GENETIC PROFILING OF FFPE SAMPLES: CHALLENGES AND CONSIDERATIONS

Alex Siebold, Ph.D.
Manager, Genomics Field Applications
USA & Canada
Diagnostics & Genomics Division
Agilent Technologies
AGENDA

1. Challenges of FFPE samples with NGS
2. Tapestation DNA Integrity Number (DIN) QC
3. Agilent’s FFPE QC kit
4. HaloPlexHS tools for FFPE Samples
AGENDA

1. Challenges of FFPE samples with NGS
2. Tapestation DNA Integrity Number (DIN) QC
3. Agilent’s FFPE QC kit
4. HaloPlexHS tools for FFPE Samples
FFPE Samples

- Formalin fixation and tissue embedding in paraffin wax is a universal approach for biopsy specimen processing prior to light microscopic evaluation.

- Major advantage: cellular and architectural morphologic detail is preserved.

- Standard formalin fixative: aqueous solution containing 37% formaldehyde and 10–15% methyl alcohol.

- Formaldehyde is a highly reactive compound that results in the formation of protein–nucleic acid, nucleic acid–nucleic acid and protein–protein crosslinks.

- FFPE samples represent the largest source of archival biological material available for large retrospective prognostic studies of human cancer.
FFPE Samples

Formalin reacts with nucleic acids to form at least four types of modifications¹:

- Formaldehyde, the active ingredient in formalin, can react with nucleic acid bases to form mono-methylol group additions (-CH2 OH). This is the most common type of damage rendered to nucleic acids through the fixation process, and it is well-known these additions can be reversed by heating nucleic acids in buffered solutions for extended periods of time.²

- N-methylol, the predominant form of formaldehyde when in solution, can undergo electrophilic attack on an amino base to form a methylene bridge between two adjacent nucleic acid bases.

- Formaldehyde treatment can generate apurinic and apyrimidinic sites via hydrolysis of the N-glycosylic bonds.

- Formaldehyde may also cause slow hydrolysis of the phosphodiester bonds leading to strand breaks.

FFPE Samples
Down stream impact on molecular analysis:

- gDNA derived from FFPE tissue is degraded due to the strand breaks and deparaffinization conditions.

- Formalin-fixed tissues exhibit a high frequency of nonreproducible sequence alteration due to cross-linking cytosine nucleotides on either strand. As a result, in PCR the Taq-DNA polymerase fails to recognize the cytosine and incorporates an adenine in the place of a guanosine, creating an artificial C-T or G-A mutation.

- Damaged DNA are known to promote jumping between templates during enzymatic amplification permitting Taq-DNA polymerase to insert an adenosine residue at the end of a template molecule, then jump to another template and continue the extension producing an artificial mutation that is subsequently amplified.

Effect of fixatives and tissue processing on the content and integrity of nucleic acids., Srinivasan M1, Sedmak D, Jewell S., Am J Pathol. 2002 Dec;161(6):1961-71
Biggest consideration in working with FFPE samples for NGS applications is **QUALITY**

- Quality of sequencing data: mappability
- Sensitivity and accuracy of variant detection: false positives/ negatives
- Allelic balance: most profound for variants of low frequency
Specific Recommendations before FFPE Library Preps

- Use an extraction method/kit specific for FFPE
- Examine the (fragment) size of the genomic DNA (gel or gDNA TapeStation). Average fragment size $\geq 500$ bp or DIN $\geq 3$ will perform best.
- Assess overall DNA quality with Nanodrop (260/280 ratio 1.75-2.0).
- Measure DNA concentration with both the Nanodrop (ssDNA+dsDNA) and Qubit (dsDNA). Ratio of Qubit/Nanodrop concentrations $\geq 0.4$
- Do not combine low quality and high quality DNA in the same target enrichment batch (sequencing bias)
Choice of Library Prep Method (FFPE)

• Transposase
  Requires two end sequences (ES) for transposases to bind. Shorter fragments in FFPE samples less likely to contain both sequences. Restriction on input ratio of DNA to transposase restricts ability to add more DNA.

• Restriction Enzyme Digestion
  Requires two separate restriction enzyme sites, but unlike transposase, the amount of DNA can be increased.

• PCR amplification
  Requires two primer sites to be present. More DNA can be added, but number of required PCR assays high (multiplexing optimization is difficult).

