The Cancer Genome Atlas

TCGA FFPE Pilot Study Progress Update

March 12th 2014

Outline

- Co-isolation of nucleic acids from FFPE
- Genomic and epigenomic characterization of analytes derived from FFPE
- Conclusions and future plans

Acknowledgements

Baylor College of Medicine Human Genome Sequencing Center
Harsha Doddapaneni
Nipun Kakkar
Liu Xi
Donna Marie Morton
Donna Marie Muzny
David Wheeler

British Columbia Cancer Agency
Andy Mungall
Andy Chu
Richard Corbett
Payal Sipahimalani

Broad Institute
Bradley Murray
Carrie Sougnez
Petar Stojanov
Michael Lawrence
Raktim Sinha

Memorial Sloan-Kettering Cancer Center
Jorge Reis-Filho
Nicholas Schultz
Cyriac Kandoth
Raymond Lim
Charlotte Ng
Marc Ladanyi

National Cancer Institute Program Office
Roy Tarnuzzer
Jean Claude Zenklusen

National Human Genome Research Institute
Carolyn Hutter
Heidi Sofia

Nationwide Children's Hospital Biospecimen Core Resource
Erik Zmuda
Natalie Bir
Jessica Frick
Jay Bowen
Julie Gastier-Foster

Oregon Health and Science University
Paul Spellman

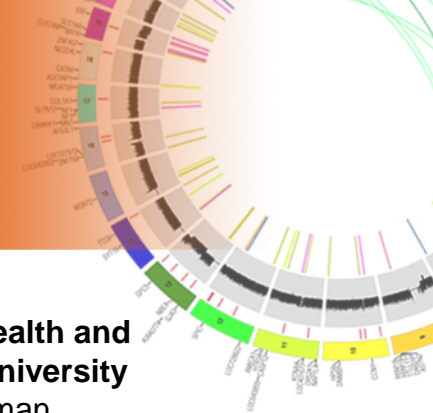**University of North Carolina School of Medicine**
Katherine Hoadley
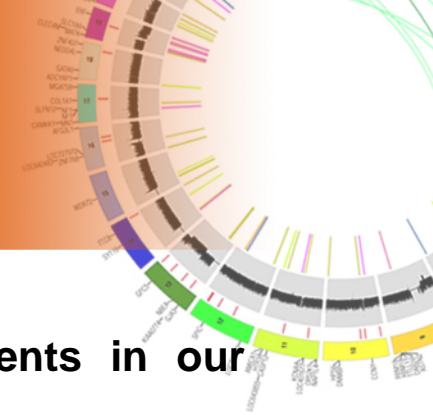Wei Zhao

University of Southern California
Dan Weisenberger
Moiz Bootwalla
Christina Curtis
Toshinori Hinoue
Peter W. Laird

Washington University
Chris Miller
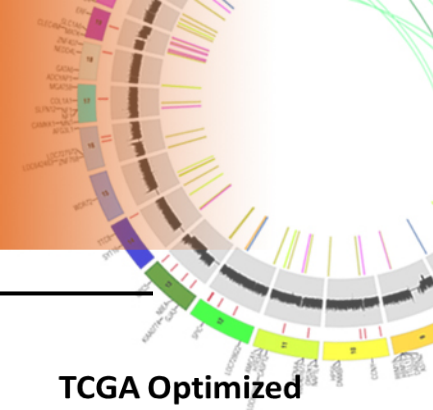Mike McLellan
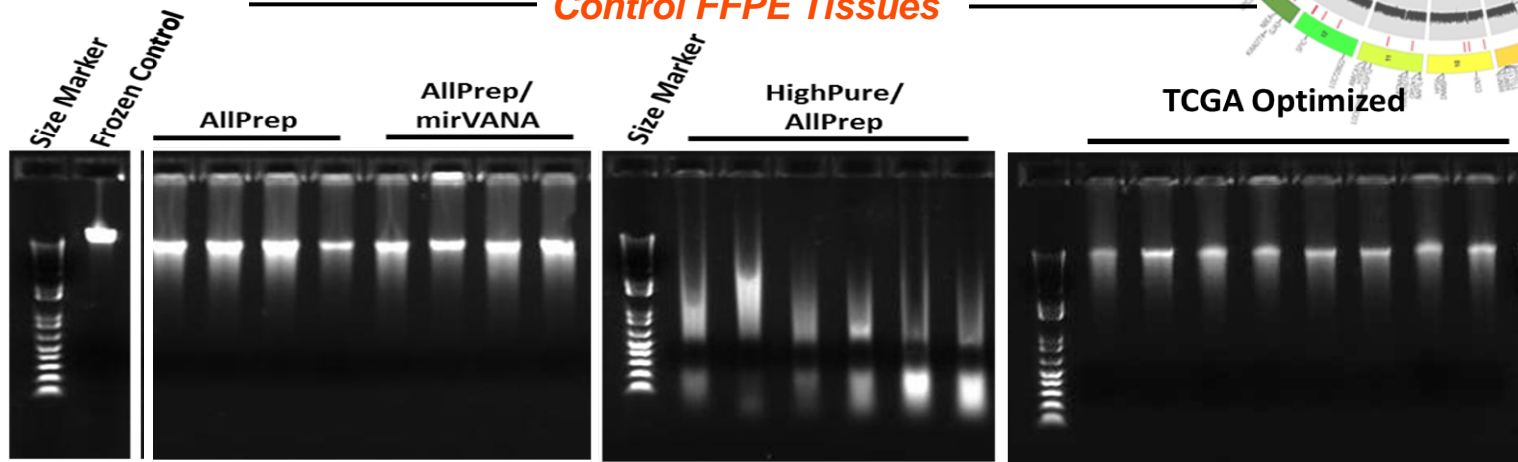Bob Fulton

The Cancer Genome Atlas

Context

- **Massively parallel sequencing has resulted in major advancements in our understanding of tumor biology**
  - new drivers of tumor progression, new therapeutic targets, and the development of a molecular-based cancer taxonomy.
  - many seminal studies drew from and were optimized for frozen tissues.

- **The concept of precision medicine involves the application of these advances to the clinical environment**
  - Challenge = diagnostic specimens are predominately formalin-fixed paraffin embedded tissues (FFPE).
  - molecular artifacts are known to be introduced by FFPE fixation.

