



# Workflow Questions Answered: Challenges with FFPE Tissue Samples

Jung Doh, Ph.D., Senior Applications Scientist, Beckman Coulter, Inc.

Next-generation sequencing (NGS) can provide in-depth analysis of short sequences, which could make it a powerful tool to use with nucleic acid isolated from formalin-fixed, paraffin-embedded (FFPE) tissue.

Senior applications scientist Jung Doh recently addressed ten common questions about working with challenging FFPE tissue in NGS sample preparation workflows. His responses and customer tips are included below.

## 1 When is using FFPE samples for NGS a better option over fresh-frozen samples?

Fresh- or snap-frozen samples are a great source for DNA, but they're costly to collect and maintain, and therefore not widely available. This low availability precludes their use in large scale retrospective studies of tumors — studies that could ultimately drive development of new cancer therapies.

In contrast, formalin-fixed paraffin-embedded (FFPE) tissue samples are routinely archived during patient treatment and care. By some estimates, there are 400 million<sup>1</sup> to more than a billion FFPE samples<sup>2</sup> in hospitals and tissue banks worldwide. Many have clinical annotations — primary diagnosis, therapeutic regimen, drug response and recurrence status — so they can be linked to clinical outcomes and long-term follow-up<sup>3</sup>.

For some types of tumor tissues, FFPE samples are often the only source of DNA<sup>4</sup>. It's these samples that can provide researchers with crucial information about some of the rarest cancers and other conditions<sup>5</sup>.

## 2 Can NGS with FFPE samples match the data quality of NGS with fresh-frozen samples?

Yes, according to a preponderance of the current literature. In fact, the power of NGS to analyze large numbers of short sequences might make it an ideal technology to apply to fragmented nucleic acids extracted from FFPE samples<sup>6</sup>.

In one recent study, for example, researchers performed whole exome sequencing (WES) from gastrointestinal stromal tumors extracted from either fresh-frozen (FF) or FFPE samples. The integrity of FFPE DNA was evaluated by a modified RAPD PCR method to help classify samples as high- or low-quality (HQ/LQ). DNA library production and exome capture were feasible for both classes of FFPE, despite the smaller yield and insert size of LQ-FFPE. WES yielded data of equal quality from FF and FFPE, with HQ-FFPE generating an amount of data comparable to FF samples<sup>7</sup>.

Other studies have revealed strong correlations between NGS data from FF and FFPE samples<sup>6,8,9,10</sup>, thereby supporting the feasibility of generating high-quality sequencing libraries and sequencing results from even the low-input DNA from FFPE tumor tissue<sup>11</sup>.

## 3 Does the FFPE process affect the yield and/or quality of DNA obtained from embedded tissues?

It can. Of course, this wasn't an issue 120 years ago, when formaldehyde (the primary component of formalin) was identified as a superior fixative for preserving tissue samples<sup>12</sup>. Pathologists and histologists have been fixing tissues with formalin for more than a century<sup>13</sup>, but it wasn't until the early 1970s that researchers discovered that formalin creates crosslinks between intracellular macromolecules such as protein and DNA<sup>14</sup>.

Another form of DNA damage induced by formalin fixation is fragmentation, which can lead to low amounts of amplifiable template for PCR amplification<sup>15</sup>. Researchers also believe long-term storage of formalin-fixed blocks can induce fragmentation due to exposure to environmental conditions<sup>16,17</sup>.

Nevertheless, numerous studies have reported the feasibility of using DNA from FFPE samples with both conventional PCR-based and NGS technologies<sup>5,18,19</sup>. Though fragmentation can be a rate-limiting factor in approaches that use longer amplicons, researchers have successfully used shorter amplicons from fragmented FFPE DNA<sup>20</sup>.

In addition, other researchers have shown that some formalin-induced modifications can be partially reversed<sup>21</sup>, a process now included in many nucleic acid extraction kits for FFPE material<sup>10</sup>. Case in point: some studies report higher instances of sequence artefacts in FFPE samples (primarily C:G>T:A base substitutions)<sup>9,22,23,24</sup> while others have seen little evidence of artefacts<sup>5,25</sup>. Regardless, C:G>T:A sequence artefacts are predominantly caused by uracil lesions, and treating FFPE DNA with uracil-DNA glycosylase prior to PCR amplification significantly reduces these artefacts without affecting true mutational sequence changes<sup>4</sup>.

