**Demodifying RNA for Transcriptomic Analyses of Archival Formalin-Fixed Paraffin-Embedded Samples**

Leah C. Wehmas\*, Charles E. Wood\*, Remi Gagne†, Andrew Williams†, Carole Yauk†, Mark M. Gosink‡, Deidre Dalmas§, Ruixin Hao¶, Raegan O’Lone||, Susan Hester\*

\*National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27709. Emails- wehmas.leah@epa.gov, wood.charles@epa.gov and hester.susan@epa.gov.

†Environmental Health Science and Research Bureau, Health Canada, Ottawa K1A 0K9, Canada. Emails- remi.gagne@canada.ca, andrew.williams@canada.ca and carole.yauk@canada.ca.

‡Pfizer Drug Safety R&D, Groton, CT 06340. Email- Mark.M.Gosink@pfizer.com

§GlaxoSmithKline, King of Prussia, PA 19406. Email- Deidre.A.Dalmas@gsk.com

¶DuPont, Newark, DE 19714. Email- ruixin.hao@dupont.com

||ILSI Health and Environmental Sciences Institute, Washington, DC 20005. Email- rolone@hesiglobal.org

To whom correspondence should be addressed at MD-B105-03, 109 T.W. Alexander Drive, U.S. Environmental Protection Agency, Research Triangle Park, NC, 27709 USA; Email: wehmas.leah@epa.gov.

**Running Title**: Demodification of RNA from FFPE Samples

**ABSTRACT**

Archival formalin-fixed paraffin-embedded (FFPE) tissue samples offer a vast but largely untapped resource for genomic research. The primary technical issues limiting use of FFPE samples are RNA yield and quality. In this study, we evaluated methods to demodify RNA highly fragmented and crosslinked by formalin fixation. Primary endpoints were RNA recovery, RNA-sequencing quality metrics, and transcriptional responses to a reference chemical (phenobarbital, PB). Frozen mouse liver samples from control and PB groups (n=6/group) were divided and preserved for 3 months as follows: frozen (FR); 70% ethanol (OH); 10% buffered formalin for 18 hours followed by ethanol (18F); or 10% buffered formalin (3F). Samples from OH, 18F, and 3F groups were processed to FFPE blocks and sectioned for RNA isolation. Additional sections from 3F received the following demodification protocols to mitigate RNA damage: short heated incubation with Tris-Acetate-EDTA buffer; overnight heated incubation with an organocatalyst using two different isolation kits; or overnight heated incubation without organocatalyst. Ribo-depleted, stranded, total RNA libraries were built and sequenced using the Illumina HiSeq 2500 platform. Overnight incubation (± organocatalyst) increased RNA yield >3-fold and RNA integrity numbers and fragment analysis values by >1.5-fold and >3.0-fold, respectively, *versus* 3F. Post-sequencing metrics also showed reduced bias in gene coverage and deletion rates for overnight incubation groups. All demodification groups had increased overlap for differentially expressed genes (77-84%) and enriched pathways (91-97%) with FR, with the highest overlap in the organocatalyst groups. These results demonstrate simple changes in RNA isolation methods that can enhance genomic analyses of FFPE samples.

**Key Words**: RNA quality, FFPE, archival resources, RNA integrity number, RNA-sequencing

**INTRODUCTION**

Clinical and research laboratories generate millions of formalin-fixed paraffin-embedded (FFPE) tissue samples each year. These specimens are routinely used for histopathological evaluations and represent the most common type of sample stored in most biorepositories. Recently, sequencing technologies have highlighted the vast potential of FFPE resources in characterizing molecular pathways associated with toxicity and different health outcomes. Potential applications range from biomarker discovery to identification of toxicity mechanisms and precision medicine approaches for individualized care (*e.g.*, National Cancer Institute, 2016) in both non-clinical species and humans. However, there are many important technical challenges when using FFPE samples for genomic analyses, most notably low RNA recovery and quality (Bass *et al.*, 2014; Greytak *et al.*, 2015; Stewart *et al.*, 2017). Formalin fixation leads to fragmentation, crosslinking, and other biochemical modifications of RNA that decrease yield, increase variability, and limit reliability of transcriptomic analyses. To date, there have been limited efforts to improve FFPE RNA for use in sequencing.

Several recent studies have examined how pre-analytical factors and RNA preparation methods influence genomic profiles obtained from FFPE samples. This work showed that RNA-sequencing (RNA-seq) of FFPE samples following ribosomal RNA (rRNA) depletion can yield transcriptional profiles highly concordant with those from matched frozen samples, but that FFPE samples may vary widely in their suitability for quantitative genomic analyses based on factors such as time-in-formalin (Webster *et al.*, 2015; Zhao *et al.*, 2014) and age-in-block (Hester *et al.*, 2016). In the latter study, for example, older FFPE samples had >90% lower total gene counts and poor concordance in global differentially expressed genes (DEGs) compared to their frozen counterparts. This work emphasizes the need for better methods to assess and improve the quality of RNA from FFPE samples prior to sequencing.

Few published studies have attempted to mitigate formalin-induced degradation of RNA in FFPE tissue. One study demonstrated that incubating formalin-fixed RNA oligonucleotides and total cellular RNA in weakly basic buffers following RNA isolation removed formalin adducts (Evers *et al.*, 2011). Another study showed that incubating FFPE RNA with a bifunctional organocatalyst during RNA isolation “demodified” damage by removing formalin adducts and reducing RNA breakage (Karmakar *et al.*, 2015). While both of these approaches indicate that FFPE RNA may be improved through demodification treatments, to date no study has specifically examined RNA-seq profiles or treatment responses using these methods to quantify the extent of improvement.

The goal of the current study was to investigate methods for improving the quality of RNA from FFPE samples for genomic analysis. We examined fixation effects on RNA yield, quality measures, and sequencing metrics, and we determined whether different RNA isolation and incubation protocols could enhance the genomic response detected in FFPE samples. Our results provide insight into the effects of formalin fixation on transcriptional profiles and demonstrate that modified protocols for RNA isolation can lead to improved RNA-seq data sets.

**MATERIALS AND METHODS**

*Experimental overview*

Archival samples for this work came from a short-term study in male B6C3F1 mice, described previously (Rooney *et al.*, 2017). Phenobarbital (PB; 99.8% purity, lot number SLBF7347V) was purchased from Sigma-Aldrich (St. Louis, MO) and administered at 0 (Con) or 600 ppm via drinking water to 10-11 week-old mice for seven days. Phenobarbital was selected for the chemical treatment in this experiment as it is a well-studied reference chemical with known biomarkers (*e.g.*, *Cyp2b10, Cyp3a11*) (Elcombe *et al.*, 2014) and transcriptional effects in the mouse liver (Geter *et al.*, 2014). Mice were reared under standard conditions within an AAALAC-accredited animal facility located in Research Triangle Park, NC, as previously described (Lake *et al.*, 2016). All procedures involving animals were approved by the U.S. Environmental Protection Agency (U.S. EPA) Institutional Animal Care and Use Committee.

At the time of collection, liver lobes were systematically trimmed, and portions of left lateral, caudate, and right medial lobes were mixed, flash frozen in liquid nitrogen, and stored at -80°C. Frozen mouse liver samples were arbitrarily selected from control (Con) and PB groups (n=6/group), divided on dry ice into four portions (~20-30 mg each), and then preserved according to the following methods (**Fig. 1**): stored at -80°C as frozen (FR); fixed in 70% ethanol for 3 months (OH); preserved in 10% buffered formalin for 18 hours followed by 70% ethanol to 3 months (18F); and fixed in 10% buffered formalin for 3 months (3F). Ethanol is a coagulant-type fixative that does not induce crosslinks and other biochemical modifications seen with formalin (Fox *et al.*, 1985). However, it often results in excessive shrinkage of morphologic features, and thus formalin is the standard fixative used for histopathological evaluation. A 10% buffered formalin solution (ThermoFisher Scientific; cat. #SF100-4; Fairlawn, NJ) contains approximately 4% formaldehyde (weight/volume).

