

Functional DNA Quality Analysis Improves the Accuracy of Next Generation Sequencing from Clinical Specimens

OVERVIEW

We have developed a novel QC assay, the SuraSeq™ DNA Quantitative Functional Index (QFI™) assay. This assay quantifies sample DNA quality by determining the number of “functional” DNA templates that are available for amplification in a sample prior to next generation sequencing (NGS). This rigorous sample characterization helps to overcome the challenge of generating high quality sequencing data from DNA of variable quality, such as DNA extracted from formalin-fixed paraffin-embedded (FFPE) specimens.

CONCLUSIONS

Our findings highlight the influence of DNA quality on NGS results, and demonstrate how sample-specific, data-driven metrics can accommodate the analysis of low quality DNA for diagnostic applications. Based on the results of this study, we recommend routine characterization of DNA extracted from clinical samples prior to PCR-based target enrichment and inclusion of sample quality metrics to guide the interpretation of NGS data. This quality control step is essential when sequencing poor quality DNA samples such as that typically extracted from FFPE specimens. The SuraSeq™ DNA QFI™ assay is a rapid, quantitative, and high throughput methodology that can be used to:

- Predict the performance of FFPE DNA for any PCR-based enrichment procedure,
- Guide changes in the DNA input to ensure a minimum number of PCR-amplifiable templates are available for enrichment,
- Assure accurate and reliable variant calling by integrating the QC information with NGS data analysis, and
- Reduces high false positive variant calls of commercial PCR enrichment cancer panels.

INTRODUCTION

It is estimated that hundreds of millions of specimens have been archived and preserved as formalin-fixed, paraffin-embedded (FFPE) procedures tissue blocks. Although methods have been developed for nucleic acid extraction from FFPE samples, and kits are commercially available, obtaining reliable molecular information from FFPE samples is complicated because the DNA is typically compromised (see Inset Box, page 3). The FFPE process causes fragmentation and chemical modifications in DNA such as cross-linking, deamination and adducts, which can strongly compromise the results of PCR-based assays¹⁻³. This is especially problematic when sequencing FFPE DNA samples based on target enrichment by PCR⁴⁻⁶.

The number of copies of DNA template that are needed for the detection of a target region by NGS depends on 1) the template availability in the sample, and 2) the sensitivity of the assay. However, the number of functional DNA templates is not accurately measured by conventional DNA quantification assays. In this report we present results from the development of the SuraSeq™ DNA QFI™ assay, an important pre-analytical component of the SuraSeq™ NGS workflow (Figure 1) for functional DNA quantification that assures accurate sequencing and variant calling of clinical specimens with poor quality DNA.

STUDY DESIGN

We evaluated how quantifying the amount of functional DNA template and adjusting the template input impacted the accuracy of NGS variant detection.

FFPE and Fine-Needle Aspiration (FNA) Specimens and Cell Lines

DNA was isolated from a set of 208 FFPE specimens and 50 FNA thyroid biopsies. FFPE specimens included six tumor types including colon, lung, skin, ovary, breast and thyroid. Samples were acquired in accordance with appropriate human subjects' regulations.

DNA Isolation and Quantification

DNA for all samples was isolated using a modification of the RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE (Life Technologies, Carlsbad, CA). The amount of DNA in a sample was measured by three assays: spectrophotometry, a fluorescent dye assay (Qubit®, Life Technologies), and the

