HUDSONALPHA DISCOVERY
SEQUENCING + BIOINFORMATICS™

Optimized Methods for Sequencing RNA from FFPE samples

Factors that affect quality and benchmarks for success

DISCOVERY LIFE SCIENCES

Delivering success one nucleotide at a time

Hidden in pathology archives and biospecimen repositories around the world is a treasure trove of information locked within formalin-fixed paraffin-embedded (FFPE) tissue samples. While FFPE is a gold standard preservation method for immunohistochemistry and other antibody-based analyses, the technique can pose difficulties for researchers wishing to extract nucleic acid-based information, especially gene expression data.

At HudsonAlpha Discovery, we've spent years optimizing this workflow, developing robust processes that include a method for dual extraction of RNA and DNA from FFPE tissues that delivers higher quality and quantities of RNA compared to other commonly used protocols, and without negatively impacting DNA quality and quantity.

In this white paper, we share a few of the important factors to consider when obtaining and sequencing RNA from FFPE tissues, factors which have been critical for HudsonAlpha Discovery's successful implementation of an FFPE RNA sequencing program.

Factor 1: Chemical modification of RNA by formaldehyde

One factor impacting the usability of RNA extracted from FFPE tissue is the reactivity of the formalin fixative with RNA. The tertiary amines found on adenines, cytosines, and, to a lesser extent, guanines, can become modified through the addition of N-methylol (Figure 1), leading to crosslinked bases via a methylene bridge[1]. Both crosslinking and the methylol adduct can impede reverse transcription and reduce binding of PCR primers, especially to poly-A tails in mRNA, leading to low yields and small fragment lengths of cDNA.

Fortunately, formalin modification of RNA can be somewhat reversed by heating in pH 8.0 buffer prior to extraction[2], although the exact conditions should be optimized for different types of samples.

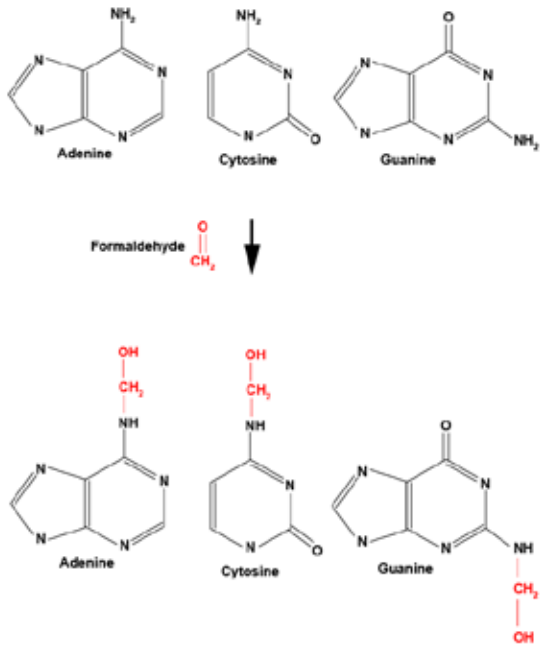


Figure 1. The tertiary amines found on adenines, cytosines, and, to a lesser extent, guanines, can become modified through the addition of N-methylol.

Factor 2: Degradation of RNA

A second factor affecting the usability of RNA from FFPE tissues is fragmentation of the RNA from hydrolysis and RNases. While RNases are an ever-present issue when extracting RNA from fresh and frozen tissues, the fixative in FFPE tissue can impair lysis leading to release of RNases[1]. In addition, fixation at low pH and prolonged tissue hypoxia can promote RNA hydrolysis. Thus, RNA extraction conditions also need to be optimized to minimize these effects.

Factor 3: Pre-analytical variables

One group of factors that can be difficult to anticipate when working with archival samples are variations in sample handling prior to- and during fixation. The length of time between resection and fixation, the temperature a sample experiences during that time, and the length of time in fixative have all been shown to affect the recovery of RNA from FFPE tissues[3-6]. However, a study examining the ability to perform RNA-seq on RNA recovered from FFPE tissues showed that anywhere from 1 - 12 hours between surgical excision and fixation had negligible effects on the reproducibility of differential mRNA and miRNA studies, although longer delays can lead to increased variability for miRNA populations and shorter time points were not tested[4]. Nevertheless, the authors of the study recommend keeping the delay to fixation time to 3 hours or less, if possible.

