specific amplification, ΔC_q- and $\Delta\Delta C_q$- values were set to the arbitrary values of 15. Inter-plate calibration was omitted, since evaluation of the C_q-values of the IPC as well as concordant technical repeats between plates indicated negligible differences between plate performances (data not shown). Statistical analyses including Kruskal-Wallis H testing and logistic regression analyses employing the native organ tissue samples and dried specimens were performed employing IBM SPSS Statistics for Windows, Version 23.0 (IBM Corp., Armonk, NY). Prediction of group affiliations by means of logistic regression analyses was then applied to the aged specimens, mixtures and mock case samples.

2.8 Compliance to the MIQE guidelines

To facilitate reliable and unequivocal interpretation of the qPCR results reported herein, information that is rated ‘essential’ according to the MIQE guidelines [39] is reported, where applicable.

3. RESULTS AND DISCUSSION

3.1 RNA quantity and quality of samples

Total RNA concentration and integrity of native samples varied notably between organ tissue types but also within groups (Supplementary Table 2). Regarding concentration, kidney and liver samples exhibited the overall highest values (429 - 1056 ng/µl and 173 - 2440 ng/µl, respectively) while muscle specimens yielded relatively low amounts of total RNA (180 - 389 ng/µl). RIN values varied between 1.8 and 3.9, with muscle samples displaying the highest (2.1 - 3.9) and kidney samples the lowest values (1.9 - 2.2). Results of the dried swabs gave a similar picture, with the exception of skin swabs, which had been obtained by swabbing the foreheads of volunteers and expectedly yielded low total RNA concentrations and RIN values (2.8 - 6.9 ng/µl and ∅ - 1.1, respectively). However, since studies suggest that low RIN values do not adversely affect RT-qPCR profiling results of

miRNAs [40,41] and no difficulties were encountered in a previous study using specimens mimicking forensic body fluid evidence that exhibited low RIN values [42], the recommendation of a minimum RIN of 5 [43,44] for samples to be proceeded was disregarded and RIN determination was omitted for mixtures and mock case samples.

Quantification results for aged specimens, mixtures and mock case samples as well as RIN values for the aged specimens are given in Supplementary Tables 3-6. Briefly, RNA concentration and integrity of samples aged for 28 days did not differ substantially from those of the samples extracted after 24 h, even in the putrefied aged native organ tissue samples. Regarding the mock case samples, concentrations ranged from 2.2 to 63.1 ng/µl and 3.5 to 88.2 ng/µl for the forward spatter and stabbing samples, respectively, while very low total RNA yields were measured for backspatter samples obtained from within the gun barrel (0.12 - 0.52 ng/µl).

Extraction negatives, RT(-)- and H_2O- controls were free of specific amplification (data not shown), indicating that the observed specific amplification truly reflects the miRNA expression status of a given sample.

3.2 Microarray data analysis and selection of candidate miRNAs for organ tissue identification

For an unbiased selection of promising miRNA candidates for organ tissue identification, microarray analyses were performed with samples from the same set of autopsy samples to be used in subsequent RT-qPCR experiments.

After the removal of miRNAs exhibiting signals above background in less than 20 % of samples, a total of 937 miRNAs remained for further analyses, of which 771 miRNAs showed differential expression (corrected p-value < 0.05) in the different tissue samples. Unsupervised clustering of the 50 miRNAs with highest standard deviations clustered the samples according to their biological origin indicating that differences between the groups are the largest contributors to expression variation (Figure 1). One heart sample was grouped between the remaining heart samples and the muscle samples and was falsely clustered among the skeletal muscle samples.

As the heterogeneity of the distinct organ tissues examined in this study might hinder accurate common normalization, quantile normalization and subsequent ranking of the miRNAs according to their expression level was also applied to each tissue type separately (data not shown). This approach was conducted in consultation with the service provider and generated a correct, scientifically sound data set, but did not allow any significance levels to be applied to the results. Hence, for candidate selection, we focused on miRNAs that were either exclusively expressed in one tissue, showed high signal intensities in one tissue type as compared to the remaining six, or displayed differential expression patterns between tissues of which the distinction is of particular interest, e.g. heart versus skeletal muscle.

The 15 candidate miRNAs selected for evaluation in qPCR experiments comprised three markers for brain (hsa-miR-9-5p, hsa-miR-124-3p and hsa-miR-219a-5p) and skeletal muscle or muscle in general (hsa-miR-1-3p, hsa-miR-133a-3p and hsa-miR-206), respectively, two markers for kidney

(hsa-miR-10b-5p and hsa-miR-204-5p), lung (hsa-miR-146b-5p and hsa-miR-233-3p), skin (hsa-miR-203a-3p and hsa-miR-205-5p) and heart (hsa-miR-208b-3p and hsa-miR-499a-5p), respectively, and a single marker for liver (hsa-miR-122-5p).

3.3 Evaluation of candidate miRNAs for organ tissue identification with RT-qPCR in pooled samples, native organ tissues and dried swabs

All miRNA markers selected from the microarray were successfully implemented in RT-qPCR (Supplementary Table 1; amplification efficiencies per marker as detected in the mixture containing all organ tissues are given in Supplementary Table 7).

The initial evaluation of the general expression levels in the investigated organ tissues was performed by means of pooled samples, where all candidate miRNAs showed the highest expression in their respective target organ tissue type, albeit with varying degrees (results summarized in Supplementary Figure 1). Subsequently, the best performing marker per organ tissue type was selected for examination in individual native samples and dry swabs, thereby adding inter-individual differences and forensically realistic sample conditions to the picture. Further, expression levels of the selected miRNAs in blood samples (n = 3; data not shown) were evaluated to test for the influence of the presence of blood, as expected in samples from violent crimes, on the performance of the respective markers.

All three potential brain specific miRNAs, hsa-miR-9-5p, -124-3p and -219a-5p, exhibited considerably higher expression in the pooled brain sample compared to the remaining six tissue types investigated, rendering them well suited markers for the identification of brain tissue, as indicated by previous studies [9,10,12–14,16,17,44,45]. Hsa-miR-219a-5p was selected for further evaluation based on its most highly pronounced separation properties with $\Delta\Delta C_q$- values greater than 9.5 between brain and the remaining tissues (corresponding to a more than 700 times higher expression of hsa-miR-219a-5p in brain than in the remaining tissues). The distinctive difference in expression levels of hsa-miR-219a-5p between brain samples and samples of the other tissues examined was verified in individual native specimens and dried swabs, although with increased intra-group expression variability in the latter (Figure 2A).

Of the miRNAs selected as potential kidney specific markers [12,16–18], hsa-miR-10b-5p displayed sufficiently different expression levels ($\Delta\Delta C_q$- values > 4) between the pooled kidney and brain, liver and heart samples, respectively, but less distinct differences to the pooled samples of muscle and skin. In contrast, hsa-miR-204-5p was only slightly lower expressed in the pooled brain sample compared to kidney, but exhibited pronounced separation properties between kidney and the remaining five organ tissues ($\Delta\Delta C_q$- values > 5). For further analyses we focused on the latter, since the presence of muscle and skin tissue should be considered in samples of violent crimes and might interfere with the identification of kidney tissue, while the differentiation of kidney and brain tissue is expected to be a rarely encountered task. The general expression pattern of hsa-miR-204-5p with a small difference in expression between kidney and brain samples and a considerably lower

expression in liver, lung, skin, heart and muscle samples was confirmed in individual native specimens and dried swabs (Figure 2B).

Hsa-miR-122-5p, the candidate marker for the identification of liver, exhibited pronounced separation properties between the pooled liver sample and the remaining pooled tissue samples, thereby confirming its previously described tissue specific expression [10,11,13–18,47]. In individual native samples and dried swabs, however, considerable intra-group expression variability in the non-target tissues became apparent (Figure 2C). But, with the exception of an outlying dry lung swab, no overlap in expression levels for hsa-miR-122-5p between liver samples and non-target tissue samples was registered.

The candidate miRNAs for the inference of lung tissue, hsa-miR-146b-5p and -223-3p, displayed only minor differences between pooled tissue samples. The latter was chosen for subsequent analyses for its moderately higher $\Delta\Delta C_q$- values (between 3 and 5). The general expression pattern was reproduced in the individual native samples (Figure 2D). In dried swabs, however, expression levels in liver and muscle samples overlapped with those in lung and two out of three skin swabs displayed even higher expression than the lung samples. Moreover, while the expression levels in blood for the other miRNAs investigated fell well within the range or were lower than those of the respective non-target tissue samples, the expression of hsa-miR-223-3p in blood was even higher than in lung samples. Hence, hsa-miR-223-3p should not be employed as a marker for the forensic identification of lung tissue. As no additional promising miRNAs had been derived from the microarray results and no undisputed candidates could be retrieved from the literature [11,14–18], the inference of lung tissue was excluded from the agenda and will not be further discussed hereinafter.

Hsa-miR-203a-3p and -205-5p were selected from the microarray results as candidate miRNAs for the identification of skin and both displayed good separation properties in the pooled samples. Due to the slightly higher $\Delta\Delta C_q$- values, hsa-miR-205-5p was investigated in individual native samples and dried swabs and displayed no overlapping expression levels between skin samples and the remaining organ tissues and a notably small intra-group variance among the skin samples (Figure 2E). Both miRNAs have previously been described as candidate markers for saliva in forensic BFI [19,20,42,48–50]. However, the herein presented results indicate that they rather are markers for the presence of epithelial cells in general. As most of the ‘saliva’ samples employed in available studies on BFI had been obtained via buccal swabs, the presence of epithelial markers is highly plausible.

Both potential heart specific markers, hsa-miR-208b-3p and -499a-5p, exhibited distinctively higher expression in the pooled samples of both muscle types (i.e. heart and skeletal muscle) than in non-muscle samples, with hsa-miR-208b-3p lacking specific amplification in the kidney, liver, lung and skin pooled samples, respectively. However, expression of hsa-miR-499a-5p differed only slightly between heart and skeletal muscle, whereas the respective $\Delta\Delta C_q$- value for hsa-miR-208b-3p was 2.7. In individual native samples and dried swabs, the difference in expression of hsa-miR-208b-3p between muscle (in general) and non-muscle samples was verified and moreover,

expression levels in heart and skeletal muscle samples were non-overlapping (Figure 2F). While both markers have been described varyingly as either heart specific or specific for muscle in general [11,15,17,51], our data suggest that hsa-miR-499a-5p is more likely a marker for muscle in general, whereas hsa-miR-208b-3p showed higher expression in heart than in skeletal muscle.

Two additional candidates for muscle in general, hsa-miR-1-3p and -133a-3p, were screened in pooled samples and displayed good separation properties, but were not further pursued as they did not improve discrimination beyond that obtained by hsa-miR-208b-3p.

The specific high expression of hsa-miR-206 in skeletal muscle, as indicated by the microarray results and previous studies [11,13,17,51], was reproduced in pooled samples with $\Delta\Delta C_q$- values greater than 9.5 between skeletal muscle and the remaining organ tissue types, including heart. The distinctly higher expression in skeletal muscle samples was further substantiated in the individual native samples and dried swabs, albeit to a lesser extend in the latter (Figure 2G).

Kruskal- Wallis testing revealed that the differences in expression between organ tissue types were statistically significant for the six miRNAs evaluated in individual samples ($p < 0.01$, data not shown).

Binary logistic regression analysis (BLR) as a means of predicting group affiliation of unknown samples (similar to the approach described by Hanson et al. [52]) was conducted per organ tissue type separately for native and dry specimens. Resulting formulae as well as the model validation parameters are given in Table 1. In all cases, except the identification of liver in dry swabs, complete separation of target versus non-target tissue was achieved. In consequence, the maximum likelihood estimation procedures gave invalid results and therefore are not reported. Yet, the calculated model χ^2- values with corresponding p- values < 0.001 indicate that the respective included markers had highly significant effects on predicting the respective target organ tissue and, as suggested by the Nagelkerke R^2, 100 % of variation in the respective data set is explained by the respective model (84.4% for the identification of liver in dry swabs). Further, the Hosmer- Lemeshow test showed that the models' predictions fit the observed data, as the null hypothesis, stating that there was no difference between the observed and the predicted values, was not rejected (p-values > 0.05). The area under the receiver operating characteristic (ROC) curve was 0.970 (95% confidence interval 0.905 – 1.000) for the identification of liver in dry swabs and 1.000 for all remaining BLR calculations (data not shown).

The lung swab wrongly predicted to be stemming from a liver sample by BLR with hsa-miR-122-5p expression (ΔC_q- values) as input variable was identified as an outlier (studentized residual -2.114, Cook distance 1.207). The native lung sample of the same individual (case 12, Supplementary Table 2) also displayed the highest expression for this marker among all non-liver samples. However, as the available information about the case yielded no indication for deviating characteristics, a contamination during the sampling process appears most likely. To determine whether this case was indeed an isolated outlier or if hsa-miR-122-5p is not suitable for identifying samples containing liver tissue due to potential false positive findings, the marker was further evaluated in aged

specimens, mixtures and mock case samples containing lung tissue.

3.4 Evaluation of expression stability in aged samples

To assess the miRNA markers' applicability in challenged samples, expression stability was examined in dry and native samples of their respective target organ tissue aged for 28 days. All investigated dry aged swabs (n = 2 per organ tissue type) expectedly yielded analyzable miRNA expression results, as previously shown with miRNA markers that were stable in dry forensic samples aged for up to 36 years [42]. When applying the respective BLR formulae for dry samples, all but one skin sample were classified correctly, including the lung samples being classified as non-liver samples (Supplementary Material 3). Further, hsa-miR-205-5p was still more strongly expressed in the wrongly classified skin sample than in all non-skin samples, so that adjusting the BLR formula by deriving it from a larger sample set will probably improve classification.

For the native specimens exposed to the elements (n = 2 per organ tissue type), analyzable results were obtained for five samples (both brain samples, both skeletal muscle samples and one skin sample) and in all cases correct classification was obtained employing the BLR formulae for native samples (Supplementary Material 3). The respective ΔC_q- values seemingly indicate an even higher expression of the organ specific miRNAs, but analysis of the raw data demonstrated that this was rather due to a markedly lower expression in the reference genes due to degradation, than higher expression of the organ specific markers (data not shown). Further, while expression of the miRNA markers was observed in 13 out of the 14 samples, reference genes did not show specific amplification in nine samples, barring expression normalization and hence analysis.

In sum, the results of this first approach on employing miRNA based organ tissue identification (OTI) on aged samples on the one hand confirm the suitability of miRNAs for the analysis of aged forensic material, even, to some extent, in decomposing tissue. On the other hand, a normalization strategy employing small nuclear and small nucleolar RNAs (snRNAs and snoRNAs, respectively) might be less suited for putrefied tissue samples as they appear to be more prone to degradation than the shorter miRNAs.

3.5 Mixtures

All selected miRNA candidates were further tested in mixtures containing two to four organ tissue types in different ratios mimicking possible mixtures as encountered in violent crimes (n = 20; Supplementary Table 4) to ascertain whether correct inference of organ tissue types could be achieved when employing the proposed BLR approach (Supplementary Material 3).

Most importantly, no false positive identifications, that could unduly incriminate a potential suspect, occurred in any of the mixtures, including correct classification of all samples containing lung tissue as "non-liver", highlighting the markers' tissue specificity even in the presence of multiple other organ tissues. Moreover, the presence of brain, liver and skeletal muscle tissue was correctly inferred in all cases, even in unbalanced mixtures with the target organ tissue being the minor component.

The positive identification of heart, kidney and skin, however, was less successful. Firstly, skin could not be identified correctly in three of the 20 mixtures. Though, as the presence of skin is usually not in question in samples from violent crimes, its identification in mixtures is of moderate interest. Secondly, for hsa-miR-204-5p, the expression level in three of the five mixtures containing kidney lay within the range of single source brain samples and were hence falsely classified as not containing kidney tissue. However, the expression was still higher than in single source samples of the remaining five organ tissues (liver, lung, skin, heart and muscle) and mixtures thereof, so that the obtained ΔC_q- values can in fact be interpreted as indicative of kidney or brain tissue in a mixture. The distinction between the presence of kidney or brain can then be determined with the BLR model employing hsa-miR-219a-5p as it is implausible for samples from violent crimes to contain material from both anatomically distant organ tissue types. Adjusting the respective BLR formula by deriving it from a larger sample set will probably also improve classification. Thirdly, identification of heart tissue was successful in mixtures containing only skin and heart tissue, but failed in those additionally containing skeletal muscle tissue.

3.6 Mock case samples

Mock case samples were prepared to further evaluate the markers' reliability and the suitability of the devised BLR models in realistic forensic type samples (Supplementary Material 3).

In the shooting experiments, organ tissues, i.e. brain, kidney, liver, lung, heart and muscle, were present in mixtures with venous blood, as expected in samples from violent crimes. In the forward spatter samples, concordant with the sufficiently high total RNA concentrations, analyzable expression results were obtained in all cases - with correct identification for all but the kidney samples. Moreover, despite very low total RNA yields, analyzable results were obtained for 13 of the 18 backspatter samples recovered from the inside of gun barrels. Brain, heart and muscle tissue were inferred correctly in all cases, while the one kidney and two liver samples with analyzable results were miss-classified.

In the mock stabbings, we focused on the aspect of mixtures of different organs as they are expected to be encountered in casework samples. Since the presence of blood did not influence the assay's performance, as shown in the shooting experiments, we did not add liquid blood as a separate component apart from residual blood in the tissues. In concordance with the findings of Lindenbergh et al. [5], we were able to show that cells from the undermost layer can remain on a blade after extracting the knife through the wound channel across the upper tissue strata including skin with adhering adipose tissue in a sufficient amount to identify the respective tissue type. Specifically, in two of three cases each, liver and heart muscle could be correctly inferred, while kidney identification again failed in all three biological replicates.

Further, consistent with the results for the mixtures, no false positive classifications for any of the inner organs occurred in samples taken from blades that were pierced through skin and muscle only. However, the presence of skin was not detected in these samples, which is most likely a result of the

vertical cut through the comparatively thin skin and hence only few cells sticking to the blade. But as discussed above the presence of skin is usually not in question in this type of samples.

The results of the mock case samples taken together confirm the general applicability of our approach to realistic forensic samples encountered at crime scenes, even if comprising only trace amounts of material.

3.7 Inferring the presence of heart, skeletal muscle or both in a mixture by multinomial logistic regression

While the identification of skeletal muscle was successful for all sample types, the identification of heart failed in six out of eight samples in which both skeletal muscle and heart were present (5/5 mixtures of native samples and 1/3 mock stabbings). Considering that a sample recovered from a trauma-inflicting object used in a violent crime which contains heart tissue inevitably also contains skeletal muscle cells and that the presence of heart tissue has high evidential value as it reflects a serious and potentially lethal injury, we examined the possibility of differentiation of heart tissue, skeletal muscle tissue, both muscle tissue types in the same sample or the sample lacking muscle tissue altogether. Plotting the expression values of hsa-miR-208b-3p and hsa-miR-206 in a two-dimensional scatter plot confirmed differential clustering of heart and skeletal muscle samples as well samples devoid of muscle tissue, while samples containing both tissues clustered in between the two muscle types (Figure 3). Multinomial logistic regression analysis with hsa-miR-208b-3p and hsa-miR-206 expression (ΔC_q- values) as input variables was hence conducted *a posteriori* with the entire sample set and 100 % of cases correctly classified (Table 2 and Supplementary Material 3). Due to the complete separation of all cases, the maximum likelihood estimation procedures put out invalid results which therefore are not reported. The calculated final model χ^2- value of 288.658 (p-value < 0.001) indicates, however, that the full model predicts the dependent variable statistically significantly better than the intercept-only model alone and Nagelkerke's R^2 implies that 100 % of variation in the data set is explained by the model.

3.8 Decision tool

For a convenient application of our findings we developed a simple Excel tool that is available as Supplementary Material (Supplementary Material 4). When fed with a sample's ΔC_q- values corresponding to the organ specific miRNA markers presented herein it gives out the decision whether a particular organ tissue is present or absent in a given sample, based on the above-described BLR models for brain, kidney, liver and skin and the MLR model for the inference of heart, skeletal muscle or both muscle types present in a mixture.

3.9 Quantification and assessment of STR typing success of co-extracted DNA

DNA co-extraction was successfully applied to the aged native specimens as well as the mock case samples.

For the aged native samples, sufficiently high DNA concentrations were detected in all cases,

ranging from 0.01 to 46.22 ng/µl (Supplementary Table 3). Although the PowerQuant System indicated DNA degradation in ten of the 14 samples, STR profiles sufficient for identification purposes were obtained even from the samples most affected by degradation (data not shown).

Mock stabbing samples and forward spatter samples from the experimental shootings exhibited only slight degradation and yielded DNA amounts reliably sufficient for STR analysis (Supplementary Table 6 and data not shown, respectively). Analogous to total RNA, DNA concentrations were quite low in most of the backspatter samples (1.1 – 64.4 pg/µl; Supplementary Table 5). Nevertheless, partial STR profiles sufficient for identification purposes were obtained in all cases (data not shown).

4. CONCLUSION

Herein, we present the first approach on forensic organ tissue identification (OTI) via miRNA expression analysis based on a thoroughly methodologically validated qPCR procedure including an empirically derived normalization strategy and unbiased statistical decision making. The method proved to be completely compatible with simultaneous forensic DNA analysis and is applicable to realistic forensic type samples including mixtures, aged and degraded material.

With hsa-miR-219a-5p, hsa-miR-122-5p, hsa-miR-205-5p, hsa-miR-208b-3p and hsa-miR-206 we describe robust markers for the inference of brain, liver, skin, and heart and skeletal muscle, respectively, with only an isolated case of false positive classification, most likely due to contamination in the sampling process. Further, hsa-miR-9-5p and hsa-miR-124-3p as well as hsa-miR-499a-5p, hsa-miR-1-3p and hsa-miR-133a-3p were found to be promising markers for the identification of brain and muscle in general, respectively.

Hsa-miR-204-5p appeared to be a promising marker for the identification of kidney in single source samples, but posed problems when applied in mixtures and forensic type samples as the expression levels of the marker are similar in samples containing kidney and brain tissue. Hence, for the inference of kidney tissue the implementation of an additional miRNA marker is advisable. As hsa-miR-223-3p and hsa-miR-146b-5p did not show sufficiently distinct expression levels between lung and the remaining tissues, a model for the identification of lung is still lacking.

Future research should comprise larger sample sets, especially regarding mixtures and samples with low RNA yields, to assess sensitivity and improve classification, and possibly expand the range of identifiable organs to potentially even organ subsections. Also, OTI in putrefied samples might be enhanced by establishing a degradation resistant normalization strategy employing miRNAs, as these are less prone to degradation than the longer snRNAs and snoRNAs.

Even greater advances in the forensic identification of cell types can be envisioned by the application of massive parallel sequencing to OTI, for example by combining miRNAs, mRNAs and methylation patterns in highly cell type specific signatures.

