# Sequencing Challenging FFPE Cancer Samples to Identify Cancer Driver Genes and Novel Biomarkers

Yasser Riazalhosseini, Ph.D., leverages key features in the Nextera® Rapid Capture Exome library prep kit to preserve integrity of low-input DNA from low-quality samples.

## Introduction

Biopsies or tissue samples from cancer patients are routinely stored in formalin-fixed, paraffin-embedded (FFPE) tissue blocks. Sequencing these samples has proved challenging, because the process partially degrades and limits the amount of genomic material. Advanced approaches, such as next-generation sequencing (NGS), generate high-quality genomic data and are suitable for analyzing these valuable samples. Yet, some sequencing preparation methods further degrade the already low-quality FFPE sample DNA.

Sequencing challenging samples is part of the job for Yasser Riazalhosseini, Ph.D., an Assistant Professor in the Department of Human Genetics at McGill University and principal investigator at McGill University and Genome Quebec Innovation. He studies unique forms of cancer to translate findings from basic science into clinical application.

Until recently, he had a backlog of stored samples yet to be analyzed due to the difficulty of working with the FFPE format and limited amount of DNA. Dr. Riazalhosseini is now accessing these samples by leveraging key features in the Nextera® Rapid Capture Exome (NRCE) library prep kit to perform exome sequencing on a HiSeq® system. Working with at least 50 ng DNA, Dr. Riazalhosseini can preserve DNA integrity and analyze disease-causing mutations and prognostic signatures.

iCommunity spoke with Dr. Riazalhosseini to learn more about his research and the benefits of using the NRCE library prep kit for exome sequencing of low-quality samples.

**Q: What are the goals of your research?**
**Yasser Riazalhosseini (YR):** My research involves generating and analyzing large cancer genomic data sets to understand molecular mechanisms involved in carcinogenesis and progression, and to identify molecules with potential clinical applications. In my lab, we rely on the generation and analysis of high-throughput genomic sequencing and transcriptome data.

**Q: What features of the NRCE kit have had the greatest impact on your research?**
**YR:** We've been very successful using the NRCE Kit. Its main advantage is that it can accommodate small DNA samples. It opens the door for us to study samples, especially biopsy



**Yasser Riazalhosseini, Ph.D.,** is an Assistant Professor in the Department of Human Genetics at McGill University, and principal investigator at the McGill University and Genome Quebec Innovation Center in Montreal.

samples, that are well-characterized clinically, but that have limited DNA. With the NRCE Kit, almost all tumor types—including ones stored as fresh-frozen or in FFPE format—can potentially be characterized. I'm using the Nextera kit for all my projects and in my collaborations at the moment, two of which have already come to publication.

**Q: What was the first study that you published using the NRCE kit?**
**YR:** The first study, involving characterization of small cell carcinoma of the ovary hypercalcemic type (SCCOHT), was led by Dr. William Foulkes at McGill University and published in *Nature Genetics* earlier this year.[1] SCCOHT are undifferentiated tumors in the ovary, usually occurring in women before the age of 40. We obtained tumor samples from colleagues in Europe, Australia, and Canada who never had a chance to look at the genomes of these samples. They provided FFPE samples, not fresh-frozen samples, which makes it difficult to analyze with NGS. We used Nextera technology to generate the exome libraries from DNA isolated from these FFPE samples and sequenced them on a HiSeq system.

**Q: What were the findings from the ovarian cancer study?**
**YR:** We identified deleterious mutations in a gene called *SMARCA4*. We were very surprised to see that almost all patients had a deficiency in *SMARCA4*. It was validated by analyzing the gene product at the protein level in tissue samples from the same patients. For the first time, we could characterize these types of tumors and begin to understand the biology behind this cancer.

Q: What type of cancer was the focus of your second study?

YR: It was a breast cancer study led by Dr. Michael Hallett at McGill University. We collaborated on this project and a paper reporting results of this study has been published in the journal *Cell Reports.*[2] The aim of the study was to validate the prognostic signatures resulting from the analysis of thousands of breast cancer samples. It's known in the literature that breast cancer heterogeneity influences the prognostic markers that are generalized to a patient. As a result, further characterization is required before extending the validity of a prognostic signature to all patients.