- **Goals of the TCGA FFPE Pilot**
  - to identify and optimize best practices for the extraction, characterization and analysis of FFPE samples.
  - to define the patterns of artifactual alterations induced by formalin fixation and paraffin embedding (i.e. molecular signature of FFPE).
  - bridge the gap to diagnostic material, and facilitate application of the emerging cancer taxonomy to clinical testing environments.
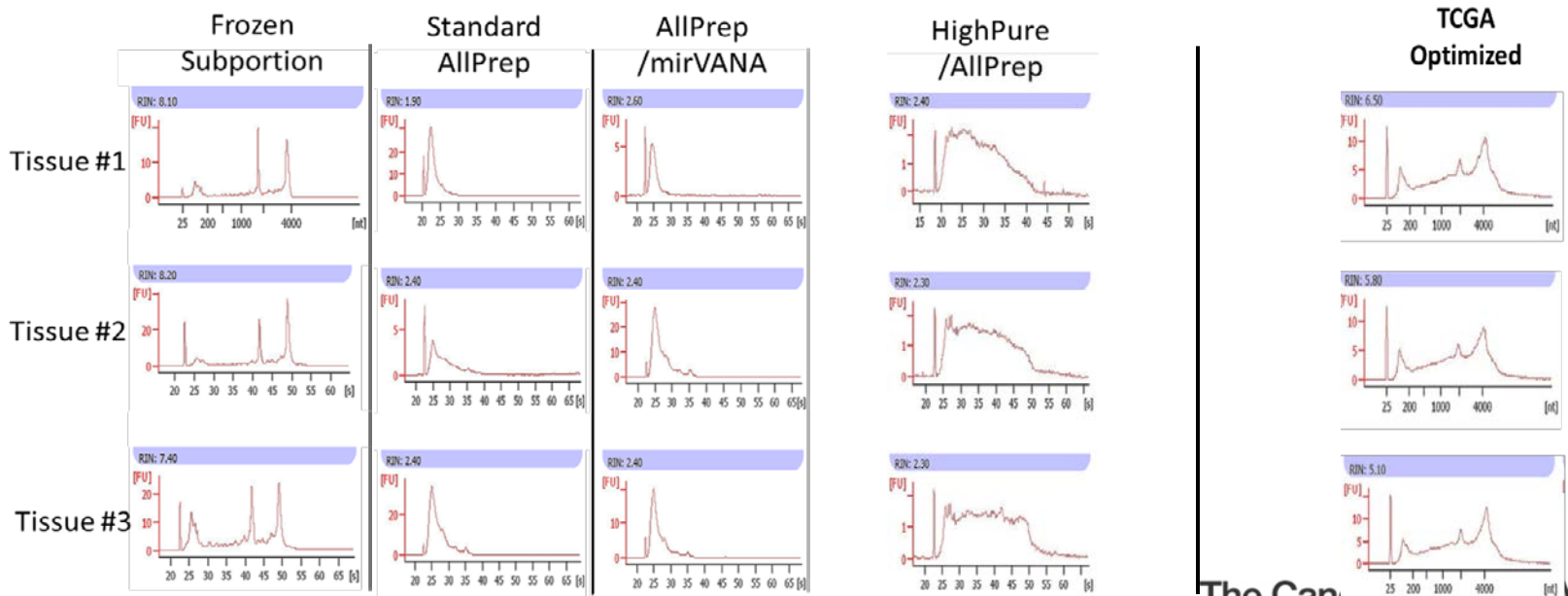
The Cancer Genome Atlas

Co-isolation of Nucleic Acids from FFPE

Control FFPE Tissues

DNA Integrity (1% agarose)

Size Marker — Frozen Control — AllPrep — AllPrep/mirVANA — Size Marker — HighPure/AllPrep — TCGA Optimized

RNA Integrity (BioAnalyzer)

Frozen Subportion — Standard AllPrep — AllPrep/mirVANA — HighPure/AllPrep — TCGA Optimized

The Cancer Genomics Atlas

4

Co-isolation of Nucleic Acids from FFPE

TCGA Frozen
Co-isolation Protocol

Tumor Portion (25-30mg)

Homogenize

All Prep DNA Column → **Tumor DNA**

Q.S. flow through to 600ul

mirVana RNA Column

Total RNA with Small RNA
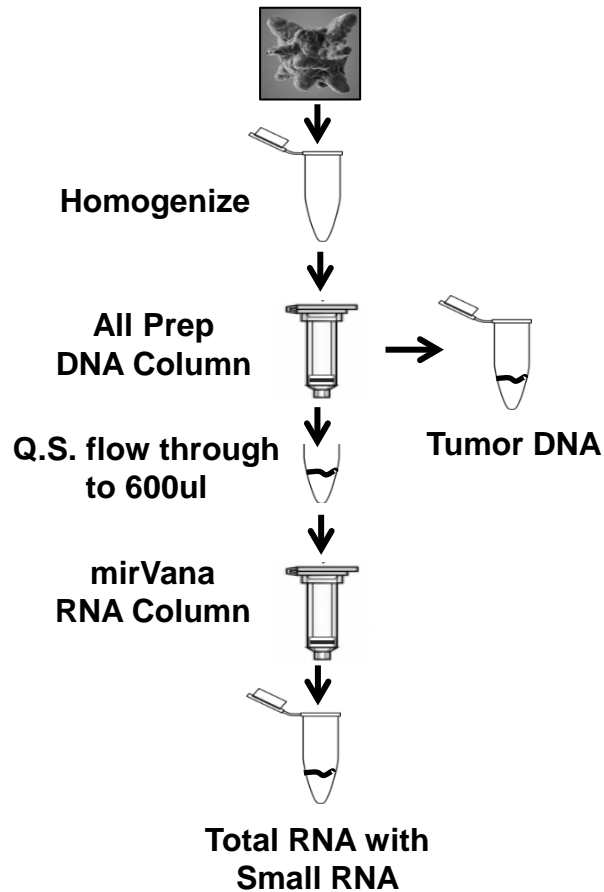
TCGA FFPE
Co-isolation Protocol

Tumor Portion ($400mm^2$ surface area)

Deparrafinize

1hr Lysis at 55 ^{0}C

Supernatant for RNA purification

Pellet for DNA purification (additional 3hr lysis)

