Federal Aviation Administration

DOT/FAA/AM-17/16
Office of Aerospace Medicine
Washington, DC 20591

Postmortem Samples from Aviation Accident Victims Maintain Tissue-Specific mRNA Expression Profiles

Dennis Burian,[1] David Hutchings,[2]
Hilary Uyhelji,[1] Austin McCauley,[3]
Daniel Williams,[4] Doris Kupfer,[1]
Vicky L. White,[1] Candyce White,[5]
Rachel Jung,[6] Savannah Willson Smith[7]

[1]Civil Aerospace Medical Institute, FAA, Oklahoma City, OK
[2]Venesco LLC, Chantilly, VA 20151
[3]University of Oklahoma, Department of Biology, Norman, OK
[4]Cytovance Biologics, Oklahoma City, OK
[5]Southern Nevada Health District, Las Vegas, NV
[6]Citgo Petroleum, Oklahoma City, OK
[7]Martin Army Community Hospital, Ft. Benning, GA

November 2017

Final Report

NOTICE

This publication and all Office of Aerospace Medicine technical reports are available in full-text from the Civil Aerospace Medical Institute's publications website: http://www.faa.gov/go/oamtechreports

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-17/16		

4. Title and Subtitle	5. Report Date
Postmortem Samples from Aviation Accident Victims Maintain Tissue-Specific mRNA Expression Profiles	November 2017
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Burian D,[1] Hutchings D,[2] Uyhelji HA,[1] McCauley A, Kupfer DM,[1] White VL,[1] White C, Jung R, Smith SW	

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
[1]FAA Civil Aerospace Medical Institute, P.O. Box 25082 Oklahoma City, OK 73125 [2]Venesco LLC, 14801 Murdock St., Ste 125, Chantilly, VA 20151	11. Contract or Grant No.

12. Sponsoring Agency name and Address	13. Type of Report and Period Covered
Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	14. Sponsoring Agency Code

15. Supplemental Notes

16. Abstract

To determine whether an aerospace medical factor such as hypoxia or sleep deprivation played a role in the cause of an aviation accident, gene expression marker panels are in development based on results from live subject studies in humans. The translation of marker panels from the discovery phase to a forensic setting requires that RNA stability and the factors that affect it be assessed in forensic samples. To assess the utility of postmortem blood samples for gene expression analysis, RNA was purified from aviation accident victim blood and tissue samples, and gene expression levels from a panel of genes assessed by qRT-PCR. Live subject data was obtained from publicly available microarray data. Correlation analyses were performed between all victim samples and tissue specific victim samples compared to live subject microarray expression data. Due to observed prokaryotic rRNA in human samples, universal primers for bacterial 16S rRNA were characterized on purified prokaryotic total RNA and victim samples. RNA integrity was variable between victim samples but there was no correlation between RNA integrity and yield or postmortem interval. Expression patterns between victims were generally concordant. Blood, brain, and muscle were correlated between live subjects and victims. The assay for bacterial rRNA was found to be sensitive for very low levels of purified bacterial RNA and highly specific for bacterial rRNA when challenged with mixtures of bacterial and human RNA. In conclusion, victim sample RNAs exhibits a wide range of degradation but qRT-PCR data maintains the relative expression values between target genes. That victim samples had similar expression patterns to live subject data suggests that marker panels developed in human subject studies can be translated to accident investigation and will allow additional information to be incorporated into medical determinations of accident causality.

17. Key Words	18. Distribution Statement
RNA integrity, RIN, messenger RNA, micro-RNA, qRT-PCR, aviation accident investigation	Document is available to the public through the Internet: http://www.faa.gov/go/oamtechreports/

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	40	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Christy Hileman for making accident records available for review, to Russell J. Lewis for his interpretation of results from toxicology analysis, and funding support from the Federal Aviation Administration Office of Aerospace Medicine.

ABBREVIATIONS

mRNA.....messenger RNA
miRNA.....micro-RNA
rRNA.....ribosomal RNA
qRT-PCR .....Quantitative Reverse Transcribed Polymerase Chain Reaction
PMI .....Postmortem interval
PMI-arr.....Postmortem interval – accident time to time of arrival at site of autopsy
PMI-aut .....Postmortem interval – accident time to autopsy time
RI .....RNA integrity
RIN .....RNA Integrity Number
bp .....base pairs of DNA product from a PCR reaction
Ct.....Threshold Cycle – the cycle number at which a qRT-PCR amplification signal plot has the same intensity across a set of reactions

Contents

POSTMORTEM SAMPLES FROM AVIATION ACCIDENT VICTIMS MAINTAIN TISSUE-SPECIFIC mRNA EXPRESSION PROFILES

Introduction 1

Materials and Methods 3

- Sample Origin and Processing 3
- mRNA 3
- qRT-PCR 4
- miRNA 5
- Live Subject Gene Expression Data Sources and Handling 5
- Descriptive Statistics and Analysis 6
- Detection of Bacterial 16S rRNA by qRT-PCR 6

Results 8

- Qualitative Assessment of Samples Collected in Serum Tubes 8
- Quantitative PCR of Serum Tube Samples 9
- Qualitative Curation of 34 Victim Samples 10
- qRT-PCR of 34 Accident Victim Samples 15
- Correlation of Tissue-specific Expression in Living vs. Victim Samples 20
- Detection of Bacterial Contamination in Accident Samples 20

Discussion 22

- Factors impacting RNA integrity 27
- Bacterial detection by qRT-PCR 28
- Implications for detection of marker panels in postmortem samples 29
- Conclusions and Future Directions 30

References 31

Appendix: A Primer on the Quantitative Polymerase Chain Reaction: Data Handling, Analysis, and Interpretation A-1

- Theory and Basics A-1
- Data Collection A-1
- Data Interpretation A-2
- Efficiency Correction A-2
- Normalization and Interplate Calibrators A-3
- Comparing qRT-PCR to Microarray Data A-3

POSTMORTEM SAMPLES FROM AVIATION ACCIDENT VICTIMS MAINTAIN TISSUE-SPECIFIC mRNA EXPRESSION PROFILES

INTRODUCTION

In postmortem samples, the RNA expression profile has been used to measure postmortem interval (PMI)—the interval from time of death to sample collection (1-3), differentiate body fluids (4-6), identify newborns (7), and determine cause of death (8). Furthermore, tissue-specific expression across 45 tissues has been shown by whole-transcriptome microarray analysis of tissues with PMIs less than 8.5 hours (9). A crucial consideration when assessing postmortem RNA expression patterns is the stability of RNA in the sample. A host of factors including PMI (10) and antemortem circumstances such as pH, hypoxia, and stress have been shown to impact postmortem RNA integrity in tissues including brain, skeletal muscle, and cardiac muscle (Reviewed in 11). RNA degradation is due to the action of RNases from the 3' end, hydrolytic chemical degradation, or exposure to ultraviolet light. Nonetheless, RNA of sufficient integrity for expression analysis by either quantitative reverse transcriptase polymerase chain reaction (qRT-PCR), or microarray analysis (12) has been purified from tissues up to a week postmortem and from dried blood up to 16 years postmortem (Reviewed in 13). It should be noted that virtually all investigations of postmortem RNA stability have either been carried out under sterile conditions with tissues, or in laboratory or hospital settings.

A convenient and widely accepted metric for the integrity of an RNA sample is the RNA Integrity Number (RIN), an algorithm implemented in the BioAnalyzer 2100 (Agilent Technologies), a “lab-on-a-chip” electrophoresis instrument that replaces polymerized polyacrylamide or agarose gels for separation of DNA, RNA, and proteins. The RIN algorithm relies on the relative signal between the large and small ribosomal RNA (rRNA), the signal between these peaks, and the signal intensity in the region indicative of RNAs smaller than the small rRNA. Eukaryotic rRNAs are known as 28S and 18S for their sedimentation rate in sucrose gradients and migrate at 48 and 43 seconds on the BioAnalyzer; prokaryotic rRNAs are smaller, 23S and 16S, and migrate at 46 and 40 seconds (Appendix, Fig. 1). The RIN falls in a range of 1 to 10. A RIN of 1 indicates completely degraded RNA, has no rRNA peaks, and high signal corresponding to RNAs less than a hundred nucleotides in length. For comparison, the small human rRNA that migrates at 43 seconds on an electropherogram is 1869 nucleotides in length (https://www.ncbi.nlm.nih.gov/gene/100008588). A sample with a RIN of 10 has well defined rRNA peaks with an approximate area-under-the-curve ratio of 2:1 between the large and small rRNA peaks and very low signal intensity in the region of RNAs smaller than the small rRNA.

There are two methods for whole-transcriptome gene expression testing currently in wide use, microarrays and the “Nex-gen” sequencing technology RNA-seq. Generally, data quality for both methods decreases with degradation of RNA. In an RNA-seq experiment on three samples, when RNA was synthetically degraded from a RIN of 9.3 to 3.8, almost all genes detected in samples with high RINs were detected in all other samples (14) although data quality decreased as RIN fell below 7. Similarly, using synthetically degraded liver, samples with RINs below 8 had dramatically lower microarray data quality scores (15). By qRT-PCR (Reviewed in Appendix and 16), a low cost single gene assay, the effects of RNA integrity on data quality were tested using artificially degraded RNAs with RINs between 8.5 to 2.5 from

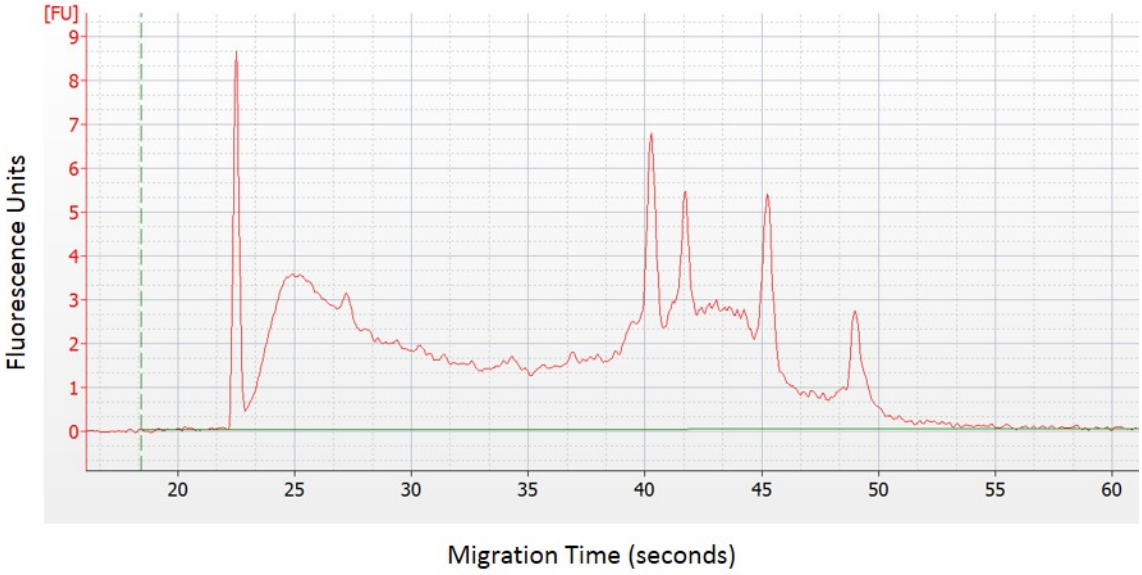

Figure 1. Electropherogram of a human sample with both pro- and eukaryotic rRNA peaks.

identically sourced tissues. There was a loss of sensitivity with decreasing RIN, as evidenced by a 3 to 4 Ct increase (See Appendix) over the range of RINs on two different mRNAs, equivalent to an apparent decrease in expression of 8- to 16-fold (17). This effect was minimized to less than 2 Ct—a 4-fold change—on samples where RIN was greater than 5. In assessing the effect of amplicon length, Cts increased considerably for amplicons greater than 400 basepairs (bp) and there was less loss of sensitivity with decreasing RIN for amplicons of 200 bp or less. In a second study, heat-degraded purified RNAs assessed by qRT-PCR for relative expression of three genes decreased less than 2-fold from a RIN of 8.9 to 6.5 (18). Ideally for qRT-PCR, samples should have a RIN greater than 5 and the amplicon length no more than 200 bp (17). Note, however, that although sensitivity decreased with decreasing RIN for all methods, RNA-seq, microarray, and qRT-PCR, relative expression patterns between genes were maintained within samples. A method comparison of qRT-PCR to microarray data assessing differential expression between cancer samples and controls found that direction-of-change was concordant between the platforms, although a few genes had slightly higher fold-change by qRT-PCR on degraded samples compared to high-quality RNA on microarrays (19). Unfortunately, RINs were not reported but rRNA peaks were completely absent in the most degraded samples. In a second method comparison study using purified RNA from three patients, there was good concordance for direction-of-change and fold-change when heat-degraded RNA was compared to untreated RNA (20). In this study, RINs ranged from an average of 9 in untreated samples to 5.5.

Biomarker discovery is ubiquitous across medicine including the need to discover and validate expression markers for sleep deprivation and hypoxia in aviation accident investigation. However, translation of markers discovered in live human subjects to accident investigation requires an understanding of post-accident molecular stability. The primary goal of this study was to assess postmortem RNA degradation and how well preserved were live subject expression patterns in victim samples. Reported herein are analyses of RNA degradation and detectable gene expression patterns using blood collected in a laboratory environment and from blood and tissue samples from aviation accident victims from across the United States. Blood and tissue samples were collected at autopsy for gene expression analysis as part of the normal workflow for aviation accident investigation. Gene expression patterns for a panel of genes were determined by qRT-PCR. Blood samples were found to have a broad range of RINs unrelated to accident conditions but good correlation of gene expression patterns between samples. Comparing samples from live

subjects and accident victims, gene expression patterns were reasonably well correlated in blood, brain, and skeletal muscle. Based on the observation of rRNA peaks in the human samples of the correct size to be bacterial, an ancillary study was performed to characterize several pairs of universal primers for the bacterial 16S rRNA.

MATERIALS AND METHODS

Sample Origin and Processing

Room-temperature RNA degradation under laboratory conditions was assessed by collecting five 10 mL serum tubes of blood from four live subjects. To prevent clotting, after collection, caps were removed from the tube, a 5 mm stainless steel ball placed in each tube, the tube recapped, and placed on a tube rotator that rotated the samples end-over-end at ~30 RPM. Two and a half milliliter aliquots were removed from each set of subject-samples at either 0, 0.5, 1, 2, 4, 8, 12, 18, 24, 36, 48, 75, 96, 168 hours (samples C and D) or 0, 1, 2, 4, 8, 12, 18, 24, 36, 48, 75, 96, 168, and 340 hours (samples E and F). Aliquots were added to PAXgene® Blood RNA tubes (PN762165, PreAnalytiX GmbH, 8634 Hombrechtikon CH), inverted ten times to thoroughly mix blood with the PAXgene solution and frozen at -20°C.

Forensic samples were taken from aviation accident victims at the time of autopsy by adding 2.5 mL of blood to PAXgene® Blood RNA tubes. Samples were shipped to the Civil Aerospace Medical Institute (CAMI), Oklahoma City, Oklahoma, in a refrigerated container with tissue samples for toxicological evaluation. From communication with the originating office and official autopsy reports, the date and time of arrival at the office performing the autopsy, the time and date of the autopsy, and the official cause of death were obtained. Two postmortem intervals were calculated, PMI-arr from time of accident to time of arrival at the site of the autopsy, whereupon remains typically are refrigerated, and PMI-aut from time of accident to time of autopsy at which time PAXgene® blood samples were collected, stopping further RNA degradation. To assess the factors that could contribute to RNA degradation, accidents were further divided into three categories of blunt force trauma: high (bftHi) where there was a puncture of the body trunk, dismemberment, or decapitation; medium (bftMed) where the body of the victim was intact without a puncture of the trunk, but where the force of the impact had resulted in an aortic rupture, typically requiring at minimum 60 g; and low (bftLo), the default condition, indicated where there was not a higher level of trauma to the victim. Accident conditions such as water landings and post-accident fire were recorded. Weather conditions including ambient temperature at the time of the crash was gathered from the National Oceanic and Atmospheric Administration and/or from National Transportation Safety Board (NTSB) reports. The date and time of the accident and details of the accident were gathered from the NTSB reports. The official toxicology report from the CAMI forensic toxicology laboratory was used to determine the presence of alcohol, and their determination of its source as ingested or postmortem production used as evidence for growth of potentially RNA degrading microorganisms in the body postmortem.

mRNA

Blood samples were purified to capture small RNAs using either the PAXgene® Blood RNA purification kit (P/N 762164; PreAnalytiX, GmbH) modified to add 1.7 mL of 100% ethanol to alter column binding conditions or with the PAXgene® miRNA purification kit (P/N 763124; PreAnalytiX, GmbH). Both methods were implemented on a Qiacube (Qiagen) sample purification robot. Soft tissue samples were purified with the miRNEasy® Mini kit on the Qiacube after tissue homogenization for 20 seconds at 20 Hz on a TissueLyser (Qiagen/Retsch P/N 20.747.2001) with a stainless steel ball. RINs were assessed on an Agilent BioAnalzer 2100 (Agilent Technologies, Life Sciences and Chemical Analysis Group, 5301 Stevens Creek Boulevard, Santa Clara, CA 95051) with RNA Nano6000 kits (Agilent Technologies; PN

5067-1511). Sample concentrations were determined on a Nanodrop 1000 (NanoDrop Products, Wilmington, DE). Samples were stored at -80°C until use.

qRT-PCR

qRT-PCR was performed on two instruments, a Stratagene MX3005P (Agilent Technologies, Santa Clara, CA) and a Fluidigm® BioMark™ (Fluidigm, South San Francisco, CA). Primer pairs for SYBR® Green reactions on the Stratagene instrument for the mRNAs *ANKRD28*, *ITGA5*, *TMEM8*, *SCAMP2*, *PLCG2*, and *SIGLEC5* were designed using Beacon Designer, version 7.0 (Premier Biosoft International, Palo Alto, CA) and purchased from Integrated DNA Technologies (IDT®, Coralville IA). The ratios of forward and reverse primer concentrations and annealing temperatures were optimized on control RNA to result in 7-point 5X standard curves with efficiencies between 90% and 105%, R2>0.995, and a single reaction product as assayed on the BioAnalyzer on DNA1000 chips (PN 5607-1504, Agilent Technologies).