Samples from OH, 18F, and 3F groups were kept at 4°C for 3 months, and then processed into paraffin blocks by standard histological methods: 80% ethanol for 30 minutes, twice; followed by 95% ethanol for 45 minutes each, twice; three changes of 100% ethanol for 45 minutes each; and three changes of xylenes (ThermoFisher Scientific, Fairlawn, NJ) for 30 minutes each; two changes of Paraplast™ Tissue Embedding Media (ThermoFisher Scientific, Fairlawn, NJ) at 58-60°C for 30 minutes each; and, finally, two changes of Paraplast™ at 58-60°C under vacuum for 1 hour each. The FR samples served as high-quality RNA controls. The 18-hour fixation time (18F) was selected to recapitulate a standard protocol for clinical and experimental samples, in particular those that may be used for immunohistochemical analyses (Fox *et al.*, 1985). The 3-month formalin fixation (3F) was selected as an arbitrary long-term fixation scenario expected to have more severe formalin-induced RNA degradation. FFPE blocks were then stored at room temperature for up to 4 months until sectioning for RNA isolation.

*RNA isolation and demodification*

Total RNA was isolated from frozen (FR) liver samples following homogenization in RNAzolRT (Molecular Research Center, Cincinnati, OH) and then purified by RNeasy MinElute column according to manufacturer recommendations (Qiagen GmbH, Hilden, Germany). For FFPE samples (OH, 18F, and 3F), total RNA was isolated from two to four 10 µm-thick paraffin sections, which were collected with a Historange or Leica microtome under RNase-free conditions, deparaffinized (Qiagen Deparaffinization Solution, cat. #19093), and digested with proteinase K prior to RNA purification using the Qiagen AllPrep® DNA/RNA kit (Qiagen, cat. #80234), as described elsewhere (Hester *et al.*, 2016).

To assess whether we could improve FFPE RNA quality for sequencing, additional sections from 3F were divided into four demodification subgroups. Total RNA was isolated from these samples using similar but slightly modified procedures to the protocol used for OH, 18F, and 3F groups. For the first demodification group (DTAE), RNA was isolated using the Qiagen AllPrep® DNA/RNA kit; however, after isolation, equal volumes of purified DTAE RNA and 2X TRIS-acetate-EDTA (TAE, final concentration 1X, pH 9.0; Sigma, St. Louis, MO) were combined and incubated for 30 minutes at 70°C. For the second demodification group (DQ), RNA was isolated and purified using the Qiagen AllPrep® DNA/RNA kit according to the manufacturer’s protocol, except the 15-minute 80°C incubation following proteinase K digestion and RNA supernatant transfer was replaced by an incubation for ~18 hours at 55°C with 40 mM NaOH-buffered 2-amino-5-methylphenyl phosphonic acid (organocatalyst, final concentration 20 mM, pH 7.0; Evans Analytical Group, Maryland Heights, MO), which catalyzes the breakage of formaldehyde-induced aminal crosslinks and hemiaminal adducts on nucleic acids (Karmakar *et al.*, 2015). For the third demodification group (DP), RNA isolation was performed using the PureLink™ FFPE Total RNA Isolation Kit (Invitrogen, Carlsbad, CA; 92008, #K1560-02). FFPE sections were placed in paraffin melting buffer, digested in proteinase K, and incubated for ~18 hours at 55°C with 40 mM organocatalyst (final concentration 20 mM, pH 7.0). Following incubation, DP RNA purification proceeded as described in the manufacturer’s protocol. As a control for organocatalyst treatment, FFPE microtome sections for the fourth group (NoD) were isolated exactly as for DQ, except that an equal volume of nuclease-free water (Ambion, Waltham, MA; cat. #AM9938) replaced the organocatalyst during the 18-hour incubation at 55°C. The concentration of RNA obtained from all samples was measured using a NanoDrop 2000c, full spectrum, UV spectrophotometer (ThermoFisher Scientific, Wilmington, DE) and Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA). Spectrophotometric-based methods calculate nucleic acid concentration indirectly through measuring UV absorbance at 260 nm; however, it cannot distinguish between measuring RNA or DNA (Thermo Fisher Scientific, 2009). Absorbance at 280 nm can provide some indications of sample purity when compared as the 260/280 ratio but contamination can also influence nucleic acid concentration values (Thermo Fisher Scientific, 2009). Fluorometric-based methods quantify RNA concentration using dyes that fluoresce only upon binding RNA making the assay insensitive to contaminants and more accurate (Molecular Probes Life Technologies, 2015). RNA integrity was evaluated by an Agilent 2100 Bioanalyzer (Agilent Technologies GmbH, Berlin, Germany). All RNA samples were stored at -80°C.

*Amplifiable RNA evaluation*

FFPE RNA was also measured by reverse transcriptase quantitative polymerase chain reaction (RT-qPCR), which is referred to as “amplifiable RNA” (*i.e.*, RNA that is not so fragmented or heavily modified by formalin adducts and crosslinks that the reverse transcriptase and polymerase are able to proceed unimpeded through cDNA synthesis and RT-qPCR reactions). Briefly, three sets of TaqMan primers and probes targeting different amplicons across the beta-actin (*Actb*) transcript were designed using Integrated DNA Technologies (IDT) PrimerQuest software and synthesized by IDT, Inc. (Coralville, IA). TaqMan primer, probe sequences and amplicon sizes can be found in **Table S1**. Oligo(dT)-primed cDNA was prepared using iScript™ Select cDNA Synthesis Kit (Bio-Rad Laboratories, Inc, Hercules, CA) according to manufacturer specifications. Triplex quantitative RT-qPCR reactions were prepared using PrimeTime® Gene Master Mix (IDT, Inc.) and run according to manufacturer instructions taking into account recommended adjustments from Nolan *et al*. (Nolan *et al.*, 2006). Copy numbers of each amplicon were obtained by running a triplex, 10-fold dilution, standard curve with each reaction.

*RNA-seq and analysis*

RNA library preparation and sequencing was completed at Expression Analysis (EA Genomic Services, Q2 Solutions—a Quintiles Quest Joint Venture, Durham, NC), as described previously (Hester *et al.*, 2016). FFPE RNA underwent reduced (or no) fragmentation (**Table S2**) during library preparation, depending on Bioanalyzer profiles. RNA was ribo-depleted and cDNA libraries were synthesized using the TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero (Illumina, San Diego, CA, cat. #RS-122-2303). Paired-end 50 base pair sequencing to at least 25 million reads per sample was performed on Illumina HiSeq 2500 instruments. Mean sequencing depth was 34.1 ± 0.3 million reads per sample with a read Phred score of 36.3 ± 0.0.

Basecall files were transformed into FASTQ files via Illumina (CASAVA v. 1.8.2). FASTQ files were demultiplexed, trimmed, and filtered for quality using EA | Q2 Solutions ea-utils (<https://github.com/ExpressionAnalysis/ea-utils/>; accessed 2017-05-20). Trimming included Illumina adapters, homopolymers at read ends, and nucleotides at read ends with Q-scores below 7. Filtering removes any read with one base at ≥95% frequency, homopolymers ≥ 4 within a read, average Q-score below 25, or length <25 bases. Reads were aligned using STAR v2.5 (parameters '--clip5pNbases 10','-sjdbGTFfile Mus\_musculus.GRCm38.84.gtf', '--quantMode TranscriptomeSAM GeneCounts') to the mouse genome using Ensembl gene annotation GRCm38 v84. QA/QC plots were generated from BAM files using Qorts v. 1.2.26 (Hartley and Mullikin, 2015) and R (R Core Team, 2016).