ACKNOWLEDGMENTS

We would like to thank Prof. Dr. Johanna Preuß-Wössner, Prof. Dr. Burkhard Madea and the forensic pathologists at the Institutes of Forensic Medicine in Kiel and Bonn for their help with sample collection as well as all volunteers for their kind donation of sample material. We are grateful to the Landeskriminalamt Schleswig-Holstein for facilitating the experimental shooting session in their premises as well Katharina Pöhls, Jörg Pirch and Lena Brandes for their helpful contributions. The expert technical assistance of Stefanie Petzel is gratefully acknowledged.

Further, we thank the Deutsche Forschungsgemeinschaft (DFG) for funding this project (CO 992/3-3).

REFERENCES

[1] A. Kimura, H. Ikeda, S. Yasuda, K. Yamaguchi, T. Tsuji, Brain tissue identification based on myosin heavy chain isoforms, Int. J. Legal Med. 107 (4) (1995) 193–196.

[2] K. Takahama, Forensic application of organ-specific antigens, Forensic Sci. Int. 80 (1-2) (1996) 63–69.

[3] Y. Seo, E. Kakizaki, K. Takahama, A sandwich enzyme immunoassay for brain S-100 protein and its forensic application, Forensic Sci. Int. 87 (2) (1997) 145–154.

[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 (1) (2004) 1–6.

[5] A. Lindenbergh, M. van den Berge, R.J. Oostra, C. Cleypool, A. Bruggink, A. Kloosterman, T. Sijen, Development of a mRNA profiling multiplex for the inference of organ tissues, Int. J. Legal Med. 127 (5) (2013) 891–900.

[6] M. van den Berge, D. Wiskerke, R.R.R. Gerretsen, J. Tabak, T. Sijen, DNA and RNA profiling of excavated human remains with varying postmortem intervals, Int. J. Legal Med. 130 (6) (2016) 1471–1480.

[7] K. Virkler, I.K. Lednev, Analysis of body fluids for forensic purposes: From laboratory testing to non-destructive rapid confirmatory identification at a crime scene, Forensic Sci. Int. 188 (1-3) (2009) 1–17.

[8] J.H. An, K.J. Shin, W.I. Yang, H. Y. Lee, Body fluid identification in forensics, BMB Rep. 45 (10) (2012) 545–553.

[9] T. Sijen, Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers, Forensic Sci. Int. Genet. 18 (2015) 21–32.

[10] O. Barad, E. Meiri, A. Avniel, R. Aharonov, A. Barzilai, I. Bentwich, U. Einav, S. Gilad, P. Hurban, Y. Karov, E.K. Lobenhofer, E. Sharon, Y.M. Shiboleth, M. Shtutman, Z. Bentwich, P. Einat, MicroRNA expression detected by oligonucleotide microarrays: System establishment and expression profiling in human tissues, Methods. 14 (12) (2004) 2486–2494.

[11] L.F. Sempere, S. Freemantle, I. Pitha-Rowe, E. Moss, E. Dmitrovsky, V. Ambros, Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation, Genome Biol. 5 (3) (2004) R13.

[12] Y. Sun, S. Koo, N. White, E. Peralta, C. Esau, N.M. Dean, R.J. Perera, Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs, Nucleic Acids Res. 32 (22) (2004) e188.

[13] S. Baskerville, D.P. Bartel, Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes, RNA. 11 (3) (2005) 241–247.

[14] P. Landgraf, M. Rusu, R. Sheridan, A. Sewer, N. Iovino, A. Aravin, S. Pfeffer, A. Rice, A.O. Kamphorst, M. Landthaler, C. Lin, N.D. Socci, L. Hermida, V. Fulci, S. Chiaretti, R. Foà, J. Schliwka, U. Fuchs, A. Novosel, R.-U. Müller, B. Schermer, U. Bissels, J. Inman, Q. Phan, M. Chien, D.B. Weir,

R. Choksi, G. De Vita, D. Frezzetti, H.-I. Trompeter, V. Hornung, G. Teng, G. Hartmann, M. Palkovits, R. Di Lauro, P. Wernet, G. Macino, C.E. Rogler, J.W. Nagle, J. Ju, F.N. Papavasiliou, T. Benzing, P. Lichter, W. Tam, M.J. Brownstein, A. Bosio, A. Borkhardt, J.J. Russo, C. Sander, M. Zavolan, T. Tuschl, A Mammalian microRNA Expression Atlas Based on Small RNA Library Sequencing, Cell. 129 (7) (2007) 1401–1414.

[15] Y. Liang, D. Ridzon, L. Wong, C. Chen, Characterization of microRNA expression profiles in normal human tissues, BMC Genomics. 8 (2007) 166.

[16] R. Bargaje, M. Hariharan, V. Scaria, B. Pillai, Consensus miRNA expression profiles derived from interplatform normalization of microarray data, RNA. 16 (1) (2010) 16–25.

[17] Z. Guo, M. Maki, R. Ding, Y. Yang, B. Zhang, L. Xiong, Genome-wide survey of tissue-specific microRNA and transcription factor regulatory networks in 12 tissues, Sci. Rep. 4 (2014) 5150.

[18] E.J. Lee, M. Baek, Y. Gusev, D.J. Brackett, G.J. Nuovo, T.D. Schmittgen, Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors, RNA. 14 (1) (2007) 35–42.

[19] 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 (4) (2013) 967–971.

[20] E.J. Omelia, M.L. Uchimoto, G. Williams, Quantitative PCR analysis of blood- and saliva-specific microRNA markers following solid-phase DNA extraction, Anal. Biochem. 435 (2) (2013) 120–122.

[21] 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.

[22] M.W. Pfaffl, A new mathematical model for relative quantification in RT-PCR, Nucleic Acids Res. 29 (9) (2001) e45.

[23] 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.

[24] S.A. Bustin, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems, J. Mol. Endocrinol. 29 (1) (2002) 23–39.

[25] S. Bustin, T. Nolan, Data analysis and interpretation, in: S. Bustin (Ed.), A-Z of Quantitative PCR, first ed., International University Line, La Jolla, CA, 2004, pp.439-492.

[26] K.J. Livak, T.D. Schmittgen, Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method, Methods. 25 (4) (2001) 402–408.

[27] M.W. Pfaffl, Validities of mRNA quantification using recombinant RNA and recombinant external calibration curves in real-time RT-PCR, Biotechnol. Lett. 23 (2001) 275–282.

[28] J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, Genome Biol. 3 (7) (2002) research0034.1.

[29] E. Sauer, I. Babion, B. Madea, C. Courts, An evidence based strategy for normalization of quantitative PCR data from miRNA expression analysis in forensic organ tissue identification, Forensic Sci. Int. Genet. 13 (2014) 217–223.

[30] C. Schyma, C. Lux, B. Madea, C. Courts, The “triple contrast” method in experimental wound ballistics and backspatter analysis, Int. J. Legal Med. 129 (5) (2015) 1027–1033.

[31] M. Grabmüller, P. Cachée, B. Madea, C. Courts, How far does it get?—The effect of shooting distance and type of firearm on the simultaneous analysis of DNA and RNA from backspatter recovered from inside and outside surfaces of firearms, Forensic Sci. Int. 258 (2016) 11–18.

[32] B.C.M. Pang, B.K.K. Cheung, Double swab technique for collecting touched evidence, Leg. Med. 9 (4) (2007) 181–184.

[33] S. Griffiths-Jones, The microRNA Registry, Nucleic Acids Res. 32 (2004) D109–111.

[34] M.E. Ritchie, J. Silver, A. Oshlack, M. Holmes, D. Diyagama, A. Holloway, G.K. Smyth, A comparison of background correction methods for two-colour microarrays, Bioinformatics. 23 (20) (2007) 2700–2707.

[35] A. Schroeder, O. Mueller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, S. Lightfoot, W. Menzel, M. Granzow, T. Ragg, The RIN: an RNA integrity number for assigning integrity values to

RNA measurements, BMC Mol. Biol. 7 (2006) 3.

[36] J.M. Ruijter, C. Ramakers, W.M. Hoogaars, Y. Karlen, O. Bakker, M.J. van den Hoff, A.F. Moorman, Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data, Nucleic Acids Res. 37 (6) (2009) e45.

[37] S. Pabinger, S. Rödiger, A. Kriegner, K. Vierlinger, A. Weinhäusel, A survey of tools for the analysis of quantitative PCR (qPCR) data, Biomol. Detect. Quantif. 1 (2014) 23–33.

[38] J.L. Guthrie, C. Seah, S. Brown, P. Tang, F. Jamieson, S.J. Drews, Use of Bordetella pertussis BP3385 to establish a cutoff value for an IS481-targeted real-time PCR assay, J. Clin. Microbiol. 46 (11) (2008) 3798–3799.

[39] S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, C.T. Wittwer, The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments, Clin. Chem. 55 (4) (2009) 611–622.

[40] M. Jung, A. Schaefer, I. Steiner, C. Kempkensteffen, C. Stephan, A. Erbersdobler, K. Jung, Robust MicroRNA stability in degraded RNA preparations from human tissue and cell samples, Clin. Chem. 56 (6) (2010) 998–1006..

[41] J.S. Hall, J. Taylor, H.R. Valentine, J.J. Irlam, A. Eustace, P.J. Hoskin, C.J. Miller, C.M. West, Enhanced stability of microRNA expression facilitates classification of FFPE tumour samples exhibiting near total mRNA degradation, Br J Cancer. 107 (4) (2012) 684–694.

[42] E. Sauer, A.-K. Reinke, C. Courts, Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR, Forensic Sci. Int. Genet. 22 (2016) 89–99.

[43] C. Becker, A. Hammerle-Fickinger, I. Riedmaier, M.W. Pfaffl, mRNA and microRNA quality control for RT-qPCR analysis, Methods. 50 (4) (2010) 237–243.

[44] 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 (19) (2006) 1601–1613.

[45] Y.K. Adlakha, N. Saini, Brain microRNAs and insights into biological functions and therapeutic potential of brain enriched miRNA-128, Mol. Cancer. 13 (2014) 33.

[46] N. Ludwig, P. Leidinger, K. Becker, C. Backes, T. Fehlmann, C. Pallasch, S. Rheinheimer, B. Meder, C. Stähler, E. Meese, A. Keller, Distribution of miRNA expression across human tissues, Nucleic Acids Res. 44 (8) (2016) 3865-3877.

[47] C. Jopling, Liver-specific microRNA-122: Biogenesis and function, RNA Biol. 9 (2) (2012) 137–142.

[48] E.K. Hanson, H. Lubenow, J. Ballantyne, Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs, Anal. Biochem. 387 (2) (2009) 303–314.

[49] C. Courts, B. Madea, Specific Micro-RNA Signatures for the Detection of Saliva and Blood in Forensic Body-fluid Identification, J. Forensic Sci. 56 (6) (2011) 1464–1470.

[50] Z. Wang, J. Zhang, W. Wei, D. Zhou, H. Luo, X. Chen, Y. Hou, Identification of Saliva Using MicroRNA Biomarkers for Forensic Purpose, J. Forensic Sci. 60 (3) (2015) 702–706.

[51] J. McCarthy, MicroRNA-206: The skeletal muscle-specific myomiR, Biochim. Biophys. Acta - Gene Regul. Mech. 1779 (11) (2008) 682–691.

[52] E.K. Hanson, M. Mirza, K. Rekab, J. Ballantyne, The identification of menstrual blood in forensic samples by logistic regression modeling of miRNA expression, Electrophoresis. 35 (21-22) (2014) 3087–3095.

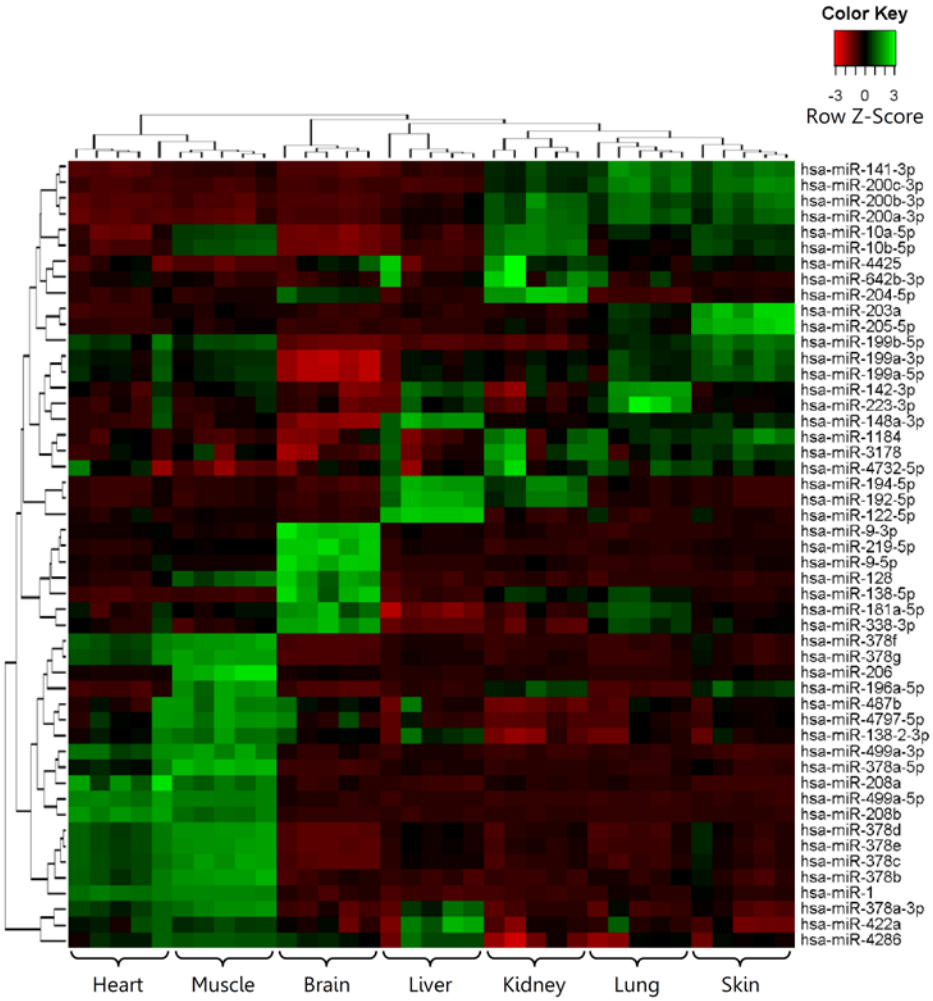


Figure 1: Heat map and unsupervised hierarchical clustering of miRNA expression in 35 samples from seven human organ tissues (five biological replicates per brain, kidney, liver, lung, skin, heart and muscle) displaying the 50 miRNAs with highest standard deviations after quantile normalization had been applied to the whole data set. The normalized log-transformed Hy3 values were used for the analysis applying the complete-linkage method together with the Euclidean distance measure. Red color represents an expression level below the reference channel while green color represents expression higher than the reference.

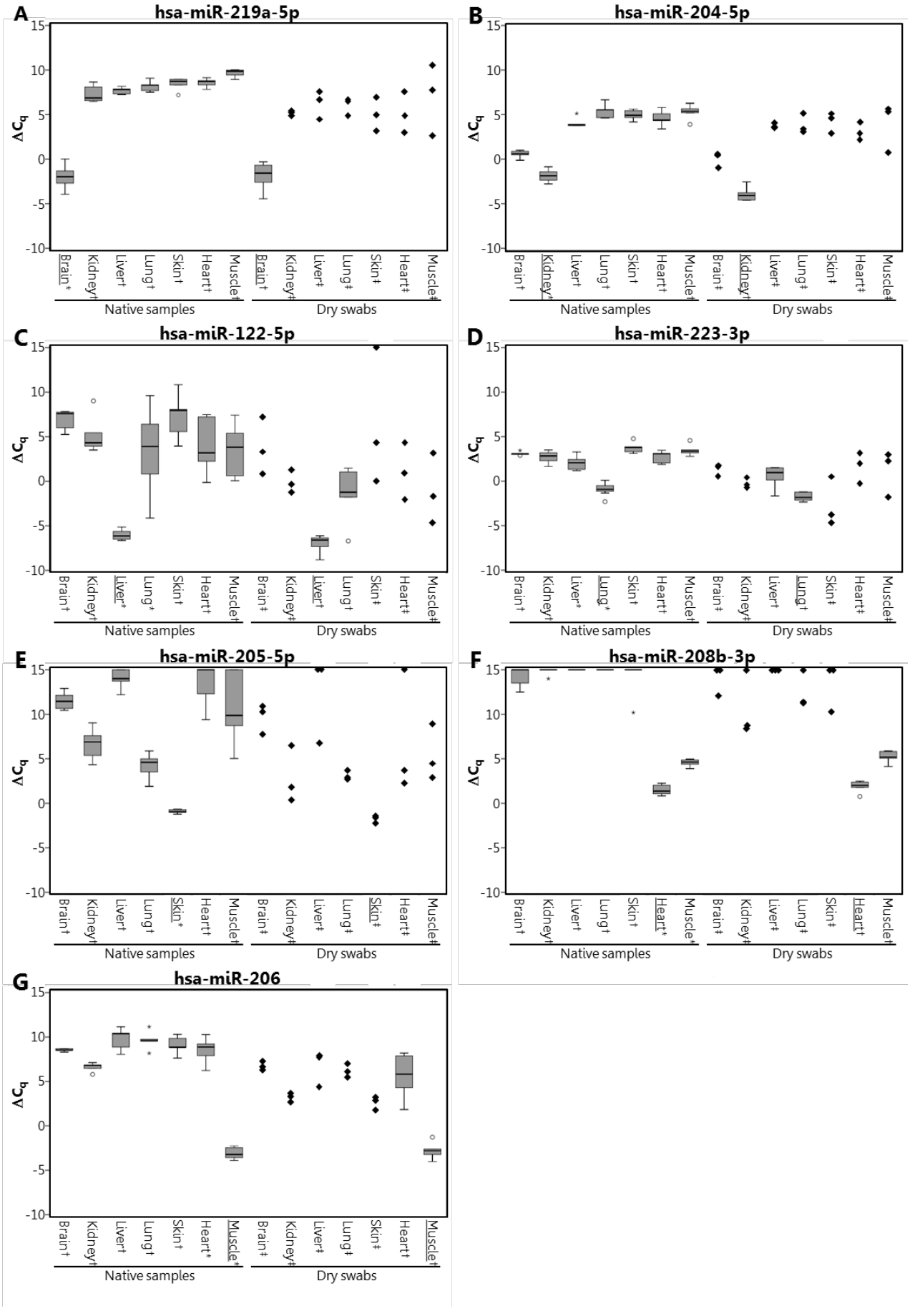

A
hsa-miR-219a-5p
B
hsa-miR-204-5p
C
hsa-miR-122-5p
D
hsa-miR-223-3p
E
hsa-miR-205-5p
F
hsa-miR-208b-3p
G
hsa-miR-206
ΔCq
Native samples
Dry swabs
Brain
Kidney
Liver
Lung
Skin
Heart
Muscle

Figure 2: Expression patterns of the seven candidate miRNAs for the identification of brain (A), kidney (B), liver (C), lung (D), skin (E), heart (F) and skeletal muscle (G). Box blots represent the lower, median and upper quartile of a respective sample set, with the whiskers indicating the (non-extreme) maximum and minimum values, and outliers (1.5x interquartile range) and extreme outliers (3x interquartile range) being depicted as circles and asterisks, respectively. For groups with three tested samples, values of individual samples are given as diamonds instead of box plots. For all samples without specific amplification ΔC_q-values were arbitrarily set to the value 15. (*, n = 10; †, n = 5; ‡, n = 3)

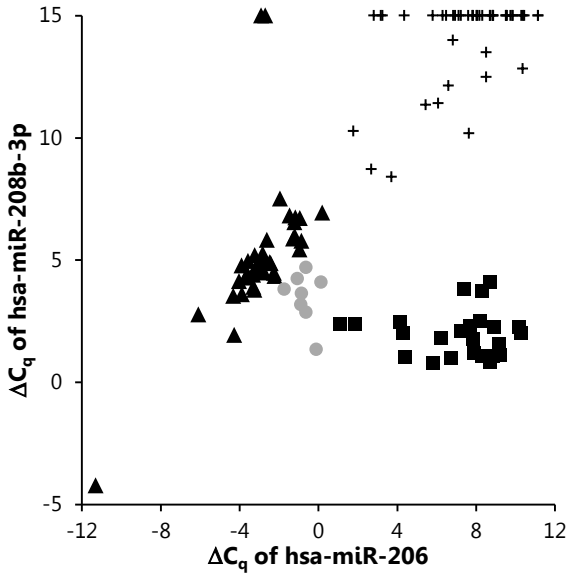


Figure 3: Scatter plot of the ΔC_q-values of hsa-miR-206 (x-axis) and hsa-miR-208b-3p (y-axis) to cluster samples containing either skeletal muscle tissue (triangles; n = 39), heart tissue (squares; n = 24), both tissue types (grey circles; n = 8) or no muscle tissue (crosses; n=45).