For experiments performed on the Stratagene, reverse transcription was performed with 100 ng total RNA with SuperScript III First-Strand Synthesis SuperMix (Catalog # 11752-050, Invitrogen Corporation, Carlsbad, CA) according to the manufacturer's recommended protocol in 20 uL reactions. Reactions were performed on an Eppendorf Mastercycler (Eppendorf North America, Hauppauge, NY) with the following program: 25°C for 10 minutes, 50°C for 50 minutes, 85°C for 5 minutes, 4°C for 15 minutes, 37°C for 20 minutes, and a 4°C hold. During the 4°C for 15 minute step, 1µL RNase H was added to each well and mixed using the pipette. The final product was used at a 5ng/µL concentration and stored at -20°C to await qPCR testing.

Final reactions contained 1 ng cDNA, 12.5µL PerfeCta® SYBR® Green FastMix LoRox™ (Cat #95074-05K, Quanta BioSciences, Gaithersburg, MD), forward and reverse primers at the optimized concentrations, and water to bring the total volume to 25µL. Reaction conditions were 95°C for 10 minutes for enzyme activation, then 40 cycles of: 95°C for 5 seconds, primer annealing at the primer-specific optimized temperature for 30 seconds, extension at 72°C for 30 seconds, and data collection following the extension step. Following amplification, melt-curve data was collected as follows: a one minute 95°C melt step, product annealing at the primer-specific annealing temperature for 30 seconds, and a 1°/sec. ramp up to 95°C with data collection every second.

Fluidigm experiments were performed using PrimeTime® (IDT®;Integrated DNA Technologies, Coralville IA) 5' nuclease qRT-PCR assays designed using Beacon Designer 7.0 (PREMIER Biosoft International, Palo Alto, CA), or TaqMan® assays. Where genes to be assayed were chosen from microarray marker discovery phase human subject studies, Affymetrix probe set identifiers annotated as specific to a single gene were converted to NCBI RefSeq identifiers (http://www.ncbi.nlm.nih.gov/refseq/) using NetAFFX (www.affymetrix.com). Reactions were performed in duplicate on Fluidigm® 96.96 Dynamic Arrays (PN# 100-0905, GE 96.96 kit, Fluidigm®, San Francisco, CA). Reverse transcription with Superscript III RT (PN# 11752-250, Thermo Fisher Scientific, Inc.) was performed with 500 ng of total RNA according to the manufacturer's recommended protocol. The resulting cDNA was diluted 1:2 and preamplified with a pooled primer/probe mix diluted 1:100 (TaqMan® assays) from manufacturer's stocks according to the Fluidigm® recommended protocol (Fluidigm®PN 68000133 Rev C) or 1:400 (PrimeTime® assays). To each well of a 96-well plate was added: 1.25 µl of cDNA, 1.25 µl diluted primer pool, and 2.5 µl Taqman® Pre-AMP and PCR master mix (PN# 4384557; Thermo Fisher Scientific, Inc.). The preamplification cycling program was: 10 minute hold at 95º C, then 14 cycles of 95ºC for 15 seconds and 60ºC for 4 minutes; preamplification product was diluted 1:5.

Individual assay mixes were made from each primer/probe stock by 1:2 dilution in Fluidigm® 2x assay loading reagent. Individual sample mixes consisted of 3.0 µl 20X Taqman® Universal PCR Master Mix (PN# 4304437; Thermo Fisher Scientific, Inc.), 0.3 µl Fluidigm® 20x GE Sample Loading Reagent, and 2.7 µl of pre-amplification product. Five µl of each assay and sample mix were loaded in the appropriate wells of a primed chip, the chip loaded on the IFC, and the chip placed into the Biomark instrument for a three-module thermal profile consisting of a thermal mix, a UNG/hot start, and a 40-cycle 2-step PCR for data acquisition. Tissue-specific normalizers were chosen from a panel of 16 genes either recommended (TATAA Biocenter human reference panel; http://www.tataa.com/products-page/reagents/gene_expression_assays_panels/human-reference-gene-panel-short-assay/) or characterized for stable expression in other studies (data not shown). Negative controls and two six-point 5X serial dilution series from control RNA—one post RT, the second post pre-amplification—were included with the samples. Efficiency correction on raw Ct data was performed using values calculated from post-RT sample data.

miRNA

RNA purifications detailed above include miRNAs in the eluates. The Exiqon miRCURY LNA Universal RT kit (PN 203300, Exiqon Inc, Woburn, MA) was used to generate an RT product that could be used for all miRNA primer sets. Fourteen µL of total RNA at a starting concentration of 1.5 ng/µL was combined with 4 µL of 5x Reaction Buffer and 2 µL of Enzyme Mix in a 96-well plate. The reactions were mixed by a gentle vortex and recovered by centrifugation. The manufacturer's recommended reaction conditions were used: 60 minutes at 42°C followed by 5 minutes at 95°C.

The SYBR® Green qPCR procedure was performed according to the instruction manual for the miRCURY LNA Universal RT kit with the SYBR® Green master mix, Universal RT (PN 203450) kit. The RT product was diluted 80x by combining 1.1µL of RT product with 86.9µL of nuclease free water. Final reaction volumes were scaled up to 23 µL from the Exiqon recommendation of 20 µL to accommodate the recommended minimum reaction volume of the Stratagene instrument. The PCR cycle was: 95°C for 10 minutes, then 40 cycles of 95°C for 10 seconds, and 60°C for 1 minute, followed by a melt curve analysis as above.

Live Subject Gene Expression Data Sources and Handling

Forty-four genes were assessed using the BioGPS dataset of tissue-specific expression (21) for highly variable expression levels between tissues. From the Gene Expression Omnibus (GEO; https://ncbi.nlm.nih.gov/geo), Affymetrix® HG-U133Plus2.0 and HG-U133A (GEO platforms GPL570 and GPL96, respectively) CEL files from datasets annotated as originating from pathologically normal brain, cardiac muscle, skeletal muscle, kidney, spleen, lung, and liver were downloaded. All samples from skeletal muscle and liver were from live biopsy samples. From brain, half were annotated as being from live laser microdissection; the remainder were either autopsy samples or unspecified. Samples from the remaining tissues were unspecified. The expression platforms were chosen due to the high concordance of probe sets on the arrays minimizing platform effects and the internal availability of blood expression data from live subjects on HG-U133Plus2.0 chips. For each Affymetrix platform within each tissue type, the R/Bioconductor package ArrayQualityMetrics was used to identify poor quality (median expression < 3) and outlier arrays which were excluded from further analysis. CEL files were RMA summarized and quantile normalized in R v3.3.1 and associated Bioconductor packages. Within each array, probe set level expression date for the 44 genes of interest were extracted from the expression tables. For genes represented by multiple probe sets, the median expression level of all probe sets was calculated. Across arrays of the

same platform, an average gene expression value was calculated. Finally, to arrive at a single tissue-specific expression value for each gene across platforms, a weighted average based on the number of arrays from each platform was calculated. Baseline samples from in-house studies collected on Affymetrix® HG-U133Plus2.0 microarry chips were the source of expression data for whole blood and handled as for tissue data. Affymetrix® probe sets and Taqman® assays were mapped to NCBI RefSeq identifiers to ensure that all data was specific to the same transcript. qRT-PCR data was collected with Taqman® assays on the Fluidigm® BioMark™. For the 44 genes, Spearman rank correlation coefficients were calculated between the microarray data from "live" subjects and qRT-PCR data from victim samples.

Descriptive Statistics and Analysis

Cts were determined within the Stratagene or Fluidigm software using ROX as a reference dye for reaction volume. Interplate calibrators were utilized where necessary to scale data by gene between plates. Data collected on different instruments for the same genes were not compared to each other or pooled for analysis due to markedly different reaction conditions. Pearson correlation coefficients were calculated in Excel 2010, or the cor function in R (ver 3.0.3). Spearman rank-order correlation coefficients were calculated in R. Where p-values are reported for correlation calculations, the coefficient and p-value were calculated using the rcorr function in the Hmisc package for R. T-tests were calculated in R. The threshold for statistical significance was $p < 0.05$)

Detection of Bacterial 16S rRNA by qRT-PCR

Primer sequences to conserved regions of the 16S rRNA were gathered from web resources (http://lutzonilab.org/16S-ribosomal-dna/) and amplicon lengths between various pairs of these previously characterized primers assessed for use in qRT-PCR reactions with SYBR® Green (Table 1A). Nineteen degenerate universal primer sequences to conserved regions of bacterial 16S rRNA were custom synthesized and a commercial primer pair for bacterial 16S detection purchased (IDT, Coralville, IA). Primers 357R, 515R, 1100R, 1185F, 1237R, 1381R, and 1391R are the reverse complement of the published sequences 357F, 515F, 16S:1100.F16, 1185mR, 1237F, 1381F and 1391R, respectively and, in combination with another of the published primers, results in an amplicon length suitable for SYBR® Green reactions. Reverse transcription was performed as above on 1 µg aliquots of *E. coli* RNA and the resulting RT product pooled. All reactions were performed with PerfeCta SYBR® Green master mix as above. After initial screening of 13 different primer pairs with unreplicated 7-point standard curves, five were further characterized in triplicate 7-point standard curves. Of these, four were tested with seven human total RNA samples that had bacterial 16S and 23S rRNA peaks by BioAnalyzer electrophergram, one sample where there were no peaks, and two that had only eukaryotic 18S and 28S rRNA peaks.

Specificity and sensitivity of the 16S:1100.F16/1237R primer pair (set #5) was tested with two different spike-in experiments. First, a 7-point standard curve was constructed in triplicate from 25 to 1.6 e-3 ng *E. coli* RT product with and without a fixed amount of RT product equivalent to 1 ng human total RNA from whole blood. Next, six sets of triplicate reactions with a fixed 5 ng of total RNA but variable ratios of 0.5 plus 4.5 ng, 1 plus 4 ng, 2 plus 3 ng, 3 plus 2 ng, 4 plus 1 ng, and 4.5 plus 0.5 ng *E. coli* and human RT product respectively, were performed. As a proof-of-principle, single reactions were performed with RT product from1 ng of total RNA from 84 aviation accident RNA samples representing all observed rRNA peak types across a variety of accident conditions.

Table 1. Characterization of bacterial 16S rRNA primers.

A. Sequence of universal primers to 16S rRNA constant regions and the pairing of those primers for initial characterization of reaction conditions; “Amplicon Length” is the expected length in basepairs (bp) of PCR product from each primer pair based on E. coli 16S rRNA sequence; “Efficiency” and “R2” are the calculated reaction efficiency and regression from 7-point 5-fold serial dilution unreplicated standard curve on RT product from purified E. coli total RNA; “Range” and “# of Dilutions” are the weight in pgs of the total RNA input amount at each end of the range used to calculate reaction efficiency and the number of dilution points within that range; “Cts from Lowest Curve Point” is the difference in Cts between the lowest curve point and the negative control.

B. Columns as in A. for four primer pairs chosen for further characterization in triplicate reactions on two different plates. Due to an unacceptable reaction efficiency, primer set 13 was dropped after the first plate in favor of primer pair 8 on the second plate.

C. Raw data and results from three triplicate standard curves with primer pair 5. The first two curves are as in B. whereas the third curve is an 11-point curve to test sensitivity at very low amounts of input material.

Table 1A. All universal 16S primer pairs initially screened for detection of bacteria and the resulting data

Pair Number	Forward Primer	Forward Primer Sequence	Reverse Primer	Reverse Primer Sequence	Amplicon Length	Efficiency	R2	Range (pg)	# of Dilutions	Cts from Lowest Curve Point
1	27F	AGAGTTTGATCMTGGCTCAG	357R	CTGCTGCCTCCCGTAGGAG	330	Curve failed				
2	357F	CTCCTACGGGAGGCAGCAG	519R	GWATTACCGCGGCKGCTG	162	Curve failed				
3	515F	GTGCCAGCMGCCGCGGTAA	907R	CCGTCAATTCMTTTRAGTTT	392	125	0.834	100 to 0.16	5	> 10 Cts
4	895F	CRCCTGGGGAGTRCRG	1100R	AGGGTTGCGCTCGTTG	205	101.3	0.985	100 to 0.16	5	> 10 Cts
5	16S:1100.F16	CAACGAGCGCAACCCT	1237R	GTWGCRCGTGTGTAGCCC	137	99.4	0.995	100 to 0.16	5	> 10 Cts
6	16S:1100.F16	CAACGAGCGCAACCCT	1185mR	GAYTTGACGTCATCCM	85	111.4	0.996	500 to 0.16	6	> 10 Cts
7	1185F	KGGATGACGTCAARTC	1237R	GTWGCRCGTGTGTAGCCC	52	Curve failed				
8	1237F	GGGCTACACACGYGCWAC	1381R	CGGTGTGTACAAGRCCYGRGA	144	96.9	0.997	500 to 0.16	6	> 10 Cts
9	1237F	GGGCTACACACGYGCWAC	1391R	GACGGGCGGTGTGTRCA	154	103.2	0.998	500 to 0.16	6	> 10 Cts
10	1237F	GGGCTACACACGYGCWAC	1492R (I)	GGTTACCTTGTTACGACTT	255	91.3	0.998	500 to 0.16	6	8.44
11	1381F	TCYCRGGYCTTGTACACACCG	1492R (I)	GGTTACCTTGTTACGACTT	111	93.8	0.995	500 to 0.16	6	3.22
12	1391F	TGYACACACCGCCCGTC	1492R (I)	GGTTACCTTGTTACGACTT	101	91.8	0.982	500 to 0.16	6	2.36
IDT ReadyMade™ 16S	forward	AGAGTTTGATCCTGGCTCAG	reverse	ACGGCTACCTTGTTACGACTT		96.3	0.981	500 to 0.16	6	No Ct
	R=A or G									
	Y=T or C									
	M=A or C									
	K=G or T									
	W=A or T									
	Curve Failed: Denotes standard curves where Ct values did not agree with the dilution series calling primer pair specificity into question									

Table 1B. Universal 16S primer pairs and data from second round of characterization

Pair Number	Forward Primer	Forward Primer Sequence	Reverse Primer	Reverse Primer Sequence	Amplicon Length	Efficiency	R2	Range (pg)	# of Dilutions	Cts from Lowest Curve Point
5	16S:1100	CAACGAGCGCAACCCT	1237R	GTWGCRCGTGTGTAGCCC	155	99.4	0.995	100 to 0.16	5	> 10 Cts
						100.2	0.998	500 to 0.032	7	9 Cts
6	16S:1100	CAACGAGCGCAACCCT	1185RmR	GAYTTGACGTCATCCM	85	106.2	0.999	500 to 0.16	6	10 Cts
						89.9	0.998	500 to 0.032	7	7 Cts
9	1237F	GGGCTACACACGYGCWAC	1391R	GACGGGCGGTGTGTRCA	154	108.7	0.999	500 to 0.032	7	10 Cts
						93.1	0.999	500 to 0.032	7	7.25 Cts
IDT ReadyMade™ 16S	forward	AGAGTTTGATCCTGGCTCAG	reverse	ACGGCTACCTTGTTACGACTT		120.2	0.998	100 to 0.032	6	No Ct
8	1237F	GGGCTACACACGYGCWAC	1381R	CGGTGTGTACAAGRCCYGRGA	144	88.8	1.000	500 to 0.032	7	6.7 Cts

Table 1C. Standard curve data for primer pair 5 characterization			
Bacterial RNA Input (pg)	Curve 1; Ct	Curve 2; Ct	Curve 3; Ct
500	7.65	6.62	6.96
100	8.74	8.74	8.74
20	10.52	10.71	11.18
4	12.78	13.08	13.73
8.00E-01	15.48	15.5	15.97
1.60E-01	17.92	18.2	18.11
3.20E-02	19.51	20.36	21.05
6.40E-03			23.56
1.28E-03			26.96
2.56E-04			28.19
5.12E-05			29.55
NoRT	30.72	29.24	29.72
Efficiency	99.4	100.2	96.6
Curve Fit (R2)	0.995	0.998	0.996
Range	100 to 0.16	Full	Full

RESULTS

Qualitative Assessment of Samples Collected in Serum Tubes

To assess RNA degradation over time under laboratory conditions, RNA was purified from four samples collected as two groups of two, C and D, and E and F, with similar time profiles (Materials and Methods). Because the first pair of samples, C and D, showed no degradation between baseline and 1 hour and had not completely degraded after one week, the second pair of samples, E and F, were not sampled at the 0.5 hour timepoint and were extended to include a 340 hour sample. Yield was negatively correlated with time (r = -0.794, Fig. 2A). RIN was positively correlated with average yield (r = 0.751, Fig. 2B) and negatively correlated with time (r = -0.984, Fig. 2C). Under these controlled laboratory conditions, yields (Fig. 2A) and RINs (Fig. 2C) remained at approximately baseline levels to 48 hours but, beyond this point, began to fall, and by 340 hours both metrics had fallen to near zero.

Figure 2. Data plots from *in vitro* samples. For four *in vitro* samples from baseline (0 hours) to 340 hours, scatter plots of A. Yield versus Time; B. RIN versus yield; C. RIN versus time.