Reads assigned to Ensembl IDs for each sample were analyzed using R. Ensembl IDs representing rRNA were removed from further analyses. The quantified, mapped reads were transformed to counts per million (CPM), normalized using TMM (Robinson and Oshlack, 2010) and filtered using both stringent (all samples within one of the treatment groups having at least 10 CPM) and less stringent (all samples within one of the treatment groups having at least 0.5 CPM) approaches within edgeR (McCarthy *et al.*, 2012; Robinson *et al.*, 2010; Robinson and Smyth, 2007; Robinson and Smyth, 2008; Zhou *et al.*, 2014). Results from the two approaches were, overall, quite similar. Data and discussion will focus on the more stringent analysis results. Differential gene expression analysis between Con and PB treatment groups was conducted on the normalized, filtered data independently for each of the sample conditions using the glmLRT function, which performs a Likelihood Ratio Test assuming the data follow a negative binomial generalized log-linear model (McCarthy *et al.*, 2012). The dispersion parameters were estimated using the empirical Bayes method for tagwise negative binomial dispersions (McCarthy *et al.*, 2012). P-values were then adjusted using the false discovery rate (FDR) approach (Benjamini and Hochberg, 1995). The criteria for DEGs were defined as those genes with a FDR-adjusted p-value <0.05 and at least a 1.5-fold change in absolute value. DEG lists and associated fold changes were uploaded into Ingenuity Pathway Analysis (IPA) Knowledgebase (Qiagen) for canonical pathway enrichment identification and Upstream Regulator and Causal Network Analysis (Kramer *et al.*, 2014) with a significance α-level <0.05. Canonical pathway p-values were converted to –log10 values. Molecule types were limited to complex, enzyme, G-protein coupled receptor, group, growth factor, ion channel, kinase, ligand-dependent nuclear receptor, other, phosphatase, peptidase, transcription regulator, translation regulator and transmembrane receptor. Hierarchical biclustering was performed in Spotfire (TIBCO Software Inc., Palo Alto, CA) on pathway –log10 p-values using the UPGMA clustering method with correlation distance measure and average weight ordering. Pathway –log10 p-values and Upstream Regulator z-scores were compared between corresponding values from FR *vs*. OH, 18F, 3F, DTAE, DQ, DP, or NoD using linear regression.

*Other statistical methods*

RNA-seq quality metrics and individual gene expression data were analyzed for normality and homogeneity of variance using the Shapiro-Wilk (R stats package v. 3.3.2) and Levene's tests (R car package v. 2.1-4), respectively. When one or both of these assumptions failed, tests were repeated following log-transformation of the data. If assumptions of normality and homogeneity of variance held, a two-way repeated measures ANOVA was performed with the Holm *post hoc* test (R stats package v. 3.3.2). The main effects model was used for testing preservation group (FR, OH, 18F, 3F, DTAE, DQ, DP, and NoD) comparisons if the treatment (PB or Con) interaction term was nonsignificant. If there was a reasonable chance of an interaction (p-value <0.1), a one-way ANOVA was run on each treatment (PB and Con) separately comparing the preservation groups within each condition using the Holm *post hoc* test. Non-normal, heteroscedastic data without preservation group and treatment interactions were analyzed by combining PB and Con data for each preservation group and conducting pairwise Wilcoxon sign rank tests between preservation groups with a Holm correction for multiple comparisons (R stats package v. 3.3.2). If there was a reasonable chance of an interaction (p-value <0.1), the pairwise Wilcoxon sign rank tests was run on paired preservation groups separated by PB and Con using a Holm correction for multiple comparisons. Statistically significant effects were identified at p-value <0.05.

**RESULTS**

*Formalin fixation reduces RNA yield and quality*

Tissue fixation significantly reduced RNA yields by 2.5- to 5.2-fold compared to FR (p-value <0.01 for OH, 18F, and 3F). Formalin fixation for 18 hours had a greater effect on yield (5.2-fold less *vs.* FR) relative to FR than OH alone (2.5-fold less *vs.* FR). Extended time in formalin (3 months) did not further decrease RNA yield compared to 18 hours (4.5-fold less *vs*. FR) (**Fig. 2A**). Fixation also significantly impacted RNA integrity number (RIN). Formalin reduced average RIN values from 5.7 ± 0.3 in FR to 2.4 ± 0.2 and 1.6 ± 0.1 in 18F and 3F groups, respectively, while ethanol effects on RIN were intermediate (3.7 ± 0.2 for OH) (**Fig. 2B**). Refer to **Table S3** for specific p-values and data summaries.

Quantification of the housekeeping gene *Actb* across the gene body by multiplex RT-qPCR revealed more substantial fixation-dependent group differences. All experimental groups (including FR) demonstrated reduction in *Actb* copies across the gene body in a 3’>5’ manner. However, within each *Actb* amplicon there were similar and significant fixation-dependent reductions in amplifiable product compared to FR (**Fig. 2C**). For *Actb*1, ethanol resulted in 4.6-fold fewer copy numbers, followed by 18-hour formalin fixation (7.0-fold fewer) and 3-month formalin fixation (16.5-fold fewer) relative to FR. This trend was magnified with further distance from the 3’ end across *Actb*2 and to some extent *Actb*3 (the amplicon closest to 5’ end). However, the signal from *Actb*3 was quite low (near the limit of detection), resulting in much more variability and indicating little amplifiable product near the 5’ end of the *Actb* transcript (**Fig 2C**). Refer to **Table S4** for specific p-values and data summaries.

*Extended heated incubation improves RNA yield and quality*

Demodification treatments had varying effects on RNA yield and quality (**Fig 2A-B**). Short-term incubation in 1X TAE (DTAE) did not improve average RNA yield (75.5-84.7 ng/µl) and RIN values (1.0 ± 0.0) compared to 3F (108.7-117.5 ng/µl and 1.6 ± 0.1, respectively). In contrast, extended incubation at 55°C with (DQ and DP) and without (NoD) the organocatalyst improved RNA yields by 2.8- to 3.6-fold and RINs by 1.5- to 2.1-fold (**Fig 2A-B**). Addition of the organocatalyst did not provide a clear benefit to total RNA yield or RIN beyond extended incubation alone. However, the RT-qPCR results showed that use of the organocatalyst significantly improved amplification of all *Actb* amplicons compared to 3F **(Fig. 2C**). For example, DQ and DP had 123.6- and 181.1-fold more *Actb*1 copies compared to 3F (3.3 ± 0.8 copies), while NoD and DTAE had more modest improvements (22.8- and 19.2-fold, respectively). This trend was similar across the other two *Actb* amplicons (**Table S4**).

*Formalin fixation adversely impacts RNA-seq metrics*

Formalin fixation negatively affected a wide range of RNA-seq metrics, while ethanol fixation resulted in more limited effects. Despite depletion of rRNA prior to sequencing, 18F and 3F groups had higher levels of rRNA (10.7-11.3% of aligned reads) compared to FR (5.0 ± 0.3%) and OH (7.5 ± 0.7%) (**Fig. 3A**), which was potentially due to decreased probe hybridization during ribo-depletion. Formalin fixation further increased 3’ bias in gene body coverage (**Fig. 3B**) but had little impact on mean % GC content which was 50.7 ± 0.3% (18F) and 51.1 ± 0.6% (3F) compared to FR (49.2 ± 0.1%).