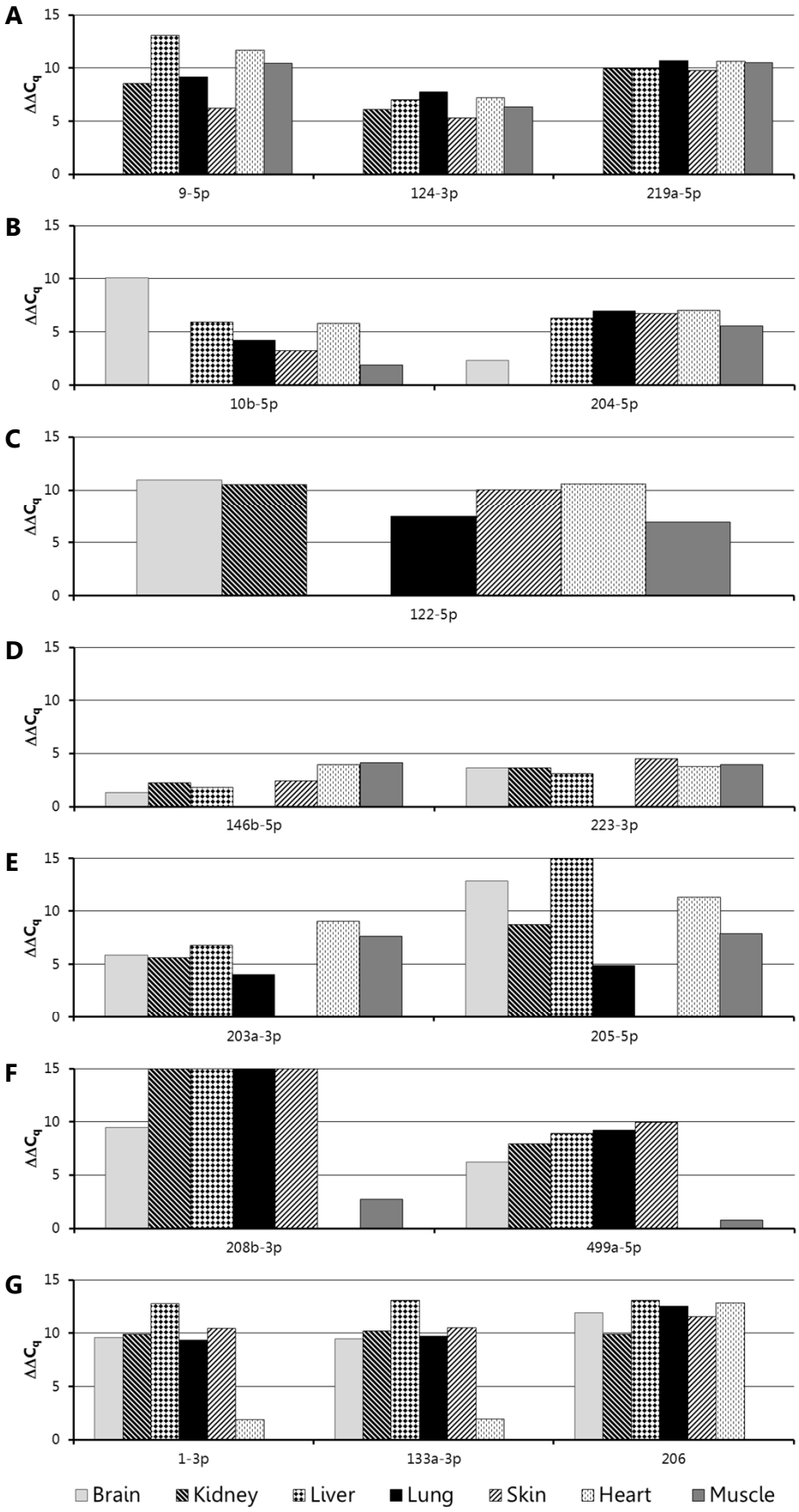
Table 1: Binary logistic regression model and model validation parameters per organ tissue type - separately for native and dry specimens.

			Model validation			Classification in the training sets		
Organ tissue	Sample type	Binary logistic regression formula	Model χ^2	Nagel-kerke R^2	Hosmer-Lemeshow test	True positive	False positive	Correct classification rate
Brain	Native	logit = 17.961 – 5.525 × ΔC_q(hsa-miR-219a-5p)	44.987*	1.000	0.000, df = 8 (p = 1.000)	10/10	0/30	100%
	Dry	logit = 13.841 – 11.504 × ΔC_q(hsa-miR-219a-5p)	24.085*	1.000	0.000, df = 8 (p = 1.000)	5/5	0/18	100%
Kidney	Native	logit = -19.347 – 39.560 × ΔC_q(hsa-miR-204-5p)	44.987*	1.000	0.000, df = 8 (p = 1.000)	10/10	0/30	100%
	Dry	logit = -34.426 – 19.946 × ΔC_q(hsa-miR-204-5p)	24.085*	1.000	0.000, df = 8 (p = 1.000)	5/5	0/18	100%
Liver	Native	logit = -127.587 – 27.521 × ΔC_q(hsa-miR-122-5p)	47.674*	1.000	0.000, df = 7 (p = 1.000)	10/10	0/35	100%
	Dry	logit = -9.879 – 1.698 × ΔC_q(hsa-miR-122-5p)	19.079*	0.844	1.054. df = 6 (p = 0.983)	5/5	1/20	96% (24/25)
Skin	Native	logit = 8.260 – 12.366 × ΔC_q(hsa-miR-205-5p)	44.987*	1.000	0.000, df = 7 (p = 1.000)	10/10	0/30	100%
	Dry	logit = -10.540 – 18.191 × ΔC_q(hsa-miR-205-5p)	17.225*	1.000	0.000, df = 8 (p = 1.000)	3/3	0/18	100%
Heart	Native	logit = 7.955 – 4.653 × ΔC_q(hsa-miR-208b-3p) + 2.915 × ΔC_q(hsa-miR-206)	47.674*	1.000	0.000, df = 7 (p = 1.000)	10/10	0/35	100%
	Dry	logit = 26.931 – 7.117 × ΔC_q(hsa-miR-208b-3p) + 4.183 × ΔC_q (hsa-miR-206)	25.020*	1.000	0.000, df = 6 (p = 1.000)	5/5	0/20	100%
Muscle	Native	logit = 8.558 – 4.451 × ΔC_q(hsa-miR-206)	47.674*	1.000	0.000, df = 7 (p = 1.000)	10/10	0/35	100%
	Dry	logit = 2.785 – 11.212 × ΔC_q (hsa-miR-206)	25.020*	1.000	0.000, df = 6 (p = 1.000)	5/5	0/20	100%

*$p < 0.001$

Table 2: Formulae of the multinomial logistic regression model for the differentiation of heart, skeletal muscle, both muscle tissue types in a mixture sample or the sample lacking muscle tissue altogether as derived *a posteriori* employing the entire sample set with 'no muscle tissue present' as the reference category.

Organ tissue	Multinomial logistic regression formulae
Heart	$\text{logit} = 145.282 - 25.416 \times \Delta C_q(\text{hsa-miR-208b-3p}) + 12.750 \times \Delta C_q(\text{hsa-miR-206})$
Muscle	$\text{logit} = -56.521 + 11.844 \times \Delta C_q(\text{hsa-miR-208b-3p}) - 48.969 \times \Delta C_q(\text{hsa-miR-206})$
Both	$\text{logit} = 153.314 - 22.926 \times \Delta C_q(\text{hsa-miR-208b-3p}) - 12.101 \times \Delta C_q(\text{hsa-miR-206})$



SUPPLEMENTARY FIGURE 1: Relative expression patterns of candidate miRNAs in pooled samples of organ tissues depicted as $\Delta\Delta C_q$-values with the respective targeted organ sample used as a calibrator sample. Candidate miRNAs for brain (A), kidney (B), liver (C), lung (D), skin (E), heart muscle (F) and skeletal muscle (G). Pooled samples consisted of ten donor samples per organ tissue type. For all samples without specific amplification $\Delta\Delta C_q$-values were arbitrarily set to 15.

SUPPLEMENTARY TABLE 1: Specifications of the 15 selected candidate miRNAs for organ tissue identification and the reference genes.

Organ tissue	Official gene symbol	miRBase-/ NCBI-mature sequence accession	Previous miRBase-IDs/ NCBI-Aliases	TaqMan Assay ID	Target sequence (amplicon length [nucleotides])
Brain	hsa-miR-9-5p	MIMAT0000441	hsa-miR-9	000583	UCUUUGGUUAUCUAGCUGUAUGA (23)
	hsa-miR-124-3p	MIMAT0000422	hsa-miR-124a; hsa-miR-124	001182	UAAGGCACGCGGUGAAUGCC (20)
	hsa-miR-219a-5p	MIMAT0000276	hsa-miR-219	000522	UGAUUGUCCAAACGCAAUUCU (21)
Kidney	hsa-miR-10b-5p	MIMAT0000254	hsa-miR-10b	002218	UACCCUGUAGAACCGAAUUUGUG (23)
	hsa-miR-204-5p	MIMAT0000265	hsa-miR-204	000508	UUCCCUUUGUCAUCCUAUGCCU (22)
Liver	hsa-miR-122-5p	MIMAT0000421	hsa-miR-122a; hsa-miR-122	002245	UGGAGUGUGACAAUGGUGUUUG (22)
Lung	hsa-miR-146b-5p	MIMAT0002809	hsa-miR-146b	001097	UGAGAACUGAAUUCCACAGGCU (22)
	hsa-miR-223-3p	MIMAT0000280	hsa-miR-223	002295	UGUCAGUUUGUCAAAUACCCCA (22)
Skin	hsa-miR-203a-3p	MIMAT0000264	hsa-miR-203a	000507	GUGAAAUGUUUAGGACCACUAG (22)
	hsa-miR-205-5p	MIMAT0000266	hsa-miR-205	000509	UCCUUCAUUCCACCGGAGUCUG (22)
Heart muscle	hsa-miR-208b-3p	MIMAT0004960		002290	AUAAGACGAACAAAAGGUUUGU (22)
	hsa-miR-499a-5p	MIMAT0002870	hsa-miR-499; hsa-miR-499-5p	001352	UUAAGACUUGCAGUGAUGUUU (21)
Muscle / Skeletal muscle	hsa-miR-1-3p	MIMAT0000416	hsa-miR-1	002222	UGGAAUGUAAAGAAGUAUGUAU (22)
	hsa-miR-133a-3p	MIMAT0000427		002246	UUUGGUCCCCUUCAACCAGCUG (22)
	hsa-miR-206	MIMAT0000462		000510	UGGAAUGUAAGGAAGUGUGUGG (22)
Reference genes	SNORD24	NR_002447	RNU24; U24	001001	ATTTGCTATCTGAGAGATGGTGATGACATTTT AAACCACCAAGATCGCTGATGCA (55)
	SNORD48	NR_002745	RNU48; U48	001006	GATGACCCCAGGTAACTCTGAGTGTGTCGCTG ATGCCATCACCGCAGCGCTCTGACC (57)
	RNU6-2	NR_002752	U6; RNU6B	001093	CGCAAGGATGACACGCAAATTCGTGAAGCGT TCCATATTTTT (42)

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

SUPPLEMENTARY TABLE 2: Specifications per sample used in the validation of candidate miRNAs with total RNA concentration and RNA integrity number.

			Age		PMI	Total RNA concentration [ng/µl] (RIN)						
Sample Type	Case	Gender	[years]	Cause of death as determined during autopsy	[days]	Brain	Kidney	Liver	Lung	Skin	Heart	Muscle
Native samples	1[P,I]	F	47	Polytrauma	4	464 (2.1)	791 (2.1)	691 (2.2)	220 (2.3)	235 (2.5)	513 (2.2)	220 (3.9)
	2[P,I]	M	30	CO intoxication	2	200 (2.2)	932 (2.1)	936 (2.6)	478 (2.5)	396 (2.4)	350 (1.8)	240 (3.1)
	3[I]	M	59	Macroscopically inconclusive	11	372 (2.2)	990 (1.9)	931 (2.5)	367 (2.3)	369 (2.0)	403 (2.2)	288 (3.0)
	4[P,I]	M	56	Drowning	4	490 (2.2)	928 (2.1)	504 (2.2)	230 (2.2)	245 (2.2)	385 (2.9)	180 (3.1)
	5[P]	M	42	Pericardial tamponade with aortic dissection	5	332 (2.4)	722 (2.0)	658 (2.2)	210 (2.3)	523 (2.1)	502 (2.0)	288 (2.1)
	6[P]	M	62	Myocardial infarction	3	293 (2.0)	647 (2.1)	1025 (2.6)	388 (2.8)	78 (2.4)	670 (2.1)	220 (3.1)
	7[P,I]	F	35	Macroscopically inconclusive	1	582 (2.0)	858 (2.1)	1455 (2.4)	235 (3.1)	486 (2.4)	480 (2.7)	325 (3.1)
	8[I]	M	26	Myocardial infarction and pericardial tamponade with coronary thrombosis	4	589 (1.9)	982 (1.9)	173 (2.3)	300 (2.7)	230 (2.2)	582 (2.2)	371 (2.9)
	9[P,I]	F	64	Coronary insufficiency	3	348 (2.0)	813 (2.2)	514 (2.9)	700 (2.5)	220 (2.1)	649 (2.7)	318 (2.8)
	10[I]	F	57	Subarachnoidal haemorrhage	6	598 (2.2)	622 (2.1)	949 (1.9)	130 (2.0)	381 (2.0)	250 (2.6)	270 (3.3)
	11[P]	F	45	Polytrauma	3	569 (2.1)	858 (2.0)	598 (2.6)	1300 (2.3)	445 (2.2)	558 (2.7)	389 (3.4)
	12[I]	M	50	Pericardial tamponade with aortic dissection	3	516 (2.1)	429 (2.0)	2440 (1.9)	685 (2.2)	390 (2.5)	801 (1.9)	279 (3.0)
	13[P]	M	44	Bolus death	3	460 (2.0)	732 (2.1)	1200 (2.5)	343 (2.6)	498 (2.7)	367 (2.7)	261 (2.9)
	14[I]	F	51	Myocardial infarction	2	476 (1.9)	676 (1.9)	509 (2.6)	245 (2.7)	190 (2.5)	549 (1.9)	195 (2.9)
	15[P]	F	41	Subarachnoidal haemorrhage	2	475 (2.0)	1056 (2.0)	700 (2.8)	301 (2.6)	392 (2.6)	471 (2.9)	251 (3.3)
	Median					*475 (2.1)*	*813 (2.1)*	*700 (2.5)*	*301 (2.5)*	*381 (2.4)*	*502 (2.2)*	*270 (3.1)*
Dry swabs	10[*]	F	57	See above	6	146.5 (2.4)	186 (2.4)	432.5 (1.4)	202 (2.3)		114 (2.9)	182 (2.4)
	11[*]	F	45	See above	3	93.9 (2.4)	712 (2.3)	514.5 (1.3)	717 (2.0)		240.5 (2.4)	79.8 (3.5)
	12[*]	M	50	See above	3	181.5 (2.2)	270.5 (2.3)	1030 (2.1)	346 (2.3)		232 (2.2)	29.6 (2.7)
	16	M	25	Macroscopically inconclusive	3	80.8 (2.4)	1145 (2.5)	200 (2.1)	222 (2.3)		126 (2.4)	96 (2.4)
	17	M	41	Macroscopically inconclusive	4	86.4 (2.3)	248 (2.3)	264 (2.3)	171 (2.3)		39.6 (2.0)	73.1 (2.6)
	18[+]	F	31	n.a.	n.a.					2.6 (-)		
	19[+]	F	24	n.a.	n.a.					6.9 (1.1)		
	20[+]	F	37	n.a.	n.a.					3.9 (-)		
	Median					*93.9 (2.4)*	*270.5 (2.3)*	*432.5 (2.1)*	*222 (2.3)*	*3.9 (-)*	*126 (2.4)*	*79.8 (2.6)*

PMI, post mortem interval; RIN, RNA integrity number; F, female; M, male; [P], samples used in pooled organ samples; [I], samples used as individual samples; [*], samples used as native samples and dry swabs; [+], swab from the forehead of informend and consenting adult volunteers; n.a., not applicable; -, not detected.

SUPPLEMENTARY TABLE 3: Total RNA concentration and integrity per aged native samples and dry swabs as well as DNA concentration, degradation status and specifications of the condition for the aged native samples.

Sample type	Case	Organ tissue	Total RNA concentration [ng/µl] (RIN)	DNA concentration [ng/µl] (PowerQuant® degradation index[+])	Description of the condition per aged native sample after 28 days in an outside location (see section 2.1.2)
Native samples	Aged 1	Brain	439 (2.3)	0.23 (1.69)	Moist/soft, brownish coloration
		Kidney	960 (2.5)	30.06 (2.64)	Dry/solid, dark brown/reddisch coloration, 1 small maggot
		Liver	943 (2.4)	46.22 (4.74)	Dry/solid, dark brown coloration, 1 small spot of white mold[‡]
		Lung	974 (2.5)	40.52 (2.47)	Dry/solid, black coloration, 2 small maggots, 1 small spot of white mold[‡]
		Skin	200 (2.4)	0.04 (32.01)	Dry, brownish coloration, extensively overgrown with black mold[‡]/bacterial colonies
		Heart	409 (2.4)	3.64 (4.05)	Moist/soft, brownish coloration, whitish mucous surface
		Muscle	787 (2.5)	1.32 (1.07)	Dry/solid, dark brown coloration, extensively overgrown with white mold[‡]
	Aged 2	Brain	657 (2.5)	14.91 (0.83)	Dry/solid, dark brown coloration, extensively overgrown with white and brownish mold[‡]
		Kidney	999 (2.6)	41.54 (4.63)	Dry/solid, dark brown coloration, extensively overgrown with white and brownish mold[‡]
		Liver	160 (1.9)	0.07 (80.66)	Moist/soft, dark brown coloration
		Lung	1035 (2.5)	30.22 (6.10)	Dry/solid, black coloration
		Skin	387 (2.4)	5.12 (1.20)	Dry, dark brown/black coloration, extensively overgrown with mold[‡] in various colors
		Heart	26 (-)	0.01 (4.39)	Moist/soft, dark brown coloration, greenish mucous surface
		Muscle	133 (1.9)	0.05 (2.56)	Moist/soft, brown/reddish coloration, greenish mucous surface
Swabs	Aged 3	Brain	228.5 (2.2)	n.d.	n.a.
		Kidney	1455 (2.4)	n.d.	n.a.
		Liver	787 (2.4)	n.d.	n.a.
		Lung	579.5 (1.6)	n.d.	n.a.
		Heart	129.5 (2.2)	n.d.	n.a.
		Muscle	80.6 (2.6)	n.d.	n.a.
	Aged 4	Brain	173.5 (2.2)	n.d.	n.a.
		Kidney	138 (2.3)	n.d.	n.a.
		Liver	148 (2.3)	n.d.	n.a.
		Lung	199 (2.2)	n.d.	n.a.
		Heart	41.4 (2.4)	n.d.	n.a.
		Muscle	47.1 (2.4)	n.d.	n.a.
Skin swabs	Aged 5*		9 (-)	n.d.	n.a.
	Aged 6*		5.6 (-)	n.d.	n.a.

RIN, RNA integrity number; -, not detected; n.d., not determined; n.a.; not applicable; *, swab from the forehead of informend and consenting volunteers; [+], a degradation index above 2 indicates degratation of the sample; [‡], determined by macroscopic visual examination.

SUPPLEMENTARY TABLE 4: Specifications of the mixture samples containing two to four organ tissue types and the respective total RNA concentration.

Sample name	Composition	Approximate ratio *(tissue input amount [mg])*	Total RNA concentration [ng/µl]
Mix 1	Skin + brain	1 : 1 *(100 + 105)*	409
Mix 1.1	Skin + (brain)	1 : 0.5 *(70 + 30)*	395
Mix 2	Skin + muscle	1 : 1 *(90 + 90)*	263
Mix 2.1	Skin + (muscle)	1 : 0.5 *(100 + 50)*	481
Mix 3	Skin + heart	1 : 1 *(100 + 100)*	481
Mix 3.1	Skin + (heart)	1 : 0.5 *(50 + 25)*	277
Mix 4	Skin + muscle + liver	1 : 1 : 1 *(100 + 100 + 95)*	439
Mix 4.1	Skin + muscle + (liver)	1 : 1 : 0.5 *(70 + 75 +35)*	716
Mix 5	Skin + muscle + lung	1 : 1 : 1 *(100 + 105 + 105)*	514
Mix 5.1	Skin + muscle + lung	1 : 1 : 2 *(90 + 90 + 185)*	660
Mix 6	Skin + muscle + kidney	1 : 1 : 1 *(95 + 95 + 100)*	928
Mix 6.1	Skin + muscle + (kidney)	1 : 1 : 0.5 *(90 + 95 + 45)*	685
Mix 7	Skin + muscle + heart	1 : 1 : 1 *(100 + 100 + 100)*	566
Mix 7.1	Skin + muscle + (heart)	1 : 1 : 0.5 *(95 + 90 + 45)*	465
Mix 8	Skin + muscle + liver + kidney	1 : 1 : 1 : 1 *(105 + 105 + 100 + 100)*	1375
Mix 8.1	Skin + muscle + liver + (kidney)	1 : 1 : 1 : 0.5 *(90 + 95 + 90 + 45)*	577
Mix 8.2	Skin + muscle + (liver) + kidney	1 : 1 : 0.5 : 1 *(90 + 90 + 45 + 95)*	1000
Mix 9	Skin + muscle + heart + lung	1 : 1 : 1 : 1 *(100 + 100 + 110 + 100)*	461
Mix 9.1	Skin + muscle + heart + lung	1 : 1 : 1 : 2 *(90 + 90 + 95 + 185)*	326
Mix 9.2	Skin + muscle + (heart) + lung	1 : 1 : 0.5 : 0.8 *(70 + 70 + 35 + 55)*	503

The minor component in an unbalanced mixture is given in parantheses; for mixtures containing lung tissue, unbalanced mixtures with lung in a higher proportion were evaluated, indicated by underlining.

SUPPLEMENTARY TABLE 5: Specifications of experimental shots and respective total RNA concentrations for forward- and backspatter samples as well as DNA concentration for backspatter samples.

Sample			Total RNA concentration [ng/µl]		DNA concentration [ng/µl]
name	Firearm	Organ tissue*	Forwardspatter	Backspatter	Backspatter
Shot 1	Ruger Speed-Six	Brain	**27.3**	**0.15**	0.0013
Shot 2		Kidney	**64.6**	0.14	0.0040
Shot 3		Liver	**100**	**0.15**	0.0013
Shot 4		Lung	**80.2**	**0.17**	0.0110
Shot 5		Heart	**26.3**	0.16	0.0014
Shot 6		Muscle	**32.5**	**0.16**	0.0041
Shot 7	Sig Sauer P 225	Brain	**36.3**	0.18	0.0016
Shot 8		Kidney	**39.5**	**0.12**	0.0039
Shot 9		Liver	**109**	0.20	0.0019
Shot 10		Lung	**3.5**	**0.52**	0.0644
Shot 11		Heart	**21.5**	**0.16**	0.0024
Shot 12		Muscle	**52.1**	**0.29**	0.0123
Shot 13	Star	Brain	**10.5**	**0.17**	0.0011
Shot 14		Kidney	**88.2**	0.12	0.0018
Shot 15		Liver	**81.0**	**0.17**	0.0020
Shot 16		Lung	**74.1**	**0.17**	0.0070
Shot 17		Heart	**40.8**	**0.17**	0.0067
Shot 18		Muscle	**88.2**	**0.13**	0.0055

Samples with analyzable miRNA results are indicated by bold print; samples for which STR profiling was performed are indicated by underlining; *, mixed with venous blood.

SUPPLEMENTARY TABLE 6: Specifications of mock stabbings and respective total RNA and DNA concentrations.