HighPure miRNA Filter

QIAamp MinElute Spin Column

Total RNA with Small RNA

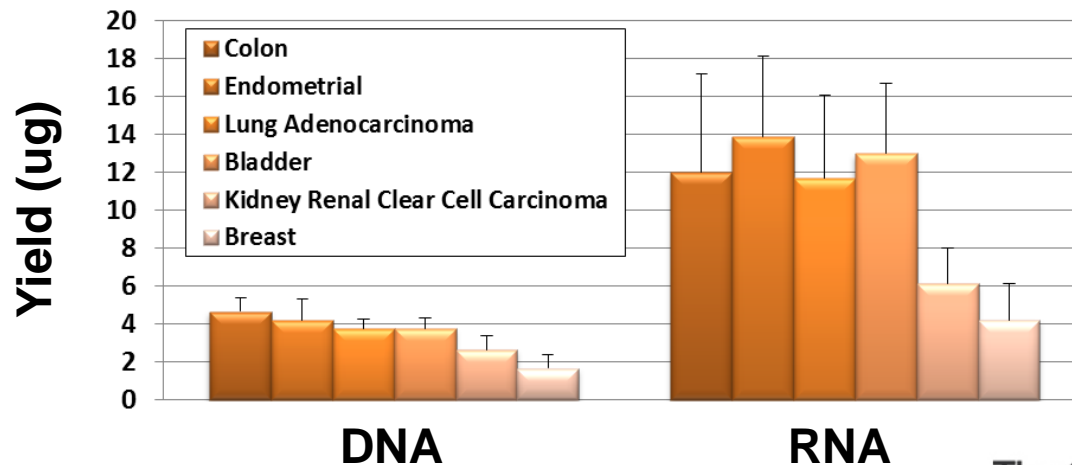**Genomic DNA**

5

The Cancer Genome Atlas
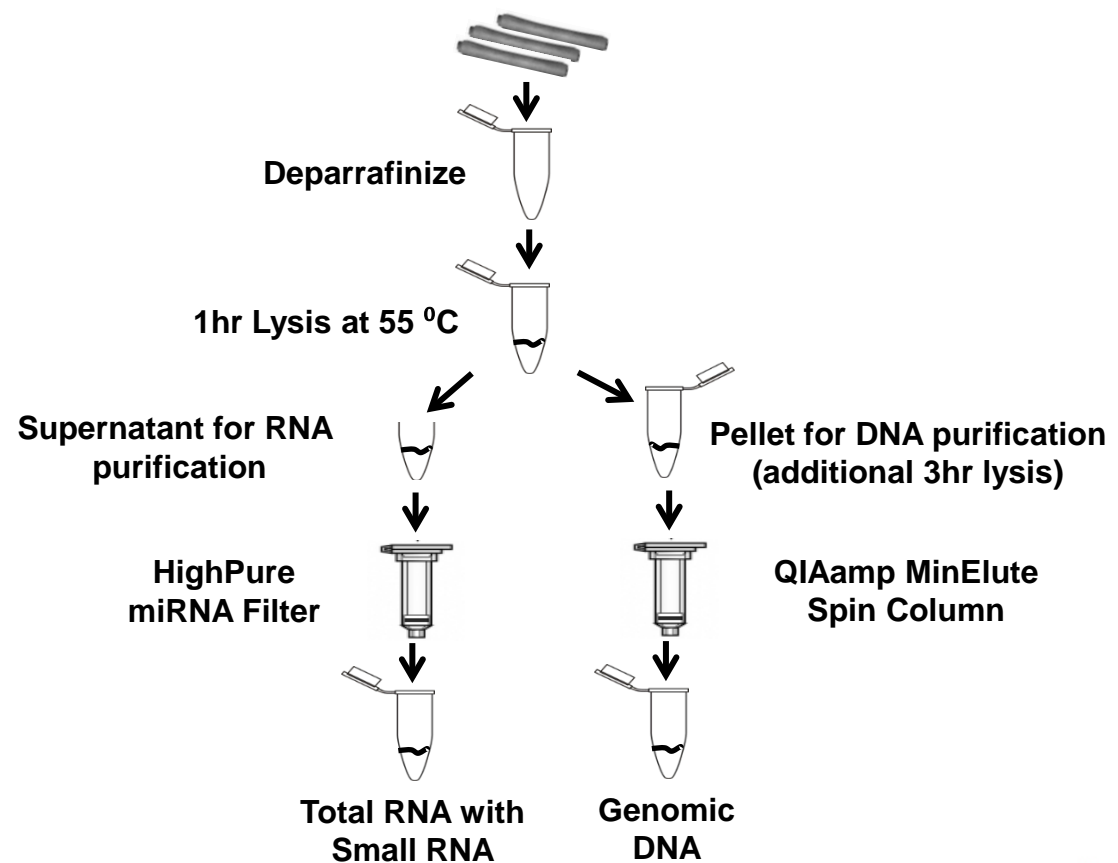
Participants and Distribution

Tumor Type	Number of Patients	Tissue Time in 10% Formalin (minutes)	Age of FFPE Tissue Block (years)	% Tumor Nuclei	% Necrosis	Number of Pooled Extractions	RNA Integrity (RIN)
Colon Adenocarcinoma	10	961.7 +/- 636	2.86 +/- 0.7	74.46 +/- 11	4.82 +/- 3	3.2 +/- 1	2.36 +/- 0.2
Endometrial Carcinoma	4	703.5 +/- 651	2.64 +/- 0.4	71.53 +/- 8	2.8 +/- 4	3.5 +/- 2	2.43 +/- 0.2
Lung Adenocarcinoma	12	780.25 +/- 562	2.97 +/- 0.6	72.64 +/- 6	5.36 +/- 5	3.17 +/- 1	2.42 +/- 0.1
Bladder Urothelial Carcinoma	3	432.33 +/- 170	2.72 +/- 0.2	89.18 +/- 5	2.49 +/- 2	4	2.33 +/- 0.1
Kidney Renal Clear Cell Carcinoma	4	437 +/- 150	2.89 +/- 0.1	89.86 +/- 6	0.83 +/- 1	5.5 +/- 3	1.9 +/- 0.5
Breast Invasive Carcinoma	5	480.8 +/- 144	2.66 +/- 0.5	74.16 +/- 5	4.03 +/- 6	4 +/- 2	2.32 +/- 0.2
Total/Average	38	716.92	2.84	76.32	4.07	3.63	2.33

Average Nucleic Acid Yield Per Extraction

The Cancer Genome Atlas

FF (tumor and normal) and FFPE derived analytes were distributed for characterization to the 5 genomic platforms listed below.

Platform	Participant																																					
Exome Sequencing	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Broad SNP 6	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	□	□	□	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
USC Methylation	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	■	□	□	■	■	■	■	■	■	■	■	■	■	■	■	■	■
BCCA miRNA Seq	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
UNC mRNA Seq	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38

Biospecimen overlap across platforms is indicated by the blue shading.

The Cancer Genome Atlas

SNP6 Array Results

- **0** FFPE SNP arrays passed QC, in part due to highly over segmented copy number profile

FFPE derived DNA gives rise to highly over segmented copy number profile.

Segment counts across sample preparations
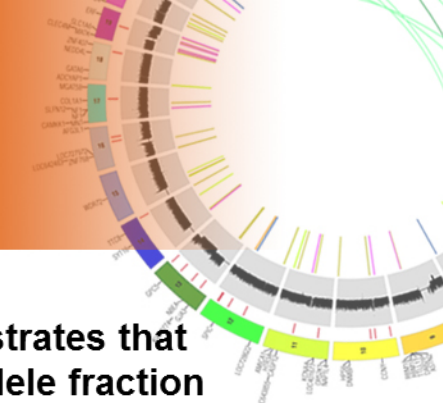

