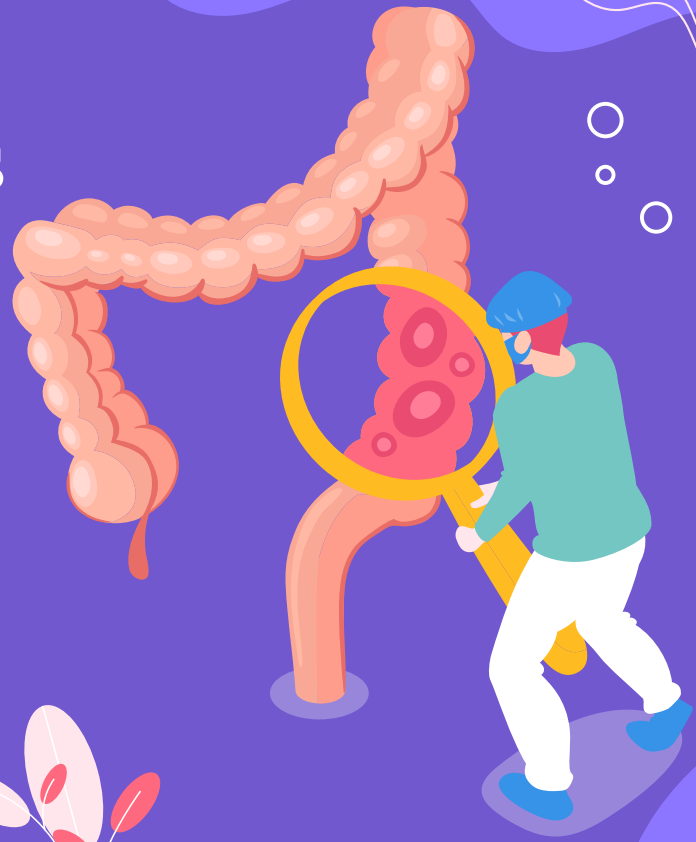
Differential Transcriptomics and Pathway Analysis of Tumor versus Normal Colorectal Tissue

Aditi Deshpande and Jessica Chen



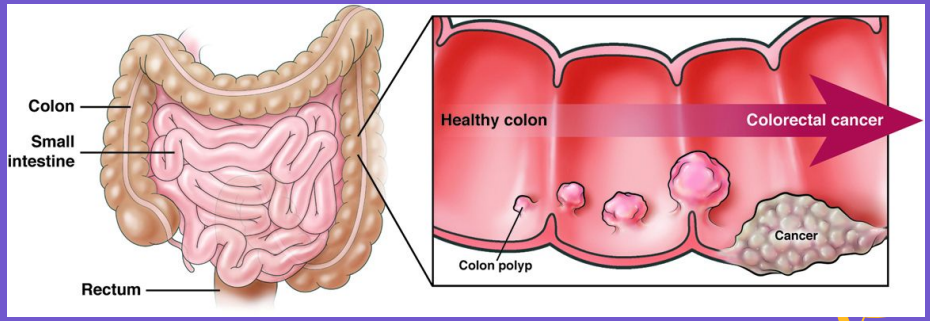
Colorectal Cancer (CRC)

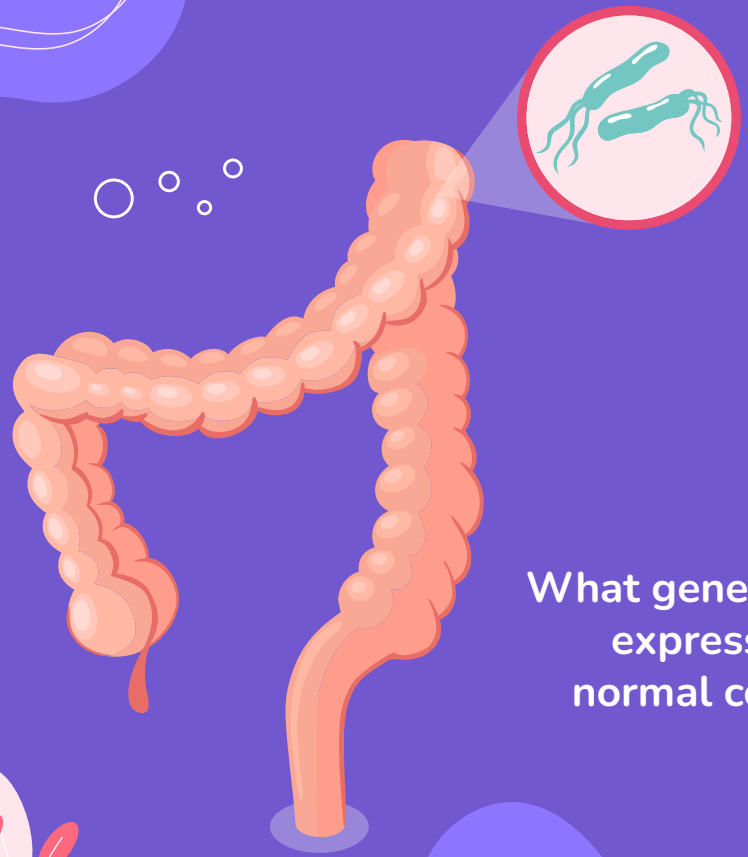
- leading cause of cancer death in individuals under 50 in the U.S.
- Incidence rates have risen rapidly since the 1990s in increasingly younger individuals
- Despite advances in treatment, survival rates remain poor for advanced-stage disease
- Early detection and understanding of disease progression are critical for improved outcomes
- **Key Challenge:** Understanding the molecular events driving CRC initiation and progression across different tumor stages