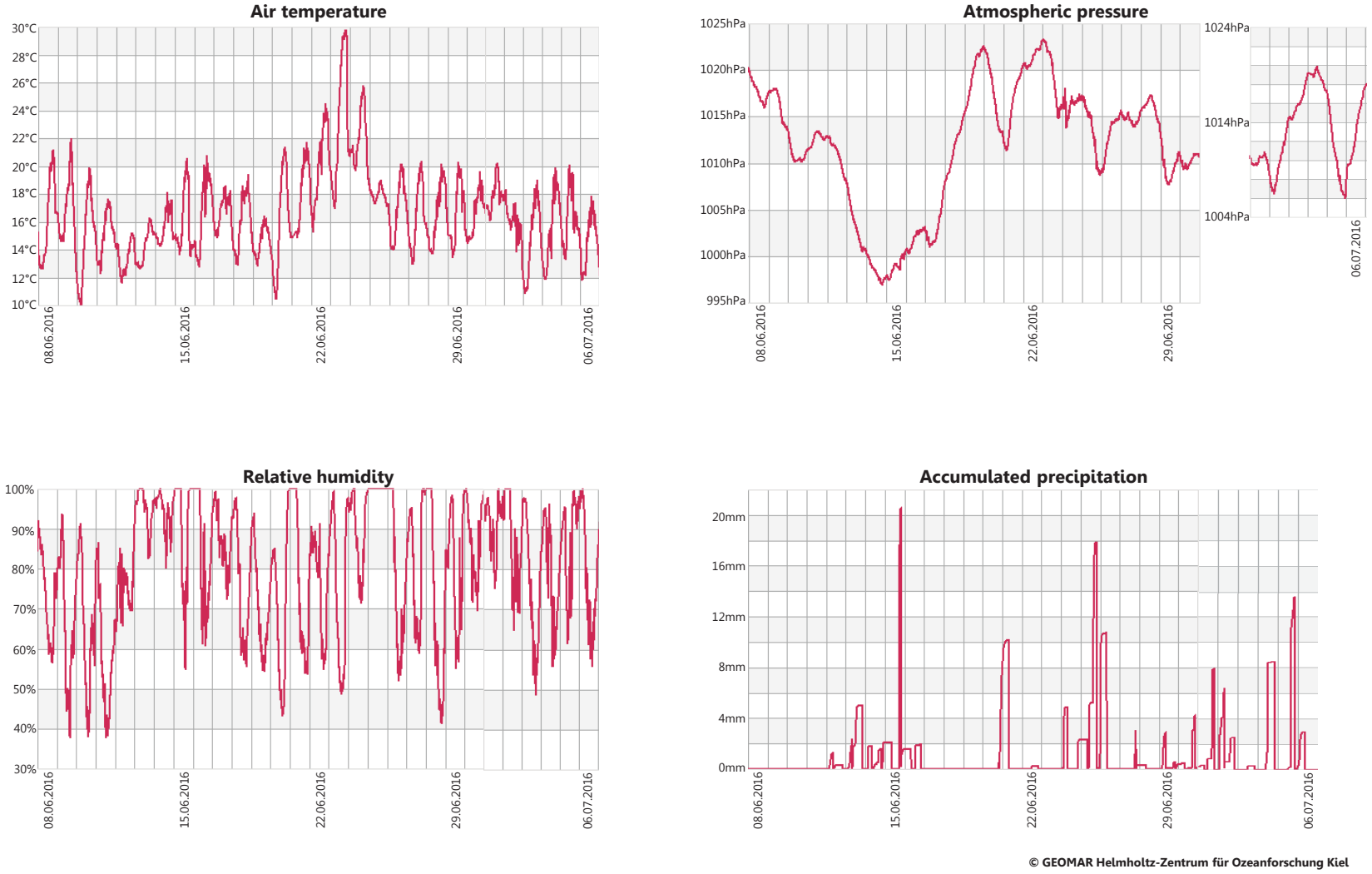
Sample name	Organ tissue strata	Total RNA concentration [ng/µl]	DNA concentration [ng/µl]
Stab 1.1	Skin/muscle	7.9	0.49
Stab 1.2	"	7.0	0.23
Stab 1.3	"	2.2	0.07
Stab 2.1	Skin/muscle/liver	34.9	1.64
Stab 2.2	"	30.0	0.63
Stab 2.3	"	63.1	1.82
Stab 3.1	Skin/muscle/kidney	45.3	3.14
Stab 3.2	"	42.5	1.93
Stab 3.3	"	45.2	3.40
Stab 4.1	Skin/muscle/lung	62.8	6.34
Stab 4.2	"	29.0	1.77
Stab 4.3	"	18.5	0.63
Stab 5.1	Skin/muscle/heart	10.0	0.81
Stab 5.2	"	5.0	0.33
Stab 5.3	"	12.0	0.91

SUPPLEMENTARY TABLE 7: Amplification efficiencies of candidate miRNAs calculated by LinRegPCR software.

		Amplification efficiency of mixture	
Organ Tissue	Gene Symbol	Mean*	SD
Brain	hsa-miR-9-5p	1.97	0.074
	hsa-miR-124-3p	1.82	0.012
	hsa-miR-219a-5p	1.82	0.025
Kidney	hsa-miR-10b-5p	1.89	0.020
	hsa-miR-204-5p	1.83	0.027
Liver	hsa-miR-122-5p	1.81	0.028
Lung	hsa-miR-146b-5p	1.85	0.022
	hsa-miR-223-3p	1.92	0.021
Skin	hsa-miR-203a-3p	1.83	0.008
	hsa-miR-205-5p	1.85	0.030
Heart	hsa-miR-208b-3p	1.87	0.015
	hsa-miR-499a-5p	1.87	0.006
Muscle in general / skeletal muscle	hsa-miR-1-3p	1.79	0.009
	hsa-miR-133a-3p	1.75	0.029
	hsa-miR-206	1.86	0.022
Reference genes	SNORD24	1.78	0.017
	SNORD48	1.69	0.017
	RNU6-2	1.77	0.033

SD, standard deviation; *, An efficiency value of 1 represents an amplification efficiency of 0 %, while 2 represents an amplification efficiency of 100%.

SUPPLEMENTARY MATERIAL 1: Weather parameters as derived from the GEOMAR website (http://www.geomar.de/service/wetter/) for the period of time during which the aged organ tissues were placed in an outside location (08.06.2016 – 06.07.2016). The GEOMAR institute is located approximately 500 m from the location of the samples.



SUPPLEMENTARY MATERIAL 2: Overview and detailed views of the experimental setup for the shooting experiments. The overview depicts the fixation of the ballistic model in a self-made contraption as well as the position of the attached plastic foil bags containing a mixture of blood and organ tissue before shooting with the Ruger Speed-Six. Also, the filter paper used to collect forward spatter is included on the far left. Contact shots were fired unimanually – represented for the Sig Sauer in the top right photograph. In the lower right detailed view, an example of the forward spatter caught on filter paper is given.

© Paul Pflüger

© Philipp Cachée

© Philipp Cachée

SUPPLEMENTARY MATERIAL 3

								Binary logistic regression						Multinomial logistic regression		
Sample type	**Sample name**	**ΔC_q (219a)**	**ΔC_q (204)**	**ΔC_q (122)**	**ΔC_q (205)**	**ΔC_q (208b)**	**ΔC_q (206)**	**Brain**	**Kidney**	**Liver**	**Skin**	**Heart**	**Muscle**	*Heart*	*Muscle*	*Heart&Muscle*
Aged native organ samples	Aged 1 Brain	-6,98						**YES**	-	-	-	-	-	-	-	-
	Aged 2 Brain	-7,23						**YES**	-	-	-	-	-	-	-	-
	Aged 1 Skin				-2,89			-	-	-	**YES**	-	-	-	-	-
	Aged 1 Muscle					2,77	-6,09	-	-	-	-	-	**YES**	-	*YES*	-
	Aged 2 Muscle					-4,23	-11,29	-	-	-	-	-	-	-	-	-
Mixtures	M1 - Skin/Brain	0,69	2,20	6,95	-0,42	12,83	10,37	**YES**	-	-	**YES**	-	-	-	-	-
	M1.1 - Skin/(Brain)	1,76	1,65	6,11	-1,31	15,00	7,82	-	-	-	**YES**	-	**YES**	-	*YES*	-
	M2 - Skin/Muscle	8,73	4,73	4,24	-0,50	7,51	-1,94	-	-	-	**YES**	-	**YES**	-	*YES*	-
	M2.1 - Skin/(Muscle)	9,14	4,65	5,21	-0,84	6,71	-0,93	-	-	-	**YES**	**YES**	-	*YES*	-	-
	M3 - Skin/Heart	7,25	4,42	4,08	-0,67	3,73	8,29	-	-	-	**YES**	**YES**	-	*YES*	-	-
	M3.1 - Skin/(Heart)	8,64	5,37	2,18	-1,22	3,81	7,38	-	-	**YES**	**YES**	-	**YES**	-	*YES*	-
	M4 - Skin/Muscle/Liver	7,17	3,24	-6,17	-0,17	5,43	-0,95	-	-	**YES**	-	-	**YES**	-	*YES*	-
	M4.1 - Skin/Muscle/(Liver)	8,49	4,00	-5,47	1,77	5,78	-0,85	-	-	-	**YES**	-	**YES**	-	*YES*	-
	M5 - Skin/Muscle/Lung	8,15	5,37	5,56	-0,71	6,54	-1,22	-	-	-	**YES**	-	**YES**	-	*YES*	-
	M5.1 - Skin/Muscle/Lung	7,88	5,10	3,37	0,08	6,82	-1,46	-	-	-	**YES**	-	**YES**	-	*YES*	-
	M6 - Skin/Muscle/Kidney	7,49	-0,07	2,12	-0,81	6,74	-1,17	-	**YES**	-	**YES**	-	**YES**	-	*YES*	-
	M6.1 - Skin/Muscle/(Kidney)	7,98	-0,51	3,57	-0,39	5,98	-1,20	-	-	-	**YES**	-	**YES**	-	-	*YES*
	M7 - Skin/Muscle/Heart	8,65	5,02	5,02	-0,48	3,17	-0,89	-	-	-	**YES**	-	**YES**	-	-	*YES*
	M7.1 - Skin/Muscle/(Heart)	8,83	5,11	5,18	-0,54	3,81	-1,73	-	-	**YES**	-	-	**YES**	-	*YES*	-
	M8 - Skin/Muscle/Liver/Kidney	5,54	-0,18	-7,10	2,35	4,43	-3,60	-	-	**YES**	-	-	**YES**	-	*YES*	-
	M8.1 - Skin/Muscle/Liver/(Kidney)	8,41	0,48	-4,98	2,01	6,93	0,20	-	**YES**	**YES**	**YES**	-	**YES**	-	*YES*	-
	M8.2 - Skin/Muscle/(Liver)/Kidney	7,44	-1,15	-5,12	0,39	4,94	-2,87	-	-	-	**YES**	-	**YES**	-	-	*YES*
	M9 - Skin/Muscle/Heart/Lung	8,39	4,99	4,10	-0,11	4,69	-0,63	-	-	-	**YES**	-	**YES**	-	-	*YES*
	M9.1 - Skin/Muscle/Heart/Lung	8,13	5,42	3,07	0,37	4,09	0,13	-	-	-	**YES**	-	**YES**	-	-	*YES*
	M9.2 - Skin/Muscle/(Heart)/Lung	8,56	4,58	0,80	-0,13	4,24	-1,07	-	-	-	-	-	-	-	-	-

(Supplementary Material 3 continued)

								Binary logistic regression						Multinomial logistic regression		
Sample type	**Sample name**	**ΔC_q (219a)**	**ΔC_q (204)**	**ΔC_q (122)**	**ΔC_q (205)**	**ΔC_q (208b)**	**ΔC_q (206)**	**Brain**	**Kidney**	**Liver**	**Skin**	**Heart**	**Muscle**	*Heart*	*Muscle*	*Heart&Muscle*
Aged dry swabs	Aged 3 Brain	-3,03						**YES**	-	-	-	-	-	-	-	-
	Aged 4 Brain	-0,80						**YES**	-	-	-	-	-	-	-	-
	Aged 3 Kidney		-5,49					-	**YES**	-	-	-	-	-	-	-
	Aged 4 Kidney		-3,86					-	-	-	-	-	-	-	-	-
	Aged 3 Liver			-6,96				-	-	**YES**	-	-	-	-	-	-
	Aged 4 Liver			-8,97				-	-	-	-	-	-	-	-	-
	Aged 3 Lung			1,24				-	-	-	-	-	-	-	-	-
	Aged 4 Lung			0,81				-	-	-	-	-	-	-	-	-
	Aged 3 Heart					1,07	8,28	-	-	-	-	**YES**	-	*YES*	-	-
	Aged 4 Heart					1,04	4,38	-	-	-	-	-	YES	-	YES	-
	Aged 3 Muscle					3,78	-3,25	-	-	-	-	-	**YES**	-	*YES*	-
	Aged 4 Muscle					3,60	-3,88	-	-	-	-	YES	**YES**	-	-	YES
	Aged 5 Skin				0,22			-	-	-	**YES**	-	-	-	-	-
	Aged 6 Skin				-1,34			-	-	-	-	-	-	-	-	-
Forwardspatter samples	FwdSpatter Brain Shot 1	-0,72						**YES**	-	-	-	-	-	-	-	-
	FwdSpatter Brain Shot 7	-0,85						**YES**	-	-	-	-	-	-	-	-
	FwdSpatter Brain Shot 13	0,15						**YES**	-	-	-	-	-	-	-	-
	FwdSpatter Kidney Shot 2		-0,33					-	-	-	-	-	-	-	-	-
	FwdSpatter Kidney Shot 8		1,54					-	-	-	-	-	-	-	-	-
	FwdSpatter Kidney Shot 14		-0,07					-	-	-	-	-	-	-	-	-
	FwdSpatter Liver Shot 3			-6,76				-	-	YES	-	-	-	-	-	-
	FwdSpatter Liver Shot 9			-7,32				-	-	YES	-	-	-	-	-	-
	FwdSpatter Liver Shot 15			-6,71				-	-	-	-	-	-	-	-	-
	FwdSpatter Lung Shot 4			-0,04				-	-	-	-	-	-	-	-	-
	FwdSpatter Lung Shot 10			2,72				-	-	-	-	-	-	-	-	-
	FwdSpatter Lung Shot 16			0,68				-	-	-	-	-	-	-	-	-
	FwdSpatter Heart Shot 5					2,07	7,25	-	-	-	-	YES	-	*YES*	-	-
	FwdSpatter Heart Shot 11					4,10	8,70	-	-	-	-	YES	-	*YES*	-	-
	FwdSpatter Heart Shot 17					2,29	7,68	-	-	-	-	-	YES	-	YES	-
	FwdSpatter Muscle Shot 6					4,87	-3,06	-	-	-	-	-	YES	-	*YES*	-
	FwdSpatter Muscle Shot 12					4,62	-3,21	-	-	-	-	-	YES	-	*YES*	-
	FwdSpatter Muscle Shot 18					3,53	-4,31	-	-	-	-	YES	YES	-	-	YES
Backspatter samples	BackSpatter Brain Shot 1	-2,27						YES	-	-	-	-	-	-	-	-
	BackSpatter Brain Shot 13	-1,46						YES	-	-	-	-	-	-	-	-
	BackSpatter Kidney Shot 8		-1,29					-	-	-	-	-	-	-	-	-
	BackSpatter Liver Shot 3			-4,69				-	-	-	-	-	-	-	-	-
	BackSpatter Liver Shot 15			-4,73				-	-	-	-	-	-	-	-	-
	BackSpatter Lung Shot 4			-0,69				-	-	-	-	-	-	-	-	-
	BackSpatter Lung Shot 10			-0,32				-	-	-	-	-	-	-	-	-
	BackSpatter Lung Shot 16			0,35				-	-	-	-	-	-	-	-	-
	BackSpatter Heart Shot 11					2,45	4,14	-	-	-	-	YES	-	*YES*	-	-
	BackSpatter Heart Shot 17					2,36	1,11	-	-	-	-	-	YES	-	YES	-
	BackSpatter Muscle Shot 6					4,49	-2,92	-	-	-	-	-	YES	-	*YES*	-
	BackSpatter Muscle Shot 12					15,00	-2,69	-	-	-	-	-	YES	-	*YES*	-
	BackSpatter Muscle Shot 18					1,93	-4,26	-	-	-	-	-	YES	-	*YES*	-
Stabbing samples	S1.1 Skin/Muscle		3,38	-1,20	1,16	15,00	-2,91	-	-	-	-	-	YES	-	*YES*	-
	S1.2 Skin/Muscle		4,97	0,63	3,35	4,34	-2,22	-	-	-	-	-	YES	-	*YES*	-
	S1.3 Skin/Muscle		4,50	-0,69	2,32	5,26	-2,82	-	-	-	-	YES	YES	-	-	YES
	S2.1 Skin/Muscle/Liver			-3,13				-	-	YES	-	-	-	-	-	-
	S2.2 Skin/Muscle/Liver			-7,37				-	-	YES	-	-	-	-	-	-
	S2.3 Skin/Muscle/Liver			-8,05				-	-	-	-	-	-	-	-	-
	S3.1 Skin/Muscle/Kidney		0,83					-	-	-	-	-	-	-	-	-
	S3.2 Skin/Muscle/Kidney		-0,48					-	-	-	-	-	-	-	-	-
	S3.3 Skin/Muscle/Kidney		-0,57					-	-	-	-	-	-	-	-	-
	S4.1 Skin/Muscle/Lung			3,76				-	-	-	-	-	-	-	-	-
	S4.2 Skin/Muscle/Lung			-0,05				-	-	-	-	-	-	-	-	-
	S4.3 Skin/Muscle/Lung			-1,81				-	-	-	-	-	-	-	-	-
	S5.1 Skin/Muscle/Heart					3,64	-0,86	-	-	-	-	YES	YES	-	-	*YES*
	S5.2 Skin/Muscle/Heart					2,86	-0,63	-	-	-	-	YES	YES	-	-	*YES*
	S5.3 Skin/Muscle/Heart					1,34	-0,11	-	-	-	-	-	-	-	-	-

SUPPLEMENTARY MATERIAL 4: OTI-Tool

Ready-to-use organ tissue inference tool NATIVE SAMPLES

	Enter respective ΔC_q-values (enter'15' if no specific amplification was observed)						**RESULTS**						
Sample name	**ΔC_q (219a)**	**ΔC_q (204)**	**ΔC_q (122)**	**ΔC_q (205)**	**ΔC_q (208b)**	**ΔC_q (206)**	**Brain**	**Kidney**	**Liver**	**Skin**	**Heart**	**Muscle**	**Heart&Muscle**
							-	-	-	-	-	-	-
							-	-	-	-	-	-	-
							-	-	-	-	-	-	-
							-	-	-	-	-	-	-
							-	-	-	-	-	-	-

Ready-to-use organ tissue inference tool DRY SAMPLES

	Enter respective ΔC_q-values (enter'15' if no specific amplification was observed)						**RESULTS**						
Sample name	**ΔC_q (219a)**	**ΔC_q (204)**	**ΔC_q (122)**	**ΔC_q (205)**	**ΔC_q (208b)**	**ΔC_q (206)**	**Brain**	**Kidney**	**Liver**	**Skin**	**Heart**	**Muscle**	**Heart&Muscle**
							-	-	-	-	-	-	-
							-	-	-	-	-	-	-
							-	-	-	-	-	-	-
							-	-	-	-	-	-	-
							-	-	-	-	-	-	-

5.3 Zusammenfassung und Diskussion

Für eine objektive Auswahl an miRNA-Kandidaten die zwischen den Organgeweben Gehirn, Lunge, Leber, Niere, Herz- und Skelettmuskel sowie Haut differentiell exprimiert werden, wurde ein miRNA Microarray-Experiment mit Proben aus Gewebematerial, welches im Rahmen rechtsmedizinischer Obduktionen entnommen worden war, durchgeführt. Wie bei den Körperflüssigkeiten bestätigte sich auch hier in einer unbeaufsichtigten, hierarchischen Clusteranalyse, dass die distinkten Organgewebe deutlich unterschiedliche miRNA-Expressionssignaturen aufwiesen, aufgrund derer eine Eingruppierung entsprechend ihrer Gewebetypen erfolgte.

Bei der Selektion geeigneter Kandidaten aus den Microarray-Ergebnissen wurden anschließend miRNAs ausgewählt, die entweder ausschließlich in einem Gewebetyp exprimiert wurden, in einem Gewebetyp ein deutlich höheres Expressionsniveau aufwiesen als in den verbleibenden sechs Geweben oder in zwei Geweben, deren Unterscheidung von besonderem Interesse ist (etwa Herz- und Skelettmuskel) unterschiedlich stark exprimiert waren.

Nachdem alle 15 ausgewählten Marker erfolgreich in die RT-qPCR implementiert werden konnten, also die entsprechenden TaqMan®-Assays spezifische Amplifikation in den Zielgeweben sowie angemessene Amplifikationseffizienz zeigten, erfolgte eine erste Evaluation der durchschnittlichen Expressionsniveaus der Kandidaten-miRNAs pro Organgewebe in vereinigten Proben mehrerer Individuen. Die Marker mit den besten Trenneigenschaften wurden anschließend sowohl in nativen Einzelproben als auch getrockneten Abrieben der Organe analysiert, wodurch das Bild durch Ergänzung zum einen um interindividuelle Unterscheide und zum anderen um den Aspekt der forensisch realistischen Beschaffenheit vervollständigt wurde.

Die miRNA hsa-miR-219a-5p zeigte dabei durchweg ausgeprägte Expressionsunterscheide zwischen Gehirnproben und den Proben der anderen Organgewebe, mit deutlich stärkerer Ausprägung in den Gehirnproben. Auch hsa-miR-9-5p und hsa-miR-124-3p wiesen vielversprechende Trenneigenschaften für die Inferenz von Gehirn auf. Beide Kandidaten-miRNAs für die Identifikation von Nierengewebe, hsa-miR-204-5p und hsa-miR-10b-5p, wurden in den Proben des Zielgewebes am stärksten exprimiert, zeigten in den verbliebenen Geweben allerdings unterschiedliche Expressionsmuster. Die miRNA hsa-miR-204-5p wurde für die weiteren Analysen ausgewählt, da diese eine gute Abgren-

zung von den Geweben Leber, Lunge, Herz-, Skelettmuskel und Haut ermöglichte und lediglich in Gehirnproben nur leicht niedrigere Expressionsniveaus als in den Nierenproben aufwies. Hsa-miR-122-5p wurde als gut geeigneter Marker für die Erkennung von Leber identifiziert, obgleich relativ große interindividuelle Unterschiede in den Nicht-Leberproben beobachtet wurden. Ebenfalls große interindividuelle Unterschiede in den Nicht-Zielproben, aber starke Ausprägung und auffallend geringe Varianzen in den Proben des Zielgewebes ergaben sich für den Hautmarker hsa-miR-205-5p. Zudem zeigte auch hsa-miR-203a-3p deutlich höhere Expression in Hautproben als in den restlichen Gewebeproben. Beide Marker waren zuvor in forensischen Studien zur Identifikation von Körperflüssigkeiten als speichelspezifisch beschrieben worden [82,137,139,140,145] (Kapitel 4). Die in dieser Arbeit erhaltenen Ergebnisse deuteten jedoch darauf hin, dass diese miRNAs vielmehr die Anwesenheit von Epithelzellen, welche auch in Speichel vorhanden sind, anzeigen. Für die Identifizierung von Herzmuskelgewebe wurde die miRNA hsa-miR-208b-3p als bestgeeignet befunden, da zum einen die Expression in den Nicht-Muskelgeweben deutlich geringer ausfiel als in den Muskelproben im allgemeinen Sinne (Herz-und Skelettmuskel) und zum anderen die Expressionsniveaus von Herz- und Skelettmuskelproben deutlich unterschiedlich waren. Hsa-miR-206 zeigte eine spezifisch hohe Ausprägung in Skelettmuskelgewebeproben, sodass dieses Zielgewebe von den verbleibenden sechs Gewebetypen, inklusive Herzmuskelgewebe, eindeutig unterschieden werden konnte. Mit hsa-miR-1-3p, hsa-miR-133a-3p und hsa-miR-499a-5p wurden weitere gut geeignete Marker für die Identifikation von Muskelgewebe im allgemeinen Sinne bestätigt. Für den Nachweis von Lungengewebe konnte kein reliabler Marker gefunden werden.