> "With the NRCE Kit, we generate libraries using low DNA amounts that result in high-quality sequencing."

Q: What were the findings of the breast cancer study?

YR: The team compared the prognostic signatures across a large collection of patients. Since some patients were almost always mispredicted, they hypothesized that such patients— especially those that ultimately experienced a recurrence— might harbor a high degree of tumoral heterogeneity. To validate this observation, we sequenced exomes of these heterogeneous breast tumors. Using the NRCE Kit we were able to sequence exomes on the HiSeq System and validate the team's hypothesis.

Q: How does the ability of the NRCE Kit to accommodate low-input DNA enable your studies?

YR: The NCRE Kit accepts 50 ng of DNA to generate an exome library for sequencing. This is critical, because it enables us to work with unique tumor samples, particularly biopsies from patients with specific conditions. We could not analyze the exomes of many samples before because of the minute amount of DNA that we had, and the absence of a technology that could generate an NGS library from such low DNA amounts. With the NRCE Kit, we generate libraries using low DNA amounts that result in high-quality sequencing. We've tried it on different samples and continue to be successful.

Q: How is the tagmentation step in the NRCE protocol beneficial to your research?

YR: Instead of shearing the DNA, or chopping it into shorter fragments to generate the library, the Nextera kit uses transposase enzymes to fragment the library. This is critical because when we use low quality FFPE samples, the DNA is already quite degraded. The shearing just destroys the DNA further, resulting in poor libraries. By leveraging the Nextera tagmentation step, we can work with FFPE samples and generate high-quality libraries.

Q: What did you learn working with the transposase enzymes?

YR: The enzymatic tagmentation is sensitive to the amount of input DNA. The size of the final library is determined by the ratio of DNA to transposase enzymes. The enzyme can also behave differently in different environments. I took advantage of this by changing the amount of input material to produce libraries with different characteristics, such as the insert size, which is important for sequencing.

Q: How do you measure assay performance?

YR: We check the quality of sequencing by looking at the number of reads and the amount of data generated, performing further analyses after alignment to measure the specificity of capture for the targeted exome. We target an average 70× depth of coverage for whole-exome sequencing. We can achieve this by sequencing three exomes on one lane of the HiSeq System. The Nextera protocol is sensitive in handling. We found that when we follow the library generation steps carefully, we get homogeneous results for all samples.

Q: How long did it take you to implement the NRCE Kit into your workflow?

YR: My lab collaborated with colleagues at the NGS unit of McGill University and the Genome Quebec Innovation Center to evaluate the beta version of the NRCE Kit. The library preparation protocol is not complicated. Due to the initial enzymatic step, it's a sensitive protocol in comparison to others.

Following the initial tests, McGill University's NGS unit established the NRCE pipeline quickly, and they are optimizing the procedures on robots for high-throughput use of the kit. My group examined different methods for DNA isolation from FFPE samples to identify a protocol that induces minimum damage to DNA during isolation, while not interfering with the library generation steps. With that process in place, it takes a matter of days to generate the libraries.

> "Using the NRCE kit, we are conducting similar studies on other less-studied cancers to understand the underlying genetic aberrations."

Q: Did the NRCE Kit support your ability to generate cancer study results that translate clinically?

YR: I think the best example is the SCCOHT study. In this cancer study, *SMARCA4* gene was recurrently mutated in all the patients included in the study. It demonstrates that this type of cancer has one main driver. SCCOHT is difficult to diagnose pathologically. The ability to identify the driver gene

in the patient genome has implications for diagnosis and clinical management in the future. Using the NRCE kit, we are conducting similar studies on other less-studied cancers to understand the underlying genetic aberrations.

## References

1. Witkowski L, Carrot-Zhang J, Albrecht S, et al. Germline and somatic *SMARCA4* mutations characterize small cell carcinoma of the ovary, hypercalcemic type. Nature Genetics 2014; 46(5): 438-43.

2. Tofigh A, Suderman M, Paquet E, et al. The prognostic ease and difficulty of invasive breast carcinoma. Cell Reports 2014. DOI: http://dx.doi.org/10.1016/j.celrep.2014.08.073.

**FOR RESEARCH USE ONLY.**

illumina®