DNA enrichment from formalin-fixed, paraffin-embedded (FFPE) samples using xGen® Lockdown® Probes

Recommendations for generating robust sequencing data

Madelyn Light[1], Steve Groenewold[2], Kristina Giorda[1], Mirna Jarosz[1]
Integrated DNA Technologies: [1] 200 Chesapeake Drive, Redwood City, CA 94063, USA; [2] 1710 Commercial Park, Coralville, IA, 52241, USA

Highlights

- Recommendations for extraction of DNA from FFPE tissue blocks, DNA quality assessment, and library construction for target enrichment using xGen® Lockdown® Probes
- Guidance for increasing coverage depth and somatic variant calling sensitivity from archived tissue samples

Introduction

Cancer genomes, commonly studied using FFPE samples, are notoriously heterogeneous, with different clonal populations each containing unique driver mutations. Additionally, the percentage of tumor DNA within these samples may be small, requiring sensitive mutation detection power to enable characterization. Therefore, accurate tumor characterization requires libraries of sufficient complexity (Figure 1) to provide a balanced representation of the sample and achieve high unique sequencing coverage.

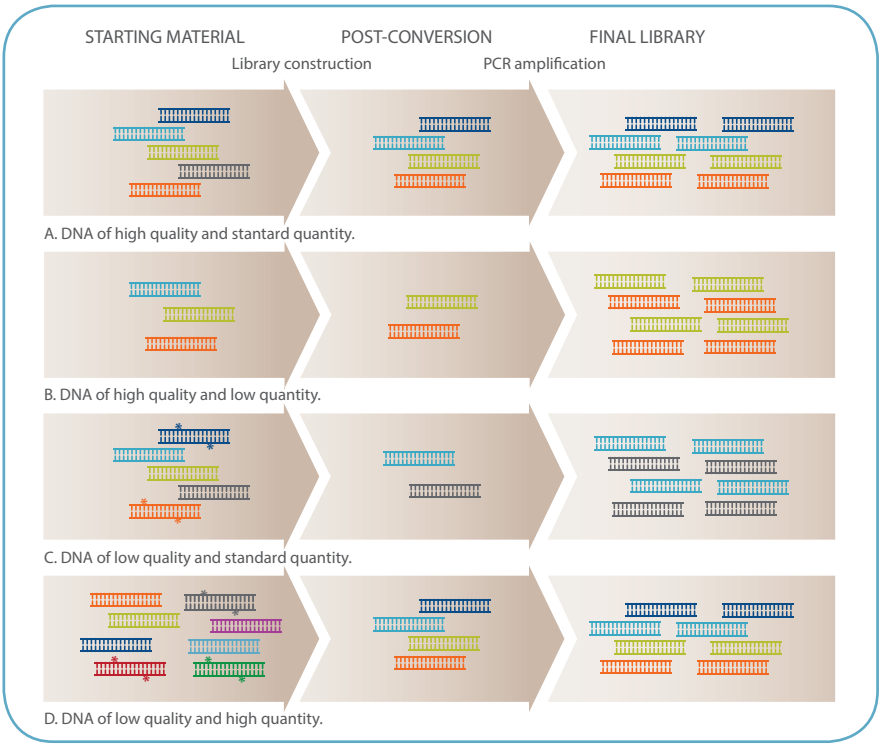

Figure 1. Effect of DNA quality on library complexity. The diversity of a sequencing library, referred to as library complexity, is dependent on the quantity and quality of the initial sample and the efficiency of library conversion. Adequate library complexity is important for achieving sensitive somatic variant detection power. Additional sequencing of low complexity libraries only increases PCR duplicate reads. **(A)** Starting with sufficient high quality DNA produces a library of high complexity that is representative of the starting sample. **(B)** An insufficient quantity of starting material produces a library of low complexity that is not completely representative of the starting sample. **(C)** A poor quality sample produces a low complexity library, because damaged DNA is not converted during library construction. **(D)** In some cases, library complexity can be improved by increasing the input amount of low quality DNA from FFPE samples.
[Asterisks indicate DNA damage.]