Um eine objektive, statistisch valide Unterscheidung der Gehirn-, Leber-, Niere-, Herzmuskel-, Skelettmuskel- und Hautproben von den jeweils anderen Gewebetypen (inklusive Lungengewebe) sicherzustellen, wurden binär logistische Regressionsanalysen (BLR) angewendet und – mit Ausnahme eines Ausreißers – wurde in allen Fällen eine vollständige Trennung der Ziel- und Nicht-Zielgewebeproben erreicht.

Die Eignung der ausgewählten miRNA-Marker und die Anwendbarkeit des erarbeiteten Klassifikationsmodells wurden im Anschluss in vier Wochen lang gealterten Organabriebproben, verwesendem Organgewebematerial, Mischungen mehrerer Organgewebetypen und simulierten Asservaten von Gewaltdelikten mit Einsatz von Stich- und Schusswaffen evaluiert.

Die aufgrund der in der Studie zur Identifikation von Körperflüssigkeiten erhaltenen guten Ergebnisse in bis zu 36 Jahre alten Blutflecken erwartete gute Übereinstimmung der Expressionsniveaus der gealterten trockenen Abriebproben mit denen nicht-gealterter Proben bestätigte sich. Selbst in bei sommerlichen Temperaturen verwesendem Organmaterial konnte trotz der für Nukleinsäuren äußerst ungünstigen feucht-warmen Bedingungen in der großen Mehrheit der Proben die entsprechenden miRNA-Marker nachgewiesen werden. Da jedoch in vielen Fällen nicht alle Referenzgene spezifische Amplifikation aufwiesen, konnten nur für eine kleine Teilmenge an Proben die Expressionsergebnisse mittels der BLR-Formeln analysiert werden – für diese war die Klassifizierung in allen Fällen erfolgreich.

Die Analyse der Mischungen aus zwei bis vier distinkten Organgeweben in unterschiedlichen Mischungsverhältnissen ergab keine falsch-positiven Befunde, welche einen Tatverdächtigen eventuell ungerechtfertigt belasten könnten, wenn beispielsweise bei einer oberflächlichen Schnitt- oder Stichwunde irrtümlich auf das Vorhandensein innerer Organe geschlossen würde. Zudem ergab sich für die hirn-, leber-, und skelettmuskelspezifischen Marker eine hohe Sensitivität, da sie selbst in Mischungen nachweisbar waren in denen das entsprechende Zielorgan den weit geringeren Mischungsanteil ausmachte. Die korrekte Zuordnung von Haut war nicht in allen Fällen erfolgreich. Deren Nachweis in Mischproben, wie sie bei Gewaltdelikten entstehen, ist allerdings auch nur von nachrangigem Interesse, da ihre Gegenwart in der Regel nicht in Frage steht. Der Nachweis von Nierengewebe in Mischungen durch die miRNA hsa-miR-204-5p war nur bedingt erfolgreich, da deren Expression in einigen Proben auf dem Niveau individueller Gehirnproben lag, sodass dieser Marker eher für das Vorhandensein von entweder Nieren- oder Gehirngewebe spricht. Die Differenzierung dieser beiden Gewebe kann dann anhand eines hirnspezifischen Markers erfolgen; aufgrund der anatomischen Entfernung der Organsysteme ist eine Vermischung beider Gewebe in realistischen Tatortspuren unplausibel. Herzmuskelgewebe wurde in allen Mischungen mit Nicht-Muskelgeweben korrekt identifiziert, in Gegenwart von Skelettmuskelgewebe in derselben Mischung gelang dies jedoch nicht in allen Fällen.

Die Untersuchung von Forwardspatter nach experimentellen Beschüssen auf mit Mischungen aus venösem Blut und Organgeweben dotierte ballistische Modelle bestätigte die bisher erhaltenen Ergebnisse. Auch für Backspatterspuren aus dem Inneren von

Waffenläufen konnten, trotz der geringen RNA-Ausbeute, in vielen Fällen analysierbare Ergebnisse erzielt und die entsprechenden Zielgewebe nachgewiesen werden. Die Stichversuche durch mehrere Lagen aus Organgeweben zur Simulation von abdominalen Stichverletzungen lieferten nicht nur Ergebnisse, die mit denjenigen der Mischungsversuche übereinstimmten, sondern belegten zudem, dass auch Zellmaterial der untersten verletzten Gewebeschicht auf einer Messerklinge, die aus einem Wundkanal wieder herausgezogen wird, haften bleiben kann.

Da der Nachweis von Herzmuskelgewebe einen hohen Beweiswert für eine potentiell letale Verletzung besitzt, jedoch mit der ermittelten BLR-Methode in Gegenwart von Skelettmuskel nicht immer erfolgreich geführt werden konnte, wurde die Herangehensweise zur Klassifizierung der Muskelgewebetypen neu evaluiert. Eine *a posteriori* durchgeführte multinomiale logistische Regressionsanalyse mit den normalisierten C_q-Werten von hsa-miR-208b-3p und hsa-miR-206 als Inputvariablen erlaubte die korrekte Klassifizierung sämtlicher analysierter Proben in die Teilmengen „Herzmuskelgewebe", „Skelettmuskelgewebe", „Mischung beider Muskelarten" und „Probe enthält kein Muskelgewebe".

In dieser ersten Studie zur miRNA-basierten Identifizierung von Organgeweben im forensischen Kontext konnten auf der Grundlage einer umfassend validierten Normalisierungsstrategie vielversprechende Marker für die Identifizierung von Gehirn-, Leber-, Nieren-, Herzmuskel- und Skelettmuskelgewebe sowie Haut identifiziert werden. Zudem wurde eine einfach anzuwendende Entscheidungshilfe entwickelt, die es gestattet, die wahrscheinliche Gewebeherkunft unbekannter forensischer Proben zu ermitteln.

6 Allgemeines Fazit und Ausblick

In den in dieser Dissertationsschrift zusammengefassten Studien konnte die Tauglichkeit der molekularen Spezies miRNA als Biomarker für die Identifizierung von Körperflüssigkeiten und erstmals auch von Organgeweben in forensisch realistischem Probenmaterial in den bisher umfassendsten Arbeiten erwiesen werden. Als Voraussetzung für diese qPCR-basierten Untersuchungen wurden hohen Standards genügende und an die jeweiligen experimentellen Bedingungen angepasste Strategien zur Datennormalisierung erarbeitet und umfassend validiert. Zusammen mit einer unvoreingenommenen, statistisch reliablen Datenauswertung wurde so die Grundlage für den Ausbau der Methode, die vollständig kompatibel mit einer simultanen Routine-Analyse von DNA ist, bis hin zur angestrebten Anwendung in der forensisch-genetischen Fallarbeit geschaffen.

In der forensischen Genetik mit typischerweise nur in geringen Mengen vorhandenem, beeinträchtigtem Probenmaterial ist die Anwendung und nachvollziehbare Dokumentation valider qPCR-Strategien essentiell, um höchsten Ansprüchen an Qualitätssicherung und Reproduzierbarkeit zu genügen. Die hier erstmalig in Studien zur miRNA-basierten Spurenartidentifikation angewendete strenge und kontinuierliche Dokumentation gemäß den MIQE-Richtlinien, findet inzwischen auch in anderen Publikationen in der einschlägigen Literatur Anwendung [296].

Ziel der hier vorgestellten Arbeiten war, mit einer möglichst kleinen Anzahl an Markern möglichst viele Spurenarten zu identifizieren. Während dieser Ansatz für den Nachweis von Sperma und Blut im allgemeinen Sinne (venöses Blut und Menstruationsblut), sowie Gehirn-, Leber-, Herzmuskel- und Skelettmuskelgewebe und Haut bereits gute Ergebnisse lieferte, ergaben sich Schwierigkeiten bei der Differenzierung von Speichel und Vaginalsekret sowie venösem Blut und Menstruationsblut und auch bei der Identifikation von Nieren- und Lungengewebe – vor allem in Mischungen. Es erscheint daher sinnvoll diesen vielversprechenden Ansatz weiter zu verfolgen, um zusätzliche miRNAs mit den bereits selektierten Markern zu spurenartspezifischen Signaturen zusammenzufassen, die in Kombination bessere Trenneigenschaften aufweisen als eine miRNA allein. Bevor die miRNA-

basierte Spurenartidentifikation routinemäßig in forensisch-genetischen Laboren eingesetzt werden kann, sollten alle ausgewählten Marker zudem in größeren Probenkollektiven eingehend untersucht und validiert werden, insbesondere hinsichtlich ihrer Sensitivität und ihres Verhaltens in Mischungen mehrerer Spurenarten. Hierzu wäre die Durchführung von Ringversuchen, etwa denjenigen entsprechend, die sich mit der Validierung von mRNA-Markern befassen (siehe Kapitel 1.2.2), empfehlenswert, um durch die Erfahrungen und Ergebnisse einer Vielzahl von Laboren die am besten geeigneten miRNAs zu identifizieren und die Klassifizierungsmethoden entsprechend anzupassen.

Auch wenn die eigentliche Funktion der einzelnen miRNAs nur von begrenzter Relevanz für die Fragestellung der forensischen Spurenartidentifikation ist, ist es wichtig, den potentiellen Einfluss von beispielsweise Geschlecht, Alter und Krankheiten auf die Ausprägung der ausgewählten Marker zu überprüfen. In den hier vorgestellten Studien sowie anderen einschlägigen Arbeiten zur forensischen Körperflüssigkeitsidentifikation [82,134–145] ergab sich bisher kein Hinweis auf alters- oder geschlechtsabhängige Expressionsvariationen der selektierten miRNA-Marker. In einigen Veröffentlichungen wurden jedoch etwa erhöhte Expressionsniveaus der leber- respektive muskelspezifischen miRNAs hsa-miR-122 beziehungsweise hsa-miR-1, hsa-miR-206 und hsa-miR-133a/b im Serum bei verschiedenen Leber- [297–299] und Muskelpathologien [300,301] oder auch nach einem Halbmarathonlauf [302] nachgewiesen. Derartige mögliche Veränderungen des Expressionsmuster einer miRNA sollten auch im forensischen Kontext untersucht und bei der Beurteilung der Ergebnisse berücksichtigt werden. Für die ausgewählten spermaspezifischen Marker sollten zudem die Expressionsniveaus in der Samenflüssigkeit vasektomierter Männer evaluiert werden.

Auch hinsichtlich des Arbeitsablaufes sind Weiterentwicklungen der Methode sinnvoll bevor diese im forensisch-genetischen Routinearbeitsablauf eingesetzt wird. Um den Verbrauch an Probenmaterial zu minimieren und die Analyse kleinster Spuren zu ermöglichen, sollten beispielsweise alle RT-Reaktionen möglichst in Multiplex-Ansätzen, wie hier bereits für die Referenzgene gezeigt, zusammengefasst werden. Auch eine Multiplexierung der qPCR-Reaktionen unter Verwendung mehrerer Fluorophore ist anzustreben, um den Kosten- und Zeitaufwand sowie den Probenmaterialverbrauch zu reduzieren – dies ist allerdings durch die begrenzte Anzahl an Farben, die durch gängige qPCR-Geräte gleichzeitig detektiert werden können, derzeit noch eingeschränkt. Des Weiteren würde

eine erneute Evaluation und Auswahl von Referenzgenen, die auch in verwesendem Material stabil erhalten bleiben, oder im gleichen Maße degradieren wie die körperflüssigkeits- und organgewebespezifischen miRNAs, die Analyse auch feucht gelagerten Materials ermöglichen und die forensische Spurenkunde bereichern.

Die Einführung und Verbreitung von massiv parallelen Sequenzierungsverfahren (MPS), auch als Sequenzierung der „nächsten Generation" (NGS, englisch: next generation sequencing) bezeichnet, birgt das Potential für enorme Fortschritte und abermalige Erweiterungen der Möglichkeiten forensischer DNA- und RNA-Analytik in naher Zukunft. Im Gegensatz zu den bisher angewendeten PCR-basierten Nachweismethoden, welche hinsichtlich ihrer Multiplex-Fähigkeiten limitiert sind, ermöglicht NGS die gezielte, parallele Analyse einer großen Zahl von DNA- und RNA-Markern. Zur Analyse der forensisch relevanten STRs mittels NGS liegen bereits eine Reihe von Studien sowie kommerziell erhältliche Kits vor [303–308]. Erste Pilotstudien haben zudem die prinzipielle Kombinierbarkeit von körperflüssigkeitsspezifischen mRNA-Markern und STRs [309] sowie die Möglichkeit der forensischen miRNA-Analyse mittels NGS aufgezeigt [310]. Die Methode besitzt darüber hinaus das Potential die Analyse der jeweils bestgeeigneten körperflüssigkeits- und organgewebespezifischen miRNA-, mRNA- und Methylierungsmarker für dieselbe Probe zu vereinigen. Auf diese Weise entstehen nicht nur besonders aussagekräftige Signaturen, sondern bei Ausfall eines Markertyps, etwa durch Degradation der längeren mRNA-Marker, könnte auf die verbleibenden Markertypen zurückgegriffen werden. Durch die Einbeziehung zusätzlicher Marker, beispielsweise zur Abschätzung des Alters einer Spur, ließe sich auf diese Weise ein immer detailreicheres Bild der Handlungsabläufe, die zur Deposition einer Spur geführt haben, zeichnen.

Literaturverzeichnis

[1] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: deciding which level to address in casework, Sci. Justice. 38 (1998) 231–239.

[2] I.W. Evett, P.D. Gill, G. Jackson, J. Whitaker, C. Champod, Interpreting small quantities of DNA: the hierarchy of propositions and the use of Bayesian networks, J. Forensic Sci. 47 (2002) 520–530.

[3] T. Sijen, Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers, Forensic Sci. Int. Genet. 18 (2015) 21–32.

[4] P. Gill, A.J. Jeffreys, D.J. Werrett, Forensic application of DNA "fingerprints," Nature 318 (1985) 577–579.

[5] A.J. Jeffreys, V. Wilson, S.L. Thein, Hypervariable "minisatellite" regions in human DNA, Nature 314 (1985) 67–73.

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

[7] M. Bauer, S. Polzin, D. Patzelt, Quantification of RNA degradation by semi-quantitative duplex and competitive RT-PCR: A possible indicator of the age of bloodstains?, Forensic Sci. Int. 138 (2003) 94–103.

[8] S. Anderson, B. Howard, G.R. Hobbs, C.P. Bishop, A method for determining the age of a bloodstain, Forensic Sci. Int. 148 (2005) 37–45.

[9] S.E. Anderson, G.R. Hobbs, C.P. Bishop, Multivariate Analysis for Estimating the Age of a Bloodstain, J. Forensic Sci. 56 (2011) 186–193.

[10] C. Hampson, J. Louhelainen, S. McColl, An RNA Expression Method for Aging Forensic Hair Samples, J. Forensic Sci. 56 (2011) 359–365.

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

[12] K. Ackermann, K.N. Ballantyne, M. Kayser, Estimating trace deposition time with circadian biomarkers: a prospective and versatile tool for crime scene reconstruction, Int. J. Legal Med. 124 (2010) 387-395.

[13] K. Lech, F. Liu, K. Ackermann, V.L. Revell, O. Lao, D.J. Skene, M. Kayser, Evaluation of mRNA markers for estimating blood deposition time: Towards alibi testing from human forensic stains with rhythmic biomarkers, Forensic Sci. Int. Genet. 21 (2016) 119–125.

[14] M. Takamiya, K. Saigusa, N. Nakayashiki, Y. Aoki, Studies on mRNA expression of basic fibroblast growth factor in wound healing for wound age determination, Int. J. Legal Med. 117 (2003) 46–50.

[15] M. Takamiya, K. Saigusa, R. Kumagai, N. Nakayashiki, Y. Aoki, Studies on mRNA expression of tissue-type plasminogen activator in bruises for wound age estimation, Int. J. Legal Med. 119 (2005) 16–21.

[16] S. Palagummi, S. Harbison, R. Fleming, A time-course analysis of mRNA expression during injury healing in human dermal injuries, Int. J. Legal Med. 128 (2014) 403–414.

[17] M.H. Gaballah, M. Fukuta, Y. Maeno, Y. Seko-Nakamura, J. Monma-Ohtaki, Y. Shibata, H. Kato, Y. Aoki, M. Takamiya, Simultaneous time course analysis of multiple markers based on DNA microarray in incised wound in skeletal muscle for wound aging, Forensic Sci. Int. 266 (2016) 357–368.

[18] S.F. Ibrahim, M. Issak, A.A. Bayoumy, D.S. Abd El-Fatah, Cutaneous (tPA) and Skeletal (TnI) mRNA as Markers of Aging in Contused Wound, J. Forensic Sci. 61 (2016) 1007–1010.

[19] M. Bauer, I. Gramlich, S. Polzin, D. Patzelt, Quantification of mRNA degradation as possible indicator of postmortem interval - A pilot study, Leg. Med. 5 (2003) 220–227.

[20] F. Sampaio-Silva, T. Magalhães, F. Carvalho, R.J. Dinis-Oliveira, R. Silvestre, Profiling of RNA Degradation for Estimation of Post Mortem Interval, PLoS One. 8 (2013) e56507.

[21] S.T. Young, J.D. Wells, G.R. Hobbs, C.P. Bishop, Estimating postmortem interval using RNA degradation and morphological changes in tooth pulp, Forensic Sci. Int. 229 (2013) 163.e1–163.e6.

[22] J. Ma, H. Pan, Y. Zeng, Y. Lv, H. Zhang, A. Xue, J. Jiang, K. Ma, L. Chen, Exploration of the R code-based mathematical model for PMI estimation using profiling of RNA degradation in rat brain tissue at different temperatures, Forensic Sci. Med. Pathol. 11 (2015) 530–537.

[23] M. Alvarez, J. Ballantyne, The identification of newborns using messenger RNA profiling analysis, Anal. Biochem. 357 (2006) 21–34.

[24] C. Weidner, Q. Lin, C. Koch, L. Eisele, F. Beier, P. Ziegler, D. Bauerschlag, K.H. Jöckel, R. Erbel, T. Mühleisen, M. Zenke, T. Brümmendorf, W. Wagner, Aging of blood can be tracked by DNA methylation changes at just three CpG sites, Genome Biol. 15 (2014) R24.

[25] S.H. Yi, L.C. Xu, K. Mei, R.Z. Yang, D.X. Huang, Isolation and identification of age-related DNA methylation markers for forensic age-prediction, Forensic Sci. Int. Genet. 11 (2014) 117–125.

[26] S.H. Yi, Y.S. Jia, K. Mei, R.Z. Yang, D.X. Huang, Age-related DNA methylation changes for forensic age-prediction, Int. J. Legal Med. 129 (2015) 237–244.

[27] C. Xu, H. Qu, G. Wang, B. Xie, Y. Shi, Y. Yang, Z. Zhao, L. Hu, X. Fang, J. Yan, L. Feng, A novel strategy for forensic age prediction by DNA methylation and support vector regression model, Sci. Rep. 5 (2015) 17788.

[28] R. Zbieć-Piekarska, M. Spólnicka, T. Kupiec, A. Parys-Proszek, Z. Makowska, A. Pałeczka, K. Kucharczyk, R. Płoski, W. Branicki, Development of a forensically useful age prediction method based on DNA methylation analysis, Forensic Sci. Int. Genet. 17 (2015) 173–179.

[29] A. Freire-Aradas, C. Phillips, A. Mosquera-Miguel, L. Girón-Santamaría, A. Gómez-Tato, M. Casares De Cal, J. Álvarez-Dios, J. Ansede-Bermejo, M. Torres-Español, P.M. Schneider, E. Pośpiech, W. Branicki, Carracedo, M. V. Lareu, Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system, Forensic Sci. Int. Genet. 24 (2016) 65–74.

[30] Y. Hamano, S. Manabe, C. Morimoto, S. Fujimoto, M. Ozeki, K. Tamaki, Forensic age prediction for dead or living samples by use of methylation-sensitive high resolution melting, Leg. Med. 21 (2016) 5–10.

[31] S.K. Mawlood, L. Dennany, N. Watson, B.S. Pickard, The EpiTect Methyl qPCR Assay as novel age estimation method in forensic biology, Forensic Sci. Int. 264 (2016) 132–138.

[32] D. Zubakov, F. Liu, I. Kokmeijer, Y. Choi, J.B.J. van Meurs, W.F.J. van IJcken, A.G. Uitterlinden, A. Hofman, L. Broer, C.M. van Duijn, J. Lewin, M. Kayser, Human age estimation from blood using mRNA, DNA methylation, DNA rearrangement, and telomere length, Forensic Sci. Int. Genet. 24 (2016) 33–43.

[33] J. Becker, P. Schmidt, F. Musshoff, M. Fitzenreiter, B. Madea, MOR1 receptor mRNA expression in human brains of drug-related fatalities-a real-time PCR quantification, Forensic Sci Int. 140 (2004) 13–20.

[34] K. Ikematsu, R. Tsuda, I. Nakasono, Gene response of mouse skin to pressure injury in the neck region, Leg. Med. 8 (2006) 128–131.

[35] D. Zhao, T. Ishikawa, L. Quan, D.R. Li, T. Michiue, C. Yoshida, A. Komatu, J.H. Chen, B.L. Zhu, H. Maeda, Tissue-specific differences in mRNA quantification of glucose transporter 1 and vascular endothelial growth factor with special regard to death investigations of fatal injuries, Forensic Sci. Int. 177 (2008) 176–183.

[36] B.L. Zhu, S. Tanaka, T. Ishikawa, D. Zhao, D.R. Li, T. Michiue, L. Quan, H. Maeda, Forensic pathological investigation of myocardial hypoxia-inducible factor-1α, erythropoietin and vascular endothelial growth factor in cardiac death, Leg. Med. 10 (2008) 11–19.

[37] U. Chung, J.S. Seo, Y.H. Kim, G.H. Son, J.J. Hwang, Quantitative analyses of postmortem heat shock protein mRNA profiles in the occipital lobes of human cerebral cortices: Implications in cause of death, Mol. Cells. 34 (2012) 473–480.

[38] Q. Wang, T. Ishikawa, T. Michiue, B.L.L. Zhu, D.W.W. Guan, H. Maeda, Intrapulmonary aquaporin-5 expression as a possible biomarker for discriminating

smothering and choking from sudden cardiac death: A pilot study, Forensic Sci. Int. 220 (2012) 154–157.