$p < 0.0001$
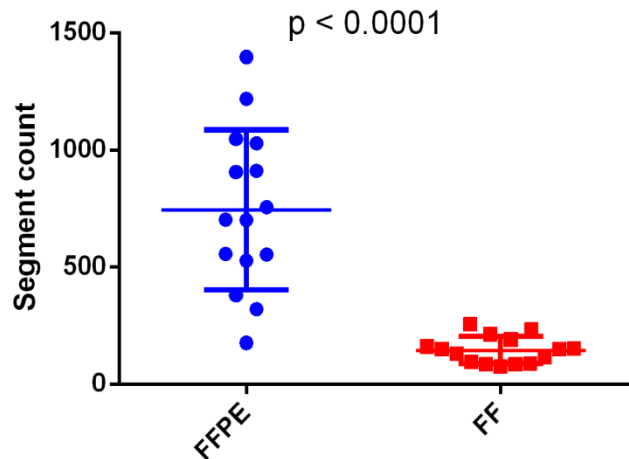
FFPE can validate the copy number profile of FF, but segmentation artifacts result in a high false discovery rate
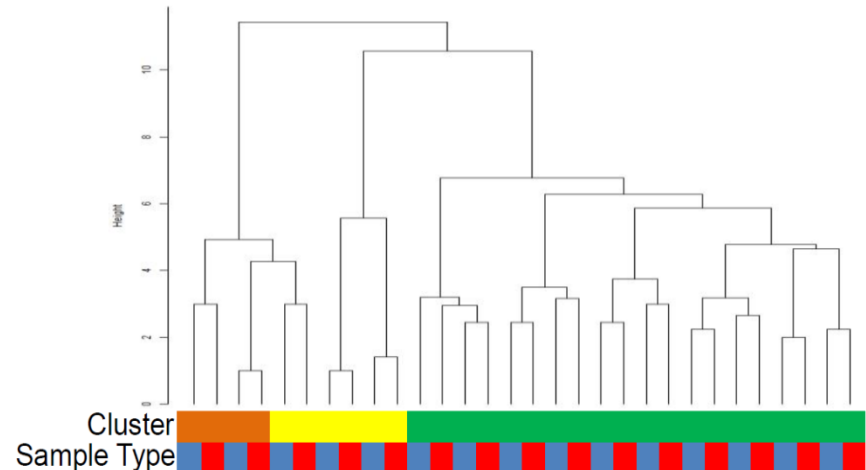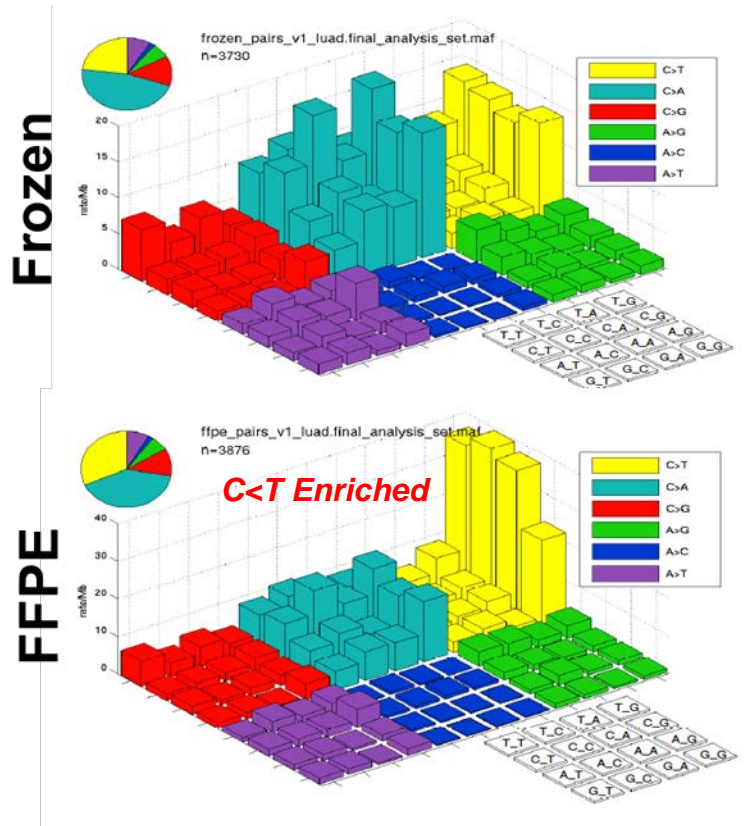

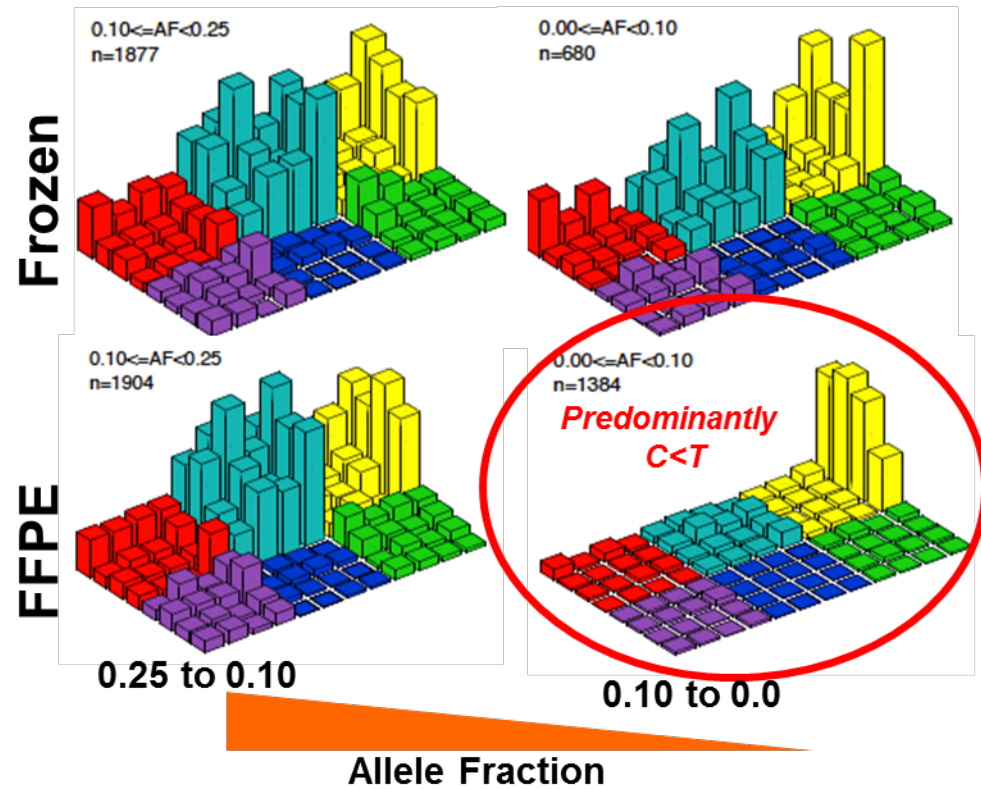*Segmentation artifacts comprise the stand-alone utility of determining SCNAs from FFPE through SNP6 array*

The Cancer Genome Atlas

Overall mutation spectrum in LUAD reveals shift towards C>T transitions in FFPE

Binning by allele fraction illustrates that FFPE effect is limited to low allele fraction

Results support use of FFPE for exome sequencing, however additional tools are needed to compensate for low allele fraction C>T SNV artifact.