Molecular Drivers of CRC

- **Known mutations:** Canonical pathways such as Wnt and p53 are well-established in CRC
- **Limitation of genetic data alone**: Mutations don't tell the full story of disease biology
- Gene expression patterns reveal:
  - Active biological processes in tumors
  - Dysregulated pathways across disease stages
  - Tissue-specific molecular signatures

Stage-specific insights: Understanding how gene expression changes from early to late-stage tumors reveals disease progression mechanisms



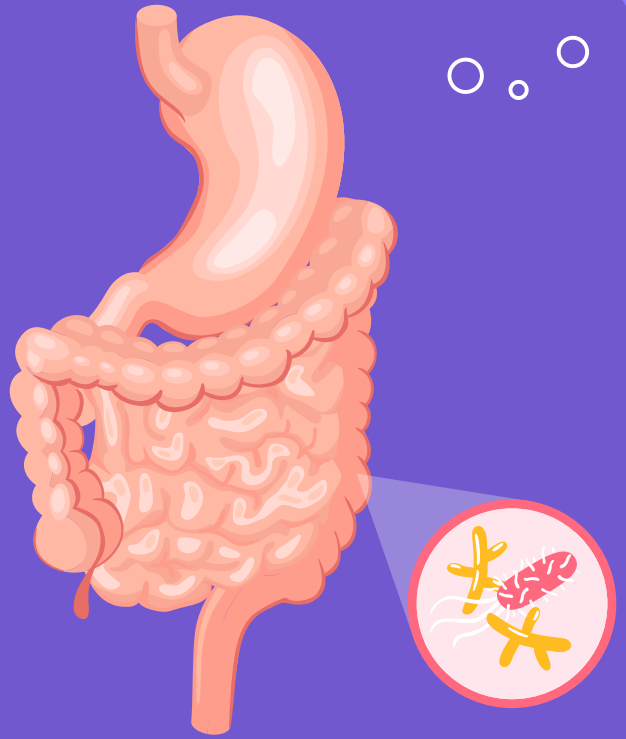


Primary Research Question!

What genes and molecular pathways are differentially expressed between primary colorectal tumors and normal colon tissue, and how do these changes vary across tumor stages (Early, Mid, Late)?

Key Goals

- Identify key transcriptional changes that distinguish cancerous from normal tissue
- Discover stage-specific molecular signatures that could inform prognosis and treatment
- Reveal dysregulated biological processes driving CRC progression
- Provide targets for potential therapeutic intervention
- Contribute to understanding of CRC biology at the molecular level



Methodology

Data Source: The Cancer Genome Atlas (TCGA) Colon Adenocarcinoma (TCGA-COAD) dataset

Sample Composition:

- Normal tissue: 41 samples
- Tumor samples: 50 samples

Early stage (I-II): 26 tumors, 24 normal
Mid stage (III): 12 tumors, 6 normal
Late stage (IV): 4 tumors, 2 normal

Data Type: RNA-seq gene-level expression counts (HTSeq-counts)

Differential Gene Expression Analysis Results

Early Stage Tumor vs Normal

Downregulated Genes	Upregulated Genes
SERPINA1	ERVE-5
ALB	LINC02418
TMIGD1	KLK6
OTOP3	ZIC5
CA1	LINC02474

