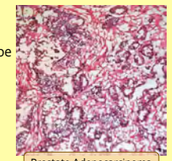
Overcoming the Challenge of Using Formalin-Fixed-Paraffin-Embedded (FFPE) Tissues for Large Scale Cancer Genome Sequencing Studies

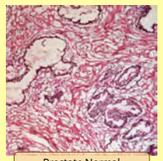
Betsy, Fernando Viloria, Clint Chalk, Brendan Blumenstiel, Matt DeFelice, Wendy Winckler, and Kristin Ardlie

Biological Samples Platform & Genetic Analysis Platform, Broad Institute, Cambridge, MA

Introduction

Cancer re-sequencing studies typically utilize DNA isolated from fresh frozen tissues. While ideal, these samples can be difficult to obtain in large numbers, especially those with high tumor percentages. Pathology departments routinely store diagnostic samples that are formalin fixed and paraffin-embedded (FFPE). Unfortunately, tissues fixed with formalin present numerous problems to researchers.


Prostate Adenocarcinoma


Prostate Normal

Formalin-Fixed-Paraffin-Embedded Tissue

FFPE tissues are tumor and normal tissues fixed in formalin (commonly known as formaldehyde) for diagnostic purposes.

- Abundancy (in hospital tissue banks) opens doorways for large scale Cancer Genome Sequencing projects.
- The fixation process poses a challenge to the recovery of nucleic acids, however, resulting in lower yields and often highly degraded DNA.
- FFPE tissues may also contain residual contaminants that may interfere with downstream molecular biology applications.

Previous Work on FFPE:

Schweiger et al. (2009) compared the performance of FFPE tissues and fresh frozen tissues:

- 5 newly-made FFPE samples
- 1 eighteen-year-old FFPE sample
- 1 fourteen-year-old FFPE sample
- Fresh frozen comparison samples

They concluded that longer ischemia &/or fixative times do not affect the sequence quality of FFPE tissues, which performed well on Illumina sequencing.

However, they found greater variation in mappable reads, a higher rate of mutations, and a lower rate of known SNPs than in fresh samples.

Broad Institute FFPE Project:

Our goal in this study is to investigate a much broader range of tissue samples that might be more representative of samples available in current U.S. and global sample banks. We obtained 161 FFPE samples from the Cooperative Human Tissue Network (CHTN) representing a range of cancer and normal tissues. 24 samples are tumor/normal paired samples .

Questions We Aim to Answer

- Is DNA derived from FFPE samples compatible with current next generation sequencing technologies (e.g. the Illumina Genome Analyzer) to detect copy number mutations and single-nucleotide-polymorphisms?
- What type of Whole Genome Amplification (WGA) methods, if any, are the most ideal to amplify the low DNA yields obtained from FFPE samples?
- What simple DNA quality control methods can best assess which FFPE samples will give successful downstream results? Or can we develop simple methods to achieve this goal?
- What proportion of banked FFPE samples might be of high enough quality to be used in cancer genome discovery projects? And what proportion might only be useful for validation studies?

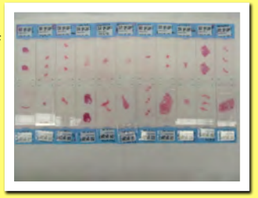
Materials and Methods

Hematoxylin & Eosin Staining

H & E staining employs hematoxylin and eosin to stain nucleic acids and the cytoplasmic parts of the cell, respectively.
- Xylene de-paraffinizes;
- Ethanol acts as a washing reagent