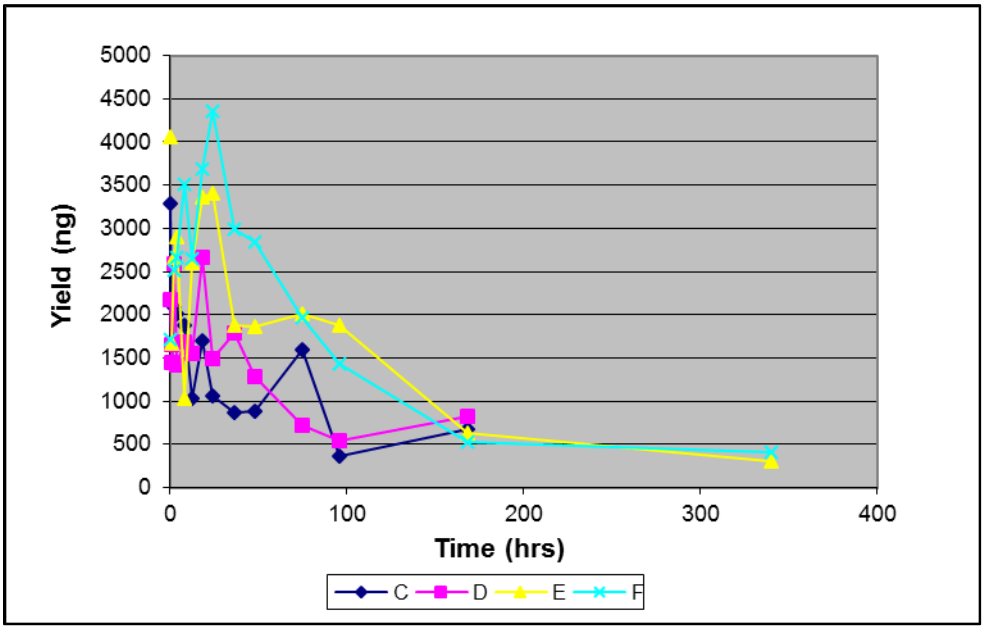

Figure 2A. Serum tube RNA Yield versus Time of sampling.

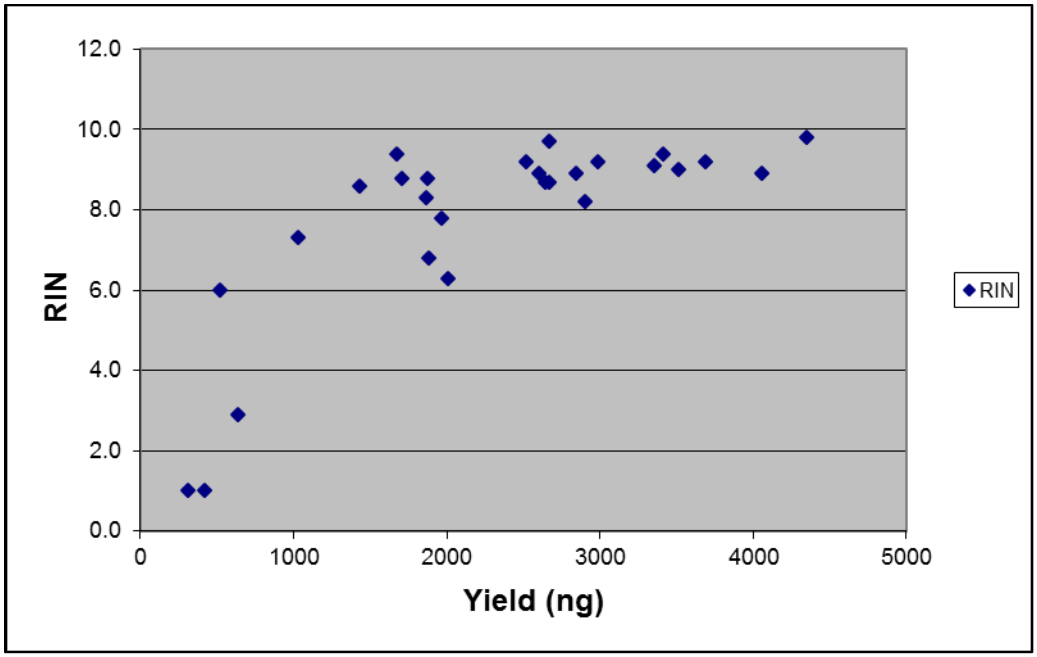

Figure 2B. Serum tube RIN versus Yield.

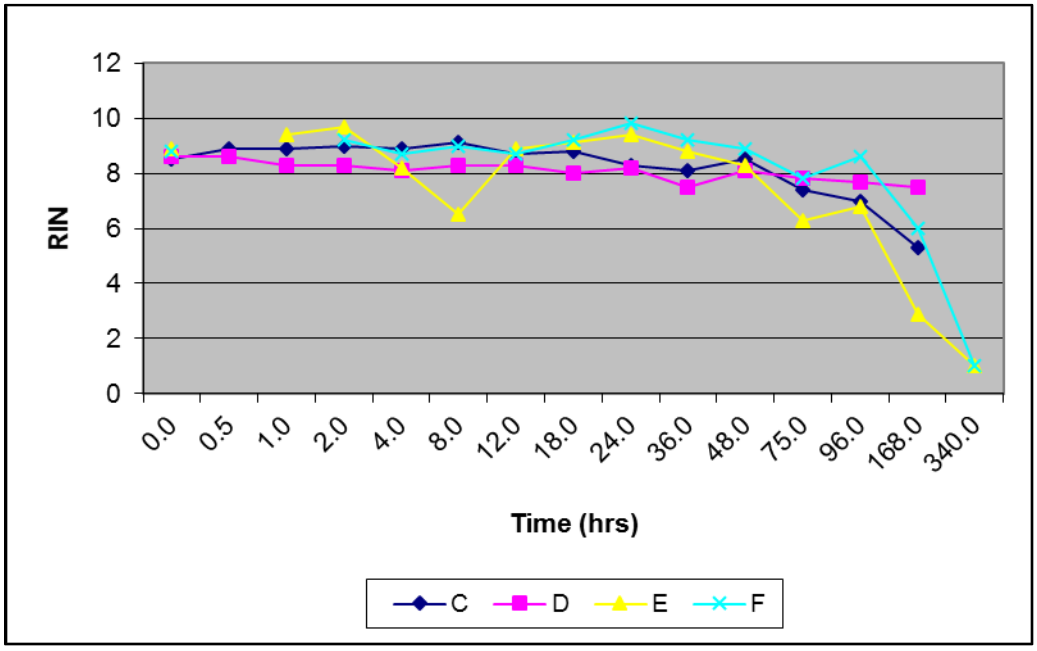

Figure 2C. Serum tube RIN versus Time of sampling.

Quantitative PCR of Serum Tube Samples

Time-dependent gene expression patterns were obtained for six messenger RNA (mRNA) and four micro-RNA (miRNA) genes by qRT-PCR on these serum tube samples. The six mRNAs were *ANKRD28*A, *TMEM8*, *PLCG2*, *ITGA5*, *SCAMP2A*, and *SIGLEC5*. In previous characterization, these genes were all easily detected in fresh blood samples by SYBR® Green based assays and had efficiencies greater than 90% (data not shown). Likewise, these genes were easily detected in these samples; however, at times greater than 96 hours, Ct's started to rise across all six genes reflecting a lower apparent expression (See Appendix). To address relative degradation rates between genes, timepoint data was averaged across the

four samples and correlation coefficients calculated between all possible pairs of genes. The coefficients averaged 0.70 (data not shown) suggesting that these six genes degraded at similar rates in all four samples. Over time, sample RINs were negatively correlated with Cts (average r = -0.85) suggesting that degradation (decreasing RIN) correlates with decreased apparent expression level (Cts increase as expression decreases; a negative correlation value demonstrates a decrease in apparent expression level associated with RNA degradation; see Appendix).

Likewise, four miRNAs were assessed in these samples, *hsa-miR16a*, *hsa-let7a*, *hsa-miR142-3p*, and *hsa-miR142-5p*. As above, correlation coefficients were calculated for all possible pairwise comparisons of miRNAs and these averaged 0.44 (data not shown) suggesting greater variability in degradation pattern of miRNAs compared to mRNAs. Due to low yields at the 340 hour timepoint, insufficient RNA was available to perform qRT-PCR with the miRNA primers; nonetheless, the last timepoint available in these assays at 168 hours did not show the increased Ct evident at this point in the mRNA samples.

Qualitative Curation of 34 Victim Samples

Next, purified RNA from 34 victim blood samples was assessed for quality metrics. To determine what other factors might affect postmortem gene expression patterns, cadaver, accident, and weather condition data was gathered (Materials and Methods; Table 2). Gene expression was assessed for six mRNAs and four miRNAs (Table 3A). Pearson correlation coefficients were calculated between RIN, RNA yield, PMI-arr, and PMI-aut. In contrast to the RNA samples purified from serum tubes, RNA samples from aviation accident victims were highly variable showing no correlation between RIN and PMI-arr (r = 0.06), RIN and PMI-aut (r = -0.08), or RIN and RNA yield (r = -0.06, Table 3B). Only PMI-arr and PMI-aut were correlated (r = 0.79). RNA yields for some samples were observed to be greater than expected for whole blood samples, which may be attributed to microbial growth in the sample and subsequent RNA purification of mixed microbial and human RNA. Poor correlation also was found between RIN and age of the victim, body type as suggested by BMI of the victim, and air temperature at time of the accident (Table 3B).

By comparing migration times on the BioAnalyzer, prokaryotic peaks corresponding to 23S and 16S rRNAs can be distinguished from eukaryotic 28S and 18S rRNAs (Fig. 1). RIN scores were not different between samples with prokaryotic and eukaryotic peak sizes by t-test ($p = 0.332$) suggesting that the RIN algorithm does not utilize migration time, only rRNA peak Area Under the Curve. Because RIN scores depend on the presence of rRNA peaks, comparison between RIN values of samples that had no rRNA peaks and those that did was not performed. RINs from the five samples with prokaryotic peaks were positively correlated to both PMI-arr (r = 0.657) and PMI-aut (r = 0.685). Conversely, samples with eukaryotic rRNA peaks showed negative correlations between RIN and both PMI-arr (r = -0.565) and PMI-aut (r = -0.602). To characterize the degree to which trauma affects RNA quality, boxplots were created of RINs by trauma category for samples with eukaryotic or no rRNA peaks. Samples gathered from victims that underwent moderate or high trauma levels had significantly lower RIN scores by t-test than samples from cadavers classified as low trauma (Fig. 3, $p < 0.02$).

In assessing individual accidents, 22 of the 34 victim samples were independent—one sampled victim per accident (Table 2). Of these, Samples 17 and 20 were from water landings. Sample 17 had no rRNA peaks and a RIN of 2; Sample 20 had a RIN of 7.7 but prokaryotic rRNA peaks. Both were positive for postmortem produced ethanol. Seven accidents that resulted in a single victim sample involved fire and the trauma level was bftLo in three, bftMed in two and bftHe in two. The two bftHi samples had RINs of 0, and 3.2, whereas only one of the remaining samples had a RIN less than 6. None of these samples had prokaryotic rRNA peaks.

Table 2. Curated accident and victim data from 34 accident victims. “Key” indexes the victim sample across all tables where that sample was assayed. “Dependent” denotes victims linked by being in the same accident except 6a and b, and 24 a and b where duplicate tubes were collected at autopsy, RNA purified and data collected. “Trauma Level,” see text. “Fire” and “Water” indicate accident conditions—1 is affirmative, 0, negative. “Yield” is the total amount of RNA from purification; “RIN” is the RNA Integrity Number from BioAnalyzer2100 electropherograms; “rRNA Peak Origin” indicates the presence or absence of rRNA peaks and the interpretation of their source based on the migration times of those peaks on BioAnalyzer electropherograms; Alcohol positives are from toxicological analysis, source of alcohol is based on data interpretation by the CAMI toxicology staff—“PM,” postmortem alcohol production. “Month,” “Date,” “AirTemp,” “Weather,” and “Time of Day” are statistics from the day of the accident potentially impacting the rate RNA degradation (year is left off intentionally to protect the identity of the victim). “PMI-arr” and “PMI-aut” are the calculated postmortem interval from time of accident to either arrival at the site of autopsy (arr) or time of the autopsy itself (aut); arrival at the site of autopsy is taken as an approximation of the time the cadaver is placed under chilled conditions slowing RNA degradation. “Notes” in last column give further details about these accidents, usually whether victims were in the same or different craft.

Sample	Dependent	Trauma Level	Fire	Water	Original conc (ng/µl)	Elution volume (ul)	Yield	RIN	rRNA Peak Origin	BMI	Alcohol Present	Month	Date	AirTemp (°F)	Age	Weather	PMI-arr	PMI-aut	Time of Day	Notes
1	1	hi	0	1	600.0	44	26400.0	2.4	none	29.8	neg	9	19	73	66	overcast	8.50	47.5	1106	Pilot 1
2	1	hi	0	1	418.6	44	18419.3	6.9	prokaryotic	no info	neg	9	19	73	30	overcast	8.50	24.4	1106	Passenger 1
3	2	hi	1	0	15.1	20	302.2	0	none	26.6	neg	10	26	66	44	severe t-storms	32.30	45.5	1142	
4	3	hi	0	0	2007.8	44	88343.2	2.4	none	no info	neg	8	24	66	63	overcast/fog	2.25	72	828	Pilot 3
5	3	hi	0	0			87234.4	2.6	none	24.6	neg	8	24	66	71	overcast/fog	2.25	73.2	828	Passenger 3
6	4	hi	0	0	1717.2	54	0.0	0	none	26.0	neg	1	29	57	37	clear/ icing	8.50	16.25	1500	
7	5	hi	1	0	276.6	80	92728.8	6.3	eukaryotic	22.7	pos - PM	3	20	70	59	clear	27.50	45.5	1145	Pilot 5: fixed wing aircraft; mid-air collision with experimental 5
8	5	med	0	0	1982.6	44	1798.7	3.9	prokaryotic	23.6	neg	3	20	70	73	clear	10.10	23.25	1145	Experimental 5: midair collision with Pilot 5
9.1	6a	hi	0	0	110.0	44	22128.0	5	eukaryotic	no info	neg	7	23	90	85	clear	10.75	19.6	1225	
9.2	6b				268.0	80		7												
10	7	hi	1	0	184.8	44	4840.0	3.2	eukaryotic	42.0	neg	11	15	48	66	overcast/fog	42.18	63.1	1710	Slight prokaryotic rRNA peaks above background
11	8	hi	0	1	252.5	44	8131.2	8.3	prokaryotic	26.3	pos - PM	8	8	75	60	clear	80.40	94	1153	Pilot 8
12	8	med	0	1	109.3	44	19100.4	8.1	prokaryotic	29.5	neg	8	8	75	32	clear	29.10	45.6	1153	Passenger 8

Sample	Dependent	Trauma Level	Fire	Water	Original conc (ng/ul)	Elution volume (ul)	Yield	RIN	rRNA Peak Origin	BMI	Alcohol Present	Month	Date	AirTemp (°F)	Age	Weather	PMI-arr	PMI-aut	Time of Day	Notes
13	9	lo	1	0	181.5	80	11110.0	2.6	eukaryotic	no info	pos - ingestion	9	7	84	34	clear	37.75	39.25	1916	Passenger 9, minor eukaryotic rRNA peaks
14	9	lo	1	0	147.4	44	6485.6	6.1	eukaryotic	30.2	neg	9	7	84	47	clear	37.75	38.25	1916	Pilot 9
15	10	lo	0	0	154.7	44	4811.0	5.3	eukaryotic	23.0	neg	9	1	81	59	clear	19.40	19.67	1425	
16	11	lo	0	0	59.5	44	14522.4	5.5	eukaryotic	35.2	neg	4	17	77	63	clear	15.35	16.6	1925	
17	12	lo	0	1	165.2	44	6806.8	2	none	28.0	pos - PM	9	30	82	61	clear	69.35	89.75	1415	
18	13	lo	1	0	234.1	44	2616.2	7.1	eukaryotic	26.9	neg	1	18	55	52	clear	4.00	23.67	1350	
19	14	lo	0	0	150.5	54	7268.8	3.2	eukaryotic	22.0	pos - ingestion	12	24	54	29	cloudy/ light rain	84.50	111.5	1100	
20	15	lo	0	1	238.3	54	10300.4	7.7	prokaryotic	28.3	pos - PM	7	18	81	52	few clouds/storms in vicinity	50.67	59.5	2330	
21	16	lo	1	0	109.9	54	8127.0	7.7	eukaryotic	25.1	neg	6	21	70	42	clear	1.75	21.3	1257	Pilot 16
22	16	lo	1	0	124.4	44	5934.6	6.7	none	24.3	neg	6	21	70	46	clear	1.90	25	1257	Passenger 16
23	17	lo	1	0	105.2	44	12868.2	6.9	eukaryotic	29.4	neg	3	30	59	59	clear/ windy	4.00	14.75	1310	
24	18	lo	0	0	2026.4	54	5471.4	5.6	eukaryotic	27.2	neg	11	14	57	53	low clouds	6.00	48	1100	
25	19	lo	1	0	177.9	80	4629.7	7.5	eukaryotic	32.5	neg	4	9	70	51	clear	5.30	5.3	756	
26	20	med	1	0	434.1	44	109425.6	2.4	none	26.1	pos - PM	7	14	84	54	overcast/storms in vicinity	2.50	15.5	2105	
27	21	med	0	0	151.3	44	14228.0	4.7	eukaryotic	28.4	neg	8	8	63	36	clear	5.25	30.33	700	
28	22	med	0	0	195.3	54	6657.2	2.4	none	25.1	neg	7	22	72	33	few clouds	3.50	24.8	1125	
29	23	med	0	0	33.3	54	10546.2	6	eukaryotic	29.5	neg	2	27	43	3	clear	2.00	15.5	1625	
30.1	24a	med	1	0	96.7	44	4253.9	6.2	eukaryotic	no info	neg	11	7	57	59	foggy	5.25	25.75	800	
30.2	24b				159.2	44	7003.0	5.7												
31	25	med	0	0	37.5	54	2023.9	1.9	none	25.2	neg	7	11	70	52	clear	14.80	38.25	2220	
32	26	med	0	0	146.3	30	4388.4	2.4	none	29.7	neg	5	9	NA	46	clear/ windy	22.10	45.25	1317	
33	27	med	0	0			0.0	0	none	26.6	neg	9	9	72	66	possible fog	8.90	12.5	1900	
34	28	med	0	0	21.7	20	434.8	0	none	24.8	neg	6	24	81	57	few clouds	3.75	22.67	1120	Slight eukaryotic rRNA peaks above background

Table 3. Numerical accident, victim and qRT-PCR data from 34 accident victims. Numerical accident and victim data from Table 2 and qRT-PCR raw Ct data for six mRNAs and four miRNAs from 34 accident victims. B. Descriptive statistics for columnar data in A. and Pearson correlation coefficients as indicated.