It's important to note that this same study found that RNA from FFPE tissues did have consistently more intronic sequences than RNA from matched snap-frozen tissues, which may complicate quantification of individual splice variants. Thus, library preparation methods that enrich for exonic target sequences or sequencing to an increased read depth is needed to compensate for the excess intronic material.

Factor 4: Size and age of the FFPE block

In the community, there is some concern that the age of the FFPE block can affect the quality of the extracted RNA. At HudsonAlpha Discovery, we find that the impact of the age of the block on RNA quality can be mitigated if the block is large enough that the exposed areas can be removed and analysis performed on internal sections.

Factor 5: Metrics of RNA quality

When preparing RNA for RNA-seq studies, the quality of the RNA matters. The way you measure quality is also important. While many researchers rely on the RNA integrity number (RIN) to evaluate quality, this metric is not as predictive of success in RNA-seq studies. Instead we recommend using DV200[6], the percentage of fragments ≥200 nucleotides, which better correlates with successful RNA-seq runs.

Benchmarks for successful RNA-seq on FFPE tissues

At HudsonAlpha Discovery, we've shown that with careful optimization, you can achieve successful RNA-seq data on 95.6% of FFPE cores and 88.6% of FFPE sections (Table 1).

Table 1. Optimization of RNA isolation from FFPE tissues increases the probability of successful RNA-seq studies*

	Standard Protocol		HudsonAlpha Discovery	
	SECTIONS	CORES	SECTIONS	CORES
All FFPE samples submitted	3171	539	2891	1745
Failed RNA extraction	1886	317	230	48
Failed Library Prep	259	45	91	28
Failed sequencing QC	20	2	10	0
Passed Primary Seq QC	1006	175	2560	1669
Success Rate**	**31.7%**	**32.5%**	**88.6%**	**95.6%**

*Data generated as part of the Oncology Research Information Exchange Network's (ORIEN) Avatar™ program, powered by M2GEN.

**Success rate = (# of samples that passed primary seq QC) / (# of samples submitted). At HudsonAlpha Discovery, QC success is based on a combination of factors including >75% of bases at Q30 or above (we typically observe >90%), duplication rate of 10% or less (we typically observe <5%), coverage > 300x for exome tumor samples, RNAseq samples with 100M paired reads or more per sample and with a mapping rate of 85% or higher.

One of the reasons RNA-seq works so well with our RNA isolation methods is because we've optimized the process to yield longer RNAs, which is demonstrated by our ability to consistently obtain high DV200 scores within the same type of tissue (Figure 2) and across different tissue types (Figure 3). Our optimized RNA extraction protocol maintains the quality and quantity of DNA (data not shown), enabling multi-omic analysis of individual samples.