Other sequencing metrics showed variable changes based on fixation. For instance, % read duplications increased from an average of 42.6 ± .05% for FR, to 51.0 ± 1.5% and 60.9 ± 3.0% for 18F and 3F groups, respectively, while the OH group (43.4 ± 0.8%) did not differ significantly from FR (**Fig. 3C**). Read deletion rates (deleted bases per million mapped bases) were higher for 3F (353.4 ± 23.6), 18F (177.5 ± 7.0), and OH (67.4 ± 2.0) groups compared to FR (49.0 ± 1.7) (**Fig. 3D**). Formalin fixation also influenced overall gene diversity. Notably, this effect was most apparent within PB-treated samples. For example, PB-treated samples from 18F and 3F had 26.3-27.2% of reads mapping to the top 10% genes (genes with the most read counts), which was significantly more than FR 18.3 ± 0.4% (**Fig. 3E**). A similar trend was seen across the Con groups, but the differences were not statistically significant. Formalin groups generally had fewer % reads mapping to the coding region of exons, with 18F at 37.1 ± 0.6% and 3F at 26.7 ± 2.2%, compared to FR (46.9 ± 0.5%). An opposite pattern was seen in reads mapping to the untranslated region (UTR) of exons. The FR group had 19.4 ± 0.2% reads mapped to exonic UTRs, whereas formalin groups had 31.2 ± 0.9% (18F) and 32.1 ± 1.1% (3F). Reads mapping to introns did not demonstrate an apparent pattern with a slight increase in OH samples (27.7 ± 0.5%) relative to FR (26.0 ± 0.5%) but a decrease with formalin fixation to 23.5 ± 0.7% (18F) and 25.2 ± 2.2% (3F). There was a small but significant increase in % reads mapping to intergenic regions across fixed samples (6.3 ± 0.3% for OH), 6.2 ± 0.2% for 18F, and 11.1 ± 1.4% for 3F) relative to FR (5.3 ± 0.2%). The largest formalin-induced changes in % read mapping occurred within exonic regions (**Fig. 3F**). Data summaries as well as additional RNA-seq quality metrics with statistical analysis are summarized in **Table S5**.

*Extended heated incubation mitigates adverse effects of formalin on RNA-seq metrics*

Demodification protocols improved several pre- and post-alignment RNA-seq parameters compared to group 3F. Changes were most prominent in extended incubation groups (DQ, DP, and NoD), while few changes were noted in the DTAE group. DTAE, DQ, and DP groups all exhibited reduced % rRNA (8.2-8.8%) compared to 3F (10.7%) (**Fig. 3A**). All demodification groups lowered % duplication rates from 61.0 ± 3.0% in 3F to 33.5-42.8%, which approached FR rates at 42.6 ± 0.5%. Use of the organocatalyst (DQ and DP) resulted in the most substantial benefits by decreasing duplication rates to 35.8 ± 0.9 and 33.5 ± 0.8% respectively (**Fig 3C**). DQ, DP, and NoD groups also showed reversal of the 3’ bias observed across gene body coverage seen in 3F samples (**Fig. 3B**). Deletion rates (deleted bases per million mapped bases) were reduce from 353.4 ± 23.6 in 3F to between 173.8 and 223.8-173.8 across all demodification groups (**Fig. 3D**), although these effects were statistically significant only for the organocatalyst groups (183.0 ± 2.4 for DQ and 173.8 ± 3.0 for DP). While the NoD group generally showed similar results to the organocatalyst groups, one exception was gene diversity. The NoD group tended to have a higher proportion of reads (28.2% in Con and 36.5% in PB) mapping to the top 10% of genes (**Fig. 3E**), although variability within preservation groups and across PB and Con treatments made these differences not significant. DQ and DP showed improvements (not significant) in gene diversity with 24.7% (Con) and 24.2% (PB) reads mapping to the top 10% of genes (**Fig. 3E**). Demodification did not appear to greatly affect the variability between preservation groups across read mapping locations. There was a trend toward increased intronic read mapping across demodification groups (34.4 - 40.3%) relative to 3F (25.2 ± 2.2%) (**Fig. 3F**). Summary tables for sequencing quality metrics with statistical analysis are summarized in **Table S5**.

*Demodification treatments improve transcriptomic response*

Quantification of mapped reads within the FR group identified 15.8 million total raw gene counts of which 20.3 ± 0.1k where unique genes. Fixation in ethanol modestly decreased counts to 14.2 million but had no effect on uniquely mapped genes (20.3 ± 0.1k). Fixation in formalin decreased total counts to 12.2 million (19.9k ± 0.1k genes detected) and 8.2 million (17.6k ± 0.3k genes detected) in 18F and 3F groups, respectively (**Table S6**). A total of 231 DEGs were identified in PB versus Con samples within the FR group. Fixation with ethanol slightly increased filtered DEGs (235), while use of formalin reduced them to 205 and 218 for 18F and 3F groups, respectively (**Table 1**). A list of all identified DEGs along with expression values can be found in **Table S7.** The degree of overlap in DEGs between FR and fixation groups demonstrated a fixation and time-in-formalin dependent decline. For instance, OH had 83.5% overlap with FR, whereas 18F and 3F had 78.4% and 68.0% overlap, respectively. This decrease in overlap with time in formalin corresponded to an increase in the false negative DEGs, which were identified in FR but not in OH, 18F, or 3F. A clear pattern across groups was not apparent for false positive DEGs found in OH, 18F, or 3F and not FR. OH had 42 false positives compared to 61 in 3F and 24 in 18F (**Table 1**). The lists of DEGs overlapping with FR and associated expression values can be found in **Table S8**.

Principal component analysis of all DEGs revealed distinct separation of samples by Con and PB treatments, followed by sample clustering by preservation group (FR, OH, 18F, 3F, DTAE, DQ, DP, and NoD) **(Fig. 4A)**. The OH group clustered nearest FR, while the 3F group clustered farthest from FR. These results paralleled the DEG overlap data in that the OH group had the most DEGs in common with FR (83.5%), whereas 3F had the least (68.0%). Heatmap hierarchical clustering of all DEGs also showed a PB treatment effect with segregation of fixation groups in a pattern similar to the PCA plot. Despite differences across the fixation groups, a conservation in gene expression patterns was observed (**Fig 4B**). This observation was further supported by the high agreement in magnitude and direction of DEG fold changes across OH, 18F, and 3F relative to FR (R2: 0.979 – 0.993) (**Fig. S1**).

We pre-selected two established PB-induced biomarkers (*Cyp2b10* and *Cyp3a11*) and a common housekeeping gene (*Gapdh*) to help identify how preservation methods impact treatment-dependent biomarker and housekeeping gene expression pre- and post-normalization. We saw a significant fixation-related decrease in treatment response with an even larger time-in-formalin effect relative to FR across raw and normalized *Cyp2b10* and *Cyp3a11* count levels **(Fig 5**). These fixation-dependent group changes were somewhat obscured following normalization. Fixation (OH) reduced PB-induced normalized gene counts by 14.9% (*Cyp2b10*) and 23.8% (*Cyp3a11*) compared to FR (2.3 ± 0.1k and 8.8 ± 1.1k, respectively). Eighteen-hours formalin-fixed samples had lower normalized gene counts by 32.6% (*Cyp2b10*) and 33.4% (*Cyp3a11*) compared to FR, while 3F samples had looked slightly worse with normalized gene counts down by 34.5% for *Cyp2b10* and 37.1% for *Cyp3a11*. Normalized gene counts for *Gapdh* were also significantly reduced across OH (14.8%), 18F (29.7%), and 3F (65.3%) groups compared to FR (241.3 ± 5.7) across PB and Con samples (**Fig. 5**). The negative fixation-related effects were not necessarily conserved across all individual genes. For additional data on *Cyp2b10*, *Cyp3a11,* and *Gapdh*, see **Table S9**.

All demodification treatments enhanced quantification of raw gene counts to varying degrees (9.9-12.4 million, of which 19.9-20.8k were unique) (**Table S6**). Use of the organocatalyst provided the greatest improvement in raw gene counts (1.5-fold) compared to 3F, whereas the organocatalyst control group (NoD) showed the least improvement (1.2-fold) (**Table S6**). These results corroborate the cumulative gene diversity data (**Fig. 3E**).

All demodification treatments increased total DEGs (218-279) compared to 3F (218) **(Table 1, Table S7).** More relevant was the improvement in overlap of DEGs with FR across all demodification treatments. All demodification groups had higher DEG overlap with FR (76.6 to 84.4%) compared to 3F (62.9%). Across the overlapping DEGs, there was excellent concordance in the magnitude and direction of fold changes (R2: 0.976 - 0.989) (Fig. S1). Demodification also resulted in lower false negatives compared to 3F (**Table 1**). The demodification groups using the organocatalyst had the highest DEG overlap with FR (79.2% for DQ and 84.4% for DP) and lower false negative DEGs (48 and 36, respectively) compared to 3F (74). NoD also had modestly higher DEG overlap with FR (76.6%) relative to 3F but showed the highest number of false positives (54) and negatives (102). Principal component analysis of the DEGs revealed clustering of DQ, DP, and NoD groups nearer 18F (within the PB and Con exposures), whereas DTAE grouped closer to 3F (**Fig 4A**). A similar pattern of clustering was observed in the heatmap (**Fig. 4B**). This was slightly different than the trend seen in overlap of demodification group DEGs with FR, which would have predicted NoD to group more closely to DTAE and 3F.