The Cancer Genome Atlas

mRNA Sequencing Results

Pairwise Pearson correlation of transcript quantification between FF and FFPE

The Cancer Genome Atlas
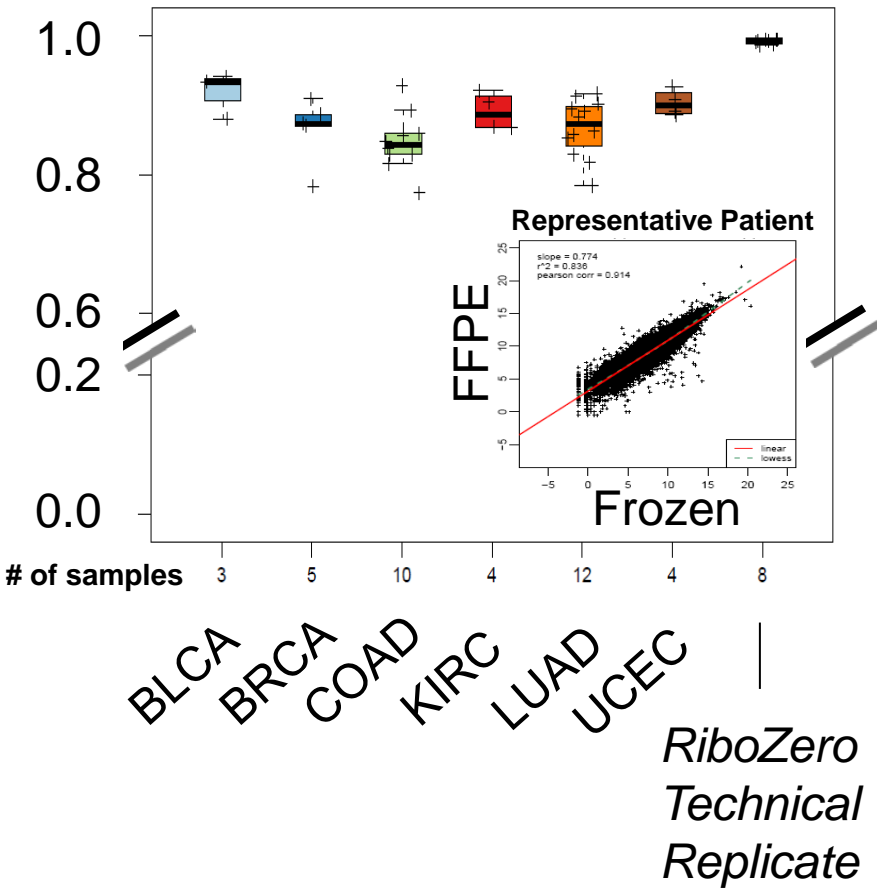
mRNA Sequencing Results

Isolating differences between FF and FFPE reveals consistent trends in quantification
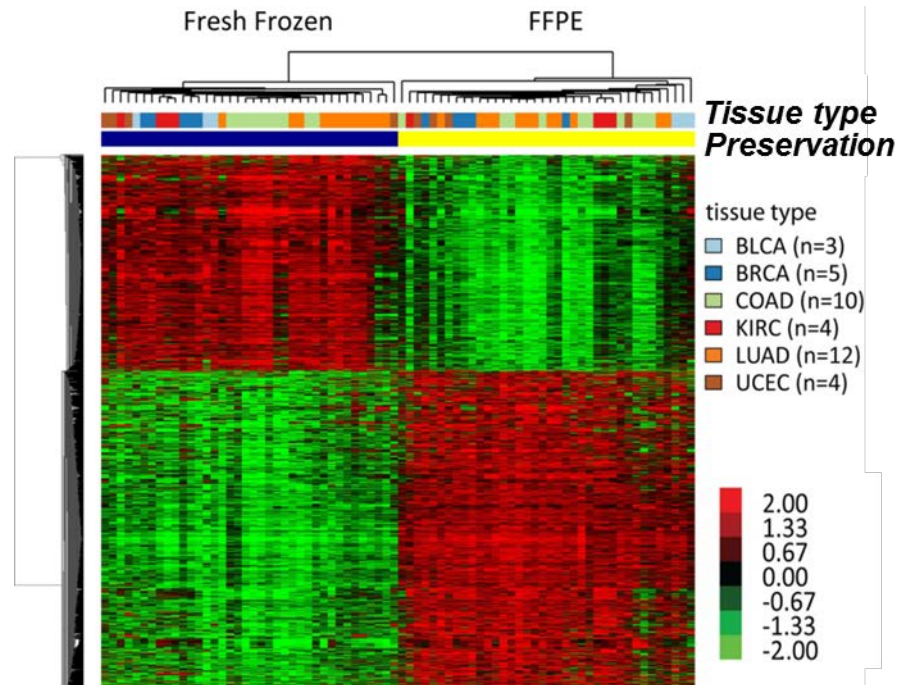
Differential transcript expression absolute level of their detection, not specific transcripts

Comparison of RSEM values
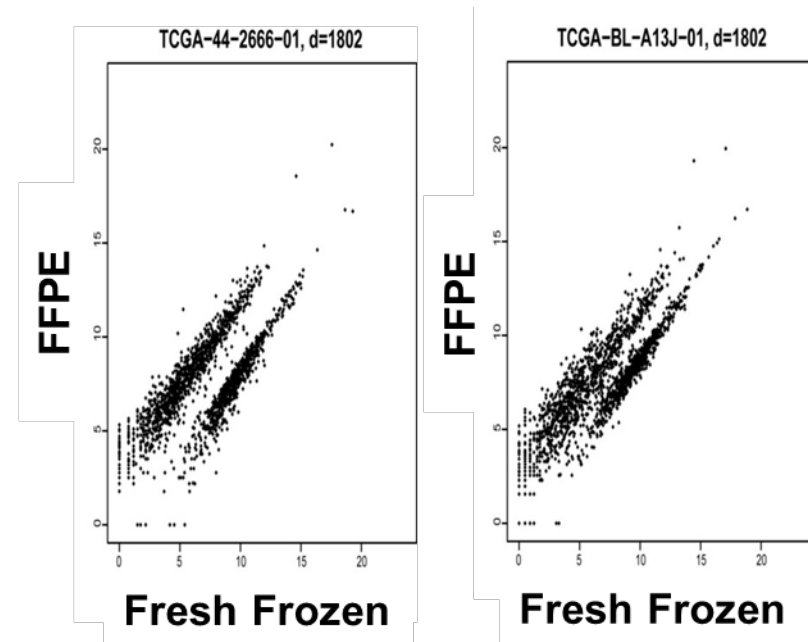
Overall - high concordance between FF and FFPE expression signatures, however additional bioinformatics steps may be required to adjust for differences in the level of expression detected in FFPE samples.