[39] Q. Wang, T. Ishikawa, T. Michiue, B.L. Zhu, D.W. Guan, H. Maeda, Molecular pathology of pulmonary edema after injury in forensic autopsy cases, Int. J. Legal Med. 126 (2012) 875–882.

[40] Q. Wang, T. Ishikawa, T. Michiue, B.L. Zhu, D.W. Guan, H. Maeda, Molecular pathology of pulmonary edema in forensic autopsy cases with special regard to fatal hyperthermia and hypothermia, Forensic Sci. Int. 228 (2013) 137–141.

[41] G.H. Son, S.H. Park, Y. Kim, J.Y. Kim, J.W. Kim, S. Chung, Y.H. Kim, H. Kim, J.J. Hwang, J.S. Seo, Postmortem mRNA Expression Patterns in Left Ventricular Myocardial Tissues and Their Implications for Forensic Diagnosis of Sudden Cardiac Death, Mol. Cells. 37 (2014) 241–247.

[42] Q. Wang, T. Ishikawa, T. Michiue, B.L. Zhu, D.W. Guan, H. Maeda, Molecular pathology of brain matrix metalloproteases, claudin5, and aquaporins in forensic autopsy cases with special regard to methamphetamine intoxication, Int. J. Legal Med. 128 (2014) 469–474.

[43] M. Vennemann, A. Koppelkamm, mRNA profiling in forensic genetics I: Possibilities and limitations, Forensic Sci. Int. 203 (2010) 71–75.

[44] D. Zhao, B.L. Zhu, T. Ishikawa, D.R. Li, T. Michiue, H. Maeda, Quantitative RT-PCR assays of hypoxia-inducible factor-1alpha, erythropoietin and vascular endothelial growth factor mRNA transcripts in the kidneys with regard to the cause of death in medicolegal autopsy, Leg. Med. (Tokyo). 8 (2006) 258–263.

[45] J. Gauvin, D. Zubakov, J. van Rhee-Binkhorst, A. Kloosterman, E. Steegers, M. Kayser, Forensic pregnancy diagnostics with placental mRNA markers, Int. J. Legal Med. 124 (2010) 13–17.

[46] P.J. Jannetto, S.H. Wong, S.B. Gock, E. Laleli-Sahin, B.C. Schur, J.M. Jentzen, Pharmacogenomics as Molecular Autopsy for Postmortem Forensic Toxicology: Genotyping Cytochrome P450 2D6 for Oxycodone Cases, J. Anal. Toxicol. 26 (2002) 438–447.

[47] E.T. Gatzidou, A.N. Zira, S.E. Theocharis, Toxicogenomics: a pivotal piece in the puzzle of toxicological research, J. Appl. Toxicol. 27 (2007) 302–309.

[48] A. Koski, I. Ojanperä, J. Sistonen, E. Vuori, A. Sajantila, A Fatal Doxepin Poisoning Associated With a Defective CYP2D6 Genotype, Am. J. Forensic Med. Pathol. 28 (2007) 259–261.

[49] Y.J. He, J. Brockmöller, H. Schmidt, I. Roots, J. Kirchheiner, CYP2D6 ultrarapid metabolism and morphine/codeine ratios in blood: was it codeine or heroin?, J. Anal. Toxicol. 32 (2008) 178–182.

[50] F. Musshoff, U.M. Stamer, B. Madea, Pharmacogenetics and forensic toxicology, Forensic Sci. Int. 203 (2010) 53–62.

[51] H. Andresen, C. Augustin, T. Streichert, Toxicogenetics - Cytochrome P450 microarray analysis in forensic cases focusing on morphine/codeine and diazepam, Int. J. Legal Med. 127 (2013) 395–404.

[52] B. Madea, Praxis Rechtsmedizin, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[53] L. McKenna, Understanding DNA results within the case context: Importance of the alternative proposition, Front. Genet. 4 (2013) 1–3.

[54] P. Gunn, S. Walsh, C. Roux, The nucleic acid revolution continues - Will forensic biology become forensic molecular biology?, Front. Genet. 5 (2014) 1–4.

[55] K. Virkler, I.K. Lednev, Analysis of body fluids for forensic purposes: From laboratory testing to non-destructive rapid confirmatory identification at a crime scene, Forensic Sci. Int. 188 (2009) 1–17.

[56] A. Lindenbergh, M. van den Berge, R.J. Oostra, C. Cleypool, A. Bruggink, A. Kloosterman, T. Sijen, Development of a mRNA profiling multiplex for the inference of organ tissues, Int. J. Legal Med. 127 (2013) 891–900.

[57] C. Lux, C. Schyma, B. Madea, C. Courts, Identification of gunshots to the head by detection of RNA in backspatter primarily expressed in brain tissue, Forensic Sci. Int. 237 (2014) 62–69.

[58] J.H. An, K.J. Shin, W.I. Yang, H.Y. Lee, Body fluid identification in forensics, BMB Rep. 45 (2012) 545–553.

[59] R.E. Gaensslen, Sourcebook in Forensic Serology, Immunology, and Biochemistry, U.S. Department of Justice, Washington, DC (1983).

[60] R.P. Spalding, Identification and characterization of blood and bloodstains, in: S.H. James, J.J. Nordby (Eds.), Forensic Science: an Introduction to Scientific and Investigative Techniques, CRC Press, Boca Raton (2003) 181–201.

[61] P. Wiegand, B. Rolf, Analyse biologischer Spuren: Funktionelle Blutspurenmorphologie, Körpersekrete, Haare, Rechtsmedizin. 13 (2003) 103–113.

[62] R.C. Shaler, Modern forensic biology, in: R. Saferstein (Ed.), Forensic Science Handbook, Prentice Hall, Upper Saddle River, NJ (2002) 529–546.

[63] F. Barni, S.W. Lewis, A. Berti, G.M. Miskelly, G. Lago, Forensic application of the luminol reaction as a presumptive test for latent blood detection, Talanta. 72 (2007) 896–913.

[64] A. Castelló, M. Alvarez, F. Verdú, Accuracy, Reliability, and Safety of Luminol in Bloodstain Investigation, Can. Soc. Forensic Sci. J. 35 (2002) 113–121.

[65] I. Quinones, D. Sheppard, S. Harbison, D. Elliot, Comparative Analysis of Luminol Formulations, Can. Soc. Forensic Sci. J. 40 (2007) 53–63.

[66] E. Johnston, C.E. Ames, K.E. Dagnall, J. Foster, B.E. Daniel, Comparison of Presumptive Blood Test Kits Including Hexagon OBTI, J. Forensic Sci. 53 (2008) 687–689.

[67] V.K. Kashyap, A Simple Immunosorbent Assay for Detection of Human Blood, J. Immunoassay. 10 (1989) 315–324.

[68] I.P. Hurley, R. Cook, C.W. Laughton, N.A. Pickles, H.E. Ireland, J.H.H. Williams, Detection of human blood by immunoassay for applications in forensic analysis, Forensic Sci. Int. 190 (2009) 91–97.

[69] A. Greenfield, M.A. Sloan, Identification of biological fluids and stains, in: S.. James, J.J. Nordby (Eds.), Forensic Science: an Introduction to Scientific and Investigative Techniques, CRC Press, Boca Raton (2003) 203–220.

[70] S. Kaye, Acid phosphatase test for identification of seminal stains, J. Lab. Clin. Med. 34 (1949) 728–732.

[71] S. Miyaishi, T. Kitao, Y. Yamamoto, H. Ishizu, T. Matsumoto, Y. Mizutani, A. Heinemann, K. Püschel, Identification of menstrual blood by the simultaneous determination of FDP-D dimer and myoglobin contents, Nihon Hoigaku Zasshi. 50 (1996) 400–403.

[72] D.J. Baker, E.A. Grimes, A.J. Hopwood, D-dimer assays for the identification of menstrual blood, Forensic Sci. Int. 212 (2011) 210–214.

[73] H. Holtkötter, L. Dierig, M. Schürenkamp, U. Sibbing, H. Pfeiffer, M. Vennemann, Validation of an immunochromatographic D-dimer test to presumptively identify menstrual fluid in forensic exhibits, Int. J. Legal Med. 129 (2015) 37–41.

[74] 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.

[75] T. Kitao, S. Miyaishi, H. Ishizu, Identification of human skeletal muscle from a tissue fragment by detection of human myoglobin using a double-sandwich ELISA, Forensic Sci. Int. 71 (1995) 205–214.

[76] T. Kitao, S. Miyaishi, Y. Yamamoto, H. Ishizu, Identification of human skin from a tissue fragment by detection of squamous cell carcinoma-related antigen using an enzyme immunoassay, Forensic Sci. Int. 83 (1996) 81–86.

[77] K. Takahama, Forensic application of organ-specific antigens, Forensic Sci. Int. 80 (1996) 63–69.

[78] 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.

[79] 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.

[80] C.A. Nichols, M.A. Sens, Recovery and evaluation by cytologic techniques of trace material retained on bullets, Am. J. Forensic Med. Pathol. 11 (1990) 17–34.

[81] J. Juusola, J. Ballantyne, Messenger RNA profiling: A prototype method to supplant conventional methods for body fluid identification, Forensic Sci. Int. 135 (2003) 85–96.

[82] 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.

[83] K. Virkler, I.K. Lednev, Raman spectroscopy offers great potential for the nondestructive confirmatory identification of body fluids, Forensic Sci. Int. 181 (2008) e1–e5.

[84] T.B. Miranda, P.A. Jones, DNA methylation: The nuts and bolts of repression, J. Cell. Physiol. 213 (2007) 384–390.

[85] M. Ehrlich, M.A. Gama-Sosa, L.H. Huang, R.M. Midgett, K.C. Kuo, R.A. McCune, C. Gehrke, Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells, Nucleic Acids Res. 10 (1982) 2709–2721.

[86] M. Stevens, J.B. Cheng, D. Li, M. Xie, C. Hong, C.L. Maire, K.L. Ligon, M. Hirst, M.A. Marra, J.F. Costello, T. Wang, Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods, Genome Res. 23 (2013) 1541–1553.

[87] A. Bird, DNA methylation patterns and epigenetic memory, Genes Dev. 16 (2002) 6–21.

[88] J. Ohgane, S. Yagi, K. Shiota, Epigenetics: The DNA Methylation Profile of Tissue-Dependent and Differentially Methylated Regions in Cells, Placenta. 29 (2008) 29–35.

[89] K. Huang, G. Fan, DNA methylation in cell differentiation and reprogramming: an emerging systematic view, Regen. Med. 5 (2010) 531–544.

[90] R.J. Klose, A.P. Bird, Genomic DNA methylation: the mark and its mediators, Trends Biochem. Sci. 31 (2006) 89–97.

[91] T. Hashimshony, J. Zhang, I. Keshet, M. Bustin, H. Cedar, The role of DNA methylation in setting up chromatin structure during development, Nat. Genet. 34 (2003) 187–192.

[92] V.K. Rakyan, T.A. Down, N.P. Thorne, P. Flicek, E. Kulesha, S. Graf, E.M. Tomazou, L. Backdahl, N. Johnson, M. Herberth, K.L. Howe, D.K. Jackson, M.M. Miretti, H. Fiegler, J.C. Marioni, E. Birney, T.J.P. Hubbard, N.P. Carter, S. Tavare, S. Beck, An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs), Genome Res. 18 (2008) 1518–1529.

[93] S. Fan, X. Zhang, CpG island methylation pattern in different human tissues and its correlation with gene expression, Biochem. Biophys. Res. Commun. 383 (2009) 421–425.

[94] K. Lokk, V. Modhukur, B. Rajashekar, K. Märtens, R. Mägi, R. Kolde, M. Koltšina, T.K. Nilsson, J. Vilo, A. Salumets, N. Tõnisson, DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns, Genome Biol. 15 (2014) R54.

[95] D. Frumkin, A. Wasserstrom, B. Budowle, A. Davidson, DNA methylation-based forensic tissue identification, Forensic Sci. Int. Genet. 5 (2011) 517–524.

[96] I. Gomes, F. Kohlmeier, P.M. Schneider, Genetic markers for body fluid and tissue identification in forensics, Forensic Sci. Int. Genet. Suppl. Ser. 3 (2011) e469–e470.

[97] J.H. An, A. Choi, K.J. Shin, W.I. Yang, H.Y. Lee, DNA methylation-specific multiplex assays for body fluid identification, Int. J. Legal Med. 127 (2013) 35–43.

[98] B.L. LaRue, J.L. King, B. Budowle, A validation study of the Nucleix DSI-Semen kit - A methylation-based assay for semen identification, Int. J. Legal Med. 127 (2013) 299–308.

[99] A. Wasserstrom, D. Frumkin, A. Davidson, M. Shpitzen, Y. Herman, R. Gafny, Demonstration of DSI-semen—A novel DNA methylation-based forensic semen identification assay, Forensic Sci. Int. Genet. 7 (2013) 136–142.

[100] H.Y. Lee, M.J. Park, A. Choi, J.H. An, W.I. Yang, K.J. Shin, Potential forensic application of DNA methylation profiling to body fluid identification, Int. J. Legal Med. 126 (2012) 55–62.

[101] T. Madi, K. Balamurugan, R. Bombardi, G. Duncan, B. McCord, The determination of tissue-specific DNA methylation patterns in forensic biofluids using bisulfite modification and pyrosequencing, Electrophoresis. 33 (2012) 1736–1745.

[102] J.L. Park, O.H. Kwon, J.H. Kim, H.S. Yoo, H.C. Lee, K.M. Woo, S.Y. Kim, S.H. Lee, Y.S. Kim, Identification of body fluid-specific DNA methylation markers for use in forensic science, Forensic Sci. Int. Genet. 13 (2014) 147–153.

[103] S. Forat, B. Huettel, R. Reinhardt, R. Fimmers, G. Haidl, D. Denschlag, K. Olek, Methylation Markers for the Identification of Body Fluids and Tissues from Forensic Trace Evidence, PLoS One. 11 (2016) e0147973.

[104] L.L. Ma, S.H. Yi, D.X. Huang, K. Mei, R.Z. Yang, Screening and identification of tissue-specific methylation for body fluid identification, Forensic Sci. Int. Genet. Suppl. Ser. 4 (2013) e37–e38.

[105] B. Alberts, A. Johnson, P. Walter, J. Lewis, M. Raff, K. Roberts, Molecular Biology of the Cell, 5th ed., Taylor & Francis, 2007.

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

[107] C. Nussbaumer, E. Gharehbaghi-Schnell, I. Korschineck, Messenger RNA profiling: A novel method for body fluid identification by Real-Time PCR, Forensic Sci. Int. 157 (2006) 181–186.

[108] J. Juusola, J. Ballantyne, mRNA Profiling for Body Fluid Identification by Multiplex Quantitative RT-PCR, J. Forensic Sci. 52 (2007) 1252–1262.

[109] M. Bauer, D. Patzelt, Identification of menstrual blood by real time RT-PCR: Technical improvements and the practical value of negative test results, Forensic Sci. Int. 174 (2008) 55–59.

[110] 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.

[111] D. Zubakov, E. Hanekamp, M. Kokshoorn, W. van IJcken, M. Kayser, Stable RNA 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.

[112] C. Haas, B. Klesser, C. Maake, W. Bär, A. Kratzer, mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR, Forensic Sci. Int. Genet. 3 (2009) 80–88.

[113] K. Sakurada, H. Ikegaya, H. Fukushima, T. Akutsu, K. Watanabe, M. Yoshino, Evaluation of mRNA-based approach for identification of saliva and semen, Leg. Med. 11 (2009) 125–128.

[114] R.I. Fleming, S. Harbison, The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids, Forensic Sci. Int. Genet. 4 (2010) 244–256.

[115] C. Haas, E. Hanson, A. Kratzer, W. Bär, J. Ballantyne, Selection of highly specific and sensitive mRNA biomarkers for the identification of blood, Forensic Sci. Int. Genet. 5 (2011) 449–458.

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

[117] M.L.L. Richard, K.A. Harper, R.L. Craig, A.J. Onorato, J.M. Robertson, J. Donfack, Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis, Forensic Sci. Int. Genet. 6 (2012) 452–460.

[118] E.K. Hanson, J. Ballantyne, Highly specific mRNA biomarkers for the identification of vaginal secretions in sexual assault investigations, Sci. Justice. 53 (2013) 14–22.

[119] S.M. Park, S.Y. Park, J.H. Kim, T.W. Kang, J.L. Park, K.M. Woo, J.S. Kim, H.C. Lee, S.Y. Kim, S.H. Lee, Genome-wide mRNA profiling and multiplex quantitative RT-PCR for forensic body fluid identification, Forensic Sci. Int. Genet. 7 (2013) 143–150.

[120] 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.

[121] J. Jakubowska, A. Maciejewska, K.P. Bielawski, R. Pawłowski, mRNA heptaplex protocol for distinguishing between menstrual and peripheral blood, Forensic Sci. Int. Genet. 13 (2014) 53–60.

[122] Y. Xu, J. Xie, Y. Cao, H. Zhou, Y. Ping, L. Chen, L. Gu, W. Hu, G. Bi, J. Ge, X. Chen, Z. Zhao, Development of Highly Sensitive and Specific mRNA Multiplex System (XCYR1) for Forensic Human Body Fluids and Tissues Identification, PLoS One. 9 (2014) e100123.

[123] K. Sakurada, T. Akutsu, H. Fukushima, K. Watanabe, M. Yoshino, Detection of dermcidin for sweat identification by real-time RT-PCR and ELISA, Forensic Sci. Int. 194 (2010) 80–84.

[124] K. Sakurada, T. Akutsu, K. Watanabe, Y. Fujinami, M. Yoshino, Expression of statherin mRNA and protein in nasal and vaginal secretions, Leg. Med. 13 (2011) 309–313.

[125] C. Haas, E. Hanson, W. Bär, R. Banemann, A.M. Bento, A. Berti, E. Borges, C. Bouakaze, A. Carracedo, M. Carvalho, A. Choma, M. Dötsch, M. Duriancikova, P. Hoff-Olsen, C. Hohoff, P. Johansen, P.A. Lindenbergh, B. Loddenkötter, B. Ludes, O. Maroñas, N. Morling, H. Niederstätter, W. Parson, G. Patel, C. Popielarz, E. Salata, P.M. Schneider, T. Sijen, B. Sviezená, L. Zatkalíková, J. Ballantyne, mRNA profiling for the identification of blood—Results of a collaborative EDNAP exercise, Forensic Sci. Int. Genet. 5 (2011) 21–26.

[126] C. Haas, E. Hanson, M.J. Anjos, W. Bär, R. Banemann, A. Berti, E. Borges, C. Bouakaze, A. Carracedo, M. Carvalho, V. Castella, A. Choma, G. De Cock, M. Dötsch, P. Hoff-Olsen, P. Johansen, F. Kohlmeier, P.A. Lindenbergh, B. Ludes, O. Maroñas, D. Moore, M.L. Morerod, N. Morling, H. Niederstätter, F. Noel, W. Parson, G. Patel, C. Popielarz, E. Salata, P.M. Schneider, T. Sijen, B. Sviežena, M. Turanská, L. Zatkalíková, J. Ballantyne, RNA/DNA co-analysis from blood stains—Results of a second collaborative EDNAP exercise, Forensic Sci. Int. Genet. 6 (2012) 70–80.

[127] C. Haas, E. Hanson, M.J. Anjos, R. Banemann, A. Berti, E. Borges, A. Carracedo, M. Carvalho, C. Courts, G. De Cock, M. Dötsch, S. Flynn, I. Gomes, C. Hollard, B. Hjort, P. Hoff-Olsen, K. Hríbiková, A. Lindenbergh, B. Ludes, O. Maroñas, N. McCallum, D. Moore, N. Morling, H. Niederstätter, F. Noel, W. Parson, C. Popielarz, C. Rapone, A.D. Roeder, Y. Ruiz, E. Sauer, P.M. Schneider, T. Sijen, D. Syndercombe Court, B. Sviežená, M. Turanská, A. Vidaki, L. Zatkalíková, J. Ballantyne, 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.

[128] C. Haas, E. Hanson, M.J. Anjos, K.N. Ballantyne, R. Banemann, B. Bhoelai, E. Borges, M. Carvalho, C. Courts, G. de Cock, K. Drobnic, M. Dötsch, R. Fleming, C. Franchi, I. Gomes, G. Hadzic, S.A. Harbison, J. Harteveld, B. Hjort, C. Hollard, P. Hoff-Olsen, C. Hüls, C. Keyser, O. Maroñas, N. McCallum, D. Moore, N. Morling, H. Niederstätter, F. Noël, W. Parson, C. Phillips, C. Popielarz, A.D. Roeder, L. Salvaderi, E. Sauer, P.M. Schneider, G. Shanthan, D. Syndercombe Court, M. Turanská, R.A.H. van Oorschot, M. Vennemann, A. Vidaki, L. Zatkalíková, J. Ballantyne, 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.

[129] M. van den Berge, A. Carracedo, I. Gomes, E.A.M. Graham, C. Haas, B. Hjort, P. Hoff-Olsen, O. Maroñas, B. Mevåg, N. Morling, H. Niederstätter, W. Parson, P.M. Schneider, D.S. Court, A. Vidaki, T. Sijen, 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.

[130] C. Haas, E. Hanson, R. Banemann, A.M. Bento, A. Berti, Carracedo, C. Courts, G. De Cock, K. Drobnic, R. Fleming, C. Franchi, I. Gomes, G. Hadzic, S.A. Harbison, B. Hjort, C. Hollard, P. Hoff-Olsen, C. Keyser, A. Kondili, O. Maroñas, N. McCallum, P. Miniati, N. Morling, H. Niederstätter, F. Noël, W. Parson, M.J. Porto, A.D. Roeder, E. Sauer, P.M. Schneider, G. Shanthan, T. Sijen, D. Syndercombe Court, M. Turanská, M. Van Den Berge, M. Vennemann, A. Vidaki, L. Zatkalíková, J. Ballantyne, 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.

[131] A. Lindenbergh, P. Maaskant, T. Sijen, Implementation of RNA profiling in forensic casework, Forensic Sci. Int. Genet. 7 (2013) 159–166.

[132] M. Alvarez, J. Juusola, J. Ballantyne, An mRNA and DNA co-isolation method for forensic casework samples, Anal. Biochem. 335 (2004) 289–298.

[133] 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.

[134] E.K. Hanson, M. Mirza, K. Rekab, J. Ballantyne, The identification of menstrual blood in forensic samples by logistic regression modeling of miRNA expression, Electrophoresis. 35 (2014) 3087–3095.

[135] D. Zubakov, A.W.M. Boersma, Y. Choi, P.F. van Kuijk, E.A.C. 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.

[136] C. Courts, B. Madea, Micro-RNA – A potential for forensic science?, Forensic Sci. Int. 203 (2010) 106–111.

[137] 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.