At an individual biomarker gene level, demodification treatments tended to show improvements compared to 3F, although in many cases these changes were not statistically significant due to variability among individual samples following normalization **(Fig. 5**). For *Cyp2b10* and *Cyp3a11*, DQ demonstrated significant improvements in raw counts by 50.0 to 60.3% relative to 3F. All demodification treatments resulted in significant improvements of raw *Gapdh* counts relative to 3F with DP and DQ performing the best (**Fig. 5**). Demodification treatments did not show improvements in *Cyp2b10* or *Cyp3a11* following normalization. Only *Gapdh* retained significant demodification-related improvements in DTAE, DQ, and DP (24.3-45.1%) compared to 3F (83.8 ± 5.6%) after gene count normalization. Use of the organocatalyst improved normalized gene counts the most (45.1 and 42.8% for DQ and DP, respectively) compared to 3F (**Fig. 5**). These results varied for each gene. For additional data on *Cyp2b10*, *Cyp3a11*, and *Gapdh*, see **Table S9**.

*Top target pathways are conserved despite formalin effects*

Analysis of significant DEG lists identified over 300 significantly enriched canonical pathways, which showed considerable overlap between different sample conditions (**Table 2**). The most highly enriched pathways across all preservation methods included Nicotine Degradation II, PXR/RXR Activation, Xenobiotic Metabolism Signaling, and Melatonin Degradation I (**Table S10**). While fixation and time-in-formalin did not change the top-ranked PB treatment-related pathways, it did reduce the number of total enriched pathways in common with the frozen group. Demodification treatments mitigated this effect (**Table 2**). Hierarchical biclustering of the –log10 p-values for canonical pathways demonstrated that the confidence in the identified pathways was most similar between the FR, OH, and 18F samples (**Fig. 6**). The samples that received demodification treatment also clustered together. Linear regression analyses showed that enriched canonical pathways were highly consistent across preservation groups and that the demodification treatments tended to produce pathway predictions slightly more consistent with that of the FR samples than 3F and NoD samples **(Fig. S2, Table S11**).

IPA Upstream Regulator *z*-scores, which are based on the directionality of gene expression rather than the confidence in prediction measured by p-values, were also compared by linear regression (**Fig. S3**). As with comparison of canonical pathways, top upstream regulators predicted by PB treatment across preservation methods were highly similar when ranked by p-value and included several well-recognized PB responsive receptorssuch as *Nr1i2*, *Rxra*, and *Nr1i3* (**Table S12**). Linear regression of *z*-scores demonstrated that the demodification groups yielded upstream regulator predictions closer to that of the FR than did the 3F and NoD samples (**Table S11**). Furthermore, causal network analysis independently identified PB as the top regulator controlling gene expression when sorted by p-value (**Table S13**).

Supplementary data are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.7c4s5.

**DISCUSSION**

Improved methods are needed to expand molecular analyses of archival formalin-fixed tissue samples. The goal of this study was to evaluate techniques for enhancing the quality of FFPE RNA used in genomic analyses. Our findings characterize negative effects of short- and long-term formalin fixation on RNA quality and RNA-seq results at a gene, whole genome, and pathway level. Incorporation of overnight heated incubation with an organocatalyst during FFPE RNA isolation partially reversed the adverse effects of formalin fixation by improving RNA yield and quality across numerous sequencing metrics, including higher gene counts and improved DEG and pathway-level enrichment with FR control. These results demonstrate several relatively simple procedures that can be incorporated into standard FFPE RNA isolation kit protocols to improve RNA-seq data and genomic analyses of archival samples; that limiting formalin fixation time to 18-24 hours will minimize formalin-induced damage of RNA resulting in improved genomic analyses results when using FFPE samples; and, finally, when considering poorer quality FFPE samples for transcriptional analysis, careful consideration should be taken to identify the kind of data needed from the samples.

Formalin has been the most widely used tissue fixative since the early 1900s, primarily due to superior preservation (less distortion) of tissue morphology compared to alcohol fixatives (Fox *et al.*, 1985). However, many of the same biochemical features that make formalin the preferred fixative for histopathology create unique challenges for retrospective molecular analyses. Our results show that formalin fixation significantly decreased RNA yields and quality in a time-dependent manner, consistent with previously published work (Chung *et al.*, 2008; Masuda *et al.*, 1999; von Ahlfen *et al.*, 2007). These effects are likely due to a combination of formalin-induced hydroxyl methyl adducts, methylene bridges, and other biomolecular crosslinks, in addition to nucleic acid fragmentation (Evers *et al.*, 2011; Masuda *et al.*, 1999). Sequencing overcomes many of the historical problems associated with RNA fragmentation but may still be limited by formalin-induced covalent modifications. In this study, formalin fixation increased % rRNA sequenced, sequencing error rates (deletions), and overall variability across samples. The formalin-induced bias toward intronic and intergenic read mapping and greater 3’ gene body coverage also supports the observed reductions in gene diversity seen with increased time-in-formalin. This shift in read mapping from exonic to intronic and intergenic regions with formalin fixation appears to be a common feature of FFPE samples (Adiconis *et al.*, 2013; Graw *et al.*, 2015; Hedegaard *et al.*, 2014; Hester *et al.*, 2016; Morlan *et al.*, 2012; Webster *et al.*, 2015) and may be due to fixation of pre-mRNA processing machinery (*i.e.*, small nuclear ribonucleoproteins) leading to a higher abundance of unprocessed mRNA in FFPE samples contributing to fewer mapped reads. Other RNA-seq effects specific to formalin fixation included increased duplication rates within reads and reduced gene diversity. These findings document specific factors contributing to losses in gene detection and DEG identification in formalin-fixed samples.

The use of TAE buffer (DTAE) as a demodification treatment provided mixed results. Initially, it appeared to worsen RNA-seq results relative to 3F based on technical and global pre-sequencing and post-sequencing quality metrics. These findings contrasted with previous literature citing its potential benefits, which included reduced formalin fixed tissue or cellular RNA fragmentation based on gel electrophoresis of RT-PCR (Masuda *et al.*, 1999) and Agilent 2100 Bioanalyzer (Evers *et al.*, 2011), increased amplifiable formalin-fixed cellular RNA based on RT-qPCR (Evers *et al.*, 2011), and reversed formalin-fixed oligo RNA modification as indicated by TOF mass spectrometry (Masuda *et al.*, 1999) and Agilent 2100 Bioanalyzer (Evers *et al.*, 2011). However, these benefits of TAE were not evaluated using RNA-seq. At a gene and pathway level, DTAE did provide improvements in DEG and significant canonical pathway overlap with FR compared to the overlap of 3F with FR. These somewhat dichotomous results are likely due to increased fragmentation of FFPE RNA from the high temperature incubation, resulting in the lower RIN values, along with increased removal of adducts (Evers *et al.*, 2011; Masuda *et al.*, 1999), resulting in more amplifiable RNA and DEG-level detection.