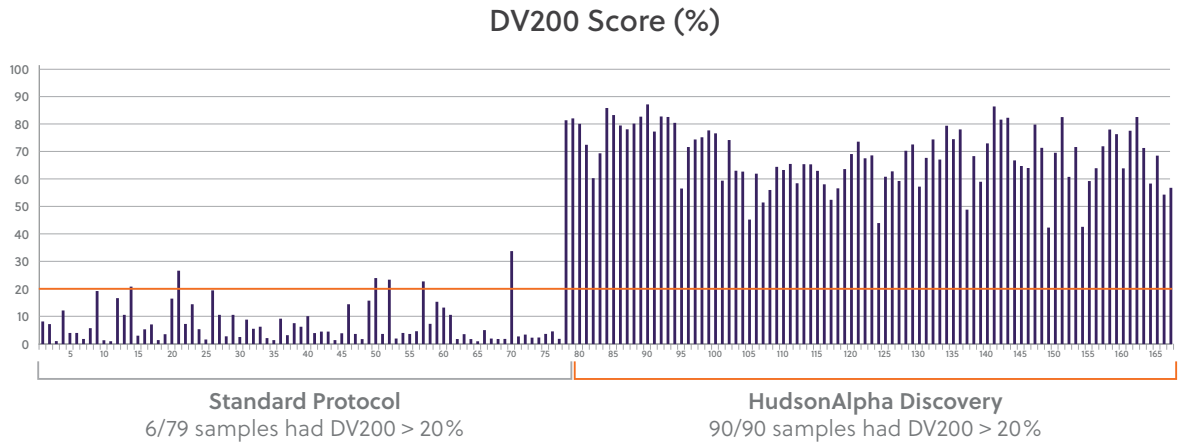


Figure 2. Our optimized protocol for extracting RNA from FFPE tissues consistently yields RNA with high DV200 scores (the smaller number of samples processed using the standard protocol is due to limitations in the amount of available tissue).

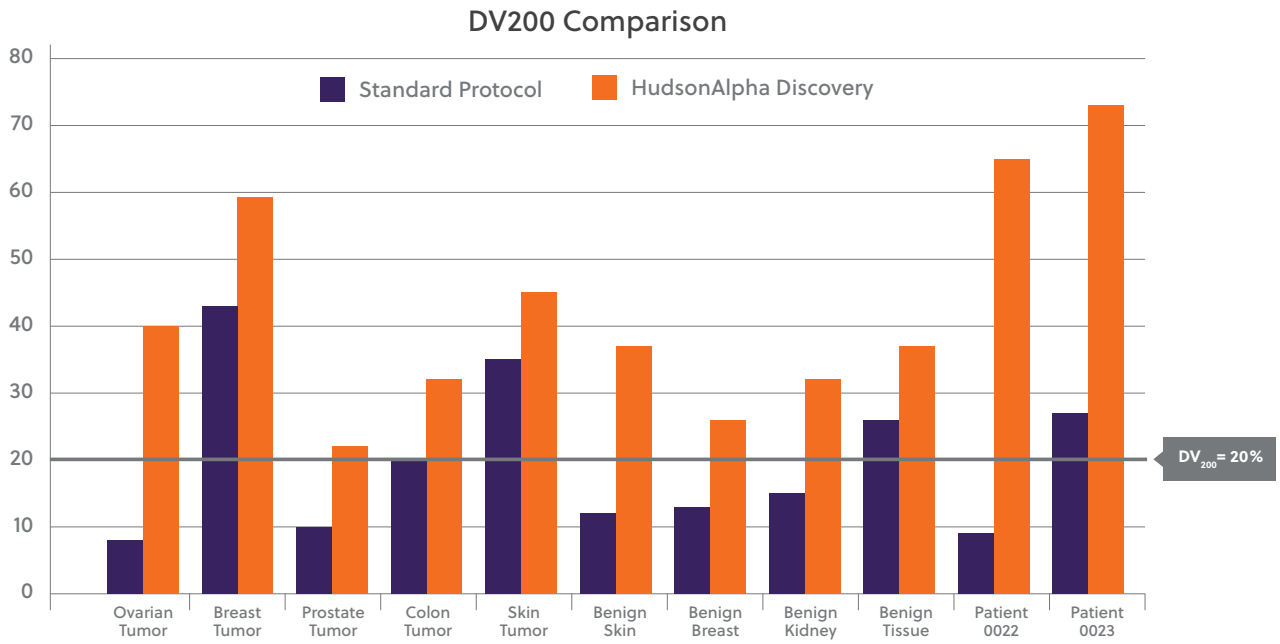


Figure 3. Our optimized protocol for extracting RNA from FFPE tissues is successful across a range of tissue types and patients.

Another reason is our use of workflows that limit technical bias (Figure 4). We use automated processes and iSeq sequencers to balance library pools before sequencing, and we distribute sequencing of balanced library pools across flow cells and sequencers to normalize variability. In a longitudinal study that tested thousands of samples, we demonstrated a plate to plate variability of <1%[7].