The Cancer Genome Atlas

miRNA Sequencing Results
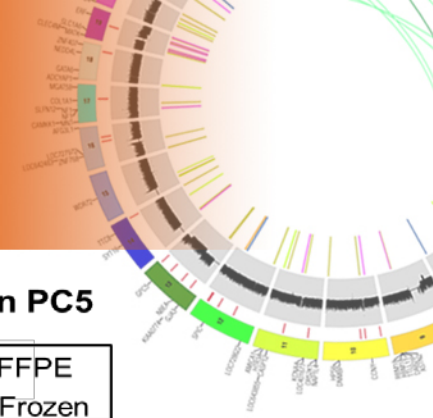
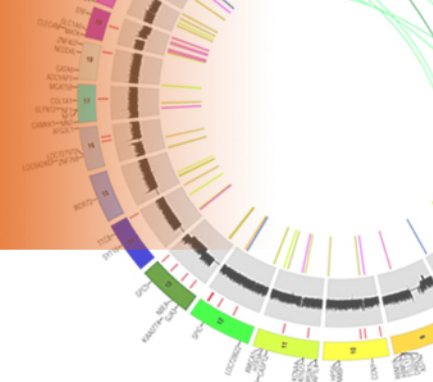Distribution of Variance between FF and FFPE among principal components
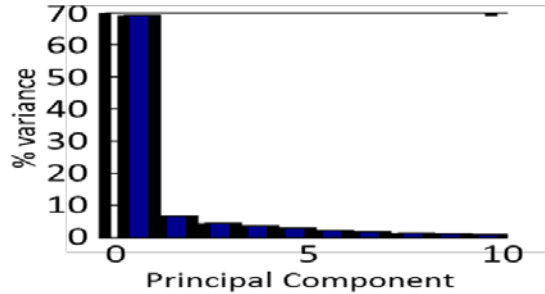
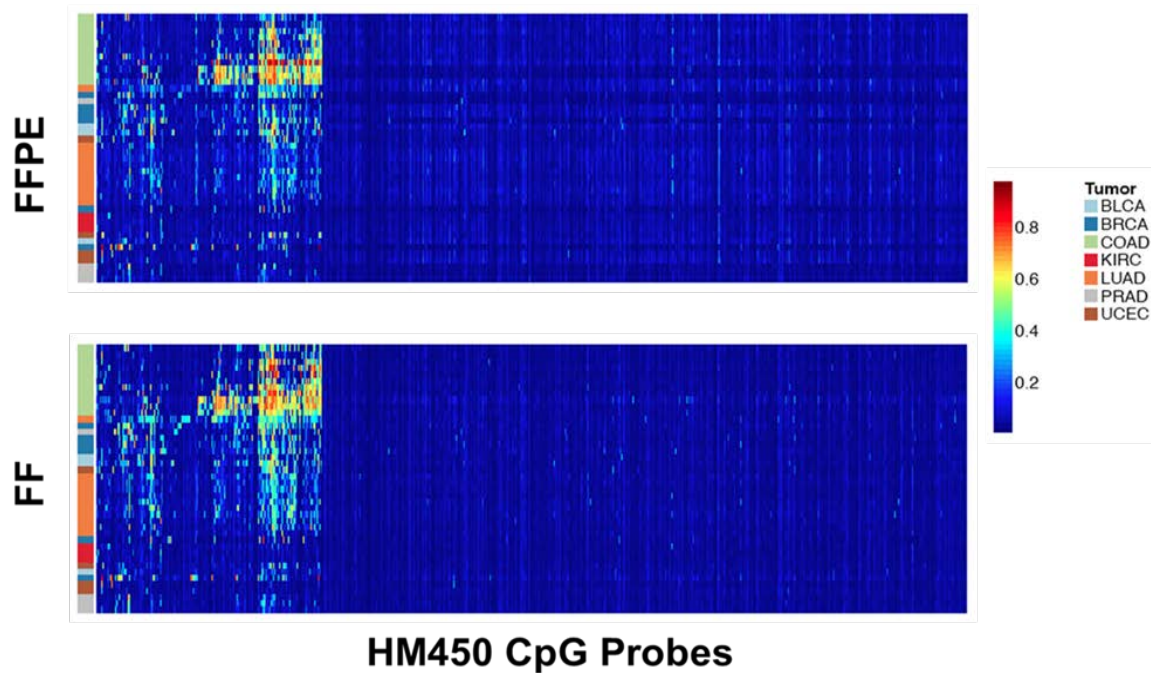Separation of FFPE from FF in PC5
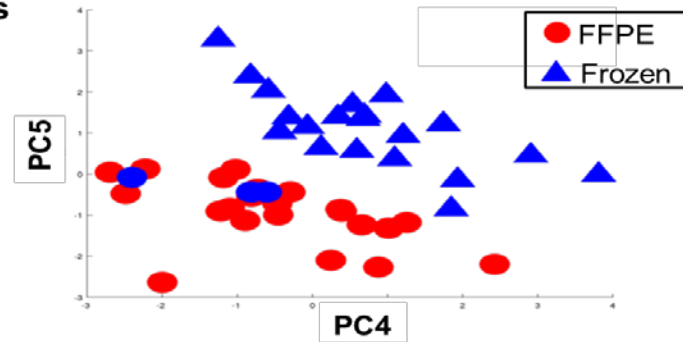
Overall- FFPE has weak effect on miRNA characterization. Additional work is needed to gain greater insight into the cause/effect of increased miRNA diversity.
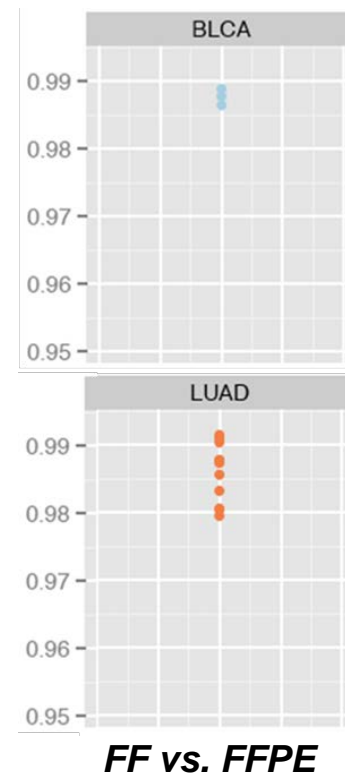
DNA Methylation Array Results