• Shearing
  Covaris instrument can shear DNA into fragments with proper size distribution. May require some tuning of settings to match size of input DNA.
AGENDA

1. Challenges of FFPE samples with NGS
2. Tapestation DNA Integrity Number (DIN) QC
3. Agilent’s FFPE QC kit
4. HaloPlexHS tools for FFPE Samples
Problems with FFPE Samples

Poor DNA quality

- Cross-linking between nucleic acid strands
- DNA adducts with histones or nucleic acid binding proteins
- DNA strand breaks
- Acid depurination of DNA
- Damage increases with amount of time after exposure to fixative
- Damage dependent on formalin fixing protocol used

DNA from blood or frozen tissues: High Molecular Weight

DNA from FFPE tissues: Low Molecular Weight

1.2% agarose gel, 20 ng DNA, Sybr Gold stain
Problems with FFPE Samples

**Poor DNA quality**

- Cross-linking between nucleic acid strands
- DNA adducts with histones or nucleic acid binding proteins
- DNA strand breaks
- Acid depurination of DNA
- Damage increases with amount of time after exposure to fixative
- Damage dependent on formalin fixing protocol used

**DNA from blood or frozen tissues:** High Molecular Weight

**DNA from FFPE tissues:** Low Molecular Weight

1.2% agarose gel, 20 ng DNA, Sybr Gold stain
Qualifying FFPE Derived DNA: DIN

DNA integrity number (DIN)

The new standard for genomic DNA sample QC

- Automated numerical software assessment of sample integrity.
- Independent of sample concentration and analyst.
- Full range of eukaryotic samples from degraded to intact gDNA.
- Easy interpretation of results.
- Ensures repeatability of experiments.
- Enables comparison of samples.
Qualifying FFPE Derived DNA: DIN

Automated numerical software assessment

Alert icons
Alert icons are visible in gel image and sample table. Different icons for warnings (yellow) vs. errors (red).

DIN
DIN is presented below gel image and as a column in the sample table.

Observations
Warnings and errors are displayed in the observations column of the result table.

The DNA integrity number (DIN) is an automated software tool designed to help scientists estimate the integrity of genomic DNA samples available with TapeStation Analysis Software version A.01.05. DIN functionality is compatible with Genomic DNA ScreenTape results generated in previous software releases.
Streamlining NGS QC of FFPE Samples with the TapeStation DIN 5991-5360EN
Tapestation 2200 gDNA tapes and DIN
Impact of DIN Value on SureSelect Enrichment Data

Table 1. The DIN and the sequencing quality criteria obtained for the six samples shown in Figure 3.

<table>
<thead>
<tr>
<th>Sample</th>
<th>DIN</th>
<th>On-target rate % &gt; 70 %</th>
<th>10x Coverage rate % &gt; 90 %</th>
<th>Deduplication rate %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.1</td>
<td>79.7</td>
<td>98.6</td>
<td>66.0</td>
</tr>
<tr>
<td>2</td>
<td>5.2</td>
<td>78.0</td>
<td>99.0</td>
<td>74.2</td>
</tr>
<tr>
<td>3</td>
<td>5.0</td>
<td>80.7</td>
<td>98.1</td>
<td>87.1</td>
</tr>
<tr>
<td>4</td>
<td>2.2</td>
<td>23.8</td>
<td>94.0</td>
<td>65.0</td>
</tr>
<tr>
<td>5</td>
<td>2.9</td>
<td>51.5</td>
<td>97.7</td>
<td>95.7</td>
</tr>
<tr>
<td>6</td>
<td>1.9</td>
<td>47.4</td>
<td>96.6</td>
<td>97.1</td>
</tr>
</tbody>
</table>

Streamlining NGS QC of FFPE Samples with the TapeStation DIN 5991-5360EN
Qualifying FFPE Derived DNA: DIN

Automated DNA Integrity Number (DIN)
Qualifying FFPE Derived DNA: DIN

On Target %

On Target % - Drop off below DIN of 3

For Research Use Only. Not for use in diagnostic procedures.
Qualifying FFPE Derived DNA: DIN

% Coverage at 10X

% Coverage at 10X – Drop off of DIN ~3
Qualifying FFPE Derived DNA: DIN

% Deduplication

% Deduplication – No correlation to DIN
AGENDA

1. Challenges of FFPE samples with NGS
2. Tapestation DNA Integrity Number (DIN) QC
3. Agilent’s FFPE QC kit
4. HaloPlexHS tools for FFPE Samples

For Research Use Only. Not for use in diagnostic procedures.
Library Preparation Workflow

WGS

1. DNA Shearing
2. End-Repair
3. A-tailing
4. Adapter Ligation
5. PCR Amplification

Target Enrichment

1. DNA Shearing
2. End-Repair
3. A-tailing
4. Adapter Ligation
5. PCR Amplification

Efficiency of adapter ligation
Need high %dsDNA

Concentration and distribution of amplified library
Need good template for PCR
Accurately qualify and quantify amplifiable DNA even in challenging samples

Optimized low input library prep workflow for improved complexity and target coverage

Complete cancer research solutions from sample to data
NGS FFPE QC Kit
The measure of quality, the measure of success