A.

									messenger RNAs						micro RNAs			
Sample	Dependent	Yield	RIN	PMI-arr	PMI-aut	BMI	Air Temp	Age	*ANKRD28*	*ITGA5*	*TMEM8*	*Scamp2a*	*PLGC2*	*SIGLEC5*	*mir16a*	*let7a*	*miR-142-3p*	*miR-142-5p*
1	1	26400.0	2.4	8.5	47.5	29.8	73	66	35.56	30.05	32.38	30.42	32.43	26.85	25.7	45	28.59	34.2
2	1	18419.3	6.9	8.5	24.4		73	30	45	40	40	32.22	33.23	34.9	23.63	37.2	29.5	33.72
3	2	302.2	0	32.3	45.5	26.6	66	44	40	45	40	40	45	NA	NA	NA	NA	NA
4	3	88343.2	2.4	4.25	72		66	63	28.57	28.52	31.56	30.1	31.52	29.09	25.92	31.97	29.69	34.24
6	4	0.0	0	8.5	16.25	26	57	37	45	45	45	45	45	45	45	45	45	45
7	5	92728.8	6.3	17.5	45.5	22.7	70	59	32.72	29.49	40	28.66	32.84	26.4	28.59	40	28.82	34.53
9.1	6a	22128.0	5	10.75	19.6		90	85	26.18	23.66	24.62	25.2	26.12	25.8	31.4	36.32	40	40
5	3	87234.4	2.6	4.25	73.2	24.6	66	71	26.73	27.05	28.65	28.52	29.21	28.62	27.68	33.29	30.11	33.89
10	7	4840.0	3.2	42.18	63.1	42	48	66	31.29	25.02	27.64	26.03	27.02	27.26	25.88	30.96	26.82	30.91
11	8	8131.2	8.3	72.3	94	26.3	75	60	29.58	29.93	30.42	28.99	30.7	30.7	18.29	28.9	21.06	27.77
13	9	11110.0	2.6	13.5	39.25		84	34	28.99	24.97	25.82	24.83	25.78	24.95	21.15	28.13	25.29	28.77
15	10	4811.0	5.4	19.35	19.67	23	81	59	26.44	22.1	21.67	21.52	24.04	23.71	24.46	35.39	29.1	31.17
16	11	14522.4	5.5	14.3	15.5	35.2	77	63	26.66	23.87	24.76	24.14	25.36	25.45	30.52	40	40	37.88
14	9	6485.6	6.1	37.75	38.25	30.2	84	47	25.65	22.48	22.23	21.41	22.77	23.61	22.92	30.45	24.1	29.13
17	12	6806.8	2	69.35	89.75	28	82	61	45	45	45	32.49	45	34.96	24.72	34.82	29.46	32.16
18	13	2616.2	7.1	4	23.67	26.9	55	52	28.33	22.64	24.54	24.13	23.99	23.83	23.38	33.25	25.8	28.88
19	14	7268.8	3.2	84.4	111.5	22	54	29	29.67	23.96	25.74	24.6	25.78	26.52	25.59	33.52	27.25	32.32
20	15	10300.4	7.7	50.67	59.5	28.3	81	52	34.42	35.55	37.58	31.08	40	33.85	20.85	32.8	26.59	31.1
21	16	8127.0	7.7	1.75	21.3	25.1	70	42	31	24.36	26.44	25.18	25.47	26.43	26.25	34.03	25.33	29.37
23	17	12868.2	6.9	4	14.75	29.4	59	59	27.52	23.62	24.3	23.6	24.81	24.74	24.51	31.78	23.2	29.42
22	16	5934.6	0	3.9	25	24.3	70	46	28	23.34	25.28	24.46	24.35	25.39	23.18	28.83	23.7	29.4
24	18	5471.4	5.6	6	48	27.2	57	53	28.99	25.77	26.75	25.89	26.64	26.75	21.94	31.12	24.75	29.26
25	19	4629.7	7.5	5.3	5.3	32.5	70	51	25.74	23.47	25.14	24.6	25.19	25.84	22.4	30.05	25.03	31.27
26	20	109425.6	2.4	2.5	15.5	26.1	84	54	33.48	32.27	32.36	29.37	30.36	29.48	29.33	35.02	31.07	34.54
27	21	14228.0	4.7	5.25	30.33	28.4	63	36	26.91	24.1	24.88	24.18	25.41	24.88	29.85	36.92	40	37.77
12	8	19100.4	8.1	29.9	45.6	29.5	75	32	26.3	26.62	25.68	24.48	25.09	25.52	19.59	28.72	20.94	28.05
28	22	6657.2	2.4	3.5	24.8	25.1	72	33	31.46	26.84	27.16	25.78	28.8	30.55	24.91	36	27.95	33.3
29	23	10546.2	6	2	15.5	29.5	43	3	45	45	45	45	45	24.79	23.62	31.53	25.97	30.6
8	5	1798.7	3.9	10.1	23.25	23.6	70	73	29.68	25.76	27.04	25.38	27.89	27.9	23.64	31.98	25.07	28.73
30.1	24a	4253.9	6.2	5.25	25.75		57	59	28.38	23.5	25.71	24.76	24.9	25.23	22.45	32.75	25.58	30.33
31	25	2023.9	1.9	14.8	38.25	25.2	70	52	28.42	24.8	26.44	24.3	25.63	26.84	27.76	40	28.64	33.04
32	26	4388.4	2.4	22.1	45.25	29.7		46	26.99	27.69	31.94	28.37	34.58	27.27	28.23	40	28.56	35.93
33	27	0.0	0	8.9	12.5	26.6	72	66	45	45	45	45	45	45	45	45	45	45
34	28	434.8		3.75	22.67	24.8	81	57	27.55	25.36	26.64	25.92	26.24	NA	NA	NA	NA	NA
9.2	6b	21440.0	7		24				26.67	24.13	25.2	24.81	25.9	25.54	32.54	39.09	34.98	45
30.2	24b	7003.0	5.7		24				28.09	23.26	25.34	24.31	24.3	24.57	22.62	31.26	25.01	29.67

B.

		Yield	RIN	PMI-arr	PMI-aut	BMI	Air Temp	Age	*ANKRD28*	*ITGA5*	*TMEM8*	*Scamp2a*	*PLGC2*	*SIGLEC5*	*mir16a*	*let7a*	*miR-142-3p*	*miR-142-5p*
									messenger RNAs						micro-RNAs			
	Median	7037.800	4.700	8.700	28.040	26.600	70.000	52.500	28.990	25.765	26.895	25.835	26.830	26.635	24.815	33.405	27.600	31.715
	Mean	18077.205	4.431	18.569	37.775	27.538	69.545	51.176	31.416	28.866	30.109	28.187	30.038	28.183	26.279	34.738	29.058	33.128
Pearson Correlations (Row X Column)	Yield		-0.057	-0.200	0.156	-0.188	0.144	0.227	-0.075	-0.010	0.114	0.004	-0.005	-0.052	0.067	0.043	0.059	0.123
	RIN			0.070	-0.072	0.157	0.020	-0.079	-0.329	-0.351	-0.317	-0.405	-0.395	-0.429	-0.577	-0.430	-0.462	-0.520
	PMI-arr				0.790	0.056	0.086	-0.018	0.092	0.142	0.125	-0.022	0.177	0.102	-0.250	-0.197	-0.220	-0.207
	PMI-aut					-0.067	-0.074	0.068	-0.014	0.049	0.090	-0.053	0.111	0.024	-0.300	-0.220	-0.281	-0.234
	ANKRD28									0.945	0.922	0.872	0.875	0.746	0.424	0.434	0.351	0.398
	ITGA5										0.956	0.931	0.961	0.791	0.439	0.399	0.380	0.447
	TMEM8											0.908	0.955	0.748	0.444	0.454	0.359	0.457
	Scamp2a												0.928	0.762	0.584	0.456	0.469	0.556
	PLGC2													0.764	0.466	0.446	0.396	0.485
	SIGLEC5														0.662	0.517	0.559	0.638
	mir16a															0.776	0.896	0.929
	let7a																0.741	0.813
	mir142-3p																	0.956

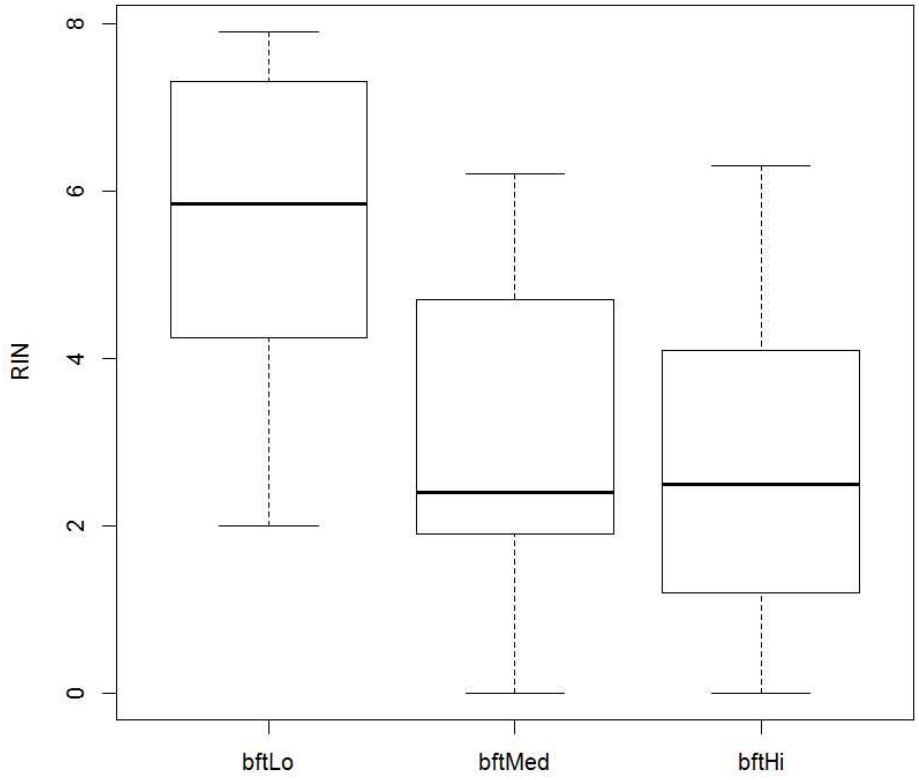

Figure 3. Boxplots of RINs categorized by degree of blunt force trauma.
Blunt force trauma (bft) was categorized for 34 victims (High, Medium, or Low; Materials and Methods) and RINs plotted by category for the RNA purified from those samples. In all boxplots, the interquartile range (25^{th} to 75^{th}) quartiles are the upper and lower bounds of the boxes, and outliers are individually indicated outside the whiskers. The bar within the interquartile box is the data median.

Six of the accidents resulted in remains from two victims ("Dependent," Table 2), which allowed an assessment of sample variability from near identical accident conditions. In five of these instances, samples were collected from individuals in a single aircraft whereas one accident was a mid-air collision between an experimental and a fixed-wing aircraft (Dependent 5) which caught fire. The sample from this pilot of this craft had a RIN of 6.3 and eukaryotic rRNA peaks whereas the sample from the other pilot had a RIN of 3.9 based on prokaryotic rRNA peaks. Dependents 1 and 8 were water landings; three of these four victims had acceptable RIN scores but prokaryotic rRNA peaks, whereas the fourth had a poor RIN of 2.4 and no rRNA peaks. Two of the single craft caught fire, Dependents 9 and 16, and three of these four victim samples had RINs greater than 6. In two Dependent cases, 1 and 8, prokaryotic peaks were present in one sample but not the other; however, in the pair of samples from Dependent 1, Sample 2, with prokaryotic peaks, had a higher RIN score than Sample 1 with no rRNA peaks. In Dependent 5, Sample 8 with prokaryotic peaks had a lower RIN than Sample 7 which had eukaryotic rRNA peaks. In Dependent 9, both samples had eukaryotic peaks but very different levels of RNA degradation by RIN score. In the remaining three Dependent pairs of samples, RINs were relatively the same and the rRNA peak types were consistent. In summary then, even within samples sourced from different individuals that underwent nearly identical accident conditions, there was wide variability in RNA quality and bacterial growth.

qRT-PCR of 34 Accident Victim Samples

The same six mRNA assays and four miRNA assays were used for qRT-PCR with the 34 accident victim samples to assess our ability to detect these genes in degraded samples. Samples from two victims, 5 and 33, yielded no RNA and could not be assayed. To address relative expression levels between samples, Pearson correlation coefficients were calculated between genes for all possible pairs of both the mRNA and miRNA assays (Table 3B, Fig. 4) across subjects. There was very high correlation within classes of molecules; correlation coefficients averaged 0.87 and 0.85 within mRNA- and miRNA-assays respectively, which was higher than observed in the serum tube samples. Close examination of the data suggested that *SIGLEC5* had a slightly different degradation pattern from other mRNAs and exclusion of this gene resulted in an increase to 0.925 for the remaining five mRNAs. *SIGLEC5* had an average coefficient of 0.762 to the other five mRNAs and, by t-test, the difference in average coefficients was statistically significant ($p = 1.03$ e-07). The average correlation coefficient between mRNA and miRNA assays was 0.467 and was significantly different than the average coefficients between all mRNA ($p = 3.5$ e-17) or miRNA assays ($p = 6.1$ e-11). From two victims, 9 and 30, two PAXgene® RNA blood tubes were included in the collected samples allowing an opportunity to assess autopsy sampling reproducibility. As expected, correlation coefficients from qRT-PCR data across all ten genes were quite high ($r > 0.99$) in sample pairs.

Next, Ct values were compared for each gene by the type of rRNA peak present in electropherograms: prokaryotic, eukaryotic, or none. With one exception, *SIGLEC5* (Fig.5A), samples with prokaryotic rRNA peaks had a Ct range approximately the same as samples that had no rRNA peaks whereas samples with eukaryotic rRNA peaks had a consistently lower range of Ct values (higher apparent expression) as represented by *ITGA5* (Fig. 5B). For *SIGLEC5,* samples with prokaryotic rRNA peaks had a higher Ct range than samples with no rRNA peaks, which in turn had a higher range of Cts than samples with eukaryotic peaks. Together, these results speak to the specificity of the assays for mRNA target molecules in addition to the presence of eukaryotic rRNA peaks as a marker for samples more likely to contain intact human mRNA species. All four miRNAs exhibited a different general pattern where samples with no rRNA peaks had consistently higher Ct ranges than samples with eukaryotic rRNA peaks, and samples with eukaryotic rRNA peaks had ranges slightly higher than those with prokaryotic peaks (represented by *hsa-miR142-5p,* Fig. 5C). The observed pattern of lower apparent expression (higher Cts) in samples with no

rRNA peaks relative to those with eukaryotic rRNA peaks was consistent with the mRNAs; however, samples with prokaryotic rRNAs showing the lowest Ct values (highest expression levels) suggests that some assays for eukaryotic miRNA genes may not be specific for their human target genes in the presence of bacterial RNA.

A further analysis by Pearson correlation coefficient between expression of individual genes and the RNA quality/accident condition metrics of RNA yield, RIN, PMI-arr, and PMI-aut, within rRNA peak types in the samples was performed. (Table 4). Negative correlation values between Cts and RIN or yield support the hypothesis that decreased degradation indicated by increasing RIN or higher yield positively affect expression detection. Positive correlation between Cts and PMI support the hypothesis that longer PMI decreases expression detection. With one exception, correlation coefficients between expression and these four metrics were poor ($-0.3 < r < 0.3$) across the three rRNA peak origin types. The exception was the mRNA *PLCG2,* which in samples without rRNA peaks, had a statistically significant positive correlation to PMI-arr ($r = 0.578$; $p = 0.038$). Note that this gene was not correlated to PMI-arr in samples with eukaryotic or prokaryotic rRNA peaks and the Pearson correlation between RIN and PMI-arr on samples with no rRNA peaks was not significant ($r = 0.047$). Three of the four miRNAs tested, *hsa-miR16a, hsa-let7a,* and *hsa-miR142-3p,* were negatively correlated to PMI-aut ($r < -0.6$) confirming the possibility that assays for human miRNAs may cross-react with bacterial RNAs.