[138] Z. Wang, H. Luo, X. Pan, M. Liao, Y. Hou, A model for data analysis of microRNA expression in forensic body fluid identification, Forensic Sci. Int. Genet. 6 (2012) 419–423.

[139] E.J. Omelia, M.L. Uchimoto, G. Williams, Quantitative PCR analysis of blood- and saliva-specific microRNA markers following solid-phase DNA extraction, Anal. Biochem. 435 (2013) 120–122.

[140] 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.

[141] 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.

[142] E.K. Hanson, K. Rekab, J. Ballantyne, Binary logistic regression models enable miRNA profiling to provide accurate identification of forensically relevant body fluids and tissues, Forensic Sci. Int. Genet. Suppl. Ser. 4 (2013) e127–e128.

[143] 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.

[144] J.L. Park, S.M. Park, O.H. Kwon, H. Lee, J. Kim, H.H. Seok, W.S. Lee, S.H. Lee, Y.S. Kim, K.M. Woo, S.Y. Kim, Microarray screening and qRT-PCR evaluation of microRNA markers for forensic body fluid identification, Electrophoresis. 35 (2014) 3062–3068.

[145] Z. Wang, J. Zhang, W. Wei, D. Zhou, H. Luo, X. Chen, Y. Hou, Identification of Saliva Using MicroRNA Biomarkers for Forensic Purpose, J. Forensic Sci. 60 (2015) 702–706.

[146] D.P. Bartel, MicroRNAs:Genomics, Biogenesis, Mechanism, and Function, Cell 116 (2004) 281–297.

[147] E. Wienholds, R.H.A. Plasterk, MicroRNA function in animal development, FEBS Lett. 579 (2005) 5911–5922.

[148] B. Zhang, Q. Wang, X. Pan, MicroRNAs and their regulatory roles in animals and plants, J. Cell. Physiol. 210 (2007) 279–289.

[149] B.J. Reinhart, F.J. Slack, M. Basson, A.E. Pasquinelli, J.C. Bettinger, A.E. Rougvie, H.R. Horvitz, G. Ruvkun, The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans, Nature 403 (2000) 901–906.

[150] R.C. Lee, R.L. Feinbaum, V. Ambros, The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14, Cell 75 (1993) 843–854.

[151] A.E. Pasquinelli, B.J. Reinhart, F. Slack, M.Q. Martindale, M.I. Kuroda, B. Maller, D.C. Hayward, E.E. Ball, B. Degnan, P. Müller, J. Spring, A. Srinivasan, M. Fishman, J. Finnerty, J. Corbo, M. Levine, P. Leahy, E. Davidson, G. Ruvkun, Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA, Nature 408 (2000) 86–89.

[152] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, Identification of novel genes coding for RNAs of small expressed RNAs, Science 294 (2001) 853–858.

[153] N.C. Lau, L.P. Lim, E.G. Weinstein, D.P. Bartel, An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans, Science 294 (2001) 858–862.

[154] R.C. Lee, V. Ambros, An extensive class of small RNAs in Caenorhabditis elegans, Science 294 (2001) 862–864.

[155] S. Griffiths-Jones, The microRNA Registry, Nucleic Acids Res. 32 (2004) D109–111.

[156] S. Griffiths-Jones, H.K. Saini, S. van Dongen, A.J. Enright, miRBase: tools for microRNA genomics, Nucleic Acids Res. 36 (2008) D154–158.

[157] A. Kozomara, S. Griffiths-Jones, miRBase: annotating high confidence microRNAs using deep sequencing data, Nucleic Acids Res. 42 (2014) D68–73.

[158] A. Rodriguez, S. Griffiths-Jones, J.L. Ashurst, A. Bradley, Identification of mammalian microRNA host genes and transcription units, Genome Res. 14 (2004) 1902–1910.

[159] V.N. Kim, J.W. Nam, Genomics of microRNA, Trends Genet. 22 (2006) 165–173.

[160] Y. Lee, K. Jeon, J.T. Lee, S. Kim, V.N. Kim, MicroRNA maturation: stepwise processing and subcellular localization, EMBO J. 21 (2002) 4663–4670.

[161] X. Cai, C.H. Hagedorn, B.R. Cullen, Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs, RNA 10 (2004) 1957–1966.

[162] Y. Lee, M. Kim, J. Han, K.H. Yeom, S. Lee, S.H. Baek, V.N. Kim, MicroRNA genes are transcribed by RNA polymerase II, EMBO J. 23 (2004) 4051–4060.

[163] M.L. Bortolin-Cavaillé, M. Dance, M. Weber, J. Cavaillé, C19MC microRNAs are processed from introns of large Pol-II, non-protein-coding transcripts, Nucleic Acids Res. 37 (2009) 3464–3473.

[164] Y. Lee, C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Rådmark, S. Kim, V.N. Kim, The nuclear RNase III Drosha initiates microRNA processing, Nature 425 (2003) 415–419.

[165] A.M. Denli, B.B.J. Tops, R.H.A. Plasterk, R.F. Ketting, G.J. Hannon, Processing of primary microRNAs by the Microprocessor complex, Nature 432 (2004) 231–235.

[166] R.I. Gregory, K.P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj, N. Cooch, R. Shiekhattar, The Microprocessor complex mediates the genesis of microRNAs, Nature 432 (2004) 235–240.

[167] J. Han, Y. Lee, K.H. Yeom, Y.K. Kim, H. Jin, V.N. Kim, The Drosha-DGCR8 complex in primary microRNA processing, Genes Dev. 18 (2004) 3016–3027.

[168] M. Landthaler, A. Yalcin, T. Tuschl, The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis, Curr. Biol. 14 (2004) 2162–2167.

[169] Y. Zeng, R. Yi, B.R. Cullen, Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha, EMBO J. 24 (2005) 138–148.

[170] R. Yi, Y. Qin, I.G. Macara, B.R. Cullen, Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs, Genes Dev. 17 (2003) 3011–3016.

[171] M.T. Bohnsack, K. Czaplinski, D. Gorlich, Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs, RNA 10 (2004) 185–191.

[172] E. Lund, S. Güttinger, A. Calado, J.E. Dahlberg, U. Kutay, Nuclear export of microRNA precursors, Science 303 (2004) 95–98.

[173] E. Bernstein, A.A. Caudy, S.M. Hammond, G.J. Hannon, Role for a bidentate ribonuclease in the initiation step of RNA interference, Nature 409 (2001) 363–366.

[174] A. Grishok, A.E. Pasquinelli, D. Conte, N. Li, S. Parrish, I. Ha, D.L. Baillie, A. Fire, G. Ruvkun, C.C. Mello, Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing, Cell 106 (2001) 23–34.

[175] G. Hutvágner, J. McLachlan, A.E. Pasquinelli, E. Bálint, T. Tuschl, P.D. Zamore, A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA, Science 293 (2001) 834–838.

[176] R.F. Ketting, S.E. Fischer, E. Bernstein, T. Sijen, G.J. Hannon, R.H. Plasterk, Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans, Genes Dev. 15 (2001) 2654–2659.

[177] T.P. Chendrimada, R.I. Gregory, E. Kumaraswamy, J. Norman, N. Cooch, K. Nishikura, R. Shiekhattar, TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing, Nature 436 (2005) 740–744.

[178] Y. Lee, I. Hur, S.Y. Park, Y.K. Kim, M.R. Suh, V.N. Kim, The role of PACT in the RNA silencing pathway, EMBO J. 25 (2006) 522–532.

[179] E. Maniataki, Z. Mourelatos, A human, ATP-independent, RISC assembly machine fueled by pre-miRNA, Genes Dev. 19 (2005) 2979–2990.

[180] J. Krol, I. Loedige, W. Filipowicz, The widespread regulation of microRNA biogenesis, function and decay, Nat. Rev. Genet. 11 (2010) 597–610.

[181] X. Liu, D.Y. Jin, M.T. McManus, Z. Mourelatos, Precursor microRNA-programmed silencing complex assembly pathways in mammals, Mol. Cell. 46 (2012) 507–517.

[182] M. Ha, V.N. Kim, Regulation of microRNA biogenesis, Nat. Rev. Mol. Cell Biol. 15 (2014) 509–524.

[183] T. Kawamata, H. Seitz, Y. Tomari, Structural determinants of miRNAs for RISC loading and slicer-independent unwinding, Nat. Struct. Mol. Biol. 16 (2009) 953–960.

[184] T. Kawamata, Y. Tomari, Making RISC, Trends Biochem. Sci. 35 (2010) 368–376.

[185] Z.S. Kai, A.E. Pasquinelli, MicroRNA assassins: factors that regulate the disappearance of miRNAs, Nat. Struct. Mol. Biol. 17 (2010) 5–10.

[186] B. Czech, G.J. Hannon, Small RNA sorting: matchmaking for Argonautes, Nat Rev Genet. 12 (2011) 19–31.

[187] P.B. Kwak, Y. Tomari, The N domain of Argonaute drives duplex unwinding during RISC assembly, Nat. Struct. Mol. Biol. 19 (2012) 145–151.

[188] A. Khvorova, A. Reynolds, S.D. Jayasena, Functional siRNAs and miRNAs exhibit strand bias, Cell 115 (2003) 209–216.

[189] D.S. Schwarz, G. Hutvágner, T. Du, Z. Xu, N. Aronin, P.D. Zamore, Asymmetry in the assembly of the RNAi enzyme complex, Cell 115 (2003) 199–208.

[190] H. Hu, Z. Yan, Y. Xu, H. Hu, C. Menzel, Y. Zhou, W. Chen, P. Khaitovich, Sequence features associated with microRNA strand selection in humans and flies, BMC Genomics. 10 (2009) 413.

[191] F. Frank, N. Sonenberg, B. Nagar, Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2, Nature 465 (2010) 818–822.

[192] H.R. Chiang, L.W. Schoenfeld, J.G. Ruby, V.C. Auyeung, N. Spies, D. Baek, W.K. Johnston, C. Russ, S. Luo, J.E. Babiarz, R. Blelloch, G.P. Schroth, C. Nusbaum, D.P. Bartel, Mammalian microRNAs: experimental evaluation of novel and previously annotated genes, Genes Dev. 24 (2010) 992–1009.

[193] J.G. Doench, Specificity of microRNA target selection in translational repression, Genes Dev. 18 (2004) 504–511.

[194] D.P. Bartel, MicroRNAs: Target Recognition and Regulatory Functions, Cell 136 (2009) 215–233.

[195] B.P. Lewis, C.B. Burge, D.P. Bartel, Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets, Cell 120 (2005) 15–20.

[196] A.M. Duursma, M. Kedde, M. Schrier, C. le Sage, R. Agami, miR-148 targets human DNMT3b protein coding region, RNA 14 (2008) 872–877.

[197] J.J. Forman, A. Legesse-Miller, H.A. Coller, A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence, Proc. Natl. Acad. Sci. 105 (2008) 14879–14884.

[198] Y. Tay, J. Zhang, A.M. Thomson, B. Lim, I. Rigoutsos, MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation, Nature 455 (2008) 1124–1128.

[199] U.A. Ørom, F.C. Nielsen, A.H. Lund, MicroRNA-10a Binds the 5'UTR of Ribosomal Protein mRNAs and Enhances Their Translation, Mol. Cell. 30 (2008) 460–471.

[200] N.P. Tsai, Y.L. Lin, L.N. Wei, MicroRNA mir-346 targets the 5'-untranslated region of receptor-interacting protein 140 (RIP140) mRNA and up-regulates its protein expression, Biochem. J. 424 (2009) 411–418.

[201] S. Yekta, I.H. Shih, D.P. Bartel, MicroRNA-directed cleavage of HOXB8 mRNA, Science 304 (2004) 594–596.

[202] F. V. Karginov, S. Cheloufi, M.M.W. Chong, A. Stark, A.D. Smith, G.J. Hannon, Diverse Endonucleolytic Cleavage Sites in the Mammalian Transcriptome Depend upon MicroRNAs, Drosha, and Additional Nucleases, Mol. Cell. 38 (2010) 781–788.

[203] D.T. Humphreys, B.J. Westman, D.I.K. Martin, T. Preiss, MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 16961–16966.

[204] R.S. Pillai, S.N. Bhattacharyya, C.G. Artus, T. Zoller, N. Cougot, E. Basyuk, E. Bertrand, W. Filipowicz, Inhibition of translational initiation by Let-7 MicroRNA in human cells, Science 309 (2005) 1573–1576.

[205] S. Nottrott, M.J. Simard, J.D. Richter, Human let-7a miRNA blocks protein production on actively translating polyribosomes, Nat. Struct. Mol. Biol. 13 (2006) 1108–1114.

[206] C.P. Petersen, M.E. Bordeleau, J. Pelletier, P.A. Sharp, Short RNAs repress translation after initiation in mammalian cells, Mol. Cell. 21 (2006) 533–542.

[207] M. Kiriakidou, G.S. Tan, S. Lamprinaki, M. De Planell-Saguer, P.T. Nelson, Z. Mourelatos, An mRNA m7G cap binding-like motif within human Ago2 represses translation, Cell 129 (2007) 1141–1151.

[208] A. Eulalio, E. Huntzinger, T. Nishihara, J. Rehwinkel, M. Fauser, E. Izaurralde, Deadenylation is a widespread effect of miRNA regulation, RNA 15 (2008) 21–32.

[209] J.E. Braun, E. Huntzinger, E. Izaurralde, A Molecular Link between miRISCs and Deadenylases Provides New Insight into the Mechanism of Gene Silencing by MicroRNAs, Cold Spring Harb. Perspect. Biol. 4 (2012) a012328–a012328.

[210] D. Baek, J. Villén, C. Shin, F.D. Camargo, S.P. Gygi, D.P. Bartel, The impact of microRNAs on protein output, Nature 455 (2008) 64–71.

[211] R.C. Friedman, K.K.H. Farh, C.B. Burge, D.P. Bartel, Most mammalian mRNAs are conserved targets of microRNAs, Genome Res. 19 (2009) 92–105.

[212] B.N. Davis-Dusenberry, A. Hata, Mechanisms of control of microRNA biogenesis, J. Biochem. 148 (2010) 381–392.

[213] F. Ozsolak, L.L. Poling, Z. Wang, H. Liu, X.S. Liu, R.G. Roeder, X. Zhang, J.S. Song, D.E. Fisher, Chromatin structure analyses identify miRNA promoters, Genes Dev. 22 (2008) 3172–3183.

[214] A.M. Monteys, R.M. Spengler, J.J. Wan, L. Tecedor, K.A. Lennox, Y. Xing, B.L. Davidson, Structure and activity of putative intronic miRNA promoters, RNA 16 (2010) 495–505.

[215] J. Han, J.S. Pedersen, S.C. Kwon, C.D. Belair, Y.K. Kim, K.H. Yeom, W.Y. Yang, D. Haussler, R. Blelloch, V.N. Kim, Posttranscriptional crossregulation between Drosha and DGCR8, Cell 136 (2009) 75–84.

[216] X. Tang, Y. Zhang, L. Tucker, B. Ramratnam, Phosphorylation of the RNase III enzyme Drosha at Serine300 or Serine302 is required for its nuclear localization, Nucleic Acids Res. 38 (2010) 6610–6619.

[217] K.M. Herbert, G. Pimienta, S.J. DeGregorio, A. Alexandrov, J.A. Steitz, Phosphorylation of DGCR8 increases its intracellular stability and induces a progrowth miRNA profile, Cell Rep. 5 (2013) 1070–1081.

[218] X. Tang, S. Wen, D. Zheng, L. Tucker, L. Cao, D. Pantazatos, S.F. Moss, B. Ramratnam, Acetylation of drosha on the N-terminus inhibits its degradation by ubiquitination, PLoS One. 8 (2013) e72503.

[219] I. Heo, C. Joo, J. Cho, M. Ha, J. Han, V.N. Kim, Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA, Mol. Cell. 32 (2008) 276–284.

[220] S.L. Ameres, P.D. Zamore, Diversifying microRNA sequence and function, Nat. Rev. Mol. Cell Biol. 14 (2013) 475–488.

[221] T.B. Hansen, T.I. Jensen, B.H. Clausen, J.B. Bramsen, B. Finsen, C.K. Damgaard, J. Kjems, Natural RNA circles function as efficient microRNA sponges, Nature 495 (2013) 384–388.

[222] E. Elkayam, C.D. Kuhn, A. Tocilj, A.D. Haase, E.M. Greene, G.J. Hannon, L. Joshua-Tor, The structure of human argonaute-2 in complex with miR-20a, Cell 150 (2012) 100–110.

[223] A. Baccarini, H. Chauhan, T.J. Gardner, A.D. Jayaprakash, R. Sachidanandam, B.D. Brown, Kinetic analysis reveals the fate of a microRNA following target regulation in mammalian cells, Curr. Biol. 21 (2011) 369–376.

[224] J. Xie, S.L. Ameres, R. Friedline, J.H. Hung, Y. Zhang, Q. Xie, L. Zhong, Q. Su, R. He, M. Li, H. Li, X. Mu, H. Zhang, J.A. Broderick, J.K. Kim, Z. Weng, T.R. Flotte, P.D. Zamore, G. Gao, Long-term, efficient inhibition of microRNA function in mice using rAAV vectors, Nat. Methods. 9 (2012) 403–409.

[225] C.C. Pritchard, H.H. Cheng, M. Tewari, MicroRNA profiling: approaches and considerations, Nat. Rev. Genet. 13 (2012) 358–369.

[226] E. Hornstein, J.H. Mansfield, S. Yekta, J.K.H. Hu, B.D. Harfe, M.T. McManus, S. Baskerville, D.P. Bartel, C.J. Tabin, The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development, Nature 438 (2005) 671–674.

[227] A.M. Krichevsky, K.S. King, C.P. Donahue, K. Khrapko, K.S. Kosik, A microRNA array reveals extensive regulation of microRNAs during brain development, RNA 9 (2003) 1274–1281.

[228] J. Lu, J. Qian, F. Chen, X. Tang, C. Li, W. V Cardoso, Differential expression of components of the microRNA machinery during mouse organogenesis, Biochem. Biophys. Res. Commun. 334 (2005) 319–323.

[229] J. Dostie, Z. Mourelatos, M. Yang, A. Sharma, G. Dreyfuss, Numerous microRNPs in neuronal cells containing novel microRNAs, RNA 9 (2003) 180–186.

[230] S. Chang, R.J. Johnston, C. Frøkjaer-Jensen, S. Lockery, O. Hobert, MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode, Nature 430 (2004) 785–789.

[231] R. Yi, D. O'Carroll, H.A. Pasolli, Z. Zhang, F.S. Dietrich, A. Tarakhovsky, E. Fuchs, Morphogenesis in skin is governed by discrete sets of differentially expressed microRNAs, Nat. Genet. 38 (2006) 356–362.

[232] E. Vigorito, K.L. Perks, C. Abreu-Goodger, S. Bunting, Z. Xiang, S. Kohlhaas, P.P. Das, E.A. Miska, A. Rodriguez, A. Bradley, K.G.C. Smith, C. Rada, A.J. Enright, K.M. Toellner, I.C.M. Maclennan, M. Turner, microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells, Immunity. 27 (2007) 847–859.

[233] J. Brennecke, D.R. Hipfner, A. Stark, R.B. Russell, S.M. Cohen, bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila, Cell 113 (2003) 25–36.

[234] A.M. Cheng, M.W. Byrom, J. Shelton, L.P. Ford, Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis, Nucleic Acids Res. 33 (2005) 1290–1297.

[235] A. Cimmino, G.A. Calin, M. Fabbri, M. V Iorio, M. Ferracin, M. Shimizu, S.E. Wojcik, R.I. Aqeilan, S. Zupo, M. Dono, L. Rassenti, H. Alder, S. Volinia, C.G. Liu, T.J. Kipps, M. Negrini, C.M. Croce, miR-15 and miR-16 induce apoptosis by targeting BCL2, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 13944–13949.

[236] J. Lu, G. Getz, E.A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B.L. Ebert, R.H. Mak, A.A. Ferrando, J.R. Downing, T. Jacks, H.R. Horvitz, T.R. Golub, MicroRNA expression profiles classify human cancers, Nature 435 (2005) 834–838.

[237] N. Rosenfeld, R. Aharonov, E. Meiri, S. Rosenwald, Y. Spector, M. Zepeniuk, H. Benjamin, N. Shabes, S. Tabak, A. Levy, D. Lebanony, Y. Goren, E. Silberschein, N. Targan, A. Ben-Ari, S. Gilad, N. Sion-Vardy, A. Tobar, M. Feinmesser, O. Kharenko, O. Nativ, D. Nass, M. Perelman, A. Yosepovich, B. Shalmon, S. Polak-Charcon, E. Fridman, A. Avniel, I. Bentwich, Z. Bentwich, D. Cohen, A. Chajut, I. Barshack, MicroRNAs accurately identify cancer tissue origin, Nat. Biotechnol. 26 (2008) 462–469.

[238] K.M. Pauley, S. Cha, E.K.L. Chan, MicroRNA in autoimmunity and autoimmune diseases J. Autoimmun. 32 (2009) 189–194.

[239] S. Santosh P, N. Arora, P. Sarma, M. Pal-Bhadra, U. Bhadra, Interaction Map and Selection of microRNA Targets in Parkinson's Disease-Related Genes, J. Biomed. Biotechnol. 2009 (2009) 1–11.

[240] J. Satoh, MicroRNAs and their therapeutic potential for human diseases: aberrant microRNA expression in Alzheimer's disease brains, J. Pharmacol. Sci. 114 (2010) 269–275.

[241] S. V Yelamanchili, H.S. Fox, Defining larger roles for "tiny" RNA molecules: role of miRNAs in neurodegeneration research, J. Neuroimmune Pharmacol. 5 (2010) 63–69.

[242] J. Hayes, P.P. Peruzzi, S. Lawler, MicroRNAs in cancer: biomarkers, functions and therapy, Trends Mol. Med. 20 (2014) 460–469.

[243] M. Hesse, C. Arenz, MicroRNA maturation and human disease, Methods Mol. Biol. 1095 (2014) 11–25.

[244] I. Basak, K.S. Patil, G. Alves, J.P. Larsen, S.G. Møller, microRNAs as neuroregulators, biomarkers and therapeutic agents in neurodegenerative diseases, Cell. Mol. Life Sci. 73 (2016) 811–827.

[245] O. Barad, E. Meiri, A. Avniel, R. Aharonov, A. Barzilai, I. Bentwich, U. Einav, S. Gilad, P. Hurban, Y. Karov, E.K. Lobenhofer, E. Sharon, Y.M. Shiboleth, M. Shtutman, Z. Bentwich, P. Einat, MicroRNA expression detected by oligonucleotide microarrays: System establishment and expression profiling in human tissues, Methods. 14 (2004) 2486–2494.

[246] R. Bargaje, M. Hariharan, V. Scaria, B. Pillai, Consensus miRNA expression profiles derived from interplatform normalization of microarray data, RNA 16 (2010) 16–25.

