Developmental validation of a miRNA panel for forensic body fluid identification

Award 2016-DN-BX-0163

Final Summary Overview

Submission Date: 01/30/2020

National Institute of Justice

FY 2016: Applied Research and Development in Forensic Science for Criminal Justice Purposes

Principal Investigator: Sarah Seashols-Williams, PhD
Assistant Professor, Department of Forensic Science
sseashols@vcu.edu

Andrea J. Publow
VCU Director of Sponsored Programs
Office of Research
Virginia Commonwealth University
(804) 828-6772
dirospa@vcu.edu

Virginia Commonwealth University
P.O Box 980568
800 East Leigh Street, Suite 3200
Richmond, VA 23298-0568

VCU PT No. FP00001631
VCU Index No. 516579

Project/Grant Period: 1/1/2017 – 12/31/2019

TABLE OF CONTENTS

MicroRNAs (miRNAs) are small, non-coding RNAs 18-25 nucleotides in length that have recently been identified and evaluated as potential markers for the identification of forensically relevant body fluids. There is significant interest in the use of miRNAs for forensic casework because of their short length and high resistance to degradation, potentially allowing for robust detection in degraded samples. They have also been shown to co-extract and be detectable in DNA extracts, which could make the use of miRNAs a more streamlined and easily implementable molecular body fluid identification method compared to other described methods.

This work is a continuation of a previously funded project utilizing high-throughput sequencing (HTS) of eight forensically relevant biological fluids to identify miRNAs with tissue-specific expression (NIJ-2012-DN-BX-K017). Through that work, candidate miRNAs were developed and expression patterns in the eight biological fluids were assessed, ultimately identifying eight miRNAs, including a pair of endogenous reference miRNAs that allow for normalization of expression without evaluation of the RNA or known input quantity. This panel uses expression detected using reverse-transcription quantitative PCR (RT-qPCR) to identify and differentiate feces, urine, peripheral blood, menstrual secretions, semen, and saliva.

In this work, we developed the initial miRNA panel further by working to identify markers for vaginal secretions and perspiration, expand the population set, develop a regression tree for body fluid identification using the miRNA markers, and evaluate the final panel for identification success in mixed samples. We also performed a comparative analysis between analysis methodologies, assessed the limit of detection, performance in DNA extracts, species and tissue specificity, and stability in compromised samples.

Project Subjects

Samples were collected from 325 donors, and a total of 505 samples were analyzed using each miRNA marker in the panel under an approved Human Subjects Research Protocol (VCU IRB HM20002931). The demographics of the population sampled were a reflection of the donor demographics in a diverse, urban University.

Project design and methods

The project was broken into two specific aims – the first aim focused on evaluating additional markers for vaginal secretions and perspiration, while developing an analysis method to objectively and accurately identify the body fluid present on evidence using the miRNA panel analysis. The second aim was focused on developmental validation of the panel using population, compromised and mixed samples.

Sample Collection & Preparation

Urine and semen were collected by donors into sterile containers, aliquoted onto a swab and allowed to dry at room temperature. Blood was collected onto a cotton swab using a finger prick from a sterilized finger. Vaginal secretions, menstrual secretions, perspiration and feces were collected on swabs by donors, returned in swab boxes, and stored at room temperature until RNA and/or DNA isolation and purification.

RNA Isolation and Analysis

RNA isolation was conducted using the Qiagen miRNeasy mini kit (Qiagen N.V., Venlo, The Netherlands). Quantitative reverse transcription was carried out via the qScript™ microRNA Quantification System (Quanta Biosciences Inc., Gaithersburg, MD, USA) according to the manufacturer's protocol using 7 µL of RNA extract. qPCR reactions were prepared in triplicate for each sample using a modified quarter reaction protocol. Each sample was

analyzed from a pooled RT reaction with 2-3 technical qPCR replicates for each sample. Each environmental chamber sample was extracted twice, and the triplicate technical analyses compared for consistency across the stain. Control samples corresponding to each biological fluid and donor were handled consistently to that of the treated samples. Use of identical substrate size (whole swab or 4 mm biopsy punch of stained material) RNA extraction, extract volume into the reverse transcription reaction, and qPCR parameters provided a consistent comparison of miRNA quality. Experiments were performed and analyzed according to MIQE guidelines and as previously described [1,2]. Data analysis was conducted using SDS software, v1.3.1 (Life Technologies, Foster City, CA, USA). DNase-treated extracts were tested for each miRNA evaluated for this project, and were less than 1 cycle different from untreated RNA extracts, demonstrating amplification of miRNA and not potential contaminating genomic DNA. Negative amplification controls were included on all qPCR plates for each miRNA primer.