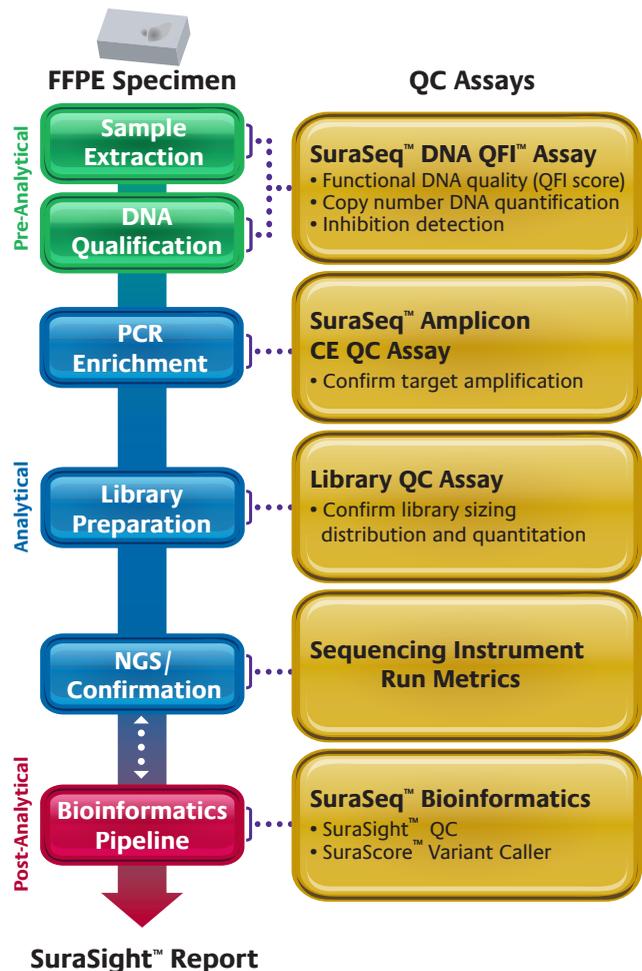


Figure 1. QC assays are critical for an accurate targeted NGS workflow. QC processes (Gold boxes) are shown for each step in the workflow (Green, blue, and red boxes). Asuragen's SuraSeq™ DNA QFI™ assay is performed before PCR amplification and NGS. Results from the assay are incorporated into the bioinformatics analysis to improve accuracy of variant calls.

SuraSeq™ DNA QFI™ assay. The SuraSeq™ DNA QFI™ assay is a novel quantitative PCR assay that measures the absolute copy number of amplifiable DNA templates. The assay amplifies an endogenous gene using an amplicon size that is matched to the median length of the PCR enrichment assay. DNA was quantitated spectrophotometrically at A260 with a NanoDrop Spectrophotometer (ThermoScientific, Wilmington, DE) and normalized to 10 ng/μL. Samples were diluted to 5 ng and evaluated by the fluorescence based Qubit® 2.0 assay and SuraSeq™ DNA QFI™ assay. In addition, PCR inhibition was assessed using a functional inhibition assay described by Nolan et al.^{6, 7}.

NGS Workflow

Target-specific PCR-enrichment was performed using Asuragen's SuraSeq™ Cancer Panels⁸ which enable sequencing of mutational hotspot regions in oncogenes for sensitive DNA variant detection in clinical specimens. The Panels provide reliable, uniform, and high depth sequencing across gene regions representing known cancer-associated mutations and are compatible with FFPE specimens.

Sequencing was then performed on either the MiSeq Personal Sequencer (Illumina, San Diego, CA) or the Ion Torrent PGM (Life Technologies, Carlsbad, CA). Samples were prepared for MiSeq analysis using the method described by Hadd et al.⁸ and processed on-board the instrument for cBot library prep and sequencing. For Ion Torrent sequencing, ~100-150 million amplicons were used for Ion OneTouch (Life Technologies) automated emulsion PCR and sequencing was performed on 318 chips.

Libraries were also prepared using the Ion AmpliSeq[®] Cancer Panel (Life Technologies), barcoded, and then sequenced on the Ion Torrent PGM, per the manufacturer's instructions.

Bioinformatics

Sequencing data from samples processed with the Ion AmpliSeq[®] Cancer Panels were pre-processed using the Torrent Suite v3.2.1 (Life Technologies), and aligned to the reference human genome sequence hg19. Variant calling was performed using the PGM plugin v3.2.4. Data from samples processed using Asuragen's SuraSeq™ Cancer Panels were analyzed using SuraScore™, Asuragen's proprietary bioinformatics pipeline and variant caller.