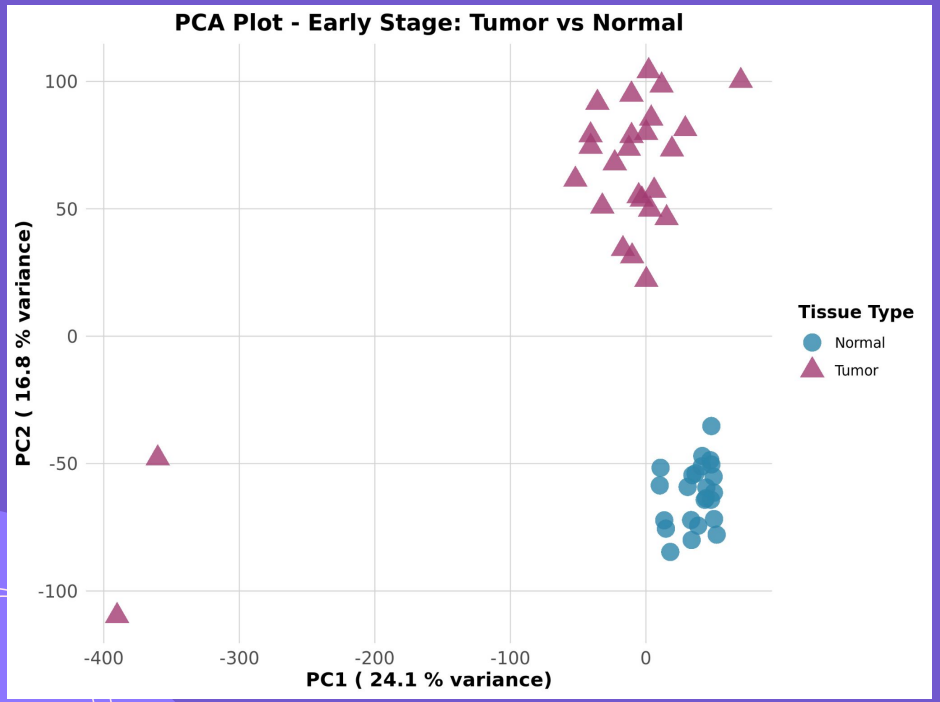
Tumor Suppressor

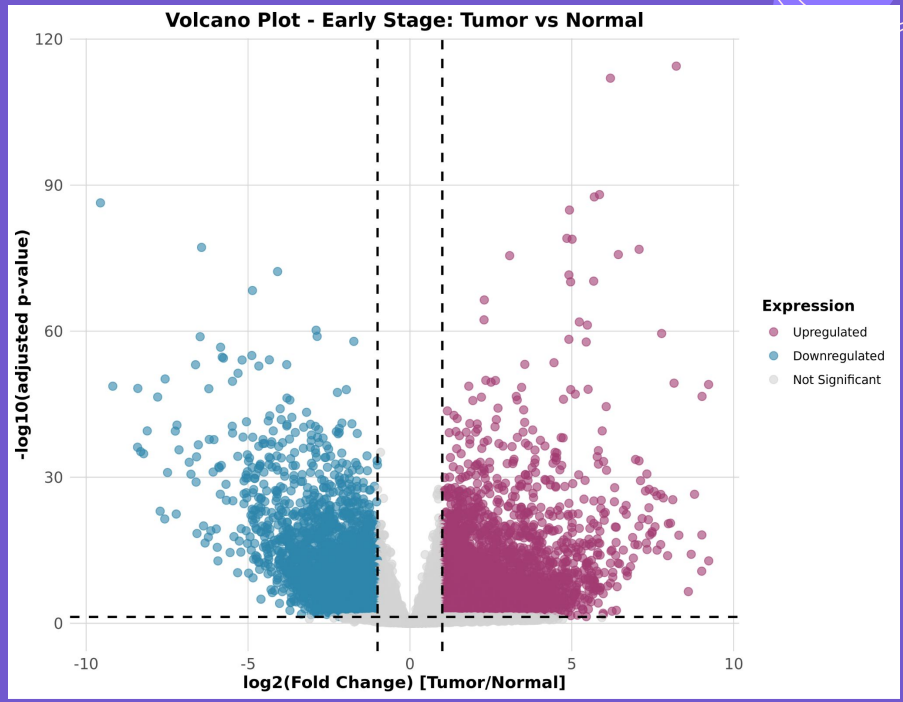
Proton Channel

Colonic epithelium

oncogenic lncRNA

Early Stage Tumor vs Normal



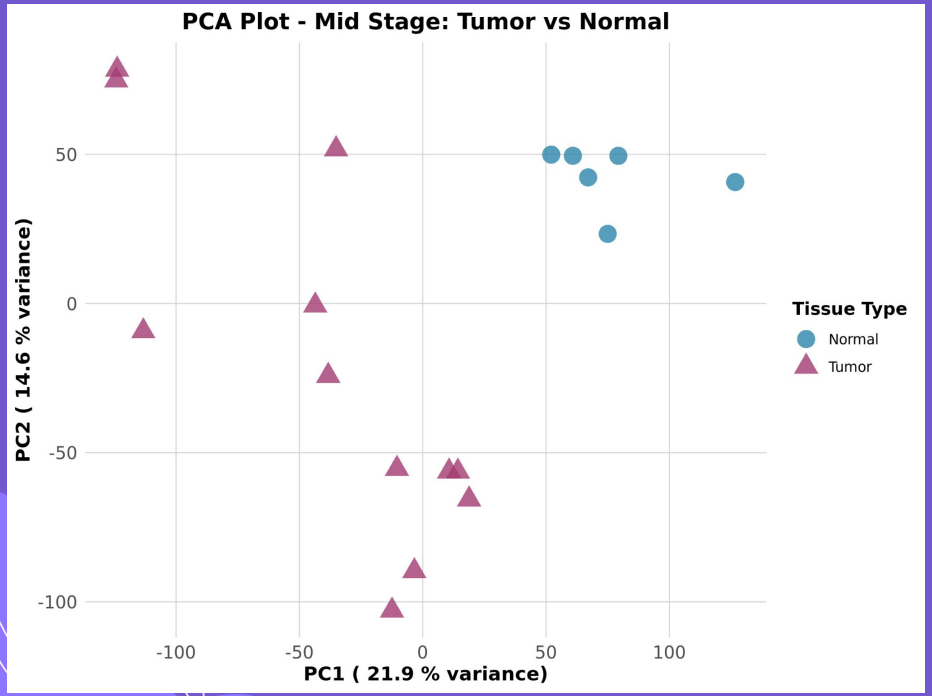


