Integrated Analysis for Identification of Risk Stratification Biomarkers for Colon Cancer

Andrea O'Hara, Pranay Vishwanath, Haythem Latif
GENEWIZ From Azenta Life Sciences, South Plainfield, NJ

Abstract

Colorectal cancer (CRC) remains the second leading cause of cancer-related deaths worldwide, underscoring the critical need for improved early detection and risk stratification methods. While polyps are detected in up to 40% of colonoscopies, most are negative for significant lesions, with only about 10% showing advanced adenomas or carcinomas. Although colonoscopy remains the gold standard for CRC screening, it has major limitations in predicting progression from benign neoplasia to advanced adenomas or carcinomas. Precision approaches that integrate molecular insights are necessary to identify biomarkers for risk stratification and better understand which individuals with colon neoplasia are at increased risk to develop advanced adenomas or carcinomas.

This study leverages a multi-modal, integrated analysis of spatial transcriptomics and single-cell RNA-sequencing, an approach that has not been comprehensively applied to risk stratification in CRC. Integrated bioinformatics analyses were used to compare and combine bulk results, along with clinical and demographic data associated with these samples, for downstream pathway and processes analysis. In addition, these results were further analyzed to cross-compare and validate single cell spatial findings.

With this integrated multiomics approach, spatial transcriptomics provides high-resolution insights into gene expression. By identifying key biomarkers and pathways associated with CRC progression, this study aims to pave the way for personalized screening strategies and targeted interventions to reduce the burden of advanced colorectal cancer.

Overview of FFPE Multiomics Workflow

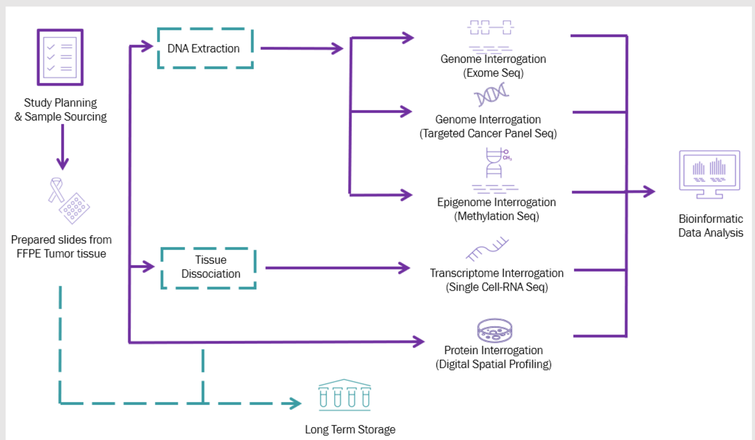
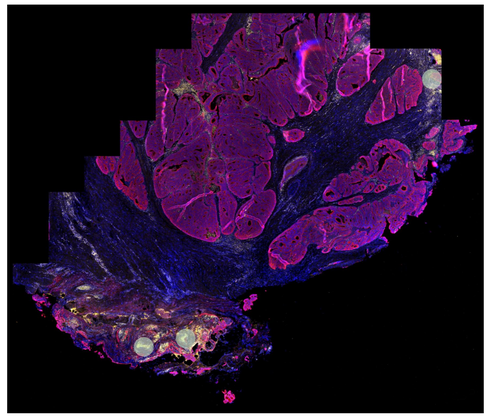


Figure 1. Schematic of multiomic sample processing workflow. FFPE archived tumor colon block was sectioned into 25µm slides. DNA was extracted followed by NGS library construction for whole exome sequencing (WES) and whole genome methylation sequencing (Methyl-Seq). For transcriptome investigation, single-cells were dissociated according to 10x Genomics protocol with Miltenyi GentleMACS dissociator and fixed single-cell gene expression libraries were constructed and analyzed by scRNA-sequencing. NanoString GeoMx Human Whole Transcriptome Atlas (WTA), Human Protein Core and Pan-Tumor panel were performed for spatial profiling.