DNA Isolation

Organic extractions, the QIAamp® DNA Investigator Kit (Qiagen) , AllPrep® DNA/RNA Mini Kit (Qiagen), DNA IQ™ (Promega), and FTA purifications (Whatman) were performed on blood, saliva, urine and semen samples according to modified manufacturer's protocols [2].

Data Analysis

Positive and negative controls were included in each set of panel plates. Cycle thresholds and background subtraction for each reaction were manually set using the QuantStudio™ Real-Time PCR Software v1.3 (Thermo Fisher, Inc.). Data was analyzed in Microsoft Excel. Differential expression for the panel miRNAs was calculated by subtracting the average Cq value of let-7g and let-7i from the Cq value of the target miRNA ($\Delta Cq = Cq_{(target)} - Cq_{(avg\ let\text{-}7g\ \&\ let\text{-}7i)}$). The differential

expression data (ΔCq values) were then used to predict the body fluid following the body fluid classification regression tree.

Statistical analyses were performed in R v3.4.2 (R Foundation for Statistical Computing, Vienna, Austria) or JMP® v14.2.0 (SAS Institute, Cary, NC, USA). Normal distribution and equal variance were confirmed for all sample sets using quantile-quantile plots and Levine's test, respectively. Student's t-tests were applied (two-tailed distribution, equal variance) in two-group comparisons (DNase treatment, DNA IQ™ wash, and Exiqon panel evaluations). In multi-group comparisons, a one-way ANOVA test was performed with a Tukey's HSD pairwise comparison. The data were presented as mean ± SD. A value of $p<0.05$ was considered statistically significant.

Results & Discussion

Specific Aim 1: Finalization of miRNA panel and forensic utility assessment

In order to identify candidate miRNAs indicative of perspiration and vaginal secretions, we first mined the high-throughput sequencing data from the 2012 NIJ project. Any candidates considered promising were evaluated using a stepwise population group of the body fluids of a small group of population samples from each body fluid. When the candidates failed, further analysis was conducted on a small population set using the miRCURY LNA™ Universal RT microRNA PCR Panels I and II (Qiagen), and from that data, miR-1208 and miR-30c-3p were identified as together distinguishing vaginal secretions from the other biological fluids through differential expression analysis. Evaluation of 20 vaginal and 5 each of blood, semen, menstrual secretions, saliva, urine and feces showed consistent discrimination. While this initial population analysis was successful for markers for vaginal secretions (miR-30c-3p and miR-1208), when we expanded the sample size to 50 individuals, we identified large variations in the population and are not able to confidently use these markers for vaginal secretion

identification. Five perspiration biomarkers identified in the previous project were evaluated using the population sample sets, and found to be inconsistent in expression between individuals. Thus, we ceased our research on miRNA markers for perspiration and vaginal secretions.

The miRNA panel was tested against an expanded population of blood, saliva, vaginal fluid, urine and semen samples, and together with previous data were used to construct a classification regression tree that provides likely sample classification based on its relative expression of each candidate. The dataset consists of 572 samples of which 81 were blood, 85 feces, 133 Menstrual Blood, 61 Semen, 70 Urine, 65 Vaginal Fluid.

Several classification models were explored to determine their ability to correctly predict the body fluid. This resulted in a 69.2% correct classification for regression tree, 71.5% correct classification for Linear Discriminant Analysis and 80.1% correct classification for Quadratic Discriminant Analysis. Table 1 shows the resulting confusion matrix for QDA and 10-fold cross validation.

Table 1: Confusion matrix for completed dataset using QDA with 10-fold cross validation using all markers. This analysis used 572 completed observations.

		Predicted						
		Blood	Feces	MB	Saliva	Semen	Urine	VF
Actual	Blood	69	0	1	0	0	0	1
	Feces	0	38	17	4	4	14	1
	MB	4	2	108	0	0	0	3
	Saliva	0	0	1	65	0	0	1
	Semen	0	5	1	1	34	4	4
	Urine	0	2	1	0	3	63	1
	VF	0	5	11	3	1	1	42

The above analysis used all markers and performed quite well. The question then becomes can a marker be removed without sacrificing much in way of correct classification rates?