Having determined that a relative expression pattern was maintained across victim samples for six mRNAs (see above), Taqman® qRT-PCR data was collected for 35 mRNAs (including the original six) on a total of 87 victim samples (including 29 of the original 35). Assays were chosen for their potential in aviation accident investigation as either markers or normalizers for hypoxia, alcohol use, or radiation exposure. Because these analyses were focused on relative gene expression levels within samples rather than absolute expression, the Spearman rank order correlation was calculated on expression data for each sample against all other samples (Table 5). When the 86 coefficients for each sample were averaged, 20 had average coefficients less than 0.6. Of these, eight were from samples with no rRNA peaks, three from samples with prokaryotic peaks, eight from samples with eukaryotic peaks, and one was a sample with both types of peaks. When the total number of samples with each peak type was considered, the three samples with average Rho < 0.6 and prokaryotic rRNA peaks represented 42.9 % of all prokaryotic peak samples whereas only 28.6% of all samples without peaks, 17.4% of eukaryotic samples, and 16.7% of samples with both peak types were below this threshold for average Rho. As expected, samples without rRNA peaks had poor RINs, averaging 1.95. The eight samples with average Rho < 0.6 and eukaryotic peaks had average RINs of 4.64 and none had average Rho < 0.5. In comparison, the 39 eukaryotic peak samples with average Rho > 0.6 had average RINs of 5.47. Coincidentally, the same average RIN was found in eukaryotic peak samples and average Rho > 0.8. Looking at all 15 samples with average Rho > 0.8, 11 had eukaryotic peaks, two had none, one had both, and one had prokaryotic.

There were victims from whom two samples were collected. Samples 9.1 and 9.2 both had eukaryotic peaks and RIN scores of 5.0 and 7.0 (Table 2), respectively. Both samples had average Rho less than 0.6 to other samples but coefficients of 0.989 to each other (Table 5) indicating that this male subject was an outlier to the population overall. Samples 30.1 and 30.2 also had eukaryotic peaks, RINs of 6.2 and 5.7 respectively, average Rho of 0.81 to all other samples, and a coefficient of 0.996 to each other.

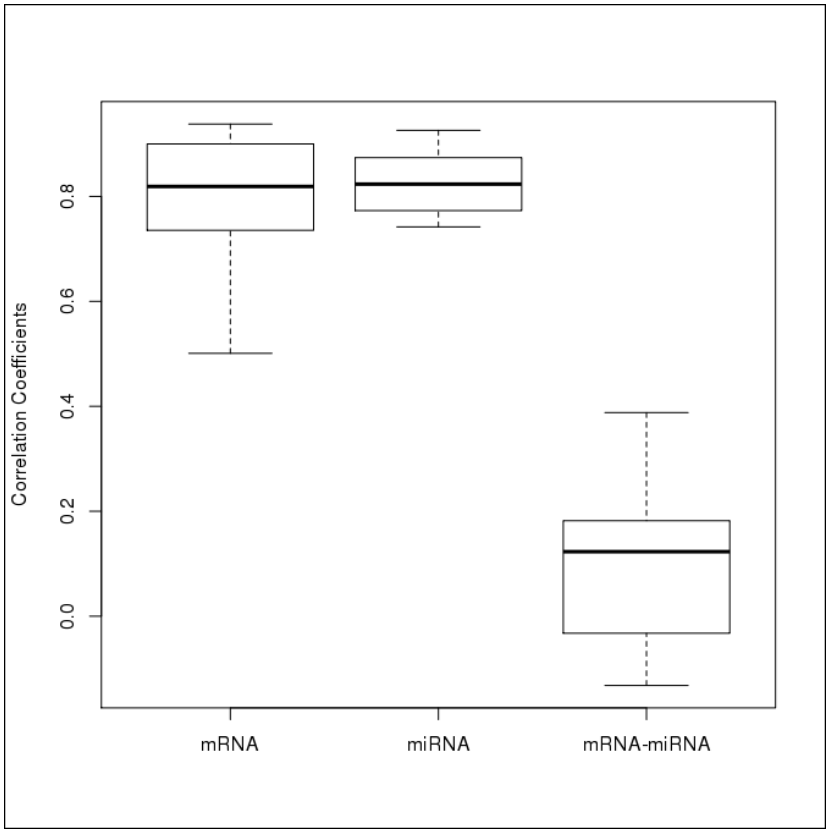

Figure 4. Pearson correlation coefficient boxplots by class comparison.
Pearson correlation values were calculated between expression values of six mRNAs and four miRNA molecules obtained by qRT-PCR on unique samples from 34 victims (Table 3). The correlation coefficients for mRNAs vs. mRNA, mRNAs vs. miRNAs, and miRNAs vs. miRNAs were plotted.

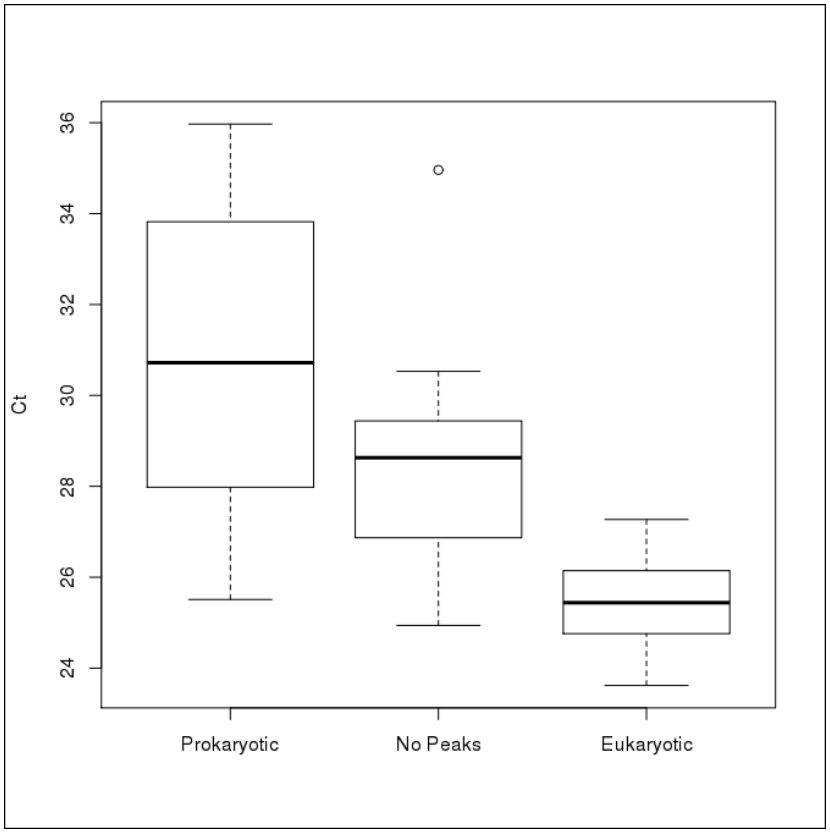

Figure 5A. Expression of three genes by rRNA peak type.
Expression data from two mRNAs, SIGLEC5 (A) and ITGA5 (B), and the miRNA, hsa-Mir142-5p (C), are boxplotted by rRNA peak type evident in BioAnalyzer data from 34 victim samples. SIGLEC5 is shown as a possible outlier to the other mRNAs; ITGA5 is shown as a representative mRNA gene in this analysis.

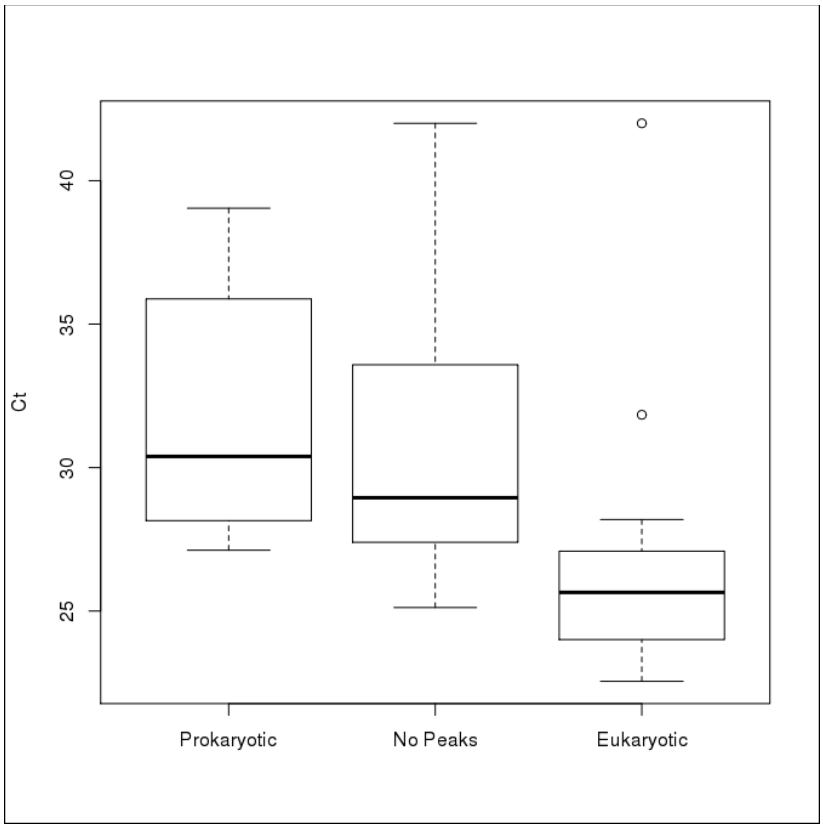

Figure 5B. Expression of three genes by rRNA peak type.
Expression data from two mRNAs, SIGLEC5 (A) and ITGA5 (B), and the miRNA, hsa-Mir142-5p (C), are boxplotted by rRNA peak type evident in BioAnalyzer data from 34 victim samples. SIGLEC5 is shown as a possible outlier to the other mRNAs; ITGA5 is shown as a representative mRNA gene in this analysis.

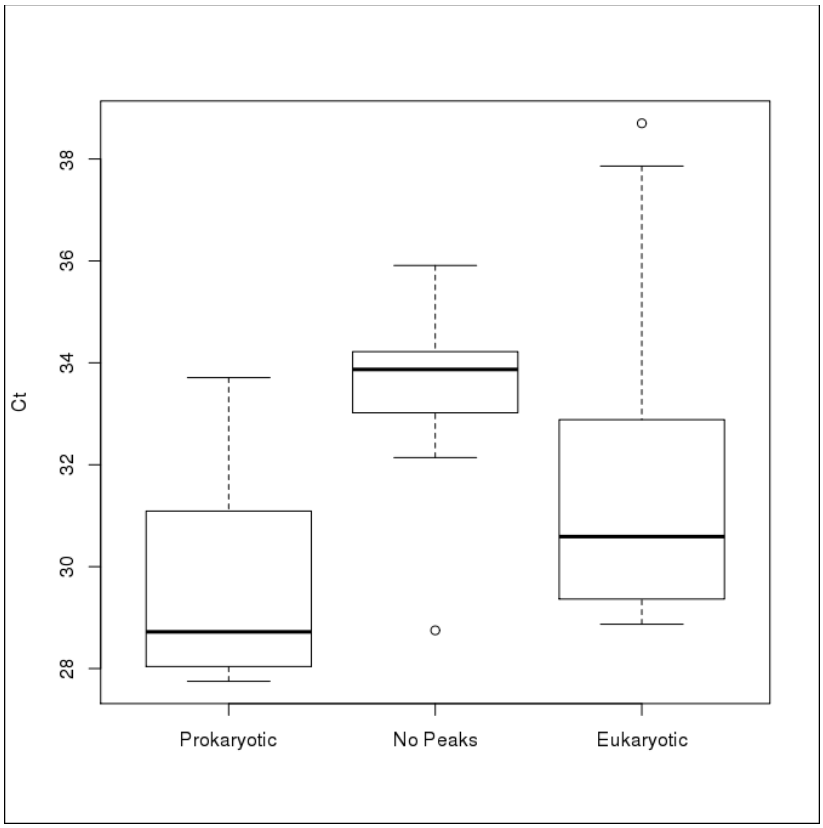

Figure 5C. Expression of three genes by rRNA peak type.
Expression data from two mRNAs, SIGLEC5 (A) and ITGA5 (B), and the miRNA, hsa-Mir142-5p (C), are boxplotted by rRNA peak type evident in BioAnalyzer data from 34 victim samples. SIGLEC5 is shown as a possible outlier to the other mRNAs; ITGA5 is shown as a representative mRNA gene in this analysis.

Table 4. Pearson correlation coefficients of mRNAs and miRNAs to RNA metrics and PMIs by rRNA peak origin.
Victims were divided by the rRNA peak origin as determined from migration time on the BioAnalyzer into A. prokaryotic, B. no peaks, or C. eukaryotic peaks and Pearson's R calculated for each of the six mRNAs and four miRNAs against RNA yield, PMI-arr, PMI-aut, and RIN.

	messenger RNAs						micro RNAs			
	ANKRD28	*ITGA5*	*PLCG2*	*SCAMP2*	*SIGLEC2*	*TMEM8*	*miR142-3p*	*miR142-5p*	*miR16A*	*let7A*
A. Prokaryotic rRNA peaks										
Yield	0.348	0.439	0.291	0.208	-0.055	0.114	-0.127	0.205	0.103	0.401
PMI-arr	-0.384	-0.079	-0.039	0.156	0.289	0.078	-0.881	-0.592	-0.557	-0.449
PMI-aut	-0.382	-0.103	-0.081	0.130	0.189	0.043	-0.898	-0.611	-0.602	-0.478
RIN	-0.033	0.282	0.207	0.297	0.216	0.167	-0.786	-0.302	-0.346	-0.040
B. No rRNA peaks										
Yield	-0.322	-0.247	-0.230	-0.215	-0.292	-0.314	-0.242	-0.446	-0.188	-0.226
PMI-arr	0.495	0.546	0.547	0.183	0.578	0.135	-0.168	-0.018	-0.085	-0.159
PMI-aut	-0.073	0.024	0.070	-0.190	0.108	-0.317	-0.524	-0.374	-0.412	-0.457
RIN	-0.537	-0.543	-0.477	-0.650	-0.485	-0.621	-0.644	-0.233	-0.572	-0.524
C. Eukaryotic rRNA peaks										
Yield	0.179	0.214	0.545	0.161	0.320	0.244	0.431	0.585	0.181	0.360
PMI-arr	-0.079	-0.180	-0.138	-0.199	-0.158	0.251	0.042	-0.047	-0.072	0.004
PMI-aut	0.050	-0.092	0.007	-0.103	-0.067	0.457	-0.024	-0.077	-0.141	-0.074
RIN	-0.005	0.036	0.077	0.049	0.026	-0.234	-0.087	0.128	-0.229	-0.182

Table 5. Spearman correlation coefficients comparing expression patterns of 35 genes between 87 victim samples. Spearman's rho is above the upper-left to lower-right diagonal, the p-value for each comparison is below the diagonal. Rho for each subject compared to all other subjects can be read from the top of the table to the diagonal, then across to the right; p-values for each correlation are read from the left to the diagonal, then down. This table is available online at https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2010s/media/table5.xls

Correlation of Tissue-specific Expression in Living vs. Victim Samples

To determine whether postmortem samples maintain tissue-specific expression patterns, relative expression patterns were determined from blood, lung, spleen, kidney, cardiac muscle (heart), skeletal muscle, brain, and liver. Spearman rank order correlation coefficients were calculated between all possible tissue sample pairs (Table 6). Significant correlations were observed in brain (Rho = -0.548, p = 1.2 e-4), lung (Rho = -0.399, p = 7.27 e-3), skeletal muscle (Rho = -0.581, p = 3.55 e-5), and blood (Rho = -0.555, p = 1.12 e-4).

Detection of Bacterial Contamination in Accident Samples

The presence of bacterial rRNA peaks in electropherograms of purified RNA from accident victim samples were motivation to develop a more sensitive bacterial detection method and then determine if that data assisted in predicting the accuracy of human mRNA assays in individual samples. From 19 custom universal 16S primers, an initial assessment was performed of 12 different amplicons and a commercial primer set, set 13 (Table 1A), using purified *E. coli* total RNA in 7-point curves across an input range of 2500 to 0.16 pg. In three primer pairs, sets 1, 2, and 7, the standard curves failed based on the observation that a sample with a starting material input amount overlapped or had a higher Ct than a sample with a lower amount of input material. Sets 10, 11, and 12 had primer 1492R (l) in common and in all three cases, the NoRT control was less than 10 Cts from the lowest point on the curve. Primer sets 3 and 4 suffered from poor curve fit, 0.834 and 0.985 respectively (Table 1A).

Based on these results, sets 5, 6, 8, and 13 were further assessed in triplicate standard curve reactions from 500 to 3.2 e-02 pg *E. coli* RT product. At this lower range of input materials, the commercial primer pair had an unacceptable efficiency of 120.2% and was dropped in favor of primer pair 8 for a second assessment (Table 1B). In these two experiments, the most consistent was primer pair 5 with efficiencies closer to the ideal of 100%, acceptable R2 values (>= 0.995), equal dynamic range to the other three sets, and a large difference between the NoRT control and the lowest input amount on all curves. To determine the lower limit, an 11 point curve was conducted (Table 1C) that resulted in a Ct of 29.55 for the lowest input amount, 50 ag, less than 1 Ct above the negative control value of 29.72. Dissociation curve analysis showed a single sharp peak suggesting a single amplification product from this primer pair. The dissociation temperature was the same in the NoRT reactions leading to the conclusion that the recombinant PCR enzyme was contaminated with bacterial genomic DNA. Because the NoRT controls had similar Cts irrespective of primer set, no template control experiments were performed with water purchased from Ambion that yielded comparable Ct values to those observed from NoRT samples (data not shown) confirming the conclusion that NoRT reaction background was due to bacterial genomic DNA in the enzyme mix or kit reagents.