Results

A



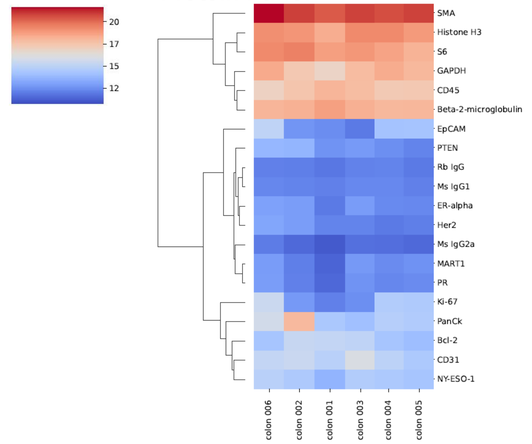
B

Top Mutations

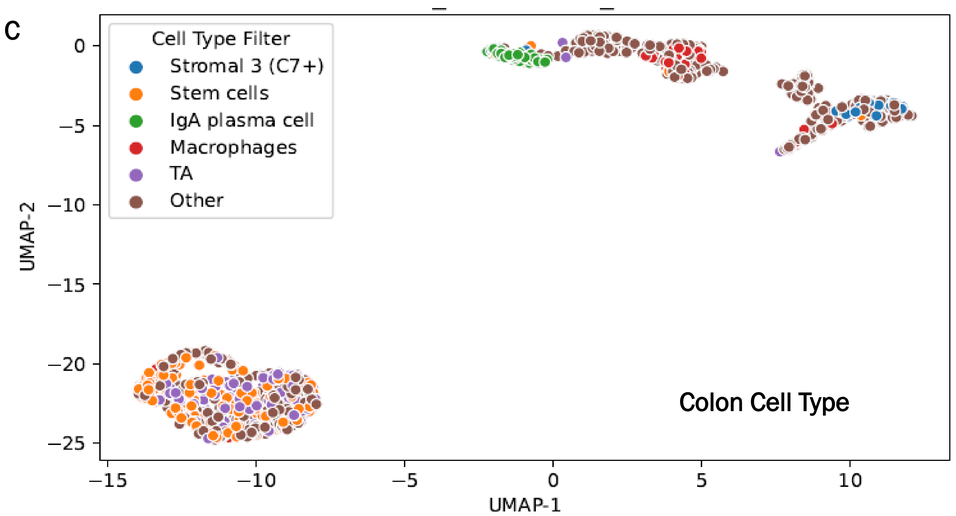
SYMBOL	Variant	Allele Depth	Freq.	Consequence	IMPACT
FOXA1	14:38061742-C>T	8,811,337	0.602	Missense	MODERATE
KMT2D	12:49434685-CCTT>C	685,874	0.560	Inframe Deletion	MODERATE
PIM1	6:37139100-T>G	11,751,273	0.520	Missense	MODERATE
NUP93	16:56855472-C>G	743,794	0.517	Missense	MODERATE
BARD1	2:215632255-CA>TG	11,481,144	0.499	Missense	MODERATE
HLA-C	6:31239378-T>G	821,711	0.464	Missense & Splice Variant	MODERATE
HLA-C	6:31239577-A>C	16,911,314	0.437	Missense	MODERATE
SH2B3	12:111884608-T>C	13,851,058	0.433	Missense	MODERATE
ERBB2	17:37880220-T>C	1,452,472	0.246	Missense	MODERATE
MGA	15:42042023-A>G	2,361,452	0.161	Missense	MODERATE
TP53	17:7578529-A>G	1,239,203	0.141	Missense	MODERATE
POLD1	19:50912061-G>C	801,119	0.130	Missense	MODERATE

D

Immuno-Oncology Protein Panel



C



E

Pathogenic Variant Expression from Whole Transcriptome Atlas

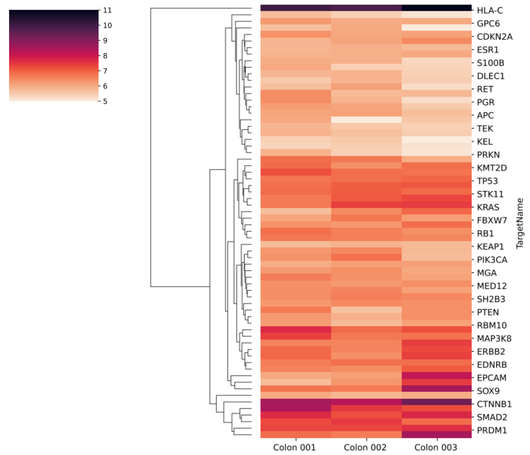


Figure 2. Colon FFPE Section (A). Pan-cancer panel sequencing (B) elucidates signature for colon cancer samples. Cancer panel library was sequenced to ~2000X raw coverage and analyzed using Sentieon TNseq pipeline. All mutations with a frequency of >10%, of moderate or high impact consequence are shown. **Single-cell gene expression results (C) identify common tissue cell types.** Single-cell sequencing was processed using the 10x Fixed RNA workflow; a defined whole transcriptome probe-based counting ideal for fixed samples. UMAP clustering of cells for colon samples, with indications for cell type based on the intestinal tract. **GeoMx DSP Immuno-Oncology protein panel (D) protein panel and whole transcriptome atlas (E) results.** NanoString GeoMx Human Whole Transcriptome Atlas (WTA), Human Protein Core and Pan-Tumor panel were performed for spatial profiling. Clustering expression analysis of Immuno-Oncology protein panel and mutated genes of interest, as described in B.

Conclusions

- Single-cell sequencing allows for a deep dive into individual cell types present in complex tissue samples. Updates to processing allows for analysis of FFPE tissues.
- Spatial profiling is ideal for assessment of specific regions of interest within the tumor microenvironment.
- Understanding tumor microenvironment is key for cancer diagnosis, prognosis and treatment considering tumor heterogeneity.
- Complex multiomics information can be collected from a single FFPE sample, the most common archived samples from tumor patients, including genomics, transcriptomics, epigenomics and proteomics investigation.