Steps:
1. Xylene
2. Ethanol
3. Hematoxylin
4. Eosin
5. Ethanol
6. Xylene

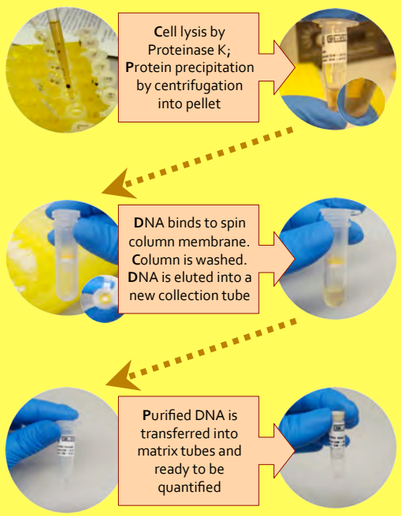
To assess the cellular makeup of a tissue sample, a glass slide of each tissue is produced and then stained (see above) to help us identify tumor-rich regions in the tissue. We prepared 95 new slides during this project.



DNA Purification via QIAGEN QIAamp/DNeasy FFPE Protocol

DNA purification was done using QIAamp MinElute columns that extract DNA free from contaminants such as proteins. Steps include:

Paraffin Removal --> Lysis --> Incubation --> DNA attachment to column membrane --> Wash --> Elution



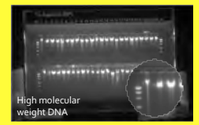
Quality Control via PicoGreen

This method quantifies DNA samples by binding double-stranded DNA fragments with flourescent PicoGreen. Varioskan, a fluorimeter, measures the absorbance of 260 nm light. DNA samples are compared with a standard curve of concentration made from a known substance.

Quality Control via Quality Gel

Materials:
- 2% Agarose Gel
- DNA Samples
- Loading Buffer
- Ladder DNA

Quality gels take advantage of the negative charge of DNA and uses a positively-charged rod to pull DNA fragments through a gel matrix. Ethidium Bromide is used to stain and visualize the DNA under UV light.

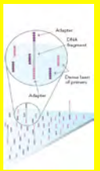


Quality Control via QC-PCR

Quality control-polymerase chain reaction (QC-PCR) assesses degraded DNA samples' ability to be amplified by polymerase chain reaction (PCR).
- 4 primers pairs designed to amplify targets of various lengths show
  1) Inhibitors and impurities that prevent amplification
  2) Overall DNA quality and extent of degradation, depending on bands that amplify

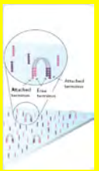
Darker bands mean more DNA of a certain length.

DNA Sequencing via Illumina

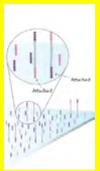


Adapter at the 5' end of a single-stranded DNA attaches to a primer on the flowcell

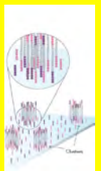
Adapter at the 3' end attaches to another primer and free nucleotides are added on



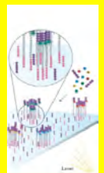
A double-stranded DNA is formed



Chemicals are flushed through the flowcell to straighten the DNA

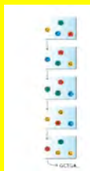


Clusters of single-stranded DNA are formed on the flowcell



Primers and labeled free nucleotides are flushed

Illumina machine takes snapshots of fluorescent sequence

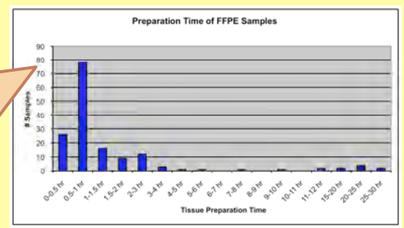


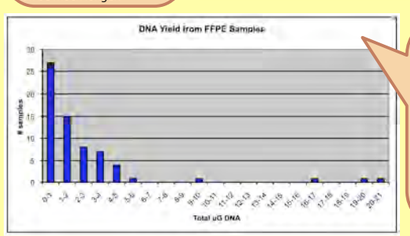
DNA sequence is recorded from continuous snapshots

[Illustrations care of www.illumina.com]

Results

- Preparation Time is measured from time of blood supply cut off to final fixation of the block:
  We found that most of the 158 FFPE samples we had data on were prepared in under 1.5 hours.
- However, there is a large range in prep times, and in cases of autopsy, prep times can be much longer.