[247] Z. Guo, M. Maki, R. Ding, Y. Yang, B. Zhang, L. Xiong, Genome-wide survey of tissue-specific microRNA and transcription factor regulatory networks in 12 tissues, Sci. Rep. 4 (2014) 5150.

[248] P. Landgraf, M. Rusu, R. Sheridan, A. Sewer, N. Iovino, A. Aravin, S. Pfeffer, A. Rice, A.O. Kamphorst, M. Landthaler, C. Lin, N.D. Socci, L. Hermida, V. Fulci, S. Chiaretti, R. Foà, J. Schliwka, U. Fuchs, A. Novosel, R.U. Müller, B. Schermer, U. Bissels, J. Inman, Q. Phan, M. Chien, D.B. Weir, R. Choksi, G. De Vita, D. Frezzetti, H.I. Trompeter, V. Hornung, G. Teng, G. Hartmann, M. Palkovits, R. Di Lauro, P. Wernet, G. Macino, C.E. Rogler, J.W. Nagle, J. Ju, F.N. Papavasiliou, T. Benzing, P. Lichter, W. Tam, M.J. Brownstein, A. Bosio, A. Borkhardt, J.J. Russo, C. Sander, M. Zavolan, T. Tuschl, A Mammalian microRNA Expression Atlas Based on Small RNA Library Sequencing, Cell 129 (2007) 1401–1414.

[249] E.J. Lee, M. Baek, Y. Gusev, D.J. Brackett, G.J. Nuovo, T.D. Schmittgen, Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors, RNA 14 (2007) 35–42.

[250] Y. Liang, D. Ridzon, L. Wong, C. Chen, Characterization of microRNA expression profiles in normal human tissues, BMC Genomics. 8 (2007) 166.

[251] Y. Sun, S. Koo, N. White, E. Peralta, C. Esau, N.M. Dean, R.J. Perera, Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs, Nucleic Acids Res. 32 (2004) e188.

[252] N. Ludwig, P. Leidinger, K. Becker, C. Backes, T. Fehlmann, C. Pallasch, S. Rheinheimer, B. Meder, C. Stähler, E. Meese, A. Keller, Distribution of miRNA expression across human tissues, (2016) 1–13.

[253] C. Chen, D.A. Ridzon, A.J. Broomer, Z. Zhou, D.H. Lee, J.T. Nguyen, M. Barbisin, N.L. Xu, V.R. Mahuvakar, M.R. Andersen, K.Q. Lao, K.J. Livak, K.J. Guegler, Real-time quantification of microRNAs by stem-loop RT-PCR, Nucleic Acids Res. 33 (2005) e179.

[254] V. Benes, M. Castoldi, Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available, Methods. 50 (2010) 244–249.

[255] S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, C.T. Wittwer, The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments, Clin. Chem. 55 (2009) 611–622.

[256] S.A. Bustin, J.F. Beaulieu, J. Huggett, R. Jaggi, F.S. Kibenge, P. a Olsvik, L.C. Penning, S. Toegel, MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments, BMC Mol. Biol. 11 (2010) 74.

[257] S.A. Bustin, Why the need for qPCR publication guidelines?—The case for MIQE, Methods. 50 (2010) 217–226.

[258] S.A. Bustin, A-Z of quantitative PCR, IUL Press, La Jolla, 2004.

[259] J. Huggett, S.A. Bustin, Standardisation and reporting for nucleic acid quantification, Accredit. Qual. Assur. 16 (2011) 399–405.

[260] M.W. Pfaffl, A new mathematical model for relative quantification in RT-PCR, Nucleic Acids Res. 29 (2001) e45.

[261] S.A. Bustin, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems, J. Mol. Endocrinol. 29 (2002) 23–39.

[262] 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.

[263] S. Bustin, T. Nolan, Data analysis and interpretation, in: S. Bustin (Ed.), A-Z of quantitative PCR, 1st ed., International University Line, La Jolla, CA (2004) 439–492.

[264] 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, in: Lipoprotein Protocols, Humana Press, New Jersey (1998) 43–62.

[265] 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.

[266] 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.

[267] K.J. Livak, T.D. Schmittgen, Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method, Methods. 25 (2001) 402–408.

[268] M.W. Pfaffl, Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR, Biotechnol. Lett. 23 (2001) 275–282.

[269] J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, Genome Biol. 3 (2002) R34.

[270] J. Huggett, K. Dheda, S. Bustin, A. Zumla, Real-time RT-PCR normalisation; strategies and considerations, Genes Immun. 6 (2005) 279–284.

[271] J.R. Chapman, J. Waldenström, With reference to reference genes: A systematic review of endogenous controls in gene expression studies, PLoS One. 10 (2015) 1–18.

[272] K. Dheda, J.F. Huggett, S.A. Bustin, M.A. Johnson, G. Rook, A. Zumla, Validation of housekeeping genes for normalizing RNA expression in real-time PCR, Biotechniques. 37 (2004) 112–119.

[273] A. Huth, B. Vennemann, T. Fracasso, S. Lutz-Bonengel, M. Vennemann, Apparent versus true gene expression changes of three hypoxia-related genes in autopsy derived tissue and the importance of normalisation, Int. J. Legal Med. 127 (2013) 335–344.

[274] A. Koppelkamm, B. Vennemann, T. Fracasso, S. Lutz-Bonengel, U. Schmidt, M. Heinrich, Validation of adequate endogenous reference genes for the normalisation of qPCR gene expression data in human post mortem tissue, Int. J. Legal Med. 124 (2010) 371–380.

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

[276] L.I. Moreno, C.M. Tate, E.L. Knott, J.E. McDaniel, S.S. Rogers, B.W. Koons, M.F. Kavlick, R.L. Craig, J.M. Robertson, Determination of an Effective Housekeeping Gene for the Quantification of mRNA for Forensic Applications, J. Forensic Sci. 57 (2012) 1051–1058.

[277] 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.

[278] Q. Wang, T. Ishikawa, T. Michiue, B.L. Zhu, D.W. Guan, H. Maeda, Molecular pathology of brain edema after severe burns in forensic autopsy cases with special regard to the importance of reference gene selection, Int. J. Legal Med. 127 (2013) 881–889.

[279] L. Wong, K. Lee, I. Russell, C. Chen, Endogenous controls for real-time quantitation of miRNA using TaqMan® microRNA assays, Applied Biosystems Application Note (2010).

[280] 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.

[281] A. Masotti, V. Caputo, L. Da Sacco, A. Pizzuti, B. Dallapiccola, G.F. Bottazzo, Quantification of Small Non-Coding RNAs Allows an Accurate Comparison of miRNA Expression Profiles, J. Biomed. Biotechnol. 2009 (2009) 1–9.

[282] 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.

[283] Y. Karlen, A. McNair, S. Perseguers, C. Mazza, N. Mermod, Statistical significance of quantitative PCR, BMC Bioinformatics. 8 (2007) 131.

[284] J.M. Ruijter, M.W. Pfaffl, S. Zhao, A.N. Spiess, G. Boggy, J. Blom, R.G. Rutledge, D. Sisti, A. Lievens, K. De Preter, S. Derveaux, J. Hellemans, J. Vandesompele, Evaluation of qPCR curve analysis methods for reliable biomarker discovery: Bias, resolution, precision, and implications, Methods. 59 (2013) 32–46.

[285] S. Pabinger, S. Rödiger, A. Kriegner, K. Vierlinger, A. Weinhäusel, A survey of tools for the analysis of quantitative PCR (qPCR) data, Biomol. Detect. Quantif. 1 (2014) 23–33.

[286] J.M. Ruijter, C. Ramakers, W.M.H. Hoogaars, Y. Karlen, O. Bakker, M.J.B. van den Hoff, A.F.M. Moorman, Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data, Nucleic Acids Res. 37 (2009) e45.

[287] J.M. Tuomi, F. Voorbraak, D.L. Jones, J.M. Ruijter, Bias in the Cq value observed with hydrolysis probe based quantitative PCR can be corrected with the estimated PCR efficiency value, Methods. 50 (2010) 313–322.

[288] J.M. Ruijter, P. Lorenz, J.M. Tuomi, M. Hecker, M.J.B. van den Hoff, Fluorescent-increase kinetics of different fluorescent reporters used for qPCR depend on monitoring chemistry, targeted sequence, type of DNA input and PCR efficiency, Microchim. Acta. 181 (2014) 1689–1696.

[289] S. Zhao, R.D. Fernald, Comprehensive Algorithm for Quantitative Real-Time Polymerase Chain Reaction, J. Comput. Biol. 12 (2005) 1047–1064.

[290] 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.

[291] 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.

[292] L.F. Sempere, S. Freemantle, I. Pitha-Rowe, E. Moss, E. Dmitrovsky, V. Ambros, Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation, Genome Biol. 5 (2004) R13.

[293] J. McCarthy, MicroRNA-206: The skeletal muscle-specific myomiR, Biochim. Biophys. Acta - Gene Regul. Mech. 1779 (2008) 682–691.

[294] C. Jopling, Liver-specific microRNA-122: Biogenesis and function, RNA Biol. 9 (2012) 137–142.

[295] Y.K. Adlakha, N. Saini, Brain microRNAs and insights into biological functions and therapeutic potential of brain enriched miRNA-128, Mol. Cancer. 13 (2014) 33.

[296] M. Sirker, R. Fimmers, P.M. Schneider, I. Gomes, Evaluating the forensic application of 19 target microRNAs as biomarkers in body fluid and tissue identification, Forensic Sci. Int. Genet. 27 (2017) 41–49.

[297] J. Xu, C. Wu, X. Che, L. Wang, D. Yu, T. Zhang, L. Huang, H. Li, W. Tan, C. Wang, D. Lin, Circulating microRNAs, miR-21, miR-122, and miR-223, in patients with hepatocellular carcinoma or chronic hepatitis, Mol. Carcinog. 50 (2011) 136–142.

[298] S. Akamatsu, C.N. Hayes, M. Tsuge, D. Miki, R. Akiyama, H. Abe, H. Ochi, N. Hiraga, M. Imamura, S. Takahashi, H. Aikata, T. Kawaoka, Y. Kawakami, W. Ohishi,

K. Chayama, Differences in serum microRNA profiles in hepatitis B and C virus infection, J. Infect. 70 (2015) 273–287.

[299] J. Krauskopf, F. Caiment, S.M. Claessen, K.J. Johnson, R.L. Warner, S.J. Schomaker, D.A. Burt, J. Aubrecht, J.C. Kleinjans, Application of high-throughput sequencing to circulating microRNAs reveals novel biomarkers for drug-induced liver injury, Toxicol. Sci. 143 (2015) 268–276.

[300] D. Cacchiarelli, I. Legnini, J. Martone, V. Cazzella, A. D'Amico, E. Bertini, I. Bozzoni, miRNAs as serum biomarkers for Duchenne muscular dystrophy, EMBO Mol. Med. 3 (2011) 258–265.

[301] K.M. Akat, D. Moore-McGriff, P. Morozov, M. Brown, T. Gogakos, J. Correa Da Rosa, A. Mihailovic, M. Sauer, R. Ji, A. Ramarathnam, H. Totary-Jain, Z. Williams, T. Tuschl, P.C. Schulze, Comparative RNA-sequencing analysis of myocardial and circulating small RNAs in human heart failure and their utility as biomarkers, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 11151–11156.

[302] C.P.C. Gomes, G.P. Oliveira-Jr, B. Madrid, J.A. Almeida, O.L. Franco, R.W. Pereira, Circulating miR-1, miR-133a, and miR-206 levels are increased after a half-marathon run, Biomarkers. 19 (2014) 585–589.

[303] M. Scheible, O. Loreille, R. Just, J. Irwin, Short tandem repeat typing on the 454 platform: strategies and considerations for targeted sequencing of common forensic markers, Forensic Sci. Int. Genet. 12 (2014) 107–119.

[304] C. Børsting, N. Morling, Next generation sequencing and its applications in forensic genetics, Forensic Sci. Int. Genet. 18 (2015) 78–89.

[305] K.B. Gettings, K.M. Kiesler, S.A. Faith, E. Montano, C.H. Baker, B.A. Young, R.A. Guerrieri, P.M. Vallone, Sequence variation of 22 autosomal STR loci detected by next generation sequencing, Forensic Sci. Int. Genet. 21 (2016) 15–21.

[306] E.H. Kim, H.Y. Lee, I.S. Yang, S.E. Jung, W.I. Yang, K.J. Shin, Massively parallel sequencing of 17 commonly used forensic autosomal STRs and amelogenin with small amplicons, Forensic Sci. Int. Genet. 22 (2016) 1–7.

[307] W. Parson, D. Ballard, B. Budowle, J.M. Butler, K.B. Gettings, P. Gill, L. Gusmão, D.R. Hares, J.A. Irwin, J.L. King, P. de Knijff, N. Morling, M. Prinz, P.M. Schneider, C. Van Neste, S. Willuweit, C. Phillips, Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements, Forensic Sci. Int. Genet. 22 (2016) 54–63.

[308] K.J. van der Gaag, R.H. de Leeuw, J. Hoogenboom, J. Patel, D.R. Storts, J.F.J. Laros, P. de Knijff, Massively parallel sequencing of short tandem repeats-Population data and mixture analysis results for the PowerSeqTM system, Forensic Sci. Int. Genet. 24 (2016) 86–96.

[309] D. Zubakov, I. Kokmeijer, A. Ralf, N. Rajagopalan, L. Calandro, S. Wootton, R. Langit, C. Chang, R. Lagace, M. Kayser, Towards simultaneous individual and tissue identification: A proof-of-principle study on parallel sequencing of STRs,

amelogenin, and mRNAs with the Ion Torrent PGM, Forensic Sci. Int. Genet. 17 (2015) 122–128.

[310] Z. Wang, D. Zhou, Y. Cao, Z. Hu, S. Zhang, Y. Bian, Y. Hou, C. Li, Characterization of microRNA expression profiles in blood and saliva using the Ion Personal Genome Machine® System (Ion PGMTM System), Forensic Sci. Int. Genet. 20 (2016) 140–146.

Weitere, in den Publikationen/dem Manuskript zitierte, in dieser Schrift nicht aufgeführte Referenzen:

S. Baskerville, D.P. Bartel, Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes, RNA 11 (2005) 241–247.

C. Becker, A. Hammerle-Fickinger, I.Riedmaier, M.W. Pfaffl, mRNA and microRNA quality control for RT-qPCR analysis, Methods 50 (2010) 237–243.

J. Bereta, M. Bereta, Stimulation of glyceraldehyde-3-phosphate dehydrogenase mRNA levels by endogenous nitric oxide in cytokine-activated endothelium, Biochem. Biophys. Res. Commun. 217 (1995) 363–369.

P. Bhatia, W.R. Taylor, A.H. Greenberg, J.A. Wright, Comparison of Glyceraldehyde-3-phosphate Dehydrogenase and 28S-Ribosomal RNA Gene Expression as RNA Loading Controls for Northern Blot Analysis of Cell Lines of Varying Malignant Potential, Anal. Biochem. 216 (1994) 223–226.

A. Brüning, F. Wiethold, Die Untersuchung und Beurteilung von Selbstmörderwaffen, Dtsch. Z Gerichtl. Med. 23 (1934) 71–82.

T.J. Chang, C.C. Juan, P.H. Yin, C.W. Chi, H.J. Tsay, Up-regulation of beta-actin, cyclophilin and GAPDH in N1S1 rat hepatoma, Oncol. Rep. 5 (1998)469–471.

J.B. de Kok, R.W. Roelofs, B.A. Giesendorf, J.L. Pennings, E.T. Waas, T. Feuth, D.W. Swinkels, P.N. Span, Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes, Lab. Invest. 85 (2005) 154–159.

S. Giampaoli, A. Berti, F. Valeriani, G. Gianfranceschi, A. Piccolella, L. Buggiotti, C. Rapone, A. Valentini, L. Ripani, V.R. Spica, Molecular identification of vaginal fluid by microbial signature, Forensic Sci. Int. Genet. 6 (2012) 559–564.

D. Goidin, A. Mamessier, M.J. Staquet, D. Schmitt, O. Berthier-Vergnes, Ribosomal 18S RNA Prevails over Glyceraldehyde-3-Phosphate Dehydrogenase and β-Actin Genes as Internal Standard for Quantitative Comparison of mRNA Levels in Invasive and Noninvasive Human Melanoma Cell Subpopulations, Anal. Biochem. 295 (2001) 17–21.

M. Grabmüller, P. Cachée, B. Madea, C. Courts, How far does it get? – The effect of shooting distance and type of firearm on the simultaneous analysis of DNA and RNA from Backspatter recovered from inside and outside surfaces of firearms, Forensic Sci. Int. 258 (2016) 11–18.

J.L. Guthrie, C. Seah, S. Brown, P. Tang, F. Jamieson, S.J. Drews, Use of Bordetella pertussis BP3385 to establish a cutoff value for an IS481-targeted real-time PCR assay, J. Clin. Microbiol. 46 (2008) 3798–3799.

J.S. Hall, J. Taylor, H.R. Valentine, J.J. Irlam, A. Eustace, P.J. Hoskin, C.J. Miller, C.M.L. West, Enhanced stability of microRNA expression facilitates classification of FFPE tumour samples exhibiting near total mRNA degradation, Br. J. Cancer 107 (2012) 684–694.

J. Hellemans, G. Mortier, A. De Paepe, 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.

M. Jung, A. Schaefer, I. Steiner, C. Kempkensteffen, C. Stephan, A. Erbersdobler, K. Jung, Robust MicroRNA Stability in Degraded RNA Preparations from Human Tissue and Cell Samples, Clin. Chem. 56 (2010) 998–1006.

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.

S.E. Larkin, S. Holmes, I.A. Cree, T. Walker, V. Basketter, B. Bickers, S Harris, S.D. Garbis, P.A. Townsend, C. Aukim-Hastie, Identification of markers of prostate cancer progression using candidate gene expression, Br. J. Cancer 106 (2012) 157–165.

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.

B.C.M. Pang, B.K.K. Cheung, Double swab technique for collecting touched evidence, Leg. Med. 9 (2007) 181–184.

M.E. Ritchie, J. Silver, A. Oshlack, M. Holmes, D. Diyagama, A. Holloway, G.K. Smyth, A comparison of background correction methods for two-colour microarrays, Bioinformatics 23 (2007) 2700–2707.

T.D. Schmittgen, B.A. Zakrajsek, Effect of experimental on housekeeping gene expression: validation by real-time, quantitative RT-PCR, J. Biochem. Biophys. Methods 46 (2000) 69–81.

A. Schroeder, O. Müller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, S. Lightfoot, W. Menzel, M. Granzow, T. Ragg, The RIN: an RNA integrity number for assessing integrity values to RNA measurements, BMC Mol. Biol. 7 (2006) 3.

C. Schyma, C. Lux, B. Madea, C. Courts, The „triple contrast“ method in experimental wound ballistics and backspatter analysis, Int. J. Legal Med. 129 (2015) 1027–1033.

P Sood, A Krek, M. Zavolan, G. Macino, N. Rajwsky, Cell-type-specific signatures of microRNAs on target mRNA expression, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 2746–2751.

W. Weimann, Über das Verspritzen von Gewebeteilen aus Einschussöffnungen und seine kriminalistische Bedeutung, Dtsch. Z Gerichtl. Med. 17 (1931) 92–105.

G. Zhu, Y. Chang, J. Zuo, X. Dong, M. Zhang, G. Hu, F. Fang, Fudenine, a C-Terminal Truncated Rat Homologue of Mouse Prominin, Is Blood Glucose-Regulated and Can Up-Regulate the Expression of GAPDH, Biochem. Biophys. Res. Commun. 281 (2001) 951–956.

Danksagung

Sehr herzlich danke ich Herrn Prof. Dr. Burkhard Madea für die Möglichkeit, meine Dissertation an seinem Institut anzufertigen, die Übernahme des Erstgutachtens und die stete Unterstützung.

Herrn Prof. Dr. Walter Witke danke ich sehr für die Übernahme des Zweitgutachtens. Außerdem danke ich Herrn Prof. Dr. Jörg Höhfeld und Herrn Prof. Dr. Thomas Martin für Ihre Bereitschaft an meiner Prüfungskommission teilzunehmen.

Bei Frau Prof. Dr. Johanna Preuß-Wössner bedanke ich mich herzlich für das entgegengebrachte Vertrauen, die Unterstützung und die Bereitschaft, mich an ihrem Institut in Kiel aufzunehmen.

Ein ganz besonderer Dank gilt natürlich Herrn PD Dr. Cornelius Courts für die hervorragende und freundschaftliche Betreuung meiner Doktorarbeit, für die vielen hilfreichen und interessanten wissenschaftlichen Diskussionen und nicht-so-wissenschaftlichen Gespräche, für die immer offene Tür, für die Ermutigungen und guten Ratschläge und insbesondere für die schier unerschöpfliche, ansteckende Begeisterung für die Forschung.

Ich danke allen aktuellen und ehemaligen Kolleginnen und Kollegen der Institute für Rechtsmedizin in Bonn und Kiel, insbesondere natürlich den Mitarbeiterinnen der beiden ForGes Marion Sauer, Renate Klemmer und Waltraud Hargesheimer sowie Ingrid Wessel, Kathi Pöhls und Steffi Petzel für die gute Arbeitsatmosphäre, die Hilfsbereitschaft und viele lustige Mittagspausen.

Alex, Konrad, Lena und Melanie danke ich sehr für die wirklich tolle Zeit mit guten Gesprächen, langen Abenden und dem ein oder anderen Kaltgetränk.

Iris, Ann-Kathrin und Antje danke ich für ihre große Einsatzbereitschaft, an Teilen dieses Projekt mitzuarbeiten.

Ich danke allen Freiwilligen für die Spende von Körperflüssigkeitsproben sowie den Rechtsmedizinerinnen und Rechtsmedizinern in Bonn und Kiel für die Entnahme der Organgewebeproben, ohne die diese Arbeit nicht möglich gewesen wäre.

Danksagung

Philipp Cachée und den Mitarbeitern des Landeskriminalamts Schleswig-Holstein danke ich für die Hilfe bei der Realisierung der experimentellen Beschüsse und Frau Dr. Uta Immel für die Bereitstellung des Probenkollektivs gealterter Blutproben.

Von ganzem Herzen bedanke ich mich bei meiner Familie, meinen Freunden und Frank. Meinen Eltern Christa und Gerd danke für ihren immerwährenden Rückhalt und ihr Vertrauen in mich und meine Pläne. Jan danke ich dafür, dass er der tollste kleine Bruder der Welt und mir in viele Dingen ein Vorbild ist. Bei Gesa, Inga, Katja, Maren, Tessa und Verena bedanke ich mich so sehr für ihre wunderbare Freundschaft, ihr Interesse und die Ablenkung. Ganz besonders danke ich Frank für seine unerschütterliche, geduldige Unterstützung, Motivation und wenn nötig Aufmunterung – einfach für alles.