DNA extracted from FFPE specimens is poor quality

- 65% of the 165 FFPE samples (Figure 2) had a QFI score of less than 6.6%.
- The median FFPE DNA template quality (10.4% QFI), was 10x lower than cell line DNA quality. Data compiled from 480 clinical FFPE specimens (Figure 5).

RESULTS

Poor DNA quality and limited DNA availability can be a challenge for accurate variant calling in FFPE specimens. We addressed the problem of DNA quality by determining if the number of functional DNA templates was directly related to the accuracy of variant calling by NGS. We developed a novel analytical tool, the SuraSeq™ DNA QFI™ assay to quantify the number of functional copies of DNA template that are available for PCR amplification prior to sequence enrichment and library preparation.

What is the optimal method to quantify FFPE DNA? Spectrophotometry is typically used to measure DNA concentrations; however, because it measures DNA absorbance, it does not reflect DNA template functionality or contaminants in the sample which can often elevate concentration measurements. We compared spectrophotometry, fluorescence, and qPCR measurements of DNA concentration. Both the QFI and the fluorescence-based assays reported DNA concentrations that were ~15 times lower on average than DNA concentrations measured by spectrophotometry (data not shown). However, the two assays differed significantly in their ability to identify the lowest quality FFPE DNA templates (Figure 2). Although both assays can be used to identify the highest quality DNA samples, only the SuraSeq™ DNA QFI™ assay was able to assess the suitability of the lowest quality DNA for PCR amplification (less than 3% QFI⁶).

How do you accurately determine the number of functional DNA templates that are compatible with PCR amplification? The SuraSeq™ DNA QFI™ assay was initially validated with an independent set of 43 FFPE specimens⁶. The QFI score for a sample is determined using this assay and is a measure of "functionality" of the template in downstream functional assays. The QFI is the percent of DNA templates that are competent for amplification compared to the amount of total DNA in a sample (determined spectrophotometrically). The SuraSeq™ DNA QFI™ assay had a limit of detection of 5 copies and the limit of quantification of 10 copies (<0.5 SD in Cq variability)⁶.

If you know the concentration of DNA and the quality of DNA (QFI), how do you calculate the minimum amount of input DNA into NGS? Examples of the minimum number of templates for optimal NGS quantification are

Table 1. Relationship between QFI quality score and minimum amplification DNA copy numbers for accurate NGS variant quantification. QFI assay quantifies as few as 10 DNA copies. *Calculations are based on the requirement for a minimum of 10 template copies of a variant for reliable quantification.

QFI (%)	Number of Amplifiable Templates in 10 ng DNA (3030 templates)	At 10 ng DNA: Variant must be represented at a minimum % for accurate NGS*	or	Adjust DNA input: Minimum DNA input needed to detect a 5% variant
6.6 =	200	5%		10 ng
3.3 =	100	10%		20 ng
1.2 =	36	28%		56 ng

shown in Table 1. For example, if at least 10 ng of DNA is required to detect a 5% mutant in a sample with a 6.6% QFI, then twice as much DNA (and thus twice the functional template copy number) is needed to detect a 5% mutant in a sample with a 3.3% QFI. The ability to adjust DNA inputs or to determine the minimum percent at which a variant can be accurately identified is directly applicable to NGS testing of clinical FFPE specimens. We determined the QFI for 165 FFPE clinical samples (Figure 2). For samples intended for AmpliSeq enrichment, 65% had a score of less than 6.6% QFI, and nearly half (79/165) had less than 3% QFI. The large fraction of samples that had poor quality DNA highlights the importance of assessing DNA quality to ensure accurate downstream NGS interpretation. Among the lowest quality samples, nearly twice as many FFPE DNA samples were suitable for enrichment using Asuragen’s SuraSeq™500 assays due to its proprietary design.