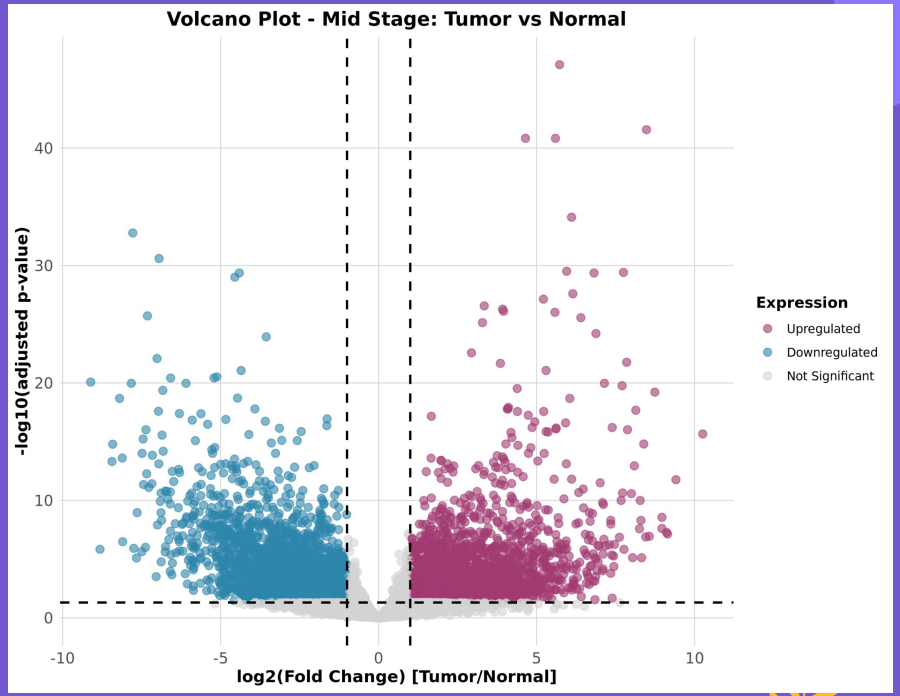
10,284 upregulated and 6,531 downregulated

Mid Stage Tumor vs Normal

Downregulated Genes	Upregulated Genes
TMIGD1	FEZF1-AS1
KRTAP13-2	LINC02418
CA1	REG1A
SCNN1G	PRSS2
OTOP3	STYXL2