Further, the FFPE preparation process, which includes long-term storage at room temperature, may generate DNA mutations and result in the identification of false single nucleotide variants (SNVs) or insertions/deletions (indels). This damage, however, appears to have a random distribution across all DNA fragments and can be corrected by choosing a high sequencing depth. Researchers therefore recommend coverage levels of at least 80× when analyzing FFPE material<sup>5</sup>.

Finally, the method used for deparaffinization can also affect the amplifiability of extracted DNA<sup>26</sup>. The higher the melting temperature used, the greater the chance of denaturing the double-stranded DNA. If the temperature is too low, however, the paraffin may not melt completely, reducing the potential nucleic acid yield. Different FFPE tissues may require different paraffin melting temperatures and times; the goal is to achieve an acceptable balance between melting temperature, DNA quality and DNA yield. The recommended maximum temperature is 90°C. Temperatures higher than that can result in a significant fraction of single-stranded DNA. Some researchers suggest that using a temperature of 75°C for five minutes — still sufficient to melt the paraffin — preserves double-stranded DNA<sup>5</sup>.

#### **4 Does the age of FFPE samples affect their ability to provide sequenceable libraries?**

In general, it does not, though success could depend on how the samples were originally fixed, the conditions under which they were stored, and the method used to create the libraries.

In a recent study, samples up to three years old yielded sequenceable libraries 94% of the time, though the success rate dropped to 50% for older samples (14–21 years)<sup>27</sup>. Other researchers have found no significant difference among macromolecules extracted from blocks stored over 11–12 years, 5–7 years, or 1–2 years when compared to current year blocks<sup>28</sup>, with some reporting successful creation of libraries from RNA isolated from 20-year-old FFPE tissues<sup>6</sup>.

#### **5 Can I really expect acceptable sequencing quality from FFPE samples that are 10 – 20 years old?**

There appears to be a negligible decrease in sequence quality from FFPE samples more than a decade old<sup>29</sup>, confirming earlier findings from a study that used 14- and 18-year-old FFPE samples<sup>25</sup>. Some credit the robustness of NGS technology for enabling molecular analyses of DNA and RNA in FFPE tissues that have been stored for up to two decades<sup>6</sup>. In any case, depending on the purpose of the analysis, nucleic acids retrieved from FFPE tissues older than 40 years can be successfully used for molecular analysis<sup>30</sup>.

#### **6 Can the quality of FFPE DNA samples affect downstream genomic applications?**

It can, though the more relevant question might be how significantly. Researchers have long known the formalin fixation process can affect downstream genomic applications due to DNA cross-linking to DNA and proteins — which can stall polymerases — as well as DNA-DNA crosslinks that can inhibit denaturation.<sup>31</sup> Much of the concern about FFPE DNA sample quality has focused on mutation screening, though many of those concerns — such as artefacts, false-negative variants and suboptimal performance of variant calling algorithms — have yet to be thoroughly investigated, so the significance of their impact remains unclear.

What is clear is that mutation detection can be impeded by DNA degradation and the presence of sequence artefacts that can be erroneously interpreted as mutations. And though the degradation challenge can be addressed by using shorter amplicons in PCR detection methods, a corroborated solution has yet to be found for the sequence artefact problem — with the possible exception of treating FFPE DNA with uracil-DNA glycosylase prior to PCR amplification<sup>4</sup>.

Despite current challenges, however, many researchers have successfully used FFPE DNA for copy number analysis and mutation detection using targeted sequencing of single genes<sup>33,34</sup>, as well as whole exome<sup>5,35</sup> and whole genome<sup>25</sup>. It has also been suggested that FFPE samples can be used successfully in place of fresh-frozen samples for gene expression studies<sup>8</sup>.