- The average DNA yield from 66 FFPE samples we have isolated thus far is only 2.5ug.
- Overall, DNA yields (and DNA concentrations) from FFPE samples are much lower than from equivalent fresh frozen samples.
- In many cases we used the entire tissue piece to prepare the DNA.

QC results for 14 FFPE Samples

- The QC-PCR gel (upper panel) shows the variability between samples in how well they amplify – some samples amplify all 4 bands, others only the smallest band or none at all. (For Results Table, see below)
- The Quality Gel (lower panel) shows that all samples show a varying degree of degradation (compare to High Molecular Weight example gel above)

Sample ID	Gel Position	Vol (uL)	Conc ng/uL	Q-Gel size range	QC-PCR Top band	Sample Type	BSP Diagnosis	Tissue Site
SM-P3K9	3	48	11.7	100-800		Normal	Normal	Skin
SM-P3K8	5	45	24.1	<100-800	663	Normal	Normal	Kidney
SM-P3K6	7	48	35.3	400-2kb	883	Tumor	Thyroid Gland Papillary Carcinoma	Lymph Node
SM-P3KC	9	48	66.0	400-2kb	883	Tumor	Adenofibroma	Ovary
SM-P3KD	11	48	45.5	400-2kb	883	Tumor	Adenocarcinoma	Colon
SM-P3KE	13	48	4.6	Too little		Tumor	Solitary Fibrous Tumor	Lung
SM-P3KF	15	48	3.9	Too little		Normal	Normal	Thyroid
SM-P3KG	17	48	21.6	<100-800	209	Normal	Normal	Bladder
SM-P3KH	19	48	14.7	200-2kb	883	Normal	Normal	Breast
SM-P3KI	21	48	64.2	100-2kb	883	Tumor	Endometrial Endometrioid Adenocarcinoma	Uterus
SM-P3KJ	23	48	36.8	100-800	883	Tumor	Carcinoma	Lymph Node
SM-P3KK	25	48	362.7	200-400	883	Tumor	Endometrial Endometrioid Adenocarcinoma	Uterus
SM-P3KL	B_2	48	62.0	100-800	663	Normal	Normal	Uterus
SM-P3KM	B_4	48	23.2	100-400	663	Normal	Normal	Uterus

Conclusions

In this study we obtained 161 FFPE samples representing a range of samples available from standard tissue banks. We produced H&E stained slides for 95 of these (for Pathology QC), and cored and isolated DNA from all samples. For the 158 samples we had Preparation Time data on, we saw a large range in times, from 0.5hr to >25 hours. We also noted variability in the "quality" of the blocks we prepped for DNA (some were hard, others were soft and not well preserved). DNA Yields were low (in many cases the entire sample was used), and typically less than the 3ug input required for current sequencing studies, so WGA methods should be investigated. All samples were degraded, ranging in size from <100bp's up to 2Kb in some cases. Our QC-PCR showed that for many samples it is possible to amplify the largest band of a multiplex QC-PCR. However, the resulting bands are often much fainter than for non-FFPE DNA, probably due to contaminants still in the sample that inhibit PCR. Next steps will be to test the performance of these samples in sequencing (both as native and WGA samples), and to determine which QC methods correlate best predict their performance.

Acknowledgments

I would like to thank Kristin Ardlie and Fernando Viloria for being such wonderful mentors, guiding and nurturing me during my time at the Broad Institute.
I would also like to thank Megan Rokop, Kate MacSwain, and Allison Martino for granting me an internship at such an amazing facility.
Finally, I would like to thank Clint Chalk, Dojna Shkoza, Julie Ann, Paula Morais, Mark Puppo, and Erin Euler for being awesome role models, for the invaluable lessons, and for all the laughter during our time together.