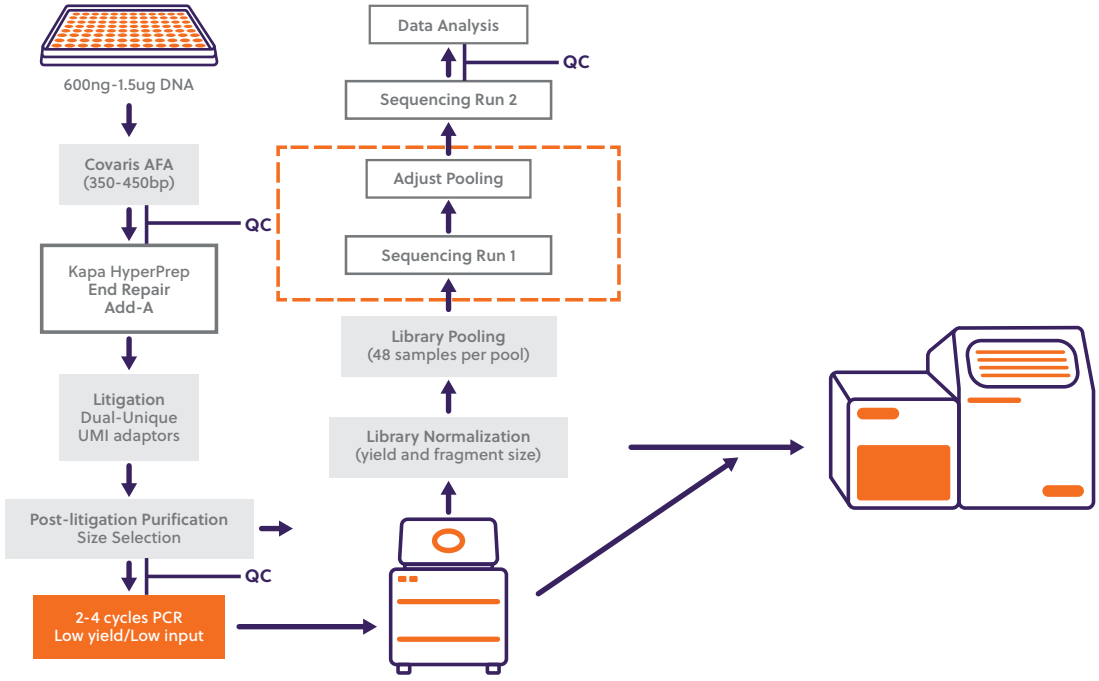


Figure 4. The HudsonAlpha Discovery workflow is optimized to limit technical bias.

Extracting disease insights from FFPE tissue

There's a great deal of optimization involved when working with FFPE samples, especially when your goal is to obtain data on gene transcript, fusion, and splice site isoform expression. At HudsonAlpha Discovery, we've already done the hard work of developing protocols that yield successful results and are ready to use our expertise and technologies, including long-read sequencing platforms like PacBio and Oxford Nanopore, to advance your search for insights into the biology of disease.

But reliably obtaining gene expression data from FFPE tissue is only a small part of the discovery journey. Working with the global Discovery Life Sciences organization, you can combine gene expression analysis with orthogonal technologies like immunohistochemistry (IHC) and flow cytometry to greatly enrich your understanding.

Just as important, we can ensure you have access to sufficient high-quality biospecimens, including normal, diseased, matched sets, and enriched subsets (e.g. dissociated tumor cells). By normalizing the collection, quality control and preservation of biospecimens, we minimize variation and ensure that genomic resolution is increased along with statistical power. Together with the Discovery Partners clinical network, HudsonAlpha Discovery can offer high-density genomic data across many sample types that drives research in a meaningful way.

References

1. Srinivasan M, Sedmak D, Jewell S. Effect of Fixatives and Tissue Processing on the Content and Integrity of Nucleic Acids. *Am J Pathol*. 2002 Dec; **161**(6): 1961–1971. PMCID: PMC1850907.