Product Features

- Upfront pre-qualification of samples for target enrichment
- Qualifies sample integrity and quantifies amplifiable template
- Quality score will tie in with recommendations on low input lib prep
  - Input based on amplifiable template
  - Pre-capture PCR cycles
  - Sequencing depth

For Research Use Only. Not for use in diagnostic procedures.
The Assay

- Enables BOTH quantitation and quality assessment
  - Quantitation: based on standard curve generated using amplicon A
  - Quality score: $\Delta \Delta C_q = \Delta C_{q_{REF}} - \Delta C_{q_{FFPE}}$

Reference DNA (intact gDNA)

$\Delta C_{q_{ref}} = C_{q_B} - C_{q_A} \approx 0$

FFPE DNA (fragmented DNA)

$\Delta C_{q_{sample}} = C_{q_B} - C_{q_A} > 0$

Cq: The cycle at which fluorescence from amplification exceeds the background fluorescence

Lower ddCq: higher quality
Higher ddCq: lower quality
Samples Tested

TapeStation traces

For Research Use Only. Not for use in diagnostic procedures.
The FFPE QC Assay generates sample quality scores comparable with DIN scores.

NOTE: Comprehensive Cancer Panel baits, 100X sequencing

For Research Use Only. Not for use in diagnostic procedures.
How do we get to these target metrics with FFPE samples?
Goal: Improve performance especially for lower quality samples

750ng Pre-cap PCR Yield

- Optimized recommendations
- Current workflow

NOTE: Comprehensive Cancer Panel baits, 100X sequencing
NGS TE Workflow Optimization

**Goal:** Improve performance especially for lower quality samples

**20x coverage**

![Graph showing comparison between optimized recommendations and current workflow](image)

**NOTE:** Comprehensive Cancer Panel baits, 100X sequencing
NGS TE Workflow Optimization

**Goal:** Improve performance especially for lower quality samples

**Qualify Samples**
- Enables analysis of wide range of FFPE

**Library Prep**
- Protocol adjustments to the SureSelectXT Low Input workflow
- Introduce 200ng of amplifiable template, not just dsDNA
- Optimized pre-capture PCR cycling to improve library complexity available for bait hybridization

**Sequencing**
- Recommendations on sequencing depth
- Adjustments on sequencing allocated per library based on quality to improve target coverage
AGENDA

1. Challenges of FFPE samples with NGS
2. Tapestation DNA Integrity Number (DIN) QC
3. Agilent’s FFPE QC kit
4. HaloPlexHS tools for FFPE Samples
HaloPlexHS Applications for FFPE

1. Digest DNA
   Sample is fragmented using restriction enzymes

2. Hybridize probes
   Probe library is added and hybridized to the targeted fragments making them form a Halo shape.

3. Purify and ligate targets
   Probe/Fragment hybrids are retrieved with magnetic streptavidin beads. The circular molecules are then closed by ligation

4. Amplify targeted fragments
   Only circular DNA targets are amplified. Sample barcodes are introduced. Final product is ready for sequencing

For Research Use Only. Not for use in diagnostic procedures.
Optimize Design for Enhanced FFPE compatibility

Probes are designed for both target fragment polarities
- Improves capture efficiency
- Eliminate strand-specific FFPE artifacts

One probe targets the sense and one targets the anti-sense strand
Optimize Design for Enhanced FFPE compatibility

Probes are designed for both target fragment polarities
  • Improves capture efficiency
  • Eliminate strand-specific FFPE artifacts

- FFPE artifact sequence (dCTP-adduct)
- True sample sequence (dGTP)
Both Captured strands are amplified independently: only one strand will reflect the C to T artifact.

Polymerase reads the C adduct as a T and incorporates an A rather than G in the newly synthesized strand.
Optimize Design for Enhanced FFPE compatibility

Because each strand is captured and amplified “independently”, the complimentary strand maintains fidelity to the true sample sequence, and you can filter mutation calls based on the ratio of its presence in both strands.
**Summary**

- Agilent provides tools for qualitative and quantitative analysis of genomic DNA from variety of sources.

- Genomic DNA screen tape is a quick and robust tool to analyze the intactness of the genomic DNA. The DNA Integrity Number (DIN) provides excellent quality metrics for the genomic DNA; which is very useful for NGS library preparation.

- Agilent FFPE QC kit provides accurate estimation of available functional DNA molecules in an FFPE sample, which is very important for optimum DNA input for a successful NGS experiment.

- In combination with our Design strategy and Molecular barcode HaloplexHS can effectively distinguish low frequency somatic mutation from chemically induced sequence errors in FFPE samples.
Alex Siebold, Ph.D.
Diagnostics & Genomics Division
Agilent Technologies, Inc.
alex_siebold@agilent.com