The specificity of primer set 5 in mixed samples was assessed. First, an *E. coli* RT product 7-point standard curve of 5-fold dilutions from 25 to 1.6 e-03 ng was performed with and without a fixed 1 ng amount of human RT-product spiked in. The reaction efficiency for the curve without human RNA was 88.4% whereas the efficiency with human RNA was 97.6% demonstrating the specificity of the primers in the presence of a wide range of relative concentrations of “contaminating” human RNA. A second set of reactions was performed with a fixed total input of 5 ng but variable ratios from one tenth bacterial to one tenth human (Materials and Methods). Again, the bacterial primers performed with good specificity with an efficiency of 104.3% and R^2 value of 0.993.

Table 6. Spearman correlation coefficients comparing tissue-specific expression patterns. All possible comparisons of live to live, live to victim, and victim to victim are shown. As in Table 5, Spearman's rho is above the upper-left to lower-right diagonal, the p-value for each comparison is below the diagonal. Rho for each tissue source can compared to all other tissue sources by reading from the top of the table to the diagonal, then across to the right; p-values for each correlation are read from the left to the diagonal, then down.

	Kidney_ Live	Brain_ Live	Heart_ Live	Liver_ Live	Lung_ Live	Muscle_ Live	Spleen_ Live	Blood_ Live	Kidney_ Victim	Brain_ Victim	Heart_ Victim	Liver_ Victim	Lung_ Victim	Muscle_ Victim	Spleen_ Victim	Blood_ Victim
Kidney_Live	1	0.566	0.907	0.755	0.665	0.747	0.812	0.362	0.021	0.176	0.171	0.159	0.2	0.239	0.002	0.355
Brain_Live	6.15E-05	1	0.563	0.491	0.654	0.69	0.498	0.356	0.236	0.548	0.418	0.444	0.338	0.508	0.331	0.391
Heart_Live	0.00E+00	6.96E-05	1	0.661	0.718	0.804	0.732	0.399	0.032	0.233	0.287	0.22	0.332	0.362	0.068	0.342
Liver_Live	3.26E-09	7.10E-04	1.04E-06	1	0.627	0.718	0.639	0.436	0.023	0.014	0.035	0.257	0.14	0.181	0.068	0.301
Lung_Live	8.69E-07	1.46E-06	4.15E-08	5.32E-06	1	0.856	0.591	0.526	0.057	0.265	0.399	0.253	0.399	0.527	0.052	0.364
Muscle_Live	5.94E-09	2.23E-07	5.06E-11	4.15E-08	1.35E-13	1	0.584	0.465	0.099	0.284	0.399	0.364	0.381	0.581	0.156	0.37
Spleen_Live	2.19E-11	5.88E-04	1.59E-08	2.99E-06	2.43E-05	3.20E-05	1	0.438	0.077	0.125	0.111	0.102	0.177	0.132	0.018	0.418
Blood_Live	1.72E-02	1.90E-02	7.97E-03	3.45E-03	2.88E-04	1.69E-03	3.34E-03	1	0.045	0.092	0.173	0.177	0.163	0.298	0.015	0.555
Kidney_Victim	8.92E-01	1.24E-01	8.38E-01	8.84E-01	7.14E-01	5.25E-01	6.21E-01	7.74E-01	1	0.461	0.618	0.807	0.695	0.475	0.874	0.414
Brain_Victim	2.53E-01	1.20E-04	1.28E-01	9.31E-01	8.17E-02	6.16E-02	4.20E-01	5.58E-01	1.64E-03	1	0.78	0.676	0.602	0.785	0.646	0.525
Heart_Victim	2.67E-01	4.70E-03	5.88E-02	8.20E-01	7.23E-03	7.27E-03	4.73E-01	2.68E-01	7.77E-06	4.26E-10	1	0.775	0.842	0.894	0.793	0.616
Liver_Victim	3.03E-01	2.54E-03	1.52E-01	9.21E-02	9.73E-02	1.51E-02	5.11E-01	2.57E-01	3.69E-11	4.66E-07	6.52E-10	1	0.768	0.743	0.851	0.609
Lung_Victim	1.93E-01	2.47E-02	2.76E-02	3.66E-01	7.27E-03	1.07E-02	2.49E-01	2.97E-01	1.65E-07	1.52E-05	7.83E-13	1.13E-09	1	0.710	0.792	0.613
Muscle_Victim	1.18E-01	4.28E-04	1.56E-02	2.40E-01	2.39E-04	3.55E-05	3.92E-01	5.26E-02	1.11E-03	2.75E-10	4.44E-16	7.87E-09	6.80E-08	1	0.62	0.585
Spleen_Victim	9.92E-01	2.81E-02	6.60E-01	6.61E-01	7.36E-01	3.11E-01	9.09E-01	9.25E-01	9.33E-15	2.16E-06	1.45E-10	2.54E-13	1.49E-10	7.04E-06	1	0.555
Blood_Victim	1.95E-02	9.44E-03	2.46E-02	4.95E-02	1.63E-02	1.45E-02	5.23E-03	1.12E-04	5.81E-03	3.05E-04	1.10E-05	1.47E-05	1.27E-05	3.72E-05	1.14E-04	1

To assess the effects of postmortem blood on bacterial expression testing, bacterial primer pairs 5, 6, 8, and 9 were used in triplicate reactions with 1 ng of RT product from ten postmortem samples (Table 7). Sample 26 had no rRNA peaks, prokaryotic peaks were found in Samples 11, 12, 20, 52, and 70, and Sample 81 had both prokaryotic and eukaryotic rRNA peaks. Subjects 2, 7, and 9 had eukaryotic peaks with RINs of 6.6, 6.3, and 5.0 respectively. For the four primers, reactions on the NoRT control had an average Ct of 30.6 and coefficient of variation (CV) of 0.26 across the four primer pairs. For the 1 ng *E. coli* positive control, Cts averaged 9.73 with a CV of 0.32. Samples with prokaryotic rRNA peaks had average Cts of 10.4, 10.2, 9.28, and 10.6 with CVs of 0.12, 0.04, 0.21, and 0.41 for the four primer pairs. These values were similar to the positive control. However, similar to these samples, Sample 26 with no rRNA peaks also had an average Ct of 8.98. Furthermore, two of the three eukaryotic peak samples, 2 and 6, had low Cts with all four primer pairs averaging 9.84 and 11.86, respectively, and Cvs of 0.07 for both. Sample 9 had mixed results; primer pair 9 had a low Ct of 11.24 whereas the remaining three pairs had Cts greater than 20. For this subject, Ct values for the primer pairs with forward primer 16S:1100:F16 were close to those for the negative controls. Finally, Sample 81 which had both rRNA peaks, also had less consistency across the four primer pairs with an average CV of 1.15 and an average Ct of 15.0. As with Sample 9, the reactions with the 16S:1100:F16 forward primer had higher Cts than those with 1237F and the Cts from those reactions had Cts intermediate to the positive and negative controls in agreement with expectations based on the presence of both types of rRNA peaks.

Primer pair 5, 16S:1100:F16 and 1237R, was used to assess 84 accident victim blood samples chosen across a range of accident conditions and the full gamut of rRNA peak types (Table 8). Samples either with direct evidence for bacterial contamination or expected to be contaminated due a water landing replaced some eukaryotic or no peak samples in this experiment. The resulting Ct values ranged from 7.62 to 31.54. Only four samples had both rRNA peak types making further inference difficult. When the remaining data was subdivided by rRNA peak type, prokaryotic, no peaks, and eukaryotic, there was considerable overlap of the interquartile ranges (Fig. 6) and no significant differences in rRNA expression values between the rRNA peak origins (data not shown). However, the medians showed the expected trend; samples with eukaryotic peaks had the highest median Ct and victims with prokaryotic peaks the lowest. When the combined samples with prokaryotic and no rRNA peaks were divided into water (N=22) and dry (N=25) landings, the water landing samples unexpectedly had a slightly higher median—a lower expression of bacterial 16S rRNA—than those in dry landings, 23.4 and 20.9, respectively (Fig. 7). When data from the landing type were further divided by rRNA peak type, the medians between samples with prokaryotic peaks (water, N=5; dry, N=6) separated by 10 Cts, 22.6 versus 12.3 (Fig. 8), although the difference was not statistically significant. The medians in samples with no rRNA peaks (water, N=17; dry, N=19) were 3 Cts apart, 24.0 versus 20.9 (Fig. 8). Among samples with data for both human mRNAs and bacterial 16S rRNA (N=68), there were no significant expression value correlations between bacterial 16S rRNA and any of the human mRNAs (data not shown).

DISCUSSION

Within the discipline of aerospace medicine, the greatest emphasis is placed on either preventing aviation accidents or determining the cause of an accident. Therefore, it is highly desirable to have gene expression marker panels for relevant factors such as hypoxia, sleep deprivation, alcohol, pathologies, and other factors that may cause impairment or in-flight incapacitation. However, the discovery of expression markers under controlled conditions in live subject studies is not sufficient for use in accident investigation; the relative expression patterns of the panels must be shown to persist postmortem. Alternatively, the circumstances under which the panels remain useful must be understood. To that end, the goal of this study was to characterize postmortem RNA stability and maintenance of detected expression patterns.

Table 7. Screening of four primer pairs on ten victim samples.
Four bacterial 16S rRNA primer pairs were compared on ten victim samples to determine the relative expression values between primers on samples with characterized rRNA peak types by BioAnalyzer analysis. RIN values are included only for samples with eukaryotic peaks. For each sample, the mean and standard deviation of Cts for the four primer pairs are in the column labelled Mean/StDev.

Sample	rRNA peak origin	Primer pair	Primer IDs	Ct (dRn)	Mean/StDev
11	Prokaryotic	8	1237F:1381R	10.84	
		9	1237F:1391R	10.68	
		6	16s:1185mR	10.36	10.41
		5	16s:1237R	9.75	0.12
26	None	8	1237F:1381R	8.77	
		9	1237F:1391R	10.02	
		6	16s:1185mR	8.32	8.98
		5	16s:1237R	8.79	0.18
70	Prokaryotic	8	1237F:1381R	10.11	
		9	1237F:1391R	10.33	
		6	16s:1185mR	10.26	10.16
		5	16s:1237R	9.93	0.04
12	Prokaryotic	8	1237F:1381R	10.51	
		9	1237F:1391R	11.00	
		6	16s:1185mR	12.74	11.24
		5	16s:1237R	10.71	0.26
20	Prokaryotic	8	1237F:1381R	8.45	
		9	1237F:1391R	10.41	
		6	16s:1185mR	9.33	9.28
		5	16s:1237R	8.93	0.21
2	Eukaryotic RIN 6.6	8	1237F:1381R	9.77	
		9	1237F:1391R	10.26	
		6	16s:1185mR	9.66	9.85
		5	16s:1237R	9.70	0.07
9	Eukaryotic RIN 5	8	1237F:1381R	20.53	
		9	1237F:1391R	11.24	
		6	16s:1185mR	28.53	21.50
		5	16s:1237R	25.68	1.90
52	Prokaryotic	8	1237F:1381R	9.63	
		9	1237F:1391R	9.89	
		6	16s:1185mR	13.07	10.64
		5	16s:1237R	9.97	0.41
7	eukaryotic RIN 6.3	8	1237F:1381R	11.64	
		9	1237F:1391R	11.70	
		6	16s:1185mR	11.86	11.86
		5	16s:1237R	12.24	0.07
81	both	8	1237F:1381R	11.44	
		9	1237F:1391R	10.83	
		6	16s:1185mR	17.21	14.96
		5	16s:1237R	20.35	1.15
Neg Ctl	NA	8	1237F:1381R	31.42	
		9	1237F:1391R	31.49	
		6	16s:1185mR	29.25	30.62
		5	16s:1237R	30.30	0.27
Pos Ctl	NA	8	1237F:1381R	10.68	
		9	1237F:1391R	10.85	
		6	16s:1185mR	9.20	9.73
		5	16s:1237R	8.17	0.32

Table 8. Bacterial rRNA detection on 84 victim samples.
Victim samples sorted by the four rRNAs peak categories, prokaryotic, eukaryotic, both types of peaks, and no peaks, were screened for relative bacterial rRNA with 16S rRNA universal primer pair 5. Ct data from the human samples, standard curve on purified *E. coli* total RNA, and a negative control are presented. Human samples were screened for a presumed higher likelihood of having prokaryotic contamination by virtue of landing in water (column label "other").

Sample	Dependent	rRNA origin	Other	Ct (dRn)
0.00032x	std curve	---		18.21
0.0016x	std curve	---		15.41
0.008x	std curve	---		13.35
0.04x	std curve	---		11.54
0.2x	std curve	---		9.34
1x	std curve	---		7.31
Neg Ctl		---		30.14
71	61	both	water	28.57
81	69	both	water	11.5
61	51	both		16.4
72	62	both		26.34
68	58	eukaryotic	water	25.24
93	80	eukaryotic	water	22.66
105	89	eukaryotic	water	30.87
10	7	eukaryotic		18.21
13	9	eukaryotic		12.83
14	9	eukaryotic		11.96
18	13	eukaryotic		7.62
19	14	eukaryotic		9.78
21	16	eukaryotic		30.76
24	18	eukaryotic		9.77
25	19	eukaryotic		14.39
27	21	eukaryotic		11.1
30.2	24	eukaryotic		10.64
56	47	eukaryotic		25.72
57	47	eukaryotic		24.67
58	48	eukaryotic		27.97
62	52	eukaryotic		8.7
67	57	eukaryotic		22.59
76	65	eukaryotic		18.89
78	67	eukaryotic		11.19
84	72	eukaryotic		25.02
85	72	eukaryotic		27.61
86	73	eukaryotic		27.15
91	78	eukaryotic		28.19
92	79	eukaryotic		27.36
94	81	eukaryotic		20.12
98	83	eukaryotic		26.42
100	85	eukaryotic		27.42
101	86	eukaryotic		26.49
102	87	eukaryotic		29.6
104	88	eukaryotic		28.02
106	90	eukaryotic		30.64

Sample	Dependent	rRNA origin	Other	Ct (dRn)
		eukaryotic		27.19
1	1	none	water	28.09
17	12	none	water	26.85
89	76	none	water	28.42
90	77	none	water	25.89
		none	water	22.84
		none	water	19.23
		none	water	26.88
		none	water	24.45
		none	water	20.61
		none	water	27.93
		none	water	8.65
		none	water	23.97
		none	water	19.44
		none	water	13.13
		none	water	18.93
		none	water	29.99
		none	water	22.58
5	3	none		15.39
22	16	none		16.31
28	22	none		9.62
31	25	none		8.93
55	47	none		22.76
60	50	none		31.54
63	53	none		26.75
69	59	none		7.69
74	64	none		20.87
75	65	none		26.44
77	66	none		16.52
79	67	none		14.36
80	68	none		11.02
82	70	none		20.92
87	74	none		27.54
88	75	none		21.27
96	82	none		25.32
97	83	none		26.28
103	87	none		24.31
2	1	prokaryotic	water	20.58
11	8	prokaryotic	water	26.26
12	8	prokaryotic	water	21.64
20	15	prokaryotic	water	22.62
52	44	prokaryotic	water	24.97
70	60	prokaryotic		12.05
83	71	prokaryotic		27.81
		prokaryotic		25.06
		prokaryotic		12.59
		prokaryotic		7.92
		prokaryotic		9.73

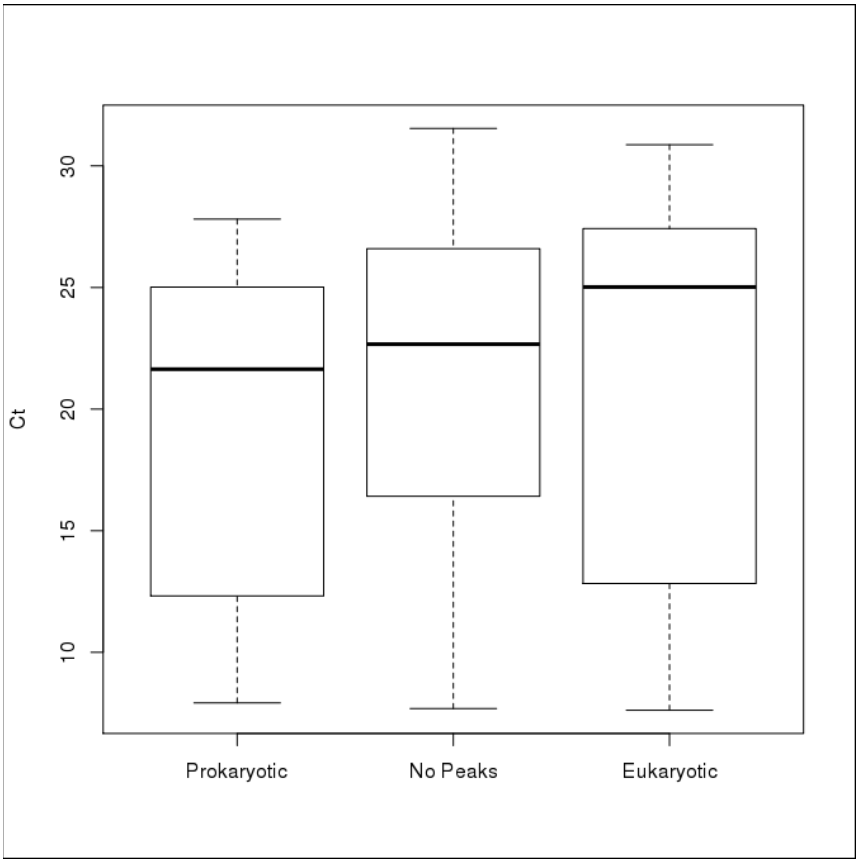

Figure 6. Boxplots of bacterial rRNA expression data from 87 victim samples by detected rRNA peak type.

Bacterial 16S expression values were collected from 87 victim samples and the data plotted by the rRNA peak type detected from the victim samples on the BioAnalyzer. Due to low sample size, samples with both peaks are not shown.