Sample enrichment with xGen Lockdown Probes allows deep coverage of targeted regions using minimal sequencing reads. However, due to the variable quality and quantity of FFPE DNA, it is important to properly assess and process these sample types to optimize library complexity to achieve better target coverage. This application note describes how commercially available DNA quality assessment kits can be used to estimate and maximize final library complexity from FFPE DNA.

Results

We extracted DNA from FFPE blocks representing tumors from various tissues (Asterand BioSciences). Measurement of DNA quality and total yield showed consistent performance among 4 commercially available extraction kits (Figure 2). The data show that DNA quality and yield are independent of extraction kit, suggesting that quality and yield are influenced by original fixation or state of the FFPE block. Next, we evaluated quality assessment methods for their reliability in predicting which samples would provide useable and reliable sequencing data (Figure 3; Table 1). Figure 4 shows comparable results when maximum mean coverage was measured as a function of Quality Score or DNA Integrity Number (DIN). Finally, we performed an input titration on control gDNA, as well as high and low quality FFPE (Figure 5) to demonstrate the relationship between input quantity and final complexity. Our data show that some FFPE libraries can be "rescued" by increasing the amount of starting material.

Evaluation of FFPE extraction kits

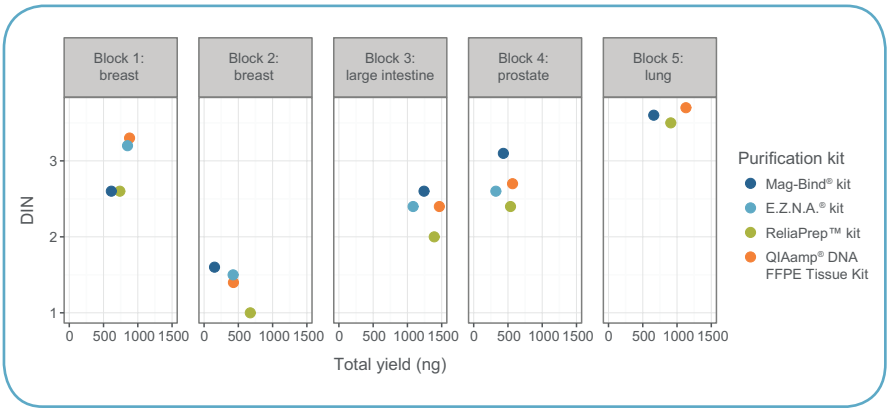

Figure 2. Extracted FFPE gDNA sample yield and quality are dependent on initial fixation. DNA was extracted from archived FFPE blocks (Asterand Bioscience) using 4 different extraction kits: Mag-Bind® FFPE DNA (Omega Bio-Tek), E.Z.N.A.® FFPE DNA Kit (Omega Bio-Tek), ReliaPrep™ FFPE gDNA MiniPrep System (Promega), and QIAamp® DNA FFPE Tissue Kit (QIAGEN). DNA extraction was performed using 10 µm sections from each block, following respective manufacturer's instructions. Total yield was assessed using the Qubit® dsDNA BR Assay Kit (Thermo Fisher Scientific) and DNA Integrity Number (DIN) was determined using the Agilent TapeStation® genomic DNA analysis tape screen.