Mid Stage Tumor vs Normal



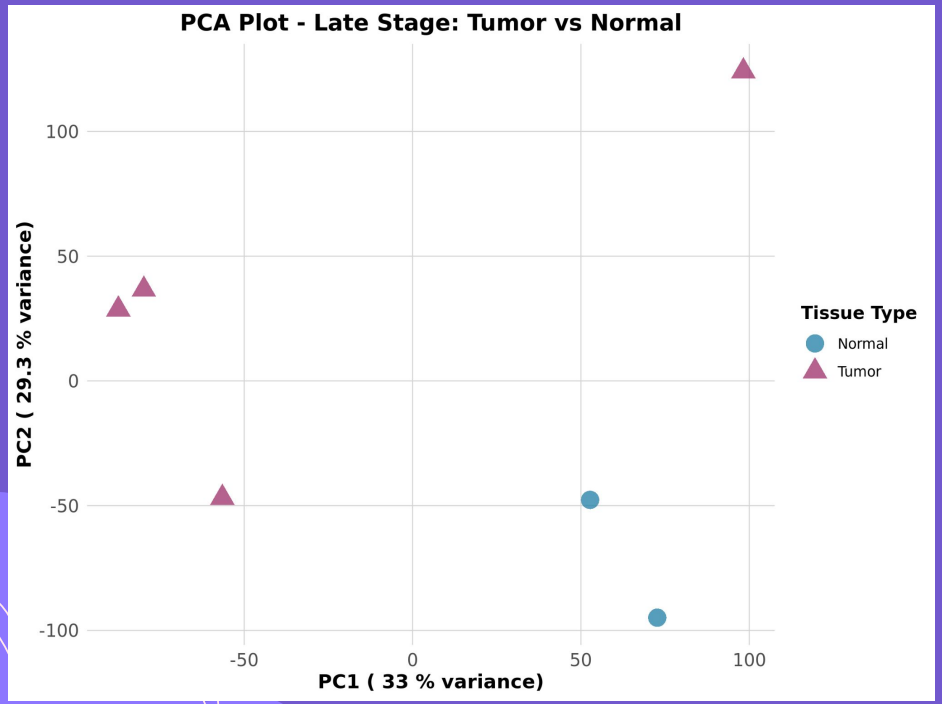


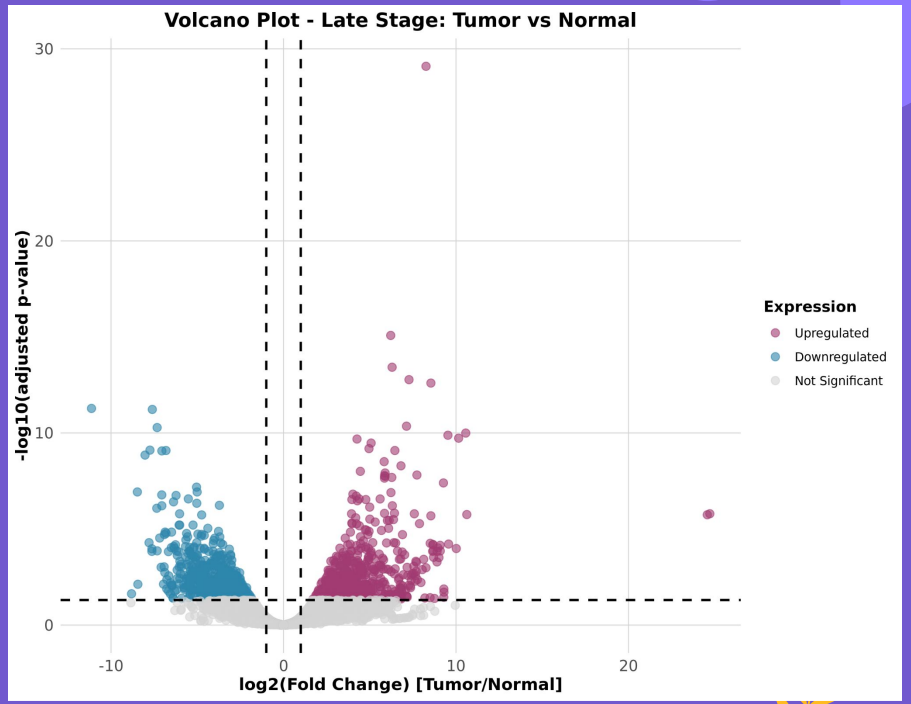
3,739 upregulated and 3,433 downregulated

Late Stage Tumor vs Normal

Downregulated Genes	Upregulated Genes
KRT24	MAGEA3
RERGL	MAGEA6
RP11	CST1
OTOP2	NOTUM
PI16	LINC02418