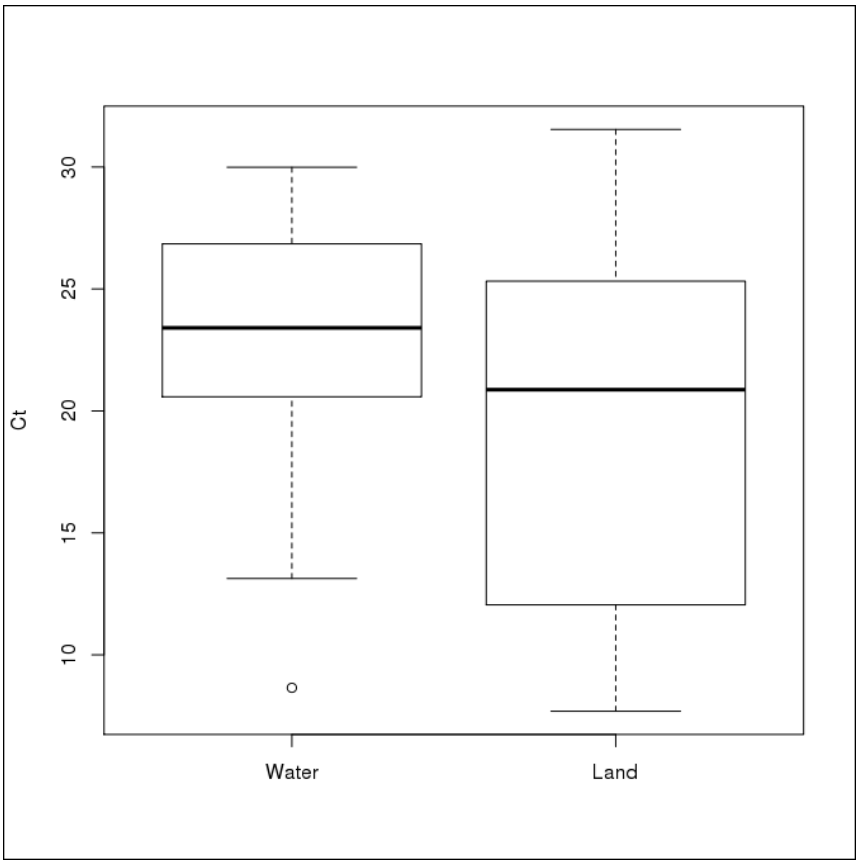

Figure 7. Boxplots of rRNA expression data from 87 victim samples by accident condition.

Bacterial 16S expression values were collected from 87 victim samples and the data plotted by the landing surface of the accident, dry land or water.

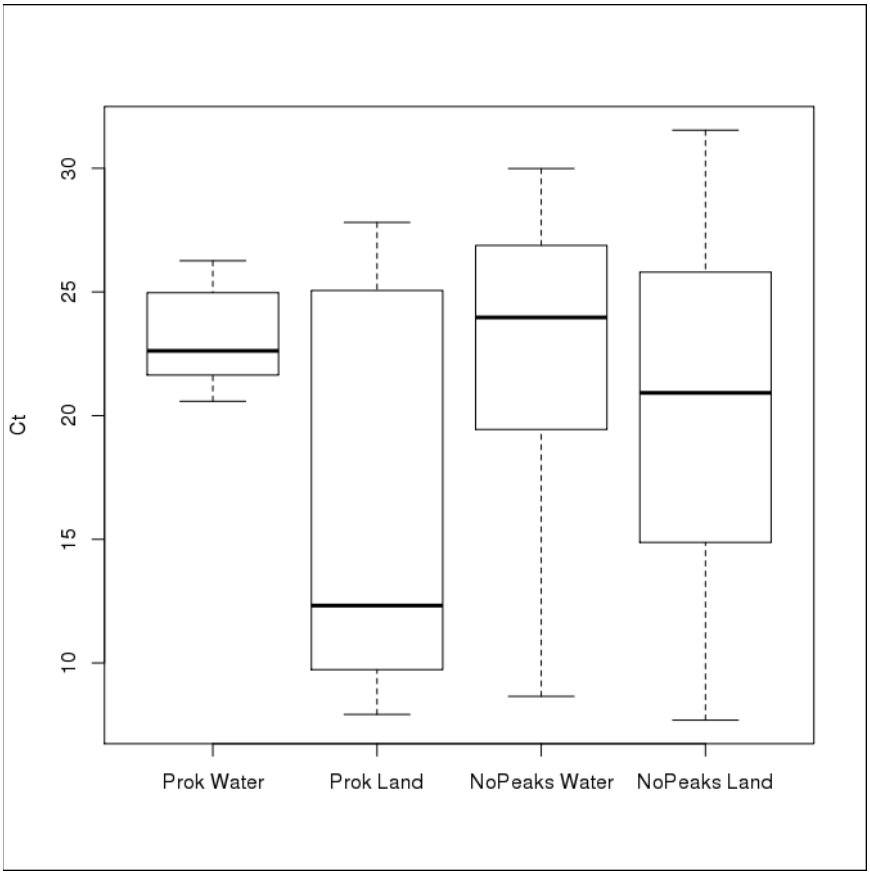

Figure 8. Box plots of rRNA expression data from victim samples with prokaryotic or no rRNA peaks divided by landing surface of the accident, dry land or water.

Factors impacting RNA integrity

Initially, blood samples were collected from four live subjects under sterile conditions in serum tubes. Blood was kept at room temperature where degradation rate would be expected to occur both enzymatically and chemically. RINs remained stably high until 48 hours (Fig. 2C) and did not decrease greatly until 96 hours. These results were generally in agreement with RNA degradation rates reported by others in tissues either by monitoring the loss of the 28S rRNA band electrophoretically (22) or more recently with RIN data (14, 15, 23). In 34 well-curated victim samples, RINs were not correlated to either PMI-arr or PMI-aut (Table 3) and, in fact, there were several examples of low RINs from samples with short PMI-arr and PMI-aut and relatively high RINs from samples with PMI-aut as great as 96 hours (Table 2). In fresh blood samples, there has been a comparison of collection in either RNAlater™ (Ambion/ThermoFisher, PN 7021) or PAXgene® tubes under various shipping conditions, followed by expression analysis of a single gene, *ATM* (24). For samples shipped at ambient temperature, RIN for PAXgene® samples was 6.9 compared to 8.2 for RNAlater™ samples. Although no difference was observed in *ATM* expression, these results suggest that the sample collection method utilized in this study may contribute to lower RINs in accident victim samples which can impact qRT-PCR assay sensitivity if RINs are less than 5 (11, 23, 25).

Presumably, RNA degradation is of two sources, either chemical hydrolysis or RNase activity. *In vivo* acidification would result in increased concentrations of donor protons speeding chemical hydrolysis. In blood, pH decreased more rapidly in rat and pig cadavers than rat and human blood *in vitro* (26), and lactate, formate, uric acid, and NADH were elevated in the *in vivo* samples. The authors attribute the effect to autolysis. They did not directly assess RNA integrity (RI), but their 96 hour timecourse correlates with our observation of rapid RNA degradation in the victim samples compared to relative stability at 96 hours *in vitro*. In brain, RI loss has been most consistently attributed to decreased pH (27-29) usually associated

with antemortem or agonal factors such as prolonged death, hypoxia, and/or a disease state. Agonal factors are unlikely to affect RI in accident victims because the period of psychological stress before an accident is usually short and the pilot population may be healthier than the population accessed in other studies. While RNA can be degraded by ultraviolet radiation, this is unlikely to be a factor in this study as the exposed skin surface would block the internal organs and cavity blood.

RNases are notably stable, resisting denaturation at high temperatures due to a high number of chemical disulfide linkages, and remain active across a range of pH and temperatures (30). An aviation accident is a violent event often accompanied by puncturing or dismemberment and possible introduction of microbes to the samples that would be used for gene expression analysis. In addition, the violent nature of these events results in a high likelihood of internal organ hemorrhage and subsequent release of the internal microbiome within the body cavity. A further complication to comparing degradation of *in vitro* to *in vivo* samples and in reconciling the wide variability of RNA integrity between accident victims was that some victims could have been under a degree of hypoxic exposure that may influence blood pH. Very few of the victims assessed here had been at altitudes greater than a few thousand feet; nonetheless, we cannot rule out variable levels of autolysis, hypoxic effects, or post-accident microbial growth as sources of variability between *in vitro* samples and victims, or RINs from victim samples with similar PMIs. Others have associated PMI with decreased RI (10, 12) but those findings were not replicated in this work. Although dried blood samples have been shown to be stable beyond the typical PMI for aviation accidents (1, 31), it should be noted that this sample type is not reproducibly available in the aviation accident setting where sample collection is performed by those providing autopsy-derived samples for forensic toxicology analylsis. By the time the cadaver has reached the site of autopsy, the damage is done. In a high-throughput environment, the convenience of collection in PAXgene® tubes and robotic RNA purification are preferred over manually filling tubes with a stabilization solution and manual purification. The high concordance of expression patterns across a wide range of RINs in accident samples suggests that validity of the marker panel, uncontrollable cadaver handling factors, and accident conditions will have a greater impact on translation of expression analysis to accident investigation than concerns that collection in PAXgene® tubes may be slightly detrimental to RI.

In addition to altitude of the craft during the fatal accident, careful attention was paid to other external factors potentially affecting loss of RI by microbial growth or enzymatic lysis including ambient temperature and weather conditions, date and time of the accident, and accident conditions such as post-crash fire or water landing. None of these factors were statistically related to RI.

Bacterial detection by qRT-PCR

The bacterial detection primer set described in this study maintained specificity for the *E. coli* 16S rRNA when samples were a mixture of *E. coli* and human total RNA. However, data gathered with this primer set in victim samples suggested bacterial growth in samples with no other evidence of bacterial contamination and otherwise high quality factors, ie., high RINs based on eukaryotic rRNA peaks. Two samples with prokaryotic rRNA peaks had Cts less than 10 which, from the standard curve, was equivalent to input amounts greater than 20 pg purified of *E. coli* RNA. Conversely, seven samples with prokaryotic rRNA peaks had Cts above 20, equivalent to less than 32 fg input, although only two of these samples had Cts above 25, approximately 1 fg input RNA. In addition, several samples with only eukaryotic peaks had Cts below 15. While these values were well within the detection limits of the assay, it is a concern that there was no evidence of prokaryotic contamination in these samples. Interestingly, the range of Cts for samples with prokaryotic rRNA peaks overlaps with the Ct range of samples that have no rRNA peaks of either kind. The bacterial rRNA assay has very high sensitivity; this high copy number gene was detected at levels down to 1 ag purified bacterial RNA input. This datapoint contributed to the standard curve, was highly

reproducible, and was one Ct above the background level. Because of the high copy number of the 16S rRNA gene, this level of sensitivity was not seen in any of the human mRNAs where standard curve datapoints typically lost reproducibility between 1.0 and 0.3 pg of purified human RNA on standard curves. Further, numerous samples with prokaryotic or four peaks had good average correlation coefficients in a human gene panel to samples with either no peaks or eukaryotic peaks (Table 5). The absence of prokaryotic peaks on the BioAnalyzer could be due to one or a combination of: 1) A lack of instrument sensitivity, 2) Low level of bacterial growth, or 3) Bacterial death followed by chemical RNA degradation. The observed background for the assay was likely due to contaminating bacterial DNA from either the PCR enzyme and/or the reagents (32), and the observed range of Cts around 35 cycles was consistent with previous observations (33). Further efforts to characterize this assay should take advantage of recently available DNA-free DNA-polymerases.

Initial normalization of samples occurs by taking a fixed quantity of total RNA for the RT reaction but quantitation is calculated from absorbance at 260 nm (A260) which does not differentiate between RNA sources confounding absolute quantification of the human RNA fraction. In a gross sense, the presence of bacterial RNA can be qualitatively determined during QA testing of purified RNA (Fig. 1). Further complicating quantitation of the human RNA fraction is the potential for fungal contamination in the samples. In short, purified RNA from victim samples may actually be a complex mixture including multiple genera of bacteria and fungi whose contributions may not be quantifiable.

Implications for detection of marker panels in postmortem samples

Tissue specificity of gene expression in forensic samples has been demonstrated by RT-PCR (34-36), microarrays (37) and RNA-seq (38). In these studies, when samples were not from live biopsies, PMIs were a few hours and could be directly calculated. All live blood samples reported here were obtained by IV catheterization and all skeletal muscle data was from biopsy samples. In the remaining tissues, to achieve an adequate number of samples, some cadaver samples were required but the PMI was typically less than 12 hours. In contrast, most victim sample qRT-PCR data reported here were from samples with a PMI between 48 and 96 hours. While it is possible that the short-PMI samples may contribute an expression pattern more like victim samples than live subjects, the data suggested otherwise. The strongest correlations were seen in two of the tissues, brain and skeletal muscle, where all live sample data were from samples of living subjects. Aside from brain, correlations in the remaining tissues were poorer whereas a confounding contribution would strengthen the correlation. Also, a recent whole transcriptome study showed highly concordant expression patterns between samples with short-PMIs and live subject samples (38). Brain remains as the lone tissue where the correlation may be artificially inflated due to the inclusion of short-PMI samples but, taking all the available data together, this is unlikely. Due to ease of sample access in human subject studies, marker development work for aerospace medical factors has focused on blood where the correlation is relatively strong. For forensic applications, blood is generally available from accident victim remains although further investigation of the concordance of expression patterns from coagulated blood is warranted.

The high concordance of gene expression patterns across many victim samples suggested that many of the sampled miRNAs (Table 3) and mRNAs (Tables 3 & 6) degrade at approximately equivalent rates and that relative expression patterns between genes within a sample were retained. Similar results have been reported (19, 23) when assaying many fewer genes although, across the entire transcriptome, degradation rates for individual mRNAs have been shown to be variable (14, 15). Further, the apparent stability of a transcript can be assay dependent as shown by dependence of probe location within the transcript upon detection of brain mRNAs on microarrays (39) where 3’ probes had decreased sensitivity.

Before translation of marker panels to an operational setting, additional characterization of qRT-PCR assays on samples of variable degradation will be necessary. The utilization of normalizer genes for relative gene quantitation in postmortem samples has been widely discussed (11, 25, 40-43) but normalizers are specific to the experimental conditions. Normalizer gene panels correct eukaryotic RNA loading errors by arithmetically placing target gene expression levels relative to the expression level of the normalizer panel rather than at their absolute level (See Appendix) as was done for comparison of victim live subject data here. Therefore, inaccurate sample quantitation due to the presence of bacterial RNA or loss of sensitivity due to degradation of target genes can be overcome by normalization to a normalizer panel with similar degradation characteristics as the target gene(s). The relative degradation rates of targets and normalizers requires experimental validation, but the choice of normalizers will be facilitated by continued sampling of normalizer panels on relatively large sample sizes with a broad range of sample degradation, the development of commercially available qRT-PCR assays with consistent performance, high-throughput instrumentation, and advances in software for qRT-PCR data analysis.

Conclusions and Future Directions

As additional data is collected from live subject studies of aerospace medical factors such as hypoxia and sleep deprivation, it must be recognized that not all statistically significant putative markers will be ideal for the operational setting in forensic samples. Obviously, of paramount importance is the predictive value of the individual markers but it is not expected that any single RNA molecule will be a “magic bullet” for a factor; rather, panels of well-characterized markers will be required that need no other information to interpret the results and assess the impact of the factor on an accident. Relative expression of the marker panel would be compared to known expression data from live subject studies and probabilities calculated based on these results. To improve the reliability of a forensic assay, other factors should be accounted for in assay development. Putative markers should be screened for those that lack degradation motifs or possess stabilizing biochemical modifications so that markers are more likely to resist enzymatic degradation as PMI increases. To improve accuracy, markers and normalizers should be screened against available whole-transcriptome expression sets to match target genes to normalizers with similar degradation rates. Where possible, longer transcripts could be chosen and paired with assays that are positioned on the transcript to improve reliable detection of markers by minimizing the effects of enzymatic degradation from the transcript ends.

The accident investigation workflow is not geared towards preservation of biological molecules for forensic testing but rather preserves the accident scene. As the utility of biomarkers becomes accepted as proven technology, the priority of biological specimen sampling should increase. With some additional training, on-site investigators could perform this task. Granted, this change in the paradigm would not be helpful in every case because, especially if the accident occurs in a remote area or is a water landing, there can be a time lag between the time of an accident and its discovery. Furthermore, counter to the observations of others, PMI was not correlated to RIN in this study. However, only 13 of the 34 samples curated in Table 2 had a PMI-aut less than 24 hours and only 3 had PMI-aut less than 15. Rapid on-site sampling would almost certainly preserve the RNA in a more intact state and considerably improve assay sensitivity and specificity.

The utility of a marker panel is in the positive detection of the gene signature for a condition. A negative or inconclusive result does not rule out that the condition existed, only that it remains undetected, a false negative which, for hypoxia and sleep deprivation, returns the accident investigation decision tree to the current situation. Of greater concern in forensic testing is the potential for calling a false positive which is highly unlikely in this type of testing due to the inherent redundancy of testing multiple genes that must have an expected expression pattern to call a positive result. Therefore, utilization of gene expression

markers panels is expected to have only positive impact on the determination of accident causality because a false negative does no harm.

REFERENCES

1. Bauer M, Gramlich I, Polzin S, Patzelt D. Quantification of mRNA degradation as possible indicator of postmortem interval--a pilot study. Legal medicine. 2003 Dec;5(4):220-7.

2. Young ST, Wells JD, Hobbs GR, Bishop CP. Estimating postmortem interval using RNA degradation and morphological changes in tooth pulp. Forensic science international. 2013 Jun 10;229(1-3):163 e1-6.