Quality assessment and the effect on library complexity

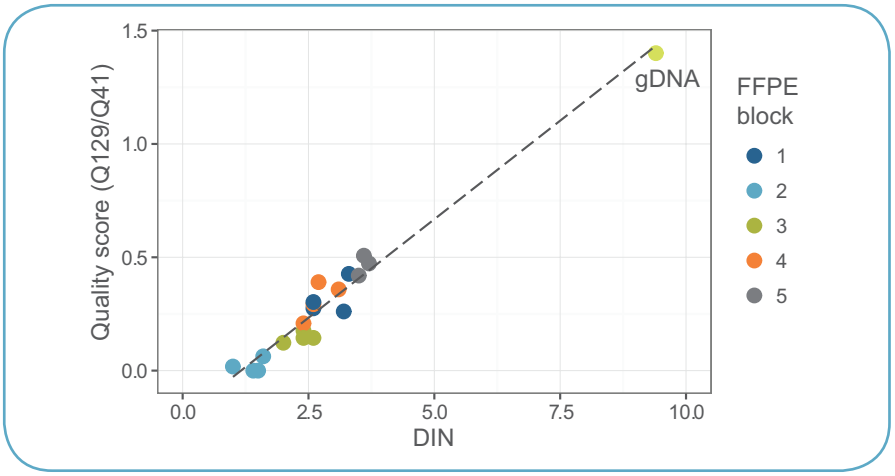

Figure 3. Quality assessment of FFPE DNA using commercially available kits. Quality scores, determined using the KAPA hgDNA Quantification and QC Kit (Kapa Biosystems), were well correlated with the DNA Integrity Number (DIN) for the extracted FFPE samples (Blocks 1–5) and a genomic DNA (gDNA) control sample (Coriell NA12878).

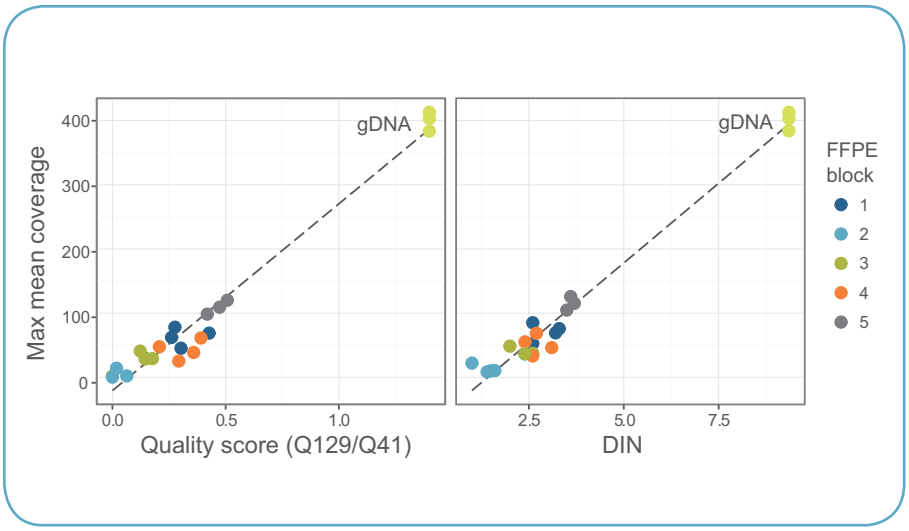

Figure 4. Relationship between quality assessment measurements and maximum mean coverage. Identical correlation is observed when maximum mean coverage is measured as a function of quality score (left) or DNA integrity number (DIN; right). For each block, sequencing data was obtained from libraries constructed from 10 ng of DNA, using the KAPA Hyper Prep Kit (Kapa Biosystems), and enriched using the xGen® AML Cancer Panel. We determined library complexity by calculating maximum mean coverage depth, the unique coverage maximum achievable for a library. Notably, we observed comparable performance when we performed library construction using the NEBNext® Ultra™ II DNA Library Prep kit (New England Biolabs; data not shown).

Effect of input DNA quantity on library complexity

Table 1. Quality qualification of DNA isolated from FFPE tissues. Quality scores (Q scores) depict DNA quality as the ratio of 129 bp vs. 41 bp amplicons. An ideal Q score is 1, indicating equivalent amplification of both amplicons and, therefore, higher DNA integrity.

Sample name	Q129/41	DIN
Low quality FFPE	<0.2	<2.5
High quality FFPE	<0.4	>3.5

Table 2. Pre-capture PCR amplification cycles performed by sample type and amount of starting material.