Does adjusting the DNA input based on the quality of DNA impact the outcome of NGS? To test the impact of DNA input on mutation calling, we performed a dilution series of DNA samples from FFPE specimens for which the oncogene mutations, fraction of mutation, and QFI scores were known. The samples were enriched with the SuraSeq™ 200 cancer gene panel and sequenced on the MiSeq Personal Sequencer. The optimal number of template copies needed prior to sequencing was determined solely from the QFI score. The results demonstrated 100% reliable and accurate calling of BRAF and PIK3CA mutations when the input copy number was at or above the minimum needed template copy number. Both false negatives and erroneously inflated variant frequencies were observed when the template numbers were reduced below this threshold (data not shown⁶).

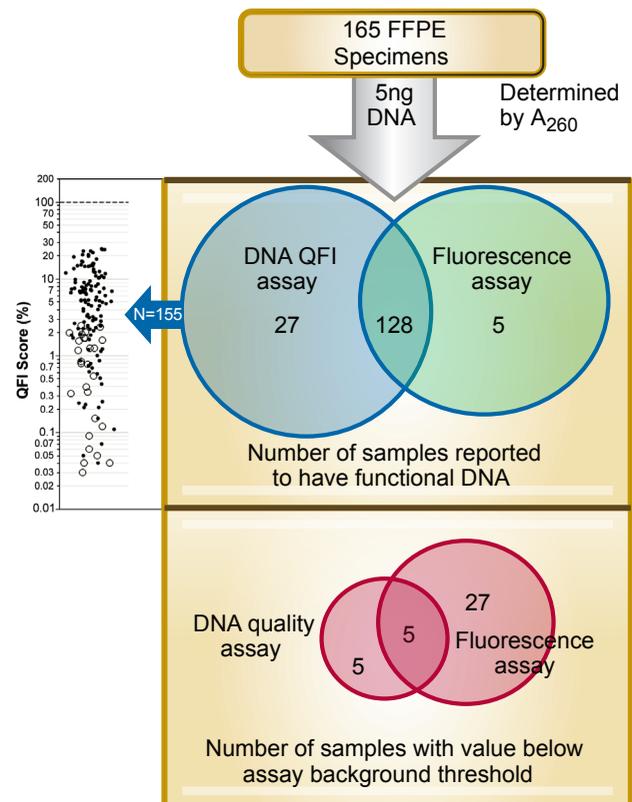


Figure 2. SuraSeq™ DNA QFI™ assay quantifies FFPE DNA samples to lower relative concentrations than a fluorescence-based assay. DNA was extracted from 165 FFPE specimens. Samples were measured spectrophotometry and 5 ng each were tested with a fluorescence dye-binding assay and the DNA QFI™ assay. Compared to the fluorescence assay, the DNA QFI™ assay detected 27 (17%) more samples with DNA templates. Most of the 27 samples that were undetected by fluorescence alone (open circles) were in the low relative DNA concentration range.

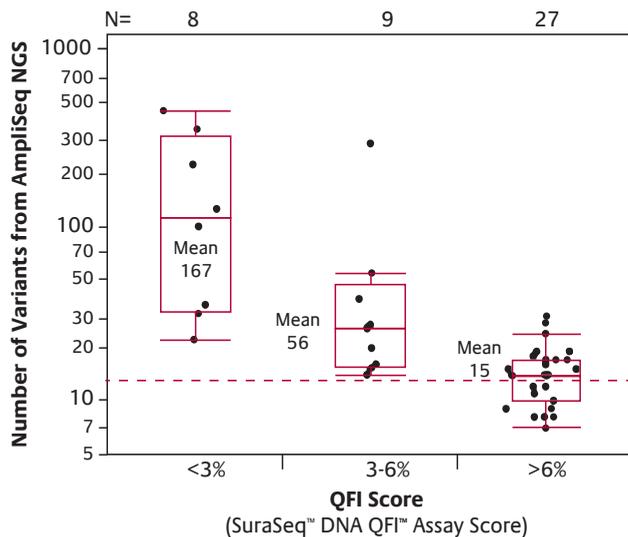


Figure 3. Samples with low QFI scores have more variants called with the Ion AmpliSeq® Cancer Panel, suggesting a highly inflated false positive call rate. Expect number of SNPs is ~13 per sample (dashed horizontal line).