## 7 How much FFPE DNA is required for successful NGS?

The answer depends largely on the type of data you want to generate. Input amounts of  $\leq 250$  ng FFPE DNA have been reported as insufficient for adequate exome coverage<sup>35</sup>, yet in a recent study of 99 FFPE samples, the authors reported “successfully” sequencing exomes from as little as 16 ng input FFPE DNA<sup>36</sup>. Several other studies have sequenced samples from as little as 10 ng, though successful whole genome sequencing analysis from 10 ng of FFPE DNA was limited to changes in copy number only<sup>37</sup>. Moreover, as little as 5 ng of template DNA from FFPE specimens has been used to generate a library of fragments for sequence analysis, resulting in a copy number karyogram indistinguishable from a karyogram generated from 1 mg of template DNA<sup>37</sup>.

Most researchers agree that much of this sequencing success is due to the ability of NGS to deliver usable data – including single-base changes, insertions/deletions and translocations – from the relatively short fragments of DNA that can be recovered from FFPE tissues<sup>29,38</sup>.

## 8 Can I successfully extract RNA and/or miRNA from FFPE samples?

Yes. RNA-sequencing from FFPE samples can, however, be a challenge because RNA is less stable than DNA. FFPE samples might be degraded when compared to RNA samples used in other applications, but it's possible to obtain high-quality sequence reads from FFPE material for miRNA profiling<sup>39</sup>. Due to their small size, miRNAs may be less prone to degradation and modification, so their analysis in FFPE specimens likely provides a more accurate replication of what would be observed in fresh tissue than that of mRNA species.

One study that included 272 independent RNA isolations from 17 tissue types and 65 FFPE blocks indicated that miRNAs are not only suitable but are likely superior analytes for the molecular characterization of compromised archived clinical specimens<sup>40</sup>. If the aim is to detect novel miRNAs, the preferred choice of platform will probably be NGS, which can deliver the most data and requires no prior knowledge of the sequences to be identified<sup>39</sup>.

## 9 Can I use FFPE tissue to detect viral DNA from past disease outbreaks?

The jury is still out on this. Novel techniques have been developed to enable detection of known viral sequences in FFPE tissue samples, which has so far included recovery of the 1918 “Spanish” influenza A/H1N1 virus. Using RNA extracted from an FFPE lung tissue sample from a victim of that famous pandemic, the virus was characterized and then recovered through reverse genetics<sup>41</sup>. In addition, sequence-independent amplification, combined with NGS, has been used to detect novel viruses from various FFPE samples, but whether the same techniques can be applied to detection of known and unknown viruses in this type of sample is currently unclear<sup>42,43,44</sup>.

## 10 Are all bioinformatics tools for NGS of FFPE samples equally reliable?

Not yet. Even as many researchers are already anticipating the potential for third-generation sequencing technology, there's still no widely accepted gold standard method for analyzing NGS data<sup>45</sup>, nor is there an agreed-upon standard for quality control of sequence data<sup>46</sup>. And although a variety of bioinformatic analysis tools have been developed for NGS data, their reproducibility still needs to improve<sup>47</sup>.

For now, all commercially available NGS platforms have different error types/rates, and users should carefully assess and address each error type to minimize the potential impact on downstream data analysis. Applying different bioinformatics strategies can also affect NGS data analysis, which is why users must also understand the principles, advantages and limitations of those tools to help ensure maximum confidence in their results<sup>46</sup>.

## REFERENCES

- 1 Sah S, Chen L, Houghton J, et al. Functional DNA quantification guides accurate next-generation sequencing mutation detection in formalin-fixed, paraffin-embedded tumor biopsies. *Genome Med* 2013;5:77.
- 2 Blow N. Tissue preparation: tissue issues. *Nature* 2007;448:959–63.
- 3 Li P, Conley A, Zhang H, et al. Whole-Transcriptome profiling of formalin-fixed, paraffin-embedded renal cell carcinoma by RNA-seq. *BMC Genomics* 2014;15:1087.
- 4 Do H, Dobrovic A. Dramatic reduction of sequence artefacts from DNA isolated from formalin-fixed cancer biopsies by treatment with uracil- DNA glycosylase. *Oncotarget* 2012;3:546–558.
- 5 Kerick M, Isau M, Timmermann B, et al. Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. *BMC Med Genomics* 2011;4:68.
- 6 Hedegaard J, Thorsen K, Lund MK, et al. 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* 2014;9:e98187.
- 7 Astolfi A, Urbini M, Indio V, et al. Whole exome sequencing (WES) on formalin-fixed, paraffin-embedded (FFPE) tumor tissue in gastrointestinal stromal tumors (GIST). *BMC Genomics* 2015;16:892.