Through systematic exploration of the markers and performing the identical analysis it was determined that miR1246 and miR320c could be omitted from the analysis with minimal loss of correct classification rate. Using QDA with 10-fold cross validation on this dataset gave an 81.4% correct classification rate and the confusion matrix found in Table 2.

Table 2: Confusion matrix for QDA using 10-fold cross validation for the reduced dataset with the following markers: miR200b, miR10b, miR26b and miR891a. This analysis used 515 completed observations.

		Predicted						
		Blood	Feces	MB	Saliva	Semen	Urine	VF
Actual	Blood	69	0	1	0	0	0	1
	Feces	0	38	17	4	4	14	1
	MB	4	2	108	0	0	0	3
	Saliva	0	0	1	65	0	0	1
	Semen	0	5	1	1	34	4	4
	Urine	0	2	1	0	3	63	1
	VF	0	5	11	3	1	1	42

We then compared miRNA detection using several different commonly used miRNA qPCR analysis platforms: qScript (Quantas), miScript (Qiagen), and TaqMan® (Life Technologies) miRNA detection methodologies. We encountered difficulties with contamination and positive signals in several batches and analyses of let-7i in miScript assays and thus ceased our analysis with miScript. We also eliminated TaqMan® Advanced miRNA assays from our platform comparison due to variabilities observed. While we found that the Taqman analysis method is the most sensitive and precise assay of the 5 different methods surveyed, we found that the Quantas method was an overall simpler and more time and cost-effective analysis method.

Specific Aim 2: miRNA multiplex marker characterization and development

The second aim of the project was to begin the developmental validation of the miRNA panel as described thus far. This required expansion of the population data set from 20 donors

of each biological fluid to 50 samples from each body fluid, each characterized against the full miRNA panel. Over 500 samples were analyzed using RT-qPCR for miRNAs let-7g, let-7i, miR-200b-3p, miR-26b, miR-320c, miR-1246, miR-10b-5p, miR-891a, miR-1208, and miR-30c-3p. The observed differential expression values were used to provide data for the above-mentioned QDA analysis.

We were also interested in evaluating the stability of the miRNA panel during compromising conditions. The manuscript for this portion of the project was published in the Journal of Forensic Sciences in the November 2019 issue. In this paper, we describe let-7g and let-7i detection stability in blood, saliva, urine and semen across both environmental and chemical/heat treatments, as well as successful predictive ability of the characteristic markers for blood and semen after compromising treatment.

Limit of detection (LOD) was explored using the qScript, TaqMan®, TaqMan® Advanced and qScript miRNA analysis methods. LOD using RNA quantity cannot be used as an across the board standard for all biological fluids, as quantities vary between biological fluids. Total RNA measurements include not only multiple species of RNA (mRNA, rRNA, miRNA, lncRNA, etc.), but also RNA from the microbiota inhabiting most biological fluids. Therefore, a previously-optimized standard curve of synthetic RNA oligos was evaluated by copy number to address sensitivity of the miRNA in question. While the TaqMan® Advanced and miScript methods were not found to be reliable and thus were not further evaluated for LOD, sensitivity of the Quantas assay was determined to be 10^5 copies/µL, and the TaqMan® assay sensitivity was found to be 10^4 copies of the synthetic miRNA in question.

MiRNA analysis across domesticated and wildlife animals that commonly interact with humans demonstrated that miRNAs are highly conserved among vertebrates, and thus much

cross-reactivity is observed. Therefore, any future recommendations for use of miRNA for body fluid identification will have a recommendation that a human specific target be used in addition to the miRNAs used for body fluid identification. Organ and tissue specificity could also complicate body fluid identification, as many of the miRNAs (miR-10b-5p, miR-26b, miR-891a, and miR-30c-3p) were observed to have similarities in differential expression values as compared to their target body fluids.

We analyzed the control miRNAs let-7g and let-7i in blood, semen, and saliva using four different DNA extraction methods commonly utilized in forensic laboratories and paired RNA extraction controls. Consistent detection of the let-7g and let-7i was observed across several DNA extraction methods in this sample set. While miRNA detection is markedly less than that of the paired RNA controls in blood and saliva, they are well within the detection range, and importantly, consistent between DNA isolation methods. Detection was also comparable between RNA and DNA methods in semen samples. Analysis of DNA extracts of blood, semen, saliva, and menstrual secretions from 3 donors resulted in 63% successful body fluid identifications using the miRNA panel. Specifically, 100% of semen samples and 2 of the 3 samples for blood and saliva were accurately identified, but all menstrual secretion samples were incorrectly eliminated as menstrual blood. This unexpected result may be correctable given a population sampling of menstrual blood in DNA extracts, and a readjustment of miR-200b-3p ΔCq range in DNA extracts. The tremendous success of this particular aim (and questions that our work developed) presents an area for further evaluation, to be pursued further in NIJ-2019-NE-BX-0005. The manuscript for this portion of the project was published in the Journal of Forensic Sciences in the November 2019 issue.