The other demodification treatments (overnight incubation ± organocatalyst) improved sequencing results at most levels, from RNA to gene and pathway. The simple addition of an extended incubation at 55°C markedly increased RNA yield, which may be an important limiting factor for smaller fixed samples such as tissue cores or biopsies or older FFPE samples with highly fragmented nucleic acid. Further studies are needed to evaluate whether this simple step also improves DNA yields from FFPE sections. The extended incubation also increased RNA integrity (RIN), although this measure does not necessarily indicate the amount of amplifiable RNA. Multiplex RT-qPCR revealed that while extended incubation increased amplifiable *Actb* RNA to some extent (~1- to 23-fold depending on the location of the *Actb* amplicon), use of the organocatalyst increased amplifiable product much more substantially (25- to 150-fold depending on the location of the *Actb* amplicon). Of note, Karmakar *et al*. (2015) also reported enhanced yield of amplifiable RNA from FFPE samples (~5- to 5.5-fold across three different genes) following an 18-hour incubation at 55°C compared to standard kit conditions and, similar to the current study, found even greater yields (~7- to 25-fold) with inclusion of the organocatalyst compared to standard kit conditions, particularly for longer amplicons (Karmakar *et al.*, 2015). Possible reasons for this latter difference may relate to the proposed mechanism of the organocatalyst. While extended heating likely helps free up covalently bound nucleic acids, the bifunctional nature of the organocatalyst allows for reversal of formaldehyde-induced aminal crosslinks and hemiaminal adducts by both general acid catalysis and nucleophilic catalysis (Karmakar *et al.*, 2015).

When only looking at RNA yield (quantified by NanoDrop/Qubit), RNA quality, and typical global RNA-seq quality metrics (% rRNA, error rates, read mapping, gene body coverage), organocatalyst treatment appeared to show little if any improvement over extended incubation alone. The primary exceptions were the RT-qPCR and gene diversity results in which the NoD group had less quantifiable *Actb* mRNA and reduced mapping of reads to diverse genes, which was not seen in groups that used the organocatalyst (DQ and DP). Benefits from use of the organocatalyst were most evident at the gene level. Improvements included significantly higher gene detection and gene diversity; higher DEG overlap with FR including reduced potential false negative DEGs; and better pathway concordance with FR, enhancing enriched signaling pathway results. These results were not always apparent at the individual gene level and biomarker level.

While extended formalin fixation had clear adverse effects on RNA, the overall quality of RNA-seq data in the 3-month formalin fixation group was still markedly better than that reported previously for older FFPE samples. For example, in Hester *et al*. (2016), 21 year-old FFPE samples had 67% fewer genes detected, 88% fewer reads aligned to the transcriptome, and no concordance in enriched pathways compared to frozen sample pairs. In the current study, there were 13% fewer genes detected, 34% fewer reads aligned to the transcriptome, and relatively high concordance in treatment responses on biomarker genes and top signaling pathways between FR and 3F groups. This difference may have resulted from several factors, including the relatively young age of the FFPE samples used here (≤4 months in block) and the fact that the tissue samples were frozen before fixation, which may have negated potential formalin effects occurring at the time of fresh tissue fixation. However, there was still evidence of transcriptional artifacts induced by embedding and increased time in formalin, which was apparent in the false negative and false positive DEGs identified in OH, 18F, and 3F groups relative to fresh frozen tissue. Demodification reduced transcriptional artifacts, but future studies are needed to further improve FFPE protocols and investigate effects of demodification procedures in older, lower quality FFPE samples fixed at the time of collection.

There is increasing emphasis on the use of mechanistic pathway-based information in toxicology, environmental health, and biomedical science (*e.g.*, Dijkstra *et al.*, 2016; Simon *et al.*, 2014). This focus is supported by recent efforts to better integrate genomic biomarkers into risk assessment and clinical decision-making. Sequencing technologies now provide unprecedented access to genetic and transcriptomic information, but improved approaches are needed to relate these molecular changes to corresponding pathologic outcomes. Information from this study should enhance use of archival tissue resources and improve quality of RNA-seq data from FFPE and other types of challenging samples. Future applications of this work include data mining, biomarker development, and quantitative models that link transcriptomic data with health outcomes of interest.

**FUNDING INFORMATION**

Funding was provided by the U.S. EPA Office of Research and Development and the Health and Environmental Science Institute (HESI) scientific initiative, which is primarily supported by in-kind contributions (from public and private sector participants) of time, expertise, and experimental effort. HESI contributions are supplemented by direct funding (that largely supports program infrastructure and management) provided by HESI corporate sponsors. A list of supporting organizations (public and private) is available at http://hesiglobal.org/application-of-genomics-to-mechanism-based-risk-assessment-technical-committee/.

**ACKNOWLEDGEMENTS**

The authors would like to thank staff at EA for generation of RNA-seq data and bioinformatics support; the HESI committee; HESI and U.S. EPA reviewers for constructive comments on this manuscript; Dr. Brian Chorley, Gail Nelson, Gleta Carswell, Jeanene Olin and Dr. Thomas Hill III for technical assistance; and Judith Schmid for biostatistical support. The research described in this article has been reviewed by the U.S. EPA and approved for publication. Approval does not signify that the contents necessarily reflect the views or policies of the Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

**REFERENCES**

Adiconis, X., Borges-Rivera, D., Satija, R., DeLuca, D. S., Busby, M. A., Berlin, A. M., Sivachenko, A., Thompson, D. A., Wysoker, A., Fennell, T., Gnirke, A., Pochet, N., Regev, A., and Levin, J. Z. (2013). Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nature methods* **10**(7), 623-9, 10.1038/nmeth.2483.

Bass, B. P., Engel, K. B., Greytak, S. R., and Moore, H. M. (2014). A review of preanalytical factors affecting molecular, protein, and morphological analysis of formalin-fixed, paraffin-embedded (FFPE) tissue: how well do you know your FFPE specimen? *Archives of pathology & laboratory medicine* **138**(11), 1520-30, 10.5858/arpa.2013-0691-RA.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(1), 289-300, citeulike-article-id:1042553 doi: 10.2307/2346101.

Chung, J. Y., Braunschweig, T., Williams, R., Guerrero, N., Hoffmann, K. M., Kwon, M., Song, Y. K., Libutti, S. K., and Hewitt, S. M. (2008). Factors in tissue handling and processing that impact RNA obtained from formalin-fixed, paraffin-embedded tissue. *J Histochem Cytochem* **56**(11), 1033-42, 10.1369/jhc.2008.951863.

Dijkstra, K. K., Voabil, P., Schumacher, T. N., and Voest, E. E. (2016). Genomics- and Transcriptomics-Based Patient Selection for Cancer Treatment With Immune Checkpoint Inhibitors: A Review. *JAMA oncology* **2**(11), 1490-1495, 10.1001/jamaoncol.2016.2214.

Elcombe, C. R., Peffer, R. C., Wolf, D. C., Bailey, J., Bars, R., Bell, D., Cattley, R. C., Ferguson, S. S., Geter, D., Goetz, A., Goodman, J. I., Hester, S., Jacobs, A., Omiecinski, C. J., Schoeny, R., Xie, W., and Lake, B. G. (2014). Mode of action and human relevance analysis for nuclear receptor-mediated liver toxicity: A case study with phenobarbital as a model constitutive androstane receptor (CAR) activator. *Critical reviews in toxicology* **44**(1), 64-82, 10.3109/10408444.2013.835786.

Evers, D. L., Fowler, C. B., Cunningham, B. R., Mason, J. T., and O'Leary, T. J. (2011). The Effect of Formaldehyde Fixation on RNA: Optimization of Formaldehyde Adduct Removal. *The Journal of Molecular Diagnostics* **13**(3), 282-288, http://dx.doi.org/10.1016/j.jmoldx.2011.01.010.

Fox, C. H., Johnson, F. B., Whiting, J., and Roller, P. P. (1985). Formaldehyde fixation. *J Histochem Cytochem* **33**(8), 845-53.

Graw, S., Meier, R., Minn, K., Bloomer, C., Godwin, A. K., Fridley, B., Vlad, A., Beyerlein, P., and Chien, J. (2015). Robust gene expression and mutation analyses of RNA-sequencing of formalin-fixed diagnostic tumor samples. *Scientific reports* **5**, 12335, 10.1038/srep12335.