- 8 Graw S, Meier R, Minn K, et al. Robust gene expression and mutation analyses of RNA-sequencing of formalin-fixed diagnostic tumor samples. *Sci Rep* 2015;5:12335.
- 9 Spencer DH, Sehn JK, Abel HJ, et al. Comparison of clinical targeted next-generation sequence data from formalin-fixed and fresh-frozen tissue specimens. *J Mol Diagn* 2013;15:623-633.
- 10 Norton N, Sun Z, Asmann YW, et al. Gene expression, single nucleotide variant and fusion transcript discovery in archival material from breast tumors. *PLoS One* 2013;8:e81925.
- 11 Munchel S, Koestler DC, Bibikova M. Targeted or whole genome sequencing of formalin fixed tissue samples: Potential applications in cancer genomics. *Oncotarget* 2015;6(28):25943-61.
- 12 Blum F: Der Formaldehyde als Hartungsmittel. *Z wiss Mikr* 1893;10:314.
- 13 Dove A. Hard-core sequencing. Science/AAAS Custom Publishing Office. 2016. Available from: URL: <http://www.sciencemag.org/custom-publishing/technology-features/hard-core-sequencing>.
- 14 Feldman MY. Reactions of nucleic acids and nucleoproteins with formaldehyde. *Prog Nucleic Acid Res Mol Biol* 1973;13:1- 49.
- 15 Sikorsky JA, Primerano DA, Fenger TW, et al. DNA damage reduces Taq DNA polymerase fidelity and PCR amplification efficiency. *Biochem Biophys Res Commun* 2007;355(2):431-437.
- 16 Bruskov VI, Malakhova LV, Masalimov ZK, et al. Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA. *Nucleic Acids Res* 2002;30(6):1354-1363.
- 17 Pfeifer GP, You YH, Besaratinia A. Mutations induced by ultraviolet light. *Mutat Res* 2005;571(1-2):19-31.
- 18 Wagle N, Berger MF, Davis MJ, et al. High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer Discov* 2012;2:82-93.
- 19 Yost SE, Smith EN, Schwab RB, et al. Identification of high-confidence somatic mutations in whole genome sequence of formalin-fixed breast cancer specimens. *Nucleic Acids Res* 2012;40(14):e107.
- 20 Wong SQ, Li J, Tan AY, et al. Sequence artefacts in a prospective series of formalin-fixed tumours tested for mutations in hotspot regions by massively parallel
- 21 Masuda N, Ohnishi T, Kawamoto S, et al. Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acids Res* 1999;27:4436-4443.
- 22 Do H, Wong SQ, Li J, Dobrovic A. Reducing sequence artifacts in amplicon-based massively parallel sequencing of formalin-fixed paraffin-embedded DNA by enzymatic depletion of uracil-containing templates. *Clin Chem* 2013;59:1376 - 83.
- 23 Anaka M, Hudson C, Lo PH, et al. Intratumoral genetic heterogeneity in metastatic melanoma is accompanied by variation in malignant behaviors. *BMC Med Genomics* 2013;6:40.
- 24 Do H, Dobrovic A. Limited copy number-high resolution melting (LCN-HRM) enables the detection and identification by sequencing of low level mutations in cancer biopsies. *Molec Cancer* 2009;8:82.
- 25 Schweiger MR, Kerick M, Timmermann B, et al. Genome-wide massively parallel sequencing of formaldehyde fixed-paraffin embedded (FFPE) tumor tissues for copy-number- and mutation-analysis. *PLoS ONE* 2009;4:e5548.
- 26 Sengüven B, Baris E, Oygur T, et al. Comparison of methods for the extraction of DNA from formalin-fixed, paraffin-embedded archival tissues. *Int J Med Sci* 2014;11(5):494-99.