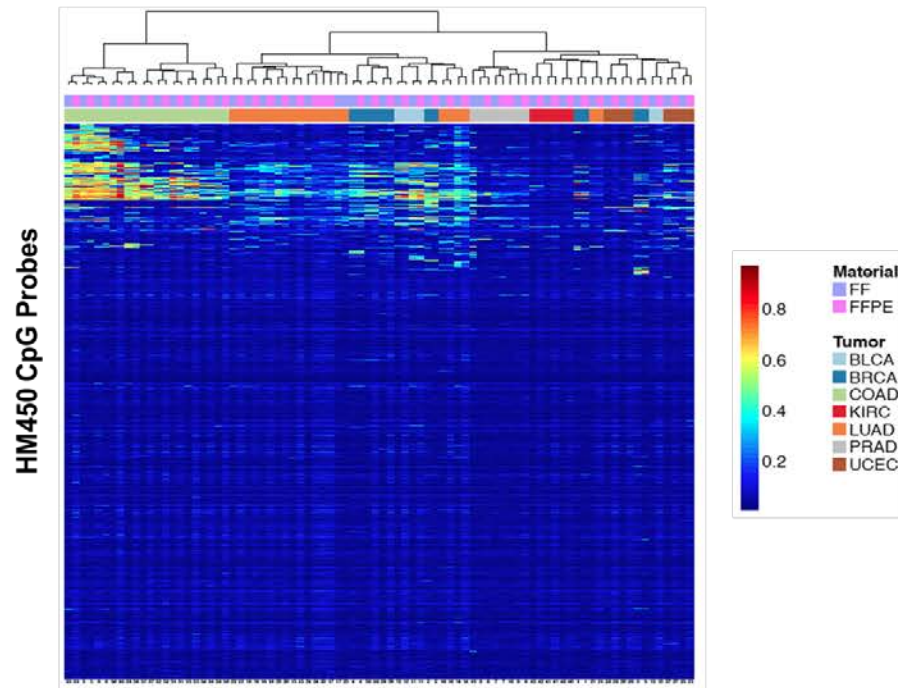Unsupervised Clustering of FF and FFPE HM450 DNA methylation datasets

HM450 CpG Probes

Pairwise Pearson Correlation Coefficients

FF vs. FFPE

The Cancer Genome Atlas

DNA Methylation Array Results

Unsupervised Clustering of HM450 DNA methylation data between FF-FFPE pairs

Overall results suggest excellent concordance in methylation signature obtained from FF and FFPE tumor specimens.

Illumina FFPE restoration protocol required

The Cancer Genome Atlas

Conclusions and Future Plans

- **Optimized a nucleic acid co-isolation method.**

- **DNA and RNA extracted from FFPE can be employed for multiple state of the art platforms.**

- **Characterization of the artifacts caused by formalin-fixation and paraffin embedding:**

  - **SNP6 arrays:** high false discovery rates due to over-segmented copy number.

  - **Exomes:** interpretable but with a low allele fraction (<0.10) C>T SNV artifact; consistent with effects of de-amination caused by formalin fixation.

  - **Methylation:** minimally affected in FFPE samples.

    - *Illumina FFPE Restoration protocol required.*

  - **mRNA-Seq:** good correlation between FF and FFPE samples, however a subset of transcripts systematically vary between FF and FFPE.

    - *RiboZero chemistry proved to be most reliable.*

  - **miRNA-Seq:** systematic increase in diversity of miRNA species from FFPE.

The Cancer Genome Atlas

Conclusions and Future Plans

- **Future efforts**
  - Analyze FFPE signature in the context of multi-center calling.
  - Delineate the influence of tumor heterogeneity in the results of this study (spatial separation exists between Frozen and FFPE portions).
  - Deeper analysis of the differences between FF and FFPE to identify potential bioinformatics mechanisms to correct of the artifacts caused by formalin fixation and paraffin embedding.

The Cancer Genome Atlas

Acknowledgements

Baylor College of Medicine Human Genome Sequencing Center
Harsha Doddapaneni
Nipun Kakkar
Liu Xi
Donna Marie Morton
Donna Marie Muzny
David Wheeler

British Columbia Cancer Agency
Andy Mungall
Andy Chu
Richard Corbett
Payal Sipahimalani

Broad Institute
Bradley Murray
Carrie Sougnez
Petar Stojanov
Michael Lawrence
Raktim Sinha

Memorial Sloan-Kettering Cancer Center
Jorge Reis-Filho
Nicholas Schultz
Cyriac Kandoth
Raymond Lim
Charlotte Ng
Marc Ladanyi

National Cancer Institute Program Office
Roy Tarnuzzer
Jean Claude Zenklusen

National Human Genome Research Institute
Carolyn Hutter
Heidi Sofia

Nationwide Children's Hospital Biospecimen Core Resource
Erik Zmuda
Natalie Bir
Jessica Frick
Jay Bowen
Julie Gastier-Foster

Oregon Health and Science University
Paul Spellman

University of North Carolina School of Medicine
Katherine Hoadley
Wei Zhao

University of Southern California
Dan Weisenberger
Moiz Bootwalla
Christina Curtis
Toshinori Hinoue
Peter W. Laird

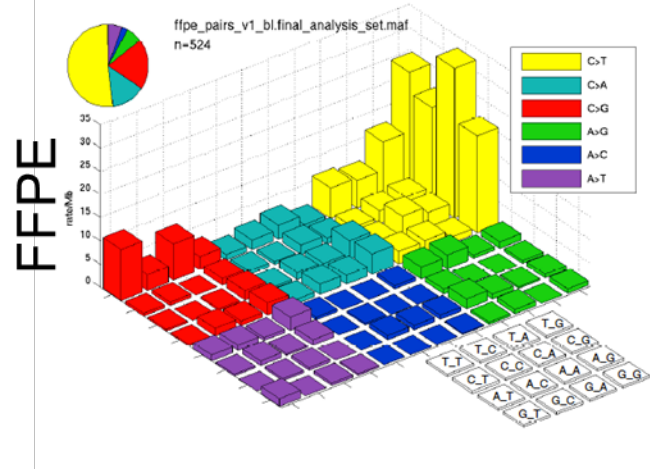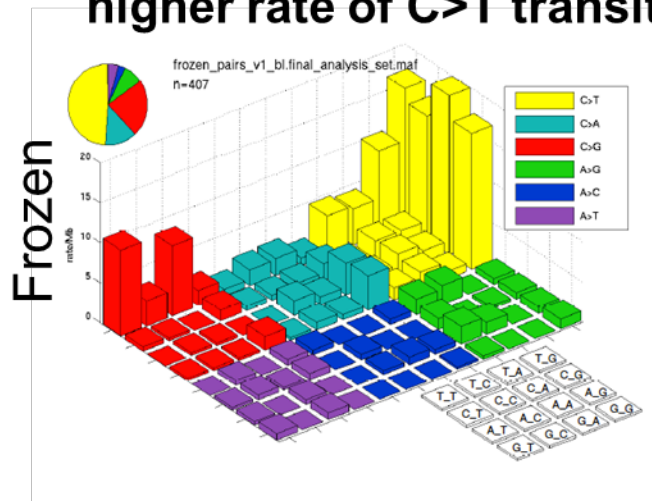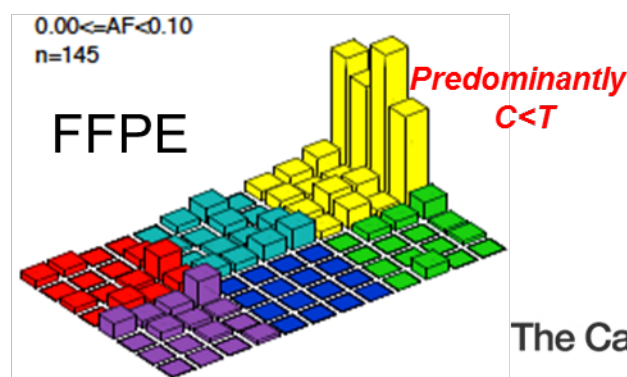**Washington University**
Chris Miller
Mike McLellan
Bob Fulton