Geter, D. R., Bhat, V. S., Gollapudi, B. B., Sura, R., and Hester, S. D. (2014). Dose-response modeling of early molecular and cellular key events in the CAR-mediated hepatocarcinogenesis pathway. *Toxicological sciences : an official journal of the Society of Toxicology* **138**(2), 425-45, 10.1093/toxsci/kfu014.

Greytak, S. R., Engel, K. B., Bass, B. P., and Moore, H. M. (2015). Accuracy of Molecular Data Generated with FFPE Biospecimens: Lessons from the Literature. *Cancer Res* **75**(8), 1541-7, 10.1158/0008-5472.CAN-14-2378.

Hartley, S. W., and Mullikin, J. C. (2015). QoRTs: a comprehensive toolset for quality control and data processing of RNA-Seq experiments. *BMC bioinformatics* **16**, 224, 10.1186/s12859-015-0670-5.

Hedegaard, J., Thorsen, K., Lund, M. K., Hein, A. M., Hamilton-Dutoit, S. J., Vang, S., Nordentoft, I., Birkenkamp-Demtroder, K., Kruhoffer, M., Hager, H., Knudsen, B., Andersen, C. L., Sorensen, K. D., Pedersen, J. S., Orntoft, T. F., and Dyrskjot, L. (2014). Next-generation sequencing of RNA and DNA isolated from paired fresh-frozen and formalin-fixed paraffin-embedded samples of human cancer and normal tissue. *PloS one* **9**(5), e98187, 10.1371/journal.pone.0098187.

Hester, S. D., Bhat, V., Chorley, B. N., Carswell, G., Jones, W., Wehmas, L. C., and Wood, C. E. (2016). Dose-Response Analysis of RNA-Seq Profiles in Archival Formalin-Fixed Paraffin-Embedded Samples. *Toxicological sciences : an official journal of the Society of Toxicology* doi: 10.1093/toxsci/kfw161, 10.1093/toxsci/kfw161.

Karmakar, S., Harcourt, E. M., Hewings, D. S., Scherer, F., Lovejoy, A. F., Kurtz, D. M., Ehrenschwender, T., Barandun, L. J., Roost, C., Alizadeh, A. A., and Kool, E. T. (2015). Organocatalytic removal of formaldehyde adducts from RNA and DNA bases. *Nat Chem* **7**(9), 752-758, 10.1038/nchem.2307

Kramer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**(4), 523-30, 10.1093/bioinformatics/btt703.

Lake, A. D., Wood, C. E., Bhat, V. S., Chorley, B. N., Carswell, G. K., Sey, Y. M., Kenyon, E. M., Padnos, B., Moore, T. M., Tennant, A. H., Schmid, J. E., George, B. J., Ross, D. G., Hughes, M. F., Corton, J. C., Simmons, J. E., McQueen, C. A., and Hester, S. D. (2016). Dose and Effect Thresholds for Early Key Events in a PPARalpha-Mediated Mode of Action. *Toxicological sciences : an official journal of the Society of Toxicology* **149**(2), 312-25, 10.1093/toxsci/kfv236.

Masuda, N., Ohnishi, T., Kawamoto, S., Monden, M., and Okubo, K. (1999). Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acids Research* **27**(22), 4436-4443.

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* **40**(10), 4288-97, 10.1093/nar/gks042.

Molecular Probes Life Technologies (2015). *Qubit® RNA BR Assay Kits For use with the Qubit® Fluorometer (all models)* doi: Thermo Fisher Scientific Inc., Carlsbad, CA.

Morlan, J. D., Qu, K., and Sinicropi, D. V. (2012). Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. *PloS one* **7**(8), e42882, 10.1371/journal.pone.0042882.

National Cancer Institute (2016). Cancer Moonshot Blue Ribbon Panel Report 2016*.*

Nolan, T., Hands, R. E., and Bustin, S. A. (2006). Quantification of mRNA using real-time RT-PCR. *Nature protocols* **1**(3), 1559-82, 10.1038/nprot.2006.236.

R Core Team. (2016). *R: A language and environment for statistical computing* doi: R Foundation for Statistical Computing, Vienna, Austria.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**(1), 139-40, 10.1093/bioinformatics/btp616.

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**(3), R25, 10.1186/gb-2010-11-3-r25.

Robinson, M. D., and Smyth, G. K. (2007). Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* **23**(21), 2881-7, 10.1093/bioinformatics/btm453.

Robinson, M. D., and Smyth, G. K. (2008). Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9**(2), 321-32, 10.1093/biostatistics/kxm030.

Rooney, J., Ryan, N., Chorley, B. N., Hester, S. D., Kenyon, E. M., Schmid, J. E., George, B. J., Hughes, M. F., Sey, Y., Tennant, A., MacMillan, D., Simmons, J. E., McQueen, C. A., Pandiri, A., Wood, C. E., and Corton, J. C. (2017). Genomic effects of androstenedione and sex-specific liver cancer susceptibility in mice. *Toxicological Sciences* doi: (In Review).

Simon, T. W., Simons, S. S., Jr., Preston, R. J., Boobis, A. R., Cohen, S. M., Doerrer, N. G., Fenner-Crisp, P. A., McMullin, T. S., McQueen, C. A., Rowlands, J. C., and Subteam, R. D.-R. (2014). The use of mode of action information in risk assessment: quantitative key events/dose-response framework for modeling the dose-response for key events. *Critical reviews in toxicology* **44 Suppl 3**, 17-43, 10.3109/10408444.2014.931925.

Stewart, J. P., Richman, S., Maughan, T., Lawler, M., Dunne, P. D., and Salto-Tellez, M. (2017). Standardising RNA profiling based biomarker application in cancer-The need for robust control of technical variables. *Biochim Biophys Acta* **1868**(1), 258-272, 10.1016/j.bbcan.2017.05.005.

Thermo Fisher Scientific (2009). *NanoDrop 2000/2000c Spectrophotometer V1.0 User Manual* doi: Thermo Fisher Scientific Inc., Wilmington, DE.

von Ahlfen, S., Missel, A., Bendrat, K., and Schlumpberger, M. (2007). Determinants of RNA quality from FFPE samples. *PloS one* **2**(12), e1261, 10.1371/journal.pone.0001261.

Webster, A. F., Zumbo, P., Fostel, J., Gandara, J., Hester, S. D., Recio, L., Williams, A., Wood, C. E., Yauk, C. L., and Mason, C. E. (2015). Mining the Archives: A Cross-Platform Analysis of Gene Expression Profiles in Archival Formalin-Fixed Paraffin-Embedded Tissues. *Toxicological sciences : an official journal of the Society of Toxicology* **148**(2), 460-72, 10.1093/toxsci/kfv195.

Zhao, W., He, X., Hoadley, K. A., Parker, J. S., Hayes, D. N., and Perou, C. M. (2014). Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC genomics* **15**, 419, 10.1186/1471-2164-15-419.

Zhou, X., Lindsay, H., and Robinson, M. D. (2014). Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic Acids Research* **42**(11), e91-e91, 10.1093/nar/gku310.

**FIGURE LEGENDS**

**Figure 1**. Experimental overview of fixation groups and demodification treatments. Abbreviations: Con-Control treatment; PB- Phenobarbital treatment; TAE- 1X Tris-Acetate EDTA.

**Figure 2.** Effects of formalin, fixation, and demodification on total RNA, amplifiable RNA, and RNA integrity. A) Total RNA yield as assessed by Nanodrop spectrophotometer and Qubit fluorometer. B) RNA integrity number as measured by Bioanalyzer. C) Amplifiable *Actb* RNA as measured by reverse transcriptase quantitative PCR of three amplicons across the gene body. \*The results of a statistical comparison between OH, 18F, and 3F *vs.* FR with a p-value <0.05. †The results of a statistical comparison between DTAE, DQ, DP, and NoD *vs.* 3F with a p-value <0.05. The lower and upper hinges of the boxplot correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 \* IQR. The lower whisker extends from the hinge to the smallest value at most 1.5 \* IQR of the hinge. Data beyond the end of the whiskers are the individually plotted points. Abbreviations: IQR- inter-quartile range or distance between the first and third quartiles.