Sample type	Input (ng)					
	1	5	10	25	50	100
gDNA	14	12	11	8	7	6
FFPE	16	14	12	9	8	7

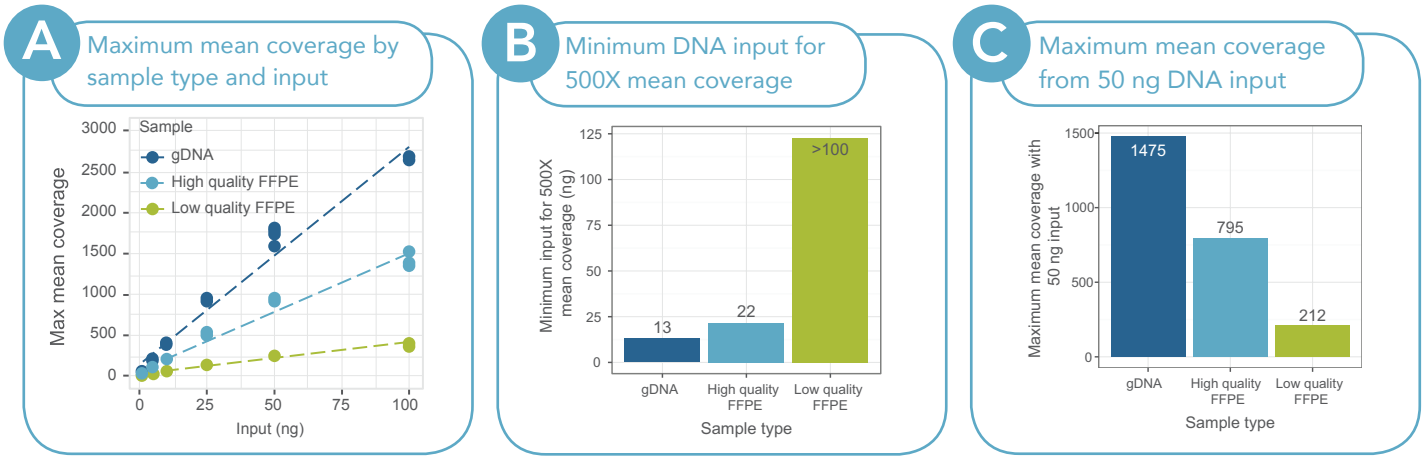

Figure 5. Variation of coverage depth with DNA input. Libraries were constructed from low and high quality FFPE DNA (as defined in Table 1) and high quality gDNA using the using the KAPA Hyper Prep Kit (Kapa Biosystems) followed by PCR amplification according to Table 2. Hybrid selection was performed on all libraries using the xGen® AML Cancer Panel and enriched libraries were sequenced on the NextSeq® System (Illumina). **(A)** Measurement of maximum mean coverage by sample type and input shows that final library complexity is heavily dependent on both input amount and sample quality (dotted lines represent linear regression). **(B)** The minimum input required to reach mean coverage of 500X was extrapolated using linear regression analysis of maximum mean coverage. **(C)** Maximum mean coverage expected with 50 ng input into library construction shows that with a fixed input mass, maximum mean coverage is influenced by sample quality.

Conclusions and final recommendations

We investigated how the quality and quantity of DNA extracted from FFPE samples affect sequencing library complexity, since these factors are of most concern when working with preserved samples. Our data demonstrate that:

1. Quality and yield of DNA extracted from FFPE samples are independent of extraction kit and more likely a factor of handling during preservation and fixation.
2. Quality score (Q score; ratio of 129 bp vs. 41 bp amplicons) and DNA integrity number (DIN) show good correlation for assessing DNA quality before library construction. Q score or DIN can be used to predict final library complexity. We found that DNA with a Q score >0.4 as measured by the KAPA hgDNA Quantification and QC Kit, or DNA integrity number >3.5 measured with the Agilent 4200 TapeStation system, was of sufficient quality to produce high complexity libraries for generating reliable sequencing data.
3. It is possible to prepare enriched libraries of adequate quality from DNA extracted from FFPE samples. However, due to the variation that occurs during handling and preserving tissue samples, it is important to perform QC on FFPE DNA to determine the minimum input required for deep coverage. Sequencing depth is dependent on DNA quality and input amount.

For more information on xGen Lockdown Probes and related products, visit **www.idtdna.com/xGen**.