3. Hunter MC, Pozhitkov AE, Noble PA. Accurate Predictions of Postmortem Interval Using Linear Regression Analyses of Gene Meter Expression Data. bioRxiv; 2016;30.

4. Juusola J, Ballantyne J. Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification. Forensic science international. 2003 Aug 12;135(2):85-96.

5. Juusola J, Ballantyne J. Multiplex mRNA profiling for the identification of body fluids. Forensic science international. 2005 Aug 11;152(1):1-12.

6. Hanson EK, Lubenow H, Ballantyne J. Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. Analytical biochemistry. 2009 Apr 15;387(2):303-14.

7. Alvarez M, Ballantyne J. The identification of newborns using messenger RNA profiling analysis. Analytical biochemistry. 2006 Oct 1;357(1):21-34.

8. Jardine D, Cornel L, Emond M. Gene expression analysis characterizes antemortem stress and has implications for establishing cause of death. Physiological genomics. 2011 Aug 24;43(16):974-80.

9. Roth RB, Hevezi P, Lee J, Willhite D, Lechner SM, Foster AC, et al. Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. Neurogenetics. 2006 May;7(2):67-80.

10. Birdsill AC, Walker DG, Lue L, Sue LI, Beach TG. Postmortem interval effect on RNA and gene expression in human brain tissue. Cell and tissue banking. 2011 Nov;12(4):311-8.

11. Vennemann M, Koppelkamm A. mRNA profiling in forensic genetics I: Possibilities and limitations. Forensic science international. 2010 Dec 15;203(1-3):71-5.

12. Catts VS, Catts SV, Fernandez HR, Taylor JM, Coulson EJ, Lutze-Mann LH. A microarray study of post-mortem mRNA degradation in mouse brain tissue. Brain research Molecular brain research. 2005 Aug 18;138(2):164-77.

13. Fordyce SL, Kampmann ML, van Doorn NL, Gilbert MT. Long-term RNA persistence in postmortem contexts. Investigative genetics. 2013;4(1):7.

14. Gallego Romero I, Pai AA, Tung J, Gilad Y. RNA-seq: impact of RNA degradation on transcript quantification. BMC biology. 2014;12:42.

15. Thompson KL, Pine PS, Rosenzweig BA, Turpaz Y, Retief J. Characterization of the effect of sample quality on high density oligonucleotide microarray data using progressively degraded rat liver RNA. BMC biotechnology. 2007;7:57.

16. A - Z of Quantitative PCR. La Jolla, CA: International University Line, 2004.

17. Fleige S, Walf V, Huch S, Prgomet C, Sehm J, Pfaffl MW. Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. Biotechnology letters. 2006 Oct;28(19):1601-13.

18. Mueller S. Optimizing real-time quantitative PCR experiments with the Agilent 2100 bioanalyzer. Agilent Technologies; 2008;8.

19. Schoor O, Weinschenk T, Hennenlotter J, Corvin S, Stenzl A, Rammensee HG, et al. Moderate degradation does not preclude microarray analysis of small amounts of RNA. BioTechniques. 2003 Dec;35(6):1192-6, 8-201.

20. Opitz L, Salinas-Riester G, Grade M, Jung K, Jo P, Emons G, et al. Impact of RNA degradation on gene expression profiling. BMC medical genomics. 2010;3:36.

21. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome biology. 2009;10(11):R130.

22. Inoue H, Kimura A, Tuji T. Degradation profile of mRNA in a dead rat body: basic semi-quantification study. Forensic science international. 2002 Dec 4;130(2-3):127-32.

23. Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. Molecular aspects of medicine. 2006 Apr-Jun;27(2-3):126-39.

24. Weber DG, Casjens S, Rozynek P, Lehnert M, Zilch-Schoneweis S, Bryk O, et al. Assessment of mRNA and microRNA Stabilization in Peripheral Human Blood for Multicenter Studies and Biobanks. Biomarker insights. 2010;5:95-102.

25. Koppelkamm A, Vennemann B, Lutz-Bonengel S, Fracasso T, Vennemann M. RNA integrity in post-mortem samples: influencing parameters and implications on RT-qPCR assays. International journal of legal medicine. 2011 Jul;125(4):573-80.

26. Donaldson AE, Lamont IL. Biochemistry changes that occur after death: potential markers for determining post-mortem interval. PloS one. 2013;8(11):e82011.

27. Atz M, Walsh D, Cartagena P, Li J, Evans S, Choudary P, et al. Methodological considerations for gene expression profiling of human brain. Journal of neuroscience methods. 2007 Jul 30;163(2):295-309.

28. Chevyreva I, Faull RL, Green CR, Nicholson LF. Assessing RNA quality in postmortem human brain tissue. Experimental and molecular pathology. 2008 Feb;84(1):71-7.

29. Stan AD, Ghose S, Gao XM, Roberts RC, Lewis-Amezcua K, Hatanpaa KJ, et al. Human postmortem tissue: what quality markers matter? Brain research. 2006 Dec 6;1123(1):1-11.

30. nucleases.html. Millipore-Aldrich; 2017 [updated 2017; cited 2017 June 15]; Available from: www.sigmaaldrich.com/life-science/metabolomicsenzyme-explorer/learning-center/nucleases.html.

31. Zubakov D, Hanekamp E, Kokshoorn M, van Ijcken W, Kayser M. Stable RNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples. International journal of legal medicine. 2008 Mar;122(2):135-42.

32. Corless CE, Guiver M, Borrow R, Edwards-Jones V, Kaczmarski EB, Fox AJ. Contamination and sensitivity issues with a real-time universal 16S rRNA PCR. Journal of clinical microbiology. 2000 May;38(5):1747-52.

33. Nadkarni MA, Martin FE, Jacques NA, Hunter N. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. Microbiology. 2002 Jan;148(Pt 1):257-66.

34. Bahar B, Monahan FJ, Moloney AP, Schmidt O, MacHugh DE, Sweeney T. Long-term stability of RNA in post-mortem bovine skeletal muscle, liver and subcutaneous adipose tissues. BMC molecular biology. 2007;8:108.

35. Gonzalez-Herrera L, Valenzuela A, Marchal JA, Lorente JA, Villanueva E. Studies on RNA integrity and gene expression in human myocardial tissue, pericardial fluid and blood, and its postmortem stability. Forensic science international. 2013 Oct 10;232(1-3):218-28.

36. Morrison PK, Bing C, Harris PA, Maltin CA, Grove-White D, Argo CM. Post-mortem stability of RNA in skeletal muscle and adipose tissue and the tissue-specific expression of myostatin, perilipin and associated factors in the horse. PloS one. 2014;9(6):e100810.

37. Hsiao LL, Dangond F, Yoshida T, Hong R, Jensen RV, Misra J, et al. A compendium of gene expression in normal human tissues. Physiological genomics. 2001 Dec 21;7(2):97-104.

38. Mele M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, et al. Human genomics. The human transcriptome across tissues and individuals. Science. 2015 May 8;348(6235):660-5.

39. Popova T, Mennerich D, Weith A, Quast K. Effect of RNA quality on transcript intensity levels in microarray analysis of human post-mortem brain tissues. BMC genomics. 2008;9:91.

40. Hruz T, Wyss M, Docquier M, Pfaffl MW, Masanetz S, Borghi L, et al. RefGenes: identification of reliable and condition specific reference genes for RT-qPCR data normalization. BMC genomics. 2011;12:156.

41. Coulson DT, Brockbank S, Quinn JG, Murphy S, Ravid R, Irvine GB, et al. Identification of valid reference genes for the normalization of RT qPCR gene expression data in human brain tissue. BMC molecular biology. 2008;9:46.

42. Vennemann M, Koppelkamm A. Postmortem mRNA profiling II: Practical considerations. Forensic science international. 2010 Dec 15;203(1-3):76-82.

43. Koppelkamm A, Vennemann B, Fracasso T, Lutz-Bonengel S, Schmidt U, Heinrich M. Validation of adequate endogenous reference genes for the normalisation of qPCR gene expression data in human post mortem tissue. International journal of legal medicine. 2010 Sep;124(5):371-80.

44. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology. 2002 Jun 18;3(7):RESEARCH0034.

45. Andersen CL, Jensen JL, Orntoft TF. 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 research. 2004 Aug 1;64(15):5245-50.

APPENDIX: A PRIMER ON THE QUANTITATIVE POLYMERASE CHAIN REACTION: DATA HANDLING, ANALYSIS, AND INTERPRETATION

Theory and Basics

The Polymerase Chain Reaction (PCR) has become a mainstay of molecular biology, allowing the specific amplification of single-digit copies of a target molecule to easily detectable levels (Comprehensively reviewed in 16). Initially developed as an endpoint detection method using a DNA-dependent DNA polymerase, the method was vastly improved by the discovery of thermostable DNA-dependent DNA polymerases in bacteria from thermal vents. As an endpoint assay, PCR cannot be used to determine the starting concentration of the template molecule because the final amount of reaction product is not dependent upon the amount of input template. The PCR strategy generally follows the following paradigm: To template material are added primers, short pieces of single-stranded DNA (ssDNA), also known as oligonucleotides or oligos; deoxynucleotides—the individual building blocks of DNA—as enzyme substrate; and enzyme in a buffer solution optimized for the enzyme in use. The reaction is heated to "melt" the template into a single stranded form to allow primer binding to complementary template sequence, followed by cooling to the proper annealing temperature for primer binding to template. The reaction temperature is adjusted to the correct temperature for polymerase to bind and extend the primer. These three steps—melting or denaturation, annealing, and extension—are repeated some number of times, usually 25 to 40 "cycles." Because the enzyme is thermostable, it is not "killed" during the denaturation step. Primer design is flexible in that a pair of longer primers, ~20 nucleotides (nt, 20-mers), can be designed to specifically amplify a single target sequence with a perfectly efficient PCR reaction doubling the amount of product every cycle. Alternatively, short primers, 6-mers or hexamers, with a random sequence can be used to non-specifically make many products with random ends from a mixed template pool.

For purposes of the gene expression assays described in this report, purified RNA is used as template for an RNA-dependent DNA polymerase, also known as Reverse Transcriptase (RT), primed with random hexamers, to make a single copy of single stranded complementary DNA (cDNA). Because RT is not thermostable, only synthesis of the first strand cDNA is possible; the melting step for the second round of synthesis would denature the enzyme. But cDNA is a template for DNA-polymerase so primer-pair specific amplification can be performed with RT product in a reaction variant known as RT-PCR for Reverse Transcriptase-Polymerase Chain Reaction. Advances in instrumentation and chemistries made it possible to detect fluorescence intensities within individual reaction chambers at each cycle of a PCR reaction. When coupled with the realization that, within a wide range of template concentrations, the initial detection of dsDNA above background was in linear correlation with the initial template concentration, Quantitative RT-PCR (qRT-PCR) was born for single gene analysis from a total RNA sample (Reviewed in \(16).

Data Collection

Most instruments for qRT-PCR use either a laser and CCD camera, or fiber-optics to excite and detect the emission from a fluorescent dye. There are two common methods for dsDNA detection – dsDNA binding dyes such as SYBR® Green, and enzymatic hydrolysis of a labelled probe. SYBR® Green binds dsDNA by slipping between, or intercalating, the stacked bases of dsDNA. Intercalation increases the emission intensity of SYBR® Green about 10-fold. Because SYBR® Green binds all the available dsDNA, great care must be taken in characterizing each set of reaction conditions to ensure that the single gene of interest is the only amplification product present.

Enzymatic hydrolysis detection takes advantage of the fact that DNA-polymerase has two types of enzymatic activity when bound to ssDNA, the synthetic activity that makes dsDNA, and a cleavage activity that allows it to remove free oligos it may encounter bound to the template strand during synthesis (a

roadblock as it were). Assays that use this strategy to assess expression levels of a gene include a third oligo that binds between the primers used for amplification. This third oligo is labelled on one end with a fluorescent detector dye and on the opposite end with a quencher dye. During the annealing step of the PCR cycle, all three oligos bind the melted template. Free intact oligo is undetected because the quencher prevents the detector from emitting signal. During the extension step, bound oligo is cleaved by polymerase, physically separating the quencher from the detector dye and allowing it to emit signal upon excitation. For both systems, SYBR® Green and enzymatic hydrolysis, instruments are set up to detect fluorescent emission at the end of the extension step at each cycle. As more product is made total fluorescent emission increases in direct relation to the total amount of product from the reaction.

Data Interpretation

Because detectors have a minimum detectable intensity, the cycle number at which fluorescence is first detected is often used as the data point for further analysis. However, instrument software allows adjustment of this "threshold" to allow for fine-tuning of data points (e.g., see Interplate Calibrators section below). The cycle number at which the amplification curve crosses the threshold is referred to as the Ct, for cycle threshold (also sometimes, Cq). Note that Ct is a linear number assigned to expression but that a well-tuned PCR reaction doubles the amount of product every cycle. Therefore, Cts are on a $\log_2$ scale in regards to the relative amount of dsDNA present in the reaction. A difference of 1 Ct reflects a 2-fold difference in expression, a 2 Ct difference reflects a 4-fold expression difference, a difference of 3 Cts reflects an expression difference of 8-fold, and so on.

Efficiency Correction

Ideally, PCR reactions perfectly double the product every cycle; however, there are several variables that affect reaction efficiency. Primers and detector oligos are typically matched for the temperature at which they anneal (bind to template) by optimizing them for nucleotide content and length but binding kinetics depends on other factors including buffer salt concentration and components which can vary by enzyme manufacturer. Synthesis of off-target product affects the availability of reactants and enzyme but can be detected by electrophoresis of reaction product to confirm a single product.

In SYBR® Green reactions, adding a melt-curve analysis to the end of the cycling protocol allows a quick assessment of the reaction product. A melt-curve takes advantage of the dye binding to dsDNA and cannot be used to assess enzyme hydrolysis reactions because the free detection dye fluoresces, not the dye that is still conjugated to the detection oligo and therefore is in close proximity to the quencher dye. A melt-curve is performed by melting the PCR product to ssDNA. Signal is detected every degree of increase to determine the temperature where the PCR product melts to ssDNA. A well-behaved reaction yields a single product of uniform length and nucleotide composition that melts at a single temperature. Reactions with off-target products may have multiple peaks in the melt curve due to differing lengths and nucleotide compositions of the products. Nonetheless, it is highly recommended that gel electrophoresis be used to confirm that a reaction is yielding a single product.

Reaction efficiencies are determined by performing a standard curve for every reaction on input material that matches the experimental input, i.e., purified whole blood control RNA for a gene expression experiment based on samples from blood, or from a pool of the experimental samples. At a minimum, the expected range of expression in the experimental samples should fall on the standard curve, and, ideally, the amount of material used as input for the curve should cover 10 logs. If instrument software does not directly calculate efficiencies, at a minimum, it will have an algorithm that graphs the standard curve data including the slope and curve fit. Efficiency is calculated according to the formula E=10^(-1/slope) where E is the amplification factor and slope is the slope of the curve on a log scale. Acceptable ranges are 1.6 <

$E < 2.1$ with $r^2 => 0.99$; $E = 2$ is ideal (16). The user must decide how much leniency to allow when reaction efficiencies fall outside the recommended ranges. In the analysis pipeline, Cts are efficiency-corrected before any further analysis is performed.

Normalization and Interplate Calibrators

The comparison of the expression value from a gene of interest to the expression value of a gene unaffected by the experimental conditions is called normalization and accounts for experimental error. In years past, it was assumed that certain "housekeeping" genes, beta-actin and GAPDH most often, had relatively constant levels of expression. It has since been unequivocally shown that expression of these genes responds to numerous conditions. If a qRT-PCR assay is to be used as validation for a whole-transcriptome experiment, the identity of a plethora of stably expressed genes can be gathered from that data. Otherwise, the user must experimentally test a panel of genes for stable expression in the experimental samples. Assays for panels of genes that are relatively stable are available (http://www.tataa.com/products-page/gene_expression_assays_panels/; http://www.primerdesign.co.uk/products/9460) and can be used as a starting point but must be empirically validated on the experimental samples.

Once the data is collected, there are two algorithms, geNorm (44) and NormFinder (45) that, used in concert, allow users to determine which genes are the most stable and how many genes are required for an appropriate level of experimental normalization. Both are available in Genex (multid.se) which includes a complete analysis pipeline that facilitates application of normalization factors to qRT-PCR data.

Large experiments that require samples and/or genes to be spread across multiple reaction plates require the use of interplate calibrators: an identical sample and reaction performed on all plates that contain data to be compared. Because thresholds can be manually set, Cts are arbitrarily different between plates. Because the calibrator reaction is identical across all plates, expression level is, as well. Global adjustment of thresholds within plates such that the calibrator has a constant Ct across plates normalizes Cts for all samples across plates. The typical workflow performs this step after efficiency correction and normalization, and before differential gene expression. Differential gene expression is performed on the transformed Ct values with the usual statistical methods for continuous numerical data based within the framework of the experimental design.

Comparing qRT-PCR to Microarray Data

Assume a PCR reaction with 100% efficiency ($E = 2.0$) and recall that a reaction of this efficiency doubles the amount of product each cycle. If the Ct is dependent on the minimum fluorescence intensity required for detection above background, and the minimum detection intensity is a fixed value, then it follows that a reaction with 1000 copies of target will reach the minimum detection intensity one Ct less than a sample that has 500 copies and 2 Cts less than a sample that has 250 copies. Therefore, the higher the Ct, the lower the expression of the target. In the body of the manuscript above is a comparison of expression values from qRT-PCR and microarrays. Expression values from microarray data are calculated from the fluorescence intensity of a spot on a chip where higher intensity is positively correlated to expression levels. As a result, when comparing expression data from qRT-PCR to that from microarrays (or RNA-seq) as performed in the above report, negative correlations demonstrate concordant expression patterns between the two types of data; positive correlations show discordance.