Late Stage Tumor vs Normal

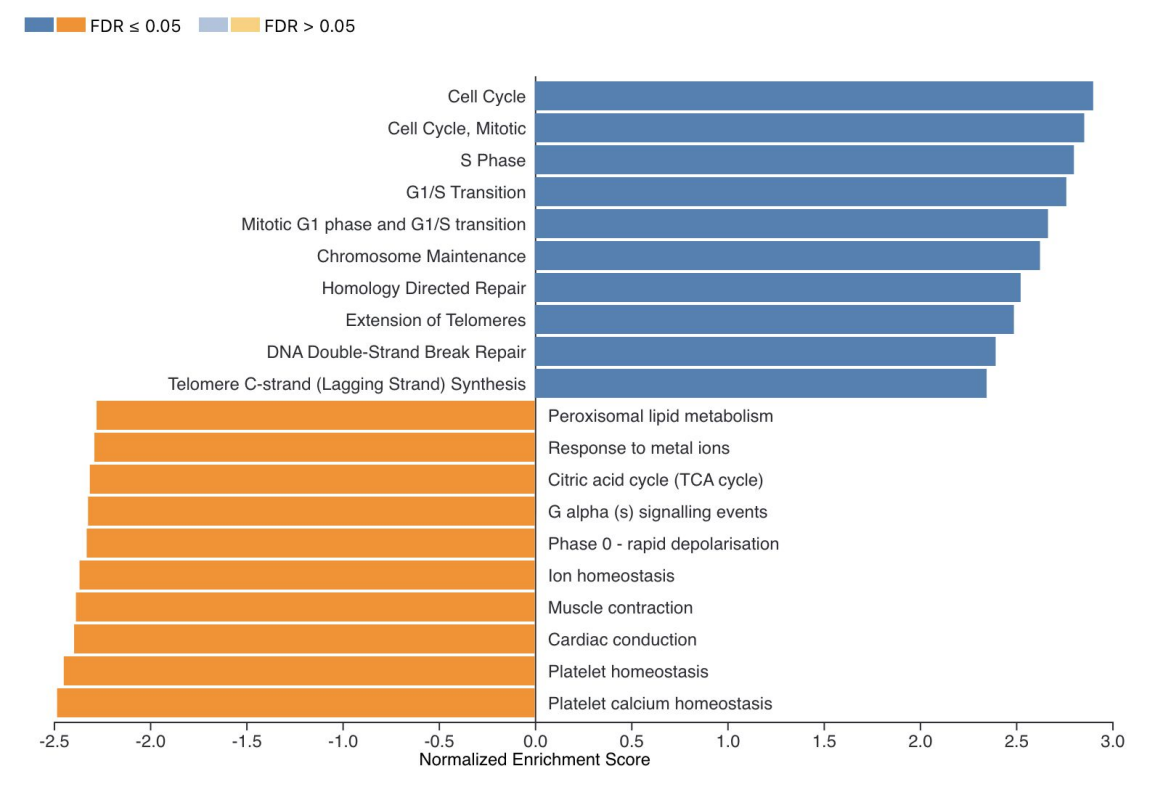




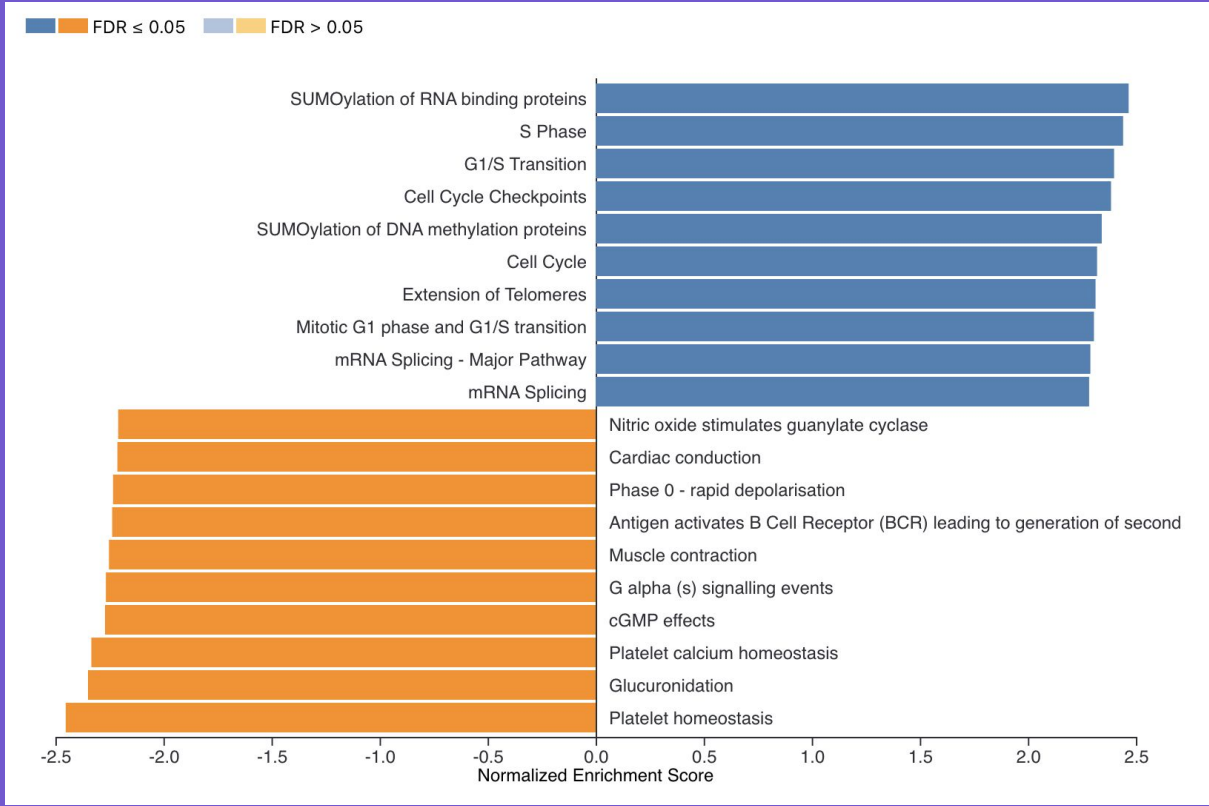
615 upregulated and 590 downregulated

GSEA Results

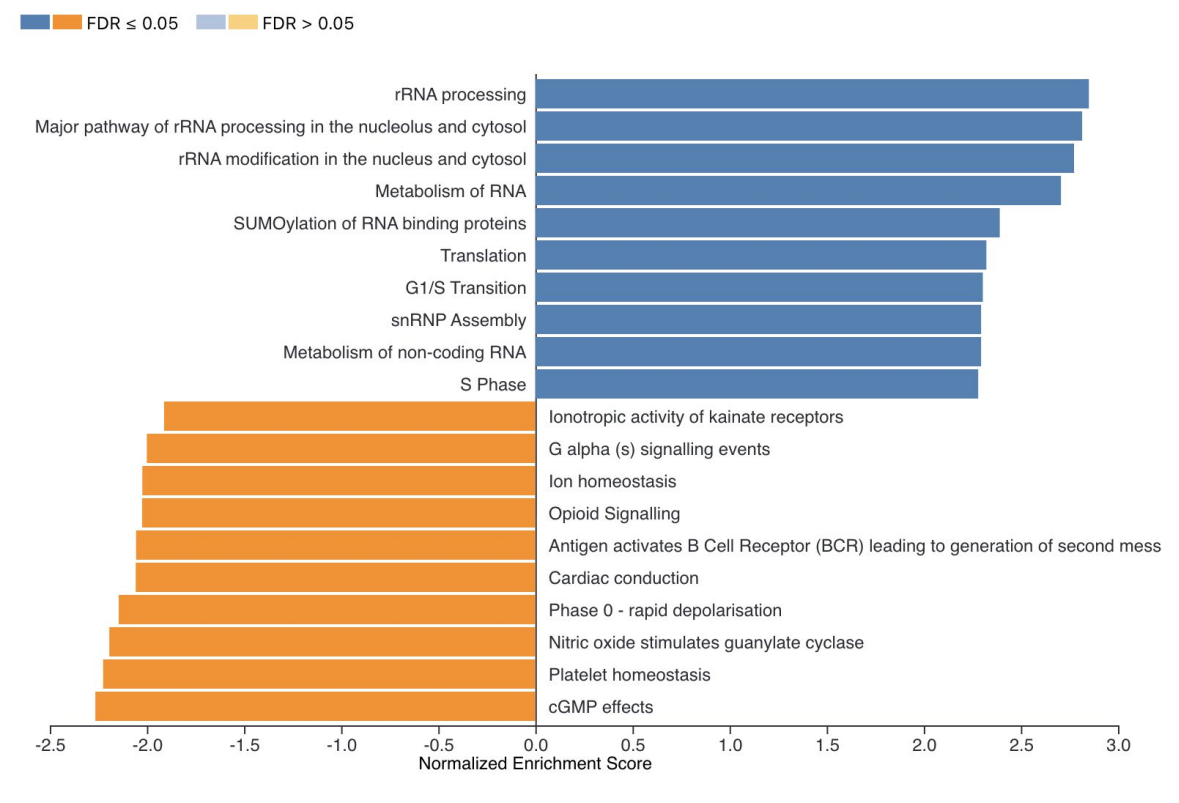
Early Stage Tumor vs Normal



Mid Stage Tumor vs Normal



Late Stage Tumor vs Normal



Enrichr Results

Early Stage Tumor vs Normal

GO Biological Processes	GO Cellular processes	GO Molecular Processes	KEGG	Reactome
Mitotic Metaphase Chromosome Alignment; Mitotic Sister Chromatid Segregation; Metaphase Chromosome Alignment	None ($p > 0.05$)	None ($p > 0.05$)	None ($p > 0.05$)	Classical Antibody-Mediated Complement Activation; Cell Cycle-Mitotic; Initial Triggering of Complement, Creation of C4 and C2 Activators