2. Evers DL, et al. The Effect of Formaldehyde Fixation on RNA: Optimization of Formaldehyde Adduct Removal. *J Mol Diagn*. 2011 May; **13**(3): 282–288. PMCID: PMC3077733.

3. Wimmer I, *et al*. Systematic evaluation of RNA quality, microarray data reliability and pathway analysis in fresh, fresh frozen and formalin-fixed paraffin embedded tissue samples. Sci Rep. 2018 Apr 20;8(1):6351. PMCID: PMC5910432.

4. Jones W, *et al*. Deleterious effects of formalin-fixation and delays to fixation on RNA and miRNA-Seq profiles. *Sci Rep*. 2019 May 6;**9**(1):6980. PMCID: PMC6502812.

5. Mathieson W, *et al*. Cold Ischemia Score: An mRNA Assay for the Detection of Extended Cold Ischemia in Formalin-Fixed, Paraffin-Embedded Tissue. *J Histochem Cytochem*. 2019 Mar;**67**(3):159-168. PMCID: PMC6393842.

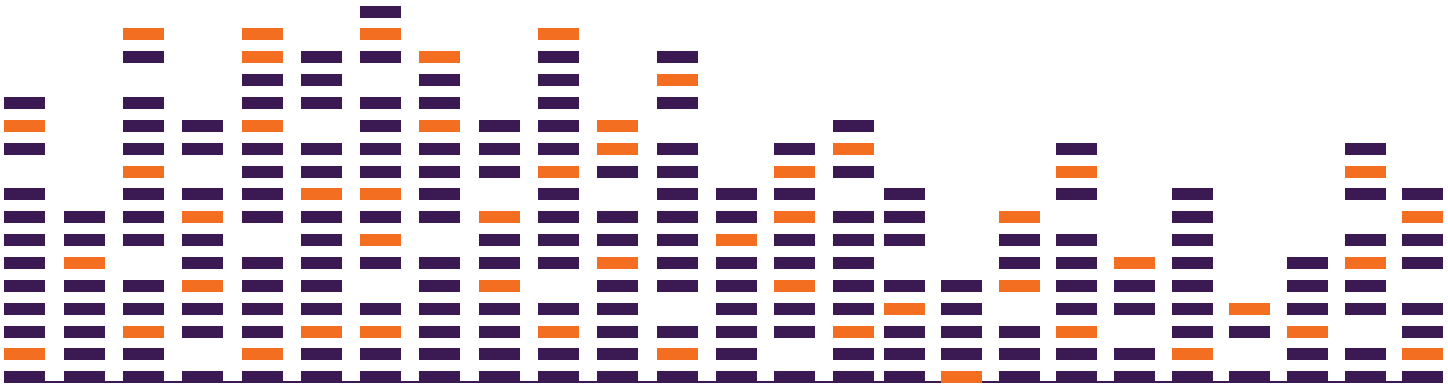
6. Carithers LJ, *et al*. The Biospecimen Preanalytical Variables Program: A Multiassay Comparison of Effects of Delay to Fixation and Fixation Duration on Nucleic Acid Quality. *Arch Pathol Lab Med*. 2019 Sep;**143**(9):1106-1118. DOI: 10.5858/arpa.2018-0172-OA.

7. Craig D, *et al*. Longitudinal Analysis of 4,500 Whole Blood Transcriptomes across over 1,500 Whole-Genome Sequenced Individuals within the Parkinson's Progression Markers Initiative. *American Society of Human Genetics* (ASHG). Oct. 2019. Houston, TX.

***Learn more about HudsonAlpha Discovery's genomic services and request a quote by visiting* dls.com/hudsonalpha-discovery.**

HUDSONALPHA
DISCOVERY
SEQUENCING + BIOINFORMATICS

Learn more about HudsonAlpha Discovery's genomic services and request a quote by visiting **dls.com/hudsonalpha-discovery.**



© 2021 Discovery Life Scie

www.dls.com

US (866) 838-2798
Europe: +359 2 87 2117