The Cancer Genome Atlas
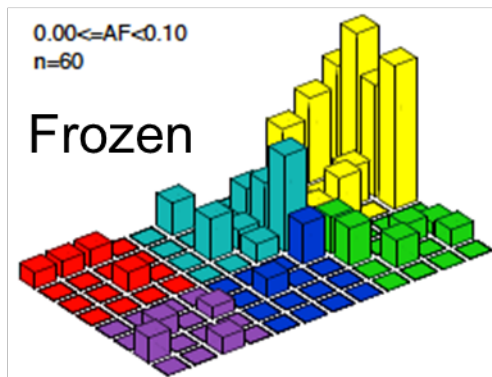
FFPE signature in BLCA is more subtle, in part due to higher rate of C>T transitions in the frozen tissue

As with LUAD, C>T transition signature of FFPE is also detected at low allele fractions in BLCA
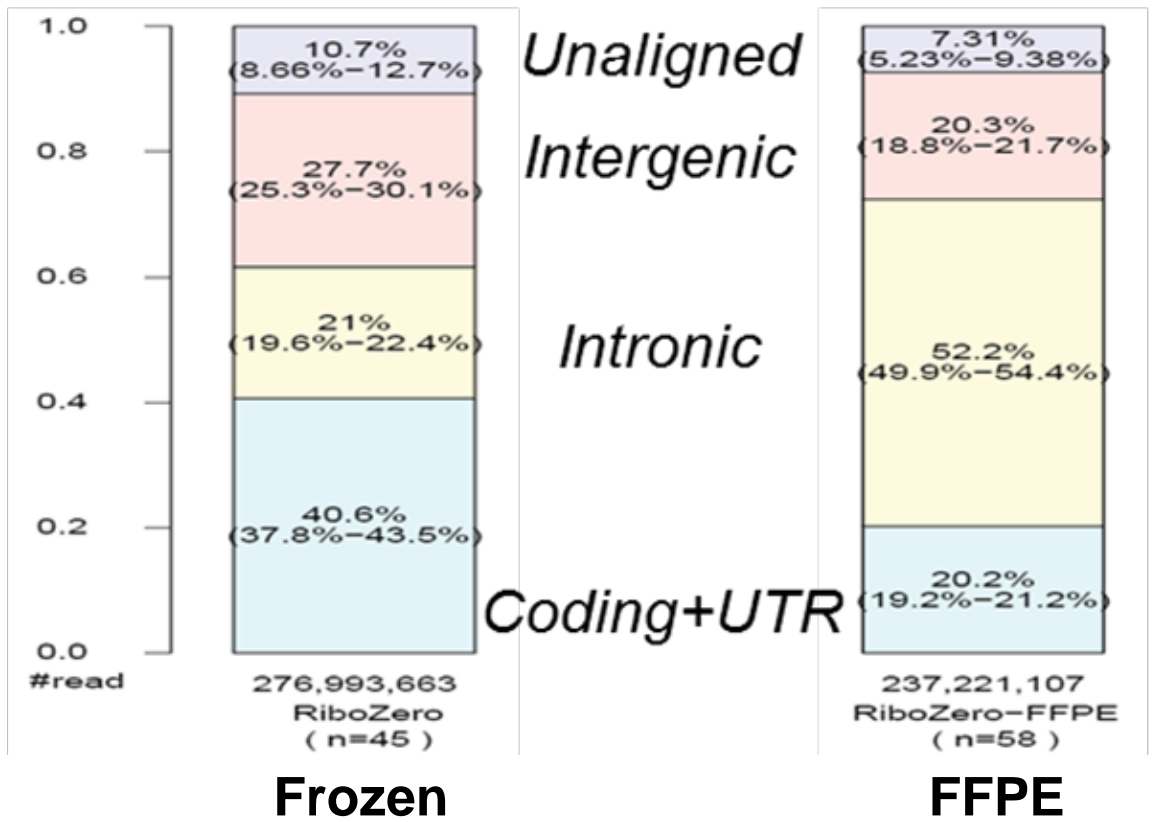
The Cancer Genome Atlas

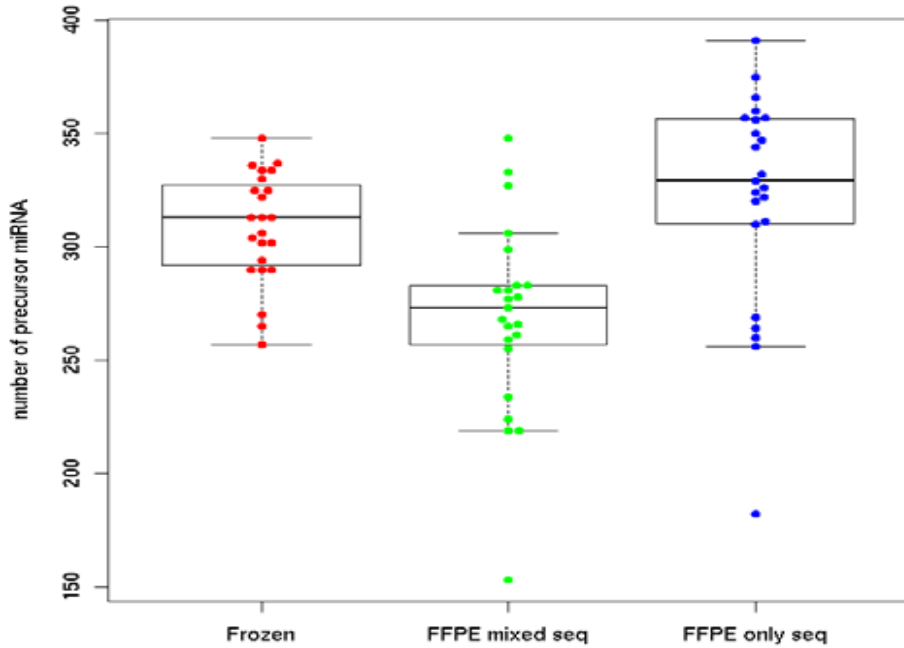Relative coverage of exons, introns, and intergenic regions

The Cancer Genome Atlas

Mapping of mRNA Sequencing Reads
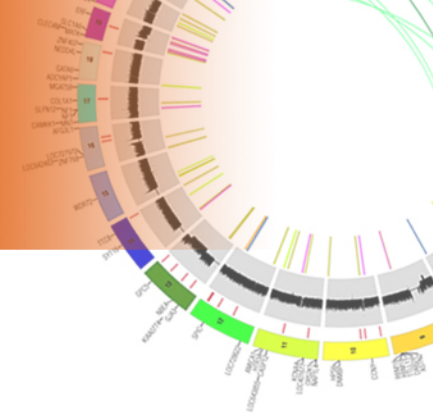
miRNA Diveristy

miRNA Yield

The Cancer Genome Atlas