Mid Stage Tumor vs Normal

GO Biological Processes	GO Cellular processes	GO Molecular Processes	KEGG	Reactome
Regulation of Cation Channel Activity	Collagen-Containing Extracellular Matrix; Small-Subunit Processome; Cell-Substrate Junction; Focal Adhesion; Sarcolemma	snoRNA Binding; Cell-Matrix Adhesion Mediator Activity	None ($p > 0.05$)	rRNA Processing in the Nucleus and Cytosol; Major Pathway of rRNA Processing in the Nucleolus and Cytosol; Extracellular Matrix Organization; Muscle Contraction; Cardiac Conduction

Late Stage Tumor vs Normal

GO Biological Processes	GO Cellular processes	GO Molecular Processes	KEGG	Reactome
Extracellular Matrix Organization, Sodium Ion Transport, Monocarboxylic Acid Transport, Lipid Transport	Collagen-Containing Extracellular Matrix, Neuron Projection	Serine-Type Endopeptidase Activity, Serine-Type Peptidase Activity, Endopeptidase Activity, Sodium Channel Activity, Thyroid Hormone Transmembrane Transporter Activity	Cytokine-cytokine receptor interaction, Bile secretion, Rheumatoid arthritis,	Scavenging of Heme From Plasma, Anti-inflammatory Response Favouring Leishmania Parasite Infection, Leishmania Parasite Growth and Survival, Role of LAT2 NTAL LAB on Calcium Mobilization, Classical Antibody-Mediated Complement Activation

Discussion

Key Findings:

- Early stage: Cell cycle, DNA repair & chromosome maintenance dysregulation
- Mid stage: Extracellular matrix & focal adhesion changes
- Late stage: rRNA processing & RNA metabolism with negative enrichment in signaling pathways
- Consistently Upregulated: LINC02418; Consistently Downregulated: TMIGD1, CA1, and OTOP3

Clinical Significance:

- Stage-specific profiles inform prognosis & treatment
- Ribosomal dysregulation = potential early-stage therapeutic target
- ECM remodeling reveals metastatic mechanisms

Conclusion

Main Findings:

- Stage-specific transcriptional signatures distinct from normal colon tissue
- Early: strong activation of mitotic cell cycle, DNA replication, and complement / immune pathways
- Mid and late: increasing dysregulation of RNA processing and ribosome biogenesis, with extracellular matrix and adhesion remodeling

Future Directions:

★ Validate signatures in independent cohorts
★ Investigate functional roles of top dysregulated genes
★ Integrate transcriptomics with DNA methylation and copy-number data
★ Test whether these signatures can nominate stage-specific therapeutic targets

Impact

- Advances CRC molecular understanding
- Foundation for stage-informed biomarkers
- Enables personalized treatment strategies

Sources

American Cancer Society. (2023). Colorectal Cancer Facts & Figures 2023–2025. American Cancer Society.

Cancer Genome Atlas Network. (2012). Comprehensive molecular characterization of human colon and rectal cancer. Nature, 487(7407), 330–337.

Clevers, H., & Nusse, R. (2012). Wnt/β-catenin signaling and disease. Cell, 149(6), 1192–1205.

De Sousa E Melo, F., Wang, X., et al. (2013). Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. Nature Medicine, 19(5), 614–618.

Guinney, J., Dienstmann, R., et al. (2015). The consensus molecular subtypes of colorectal cancer. Nature Medicine, 21(11), 1350–1356.

Jones, P. A., & Baylin, S. B. (2007). The epigenomics of cancer. Cell, 128(4), 683–692.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 15, 550.

Rawla, P., Sunkara, T., & Barsouk, A. (2019). Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. Przegląd Gastroenterologiczny, 14(2), 89–103.

Siegel, R. L., Miller, K. D., & Goding Sauer, A. (2023). Colorectal cancer statistics, 2023. CA: A Cancer Journal for Clinicians, 73(3), 233–250.

Toyota, M., et al. (1999). CpG island methylator phenotype in colorectal cancer. Proceedings of the National Academy of Sciences, 96(15), 8681–8686.

Vogelstein, B., Fearon, E. R., et al. (1988). Genetic alterations during colorectal-tumor development. New England Journal of Medicine, 319, 525–532.

Vousden, K. H., & Prives, C. (2009). Blinded by the light: The growing complexity of p53. Cell, 137(3), 413–431.