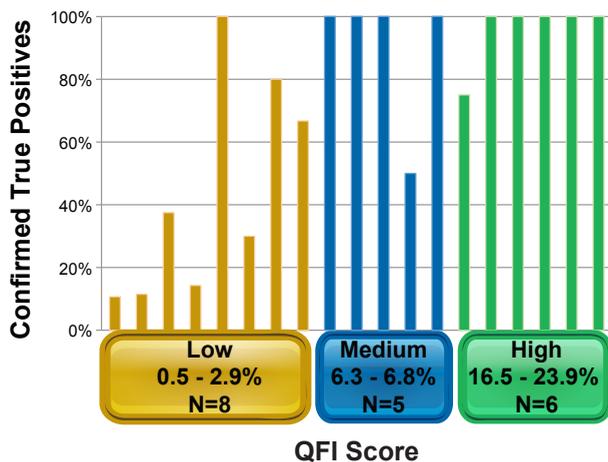


Figure 4. Mutation detection of true positives is highly correlated with FFPE DNA quality.

Does the quality of DNA impact sequencing results for a common commercial NGS enrichment method? A set of 44 FFPE DNA samples with known mutations and a range of QFI scores were prepared with the Ion AmpliSeq® Cancer Panel (a commercial, multiplexed PCR method that enriches loci in 190 amplicons from 46 cancer genes) and sequenced on the PGM. Analyses of the data revealed an inverse relationship between the QFI score and the number of variants detected (Figure 3). The mean number of variants called for samples with greater than a 6% QFI score was ~15, a number that is in good agreement with the expected number of SNP variants (Figure 3, dashed horizontal line)⁹. However, lower quality samples with

less than a 3% QFI score produced a mean of 167 variants (Figure 3). These results suggested a highly elevated false positive call rate for samples with low quality DNA.

To determine the actual false positive call rate for low quality samples, the NGS outcomes of two targeted enrichment methods, the Ion AmpliSeq® Cancer Panel or the SuraSeq™ 200 cancer panel, were compared. FFPE DNA samples with a range of QFI scores and known mutations were prepared with both methods and sequenced on the PGM. The SuraSeq™ data were analyzed with Asuragen’s bioinformatics pipeline designed to maximize positive predictive value (PPV) and sensitivity for FFPE specimens. Only variants that were covered by both panels were analyzed. The AmpliSeq® data was filtered to include only variants with more than 5% mean frequency and removal of systematic PIK3CA false positives. Less than half of the detected variants could be confirmed for samples with less than 3% QFI (Figure 4, orange bars). In contrast, 90% (blue bars) and 96% (green bars) of variants were confirmed with samples that had a QFI of 6-7% and at least 16.5% QFI, respectively. As a result, a threshold of 3-6% QFI was identified to support accurate in variant calling in clinical samples, consistent with theoretical calculations⁹.

What percent of my FFPE DNA samples will be suitable for NGS? The percent of FFPE samples that can be analyzed will be based on the quality of the samples and the need to balance false positive and false negative results. In our experience with analysis of FFPE DNA samples that have been collected from different organizations and extracted with different extraction protocols, we find a broad range of DNA

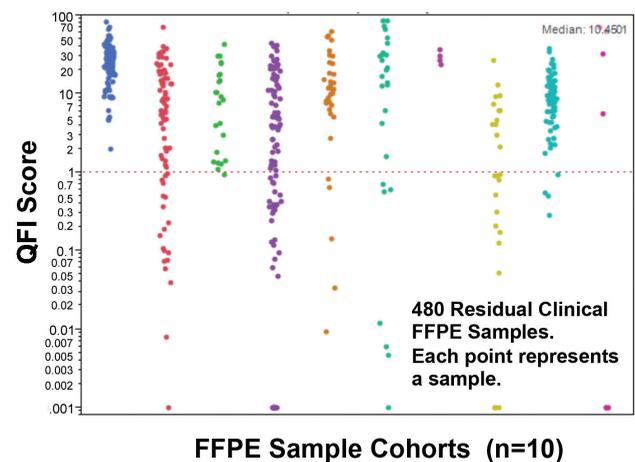


Figure 5. FFPE DNA samples have a broad range of DNA quality as measured by the DNA QFI™ assay. The median QFI was 10.4%, which was 10 fold lower than cell line DNA quality (not shown).