**Figure 3**. Effects of formalin and demodification on global RNA-sequencing quality metrics. A) Percent of reads aligning to ribosomal RNAs. B) Gene body coverage from 5' to 3' for all gene transcripts scaled to percentiles. C) Mean percent of reads 1 and 2 that have the same first 35 nucleotides. Genes overlapping with other genes and reads mapping to more than one gene were excluded. D) Deletion rate per million reads along the sequencing cycle with median point values from cycles 25 to 40. E) Gene diversity from the top 1000 genes depicting the median group percentage of reads per library *vs*. the number of genes discovered. F) The percentage of reads mapping to coding, intronic, intergenic, and ambiguous regions. \*The results of a statistical comparison between OH, 18F, and 3F *vs.* FR with a p-value <0.05. †The results of a statistical comparison between DTAE, DQ, DP, and NoD *vs.* 3F with a p-value <0.05. The lower and upper hinges of the boxplot correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 \* IQR. The lower whisker extends from the hinge to the smallest value at most 1.5 \* IQR of the hinge. Data beyond the end of the whiskers are the individually plotted points. Abbreviations: IQR- inter-quartile range or distance between the first and third quartiles.

**Figure 4**. Gene-level effects of formalin and demodification treatments on differential gene expression analysis. A) Principal component analysis of common DEGs across experimental groups from Con (left) and PB (right) treatments. B) Hierarchical cluster analysis on median count data (green indicates higher counts and red indicates lower counts) from DEGs identified in any group with FDR-adjusted p-value <0.05 and absolute fold change >1.5. Abbreviations: DEGs - differentially expressed genes, PB - phenobarbital, Con - control.

**Figure 5**. Effects of fixation and demodification on individual gene expression profiles of two PB-induced biomarkers (*Cyp2b10* and *Cyp3a11*) and one housekeeping gene (*Gapdh*) with raw count data (left) and TMM normalized count data (right). Each point and individual line on the figure represents a biological replicate (mouse) across the different preservation groups. Red dots indicate replicates from the Con exposure while blue dots indicate replicates from the PB exposure. \*The results of a statistical comparison between OH, 18F, and 3F *vs.* FR with a p-value <0.05. †The results of a statistical comparison between DTAE, DQ, DP, and NoD *vs.* 3F with a p-value <0.05. Abbreviations: DEGs - differentially expressed genes, FDR - false discovery rate, PB -phenobarbital, Con - Control, TMM - weighted trimmed mean of the log expression ratios.

**Figure 6.** Pathway-level comparison of preservation procedures. Hierarchical biclustering of canonical pathway –log10 of p-values. On the color scale, red indicates the maximum value, blue the minimum value, light gray indicates the average value and dark gray represents NA.

**TABLES**

Table 1. Significant differentially expressed genes and overlap with FR (FDR<0.05, ±1.5 fold change cut-off)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Group** | **Total** |  | **Comparison** | **# Overlap** | **FP** | **FN** | **% Overlap** |
| FR | 231 |  | - | - | - | - | - |
| OH | 235 |  | FR vs OH | 193 | 42 | 38 | 83.5 |
| 18F | 205 |  | FR vs 18F | 181 | 24 | 50 | 78.4 |
| 3F | 218 |  | FR vs 3F | 157 | 61 | 74 | 68.0 |
| DTAE | 229 |  | FR vs DTAE | 181 | 48 | 50 | 78.4 |
| DQ | 251 |  | FR vs DQ | 183 | 68 | 48 | 79.2 |
| DP | 239 |  | FR vs DP | 195 | 44 | 36 | 84.4 |
| NoD | 279 |   | FR vs NoD | 177 | 102 | 54 | 76.6 |

FDR - False discovery rate

FP - False positive

FN - False negative

Table 2. Significantly enriched IPA canonical pathways and overlap with FR (p-value <0.05, ±1.5 fold-change cutoff)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Group** | **Significant Canonical Pathways** |  | **Comparison** | **# Overlap** | **% Overlap** |
| FR | 303 |  | - | - | - |
| OH | 309 |  | FR vs OH | 282 | 93.1 |
| 18F | 292 |  | FR vs 18F | 269 | 88.8 |
| 3F | 324 |  | FR vs 3F | 271 | 89.4 |
| DTAE | 342 |  | FR vs DTAE | 293 | 96.7 |
| DQ | 338 |  | FR vs DQ | 281 | 92.7 |
| DP | 356 |  | FR vs DP | 294 | 97.0 |
| NoD | 333 |   | FR vs NoD | 275 | 90.8 |

IPA - Ingenuity Pathway Analysis

**SUPPLEMENTARY FIGURE LEGENDS**

**Figure S1.** Correlation analysis by log fold change of the differentially expressed genes found within each preservation condition and corresponding frozen group. Abbreviations: FR - frozen, OH - ethanol, 18F – 18hr formalin, 3F - 3m formalin, DTAE - Demod Tris-Acetate EDTA, DQ - Demod 18hr Qiagen, DP - Demod 18hr PureLink, and NoD - Control no catalyst.

**Figure S2**. Plot of –log10 of pathway p-values from the frozen (FR) sample versus pathway predictions from the other preservation procedures. Best fit linear regression line is shown and line expression is shown in the upper right side of each panel. Panels are: A) FR *versus* OH; B) FR *versus* 18F; C) FR *versus* DTAE; D) FR *versus* DP; E) FR *versus* DQ; F) FR *versus* NoD; G) FR *versus* 3F. Abbreviations: FR - frozen, OH - ethanol, 18F – 18hr formalin, 3F - 3m formalin, DTAE - Demod Tris-Acetate EDTA, DQ - Demod 18hr Qiagen, DP - Demod 18hr PureLink, and NoD - Control no catalyst.

**Figure S3**. Plot of Upstream Regulator activation *z*-scores from the frozen (FR) sample versus Upstream Regulator activation *z*-scores from the other isolation procedures. Best fit linear regression line is shown and line expression is shown in the upper right side of each panel. Panels are: A) FR *versus* DP; B) FR *versus* 18F; C) FR *versus* DTAE; D) FR *versus* OH; E) FR *versus* DQ; F) FR *versus* 3F; G) FR *versus* NoD. Abbreviations: FR - frozen, OH - ethanol, 18F – 18hr formalin, 3F - 3m formalin, DTAE - Demod Tris-Acetate EDTA, DQ - Demod 18hr Qiagen, DP - Demod 18hr PureLink, and NoD - Control no catalyst.

**SUPPLEMENTARY TABLE CAPTIONS**

**Table S1.** Taqman primer and probe sequences for RT-qPCR mRNA integrity assay targeting Actb (NM\_007393.5)

**Table S2.** FFPE RNA fragmentation duration for library preparation.

**Table S3.** RNA yield and RNA integrity numbers.

**Table S4.** Summary of *Actb* RNA copies by amplicon across different preservation conditions.

**Table S5.** RNA-sequencing quality summary for additional metrics that A) demonstrate no chemical exposure related interaction with preservation groups and B) those that do.

**Table S6.** Summary of total gene counts and quantity of unique genes identified prior to count normalization.

**Table S7**. All significant differentially expressed genes identified between preservation groups.

**Table S8**. Significant differentially expressed genes list from each preservation group overlapping with FR

**Table S9:** TMM normalized and raw biomarker (*Cyp2b10, Cyp3a11*) and housekeeping gene

(*Gapdh*) count summaries.

**Table S10**. Significantly enriched canonical pathways from IPA ranked by p-value or *z-*score relative to FR.

**Table S11.** Linear regression analyses comparing –log10 of pathway p-values and upstream regulator *z-*scores from the frozen (FR) sample versus predictions from respective preservation procedures.

**Table S12**. Ingenuity Pathway Analysis upstream regulator comparison ranked relative to FR p-value.

**Table S13**. Casual network analysis comparing -log10(p-values) across preservation groups.