Implications for criminal justice policy and practice in the United States

The development of candidate miRNAs and initial validation completed in the previous body of work was a vital first step towards an eventual commercial assay for body fluid identification that is robust and reliable in the hands of practitioners. The miRNA panel as validated herein provides quantifiable confidence in the body fluid(s) present in the sample. The increased efficiency could reduce or eliminate analyst time on serological assays that yield poor confidence in the reported body fluid (except for microscopic sperm cell identification). If future work on the miRNA panel using DNA extracts is successful, a significant barrier to implementation is removed – that of additional analyst time, reagent costs, and sample consumption required for a separate RNA isolation method. Much of the historical resistance to a novel body fluid identification method such as mRNA or miRNAs has been due to the additional isolation methods required; therefore using a DNA extract for body fluid identification combined with analysis methods that utilize existing equipment in a forensic laboratory could lead to rapid, large-scale implementation into the forensic DNA analysis workflow.

Dissemination of Research Findings

Publications

1. Lewis CA, Layne TR, Seashols-Williams SJ. Detection of microRNAs in DNA Extractions for Forensic Biological Source Identification. J Forensic Sci. 2019 Nov;64(6):1823-1830. doi: 10.1111/1556-4029.14070.

2. Layne TR, Green RA, Lewis CA, Nogales F, Dawson Cruz TC, Zehner ZE, Seashols-Williams SJ. microRNA Detection in Blood, Urine, Semen, and Saliva Stains After Compromising Treatments. J Forensic Sci. 2019 Nov;64(6):1831-1837. doi: 10.1111/1556-4029.14113.

Conference Posters and Presentations:

1. AC Campbell, J Szekely, CA Lewis, RA Green, T Dawson Cruz, SJ Seashols Williams. The Developmental Validation of a MicroRNA (miRNA) Panel for Forensic Body Fluid Identification. Talk, presented at the American Academy of Forensic Sciences Annual Meeting, February 2019.
2. SJ Seashols-Williams. Exploring the utility of miRNAs as biomarkers for body fluids. Invited lecture, Body Fluid Identification workshop, International Symposium on Human Identification, October 2017.
3. CA Lewis, TR Layne, SJ Seashols-Williams. MicroRNA detection in DNA extraction methods commonly used for forensic casework. Poster, presented at the International Symposium on Human Identification, October 2017.
4. J Szekely, CA Lewis, RA Green, MK Valle, SJ Seashols-Williams. Developmental validation of a miRNA panel for forensic body fluid identification. Poster presented at the NIJ Forensic Science Research & Development Poster Session, Pittcon, February 2018.
5. C Lewis, T Layne, SJ Seashols-Williams. An Evaluation of the Forensic Detection of MicroRNAs in DNA Extractions for Biological Source Identification. Talk, presented at the Mid-Atlantic Association of Forensic Scientists Annual Meeting, May 2018.
6. J Szekely, CA Lewis, SJ Seashols-Williams. Identification and evaluation of specific microRNA markers in vaginal secretions and perspiration for forensic body fluid identification. Poster presented at the Gordon Research Conference of Forensic Analysis of Human DNA, Sunday River, Maine, June 2018.
7. CA Lewis, TR Layne, SJ Seashols-Williams. Detection of microRNAs in DNA Extractions for Biological Source Identification. Poster presented at the Gordon Research Conference of Forensic Analysis of Human DNA, Sunday River, Maine, June 2018.

References

[1] 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. doi:10.1373/clinchem.2008.112797 [doi].

[2] S. Seashols-Williams, C. Lewis, C. Calloway, N. Peace, A. Harrison, C. Hayes-Nash, S. Fleming, Q. Wu, Z.E. Zehner, High-throughput miRNA sequencing and identification of biomarkers for forensically relevant biological fluids, Electrophoresis. 37 (2016) 2780–2788. doi:10.1002/elps.201600258.