- 27 Bolognesi C, Forcato C, Buson G, et al. Digital Sorting of Pure Cell Populations Enables Unambiguous Genetic Analysis of Heterogeneous Formalin-Fixed Paraffin-Embedded Tumors by Next Generation Sequencing. *Sci Rep* 2016;6:20944.
- 28 Kokkat TJ, Patel MS, McGarvey D, et al. Archived formalin-fixed paraffin-embedded (FFPE) blocks: A valuable underexploited resource for extraction of DNA, RNA, and protein. *Biopreserv and Biobank* 2013;11(2):101-06.
- 29 Corless CL, Spellman PT. Tackling formalin-fixed, paraffin-embedded tumor tissue with next-generation sequencing. *Cancer Discov* 2012;2:23-24.
- 30 Ludyga N, Grünwald B, Azimzadeh O, et al., Nucleic acids from long-term preserved FFPE tissues are suitable for downstream analyses. *Virchows Arch* 2012;460(2):131-40.
- 31 Gilbert MT, Haselkorn T, Bunce M, et al. The isolation of nucleic acids from fixed, paraffin-embedded tissues-which methods are useful when? *PLoS One* 2007;2(6):e537.
- 32 Wang M, Escudero-Ibarz L, Moody S, et al. Formalin-Fixed, Paraffin-Embedded Tissues by Fluidigm Multiplex PCR and Illumina Sequencing. *J Mol Diagn* 2015;17:521e532.
- 33 Ausch C, Buxhofer-Ausch V, Oberkanins C, et al. Sensitive detection of KRAS mutations in archived formalin-fixed paraffin-embedded tissue using mutant-enriched PCR and reverse-hybridization. *J Mol Diagn* 2009;11:508-513.
- 34 Solassol J, Ramos J, Crapez E, et al. KRAS Mutation Detection in Paired Frozen and Formalin-Fixed Paraffin-Embedded (FFPE) Colorectal Cancer Tissues. *Internat J Mol Sci* 2011;12:3191-3204.
- 35 Beltran H, Yelensky R, Frampton GM, et al. Targeted Next-generation Sequencing of Advanced Prostate Cancer Identifies Potential Therapeutic Targets and Disease Heterogeneity. *Eur Urol* 2013;63:920-26.
- 36 Van Allen EM, Wagle N, Stojanov P, et al. Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. *Nat Med* 2014;20:682-8.
- 37 Wood HM, Belvedere O, Conway C, et al. Using next-generation sequencing for high resolution multiplex analysis of copy number variation from nanogram quantities of DNA from formalin-fixed paraffin-embedded specimens. *Nucleic Acids Res* 2010;38:e151.
- 38 Duncavage EJ, Magrini V, Becker N, et al. Hybrid capture and next-generation sequencing identify viral integration sites from formalin-fixed, paraffin-embedded tissue. *J Mol Diagn* 2011;13:325-33.
- 39 Chatterjee A, Leichter AL, Fan V, et al. A cross comparison of technologies for the detection of microRNAs in clinical FFPE samples of hepatoblastoma patients. *Sci Rep* 2015;5:10438.
- 40 Doleshal M, Magotra AA, Choudhury B, et al. Evaluation and validation of total RNA extraction methods for microRNA expression analyses in formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* 2008;10(3):203-11.
- 41 Tumpey TM, Basler CF, Aguilar PV, et al. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* 2005;310:77-80.
- 42 van den Brand JM, van Leeuwen M, Schapendonk CM, et al. Metagenomic analysis of the viral flora of pine marten and European badger feces. *J Virol* 2012;86:2360-65.
- 43 Bodewes R, van de Bildt MW, Schapendonk CM, et al. Identification and characterization of a novel adenovirus in the cloacal bursa of gulls. *Virology* 2013a;440:84-88.
- 44 Bodewes R, van Run P, Schürch AC, et al. Virus characterization and discovery in formalin-fixed paraffin-embedded tissues. *J Virol Meth* 2015;214:54-59.