quality. Figure 5 shows the QFI scores from 480 FFPE DNA samples. The range of QFI scores varies widely between cohorts. The SuraSeq™ workflow is optimized for samples with low quality DNA, and 82% of the 480 FFPE DNA samples met the research criteria for analysis.

Is the QFI score useful for other types of clinical specimens? DNA samples were extracted from 50 FNA specimens, enriched for NGS with SuraSeq™ cancer panels, and sequenced on the PGM. Eight of the 50 samples were excluded from analysis: 5 of the excluded samples had QFI scores of less than 3%, and 3 of the excluded samples failed to meet SuraSight™ Reporting Tool QC criteria for bioinformatics analysis. Common variants were confirmed by a liquid bead array and/or Sanger sequencing. Overall, FNA DNA samples with QFI scores above threshold demonstrated 100% sensitivity and 94% specificity with orthogonal assays (Figure 6).

Concordance table for 42 FNA samples		SuraSeq™ NGS		
		Positive	Negative	Total
Orthogonal Mutation Assays	Positive	25	0	25
	Negative	1*	16	17
	Total	26	16	42

Figure 6. Concordance between NGS and orthogonal methods was obtained for FNA tumor biopsy specimens.

CONCLUSIONS

This report illustrates the importance of a pre-analytical QC test for DNA quantification to improve NGS outcomes. The

SuraSeq™ DNA QFI™ assay can be used to calculate the amount of DNA input required to optimize NGS results by measuring the percentage of DNA templates that are competent for PCR amplification. Many samples with low quality DNA can be “rescued” by adjusting the DNA input as guided by the assay results. In addition to targeted NGS, this functional evaluation would also be valuable for other assays that rely on PCR amplification.

Importantly, we found that a common commercial method for multiplexed PCR enrichment for targeted sequencing of cancer genes identified false positives at a high rate if the functional DNA copy number was low. Our results show that the quality of a sample not only impacts the sensitivity of NGS, but also the number of and accuracy of variants that are reported. By offering a quantitative foundation to define pre-analytical sample quality, assess template complexity, and inform DNA input the SuraSeq™ DNA QFI™ assay can help guide reliable variant identification and meet the challenge of accurate NGS testing of clinical specimens with low quality DNA.

REFERENCES

1. Ben-Ezra, J., et al., J Histochem Cytochem, 1991. 39(3): p. 351-4.
2. Srinivasan, M., D. Sedmak, and S. Jewell, Am J Pathol, 2002. 161(6): p. 1961-71.
3. Williams, C., et al., Am J Pathol, 1999. 155(5): p. 1467-71.
4. Akbari, M., et al., J Mol Diagn, 2005. 7(1): p. 36-9.
5. Didelot, A., et al., Clin Chem, 2013.
6. Sah, S., et al., submitted, 2013.
7. Nolan, T., et al., Anal Biochem, 2006. 351(2): p. 308-10.
8. Hadd, A.G., et al., J Mol Diagn, 2013. 15(2): p. 234-47.
9. 1000 Genomes Project Consortium, et al., Nature, 2010. 467(7319): p. 1061-73..

About Asuragen

Asuragen is a molecular diagnostics company using genomics to drive better patient management through best-in-class clinical solutions. The company uses a breadth of technologies and talent to discover, develop, and commercialize diagnostic products and clinical testing services with efficiency and flexibility both internally and for our companion diagnostic partners. Asuragen’s products, services, and technologies drive countless patient management decisions across oncology, genetic disease and other molecular testing modalities.