

Transcriptomic Analysis to Identify Links Between the Consensus Molecular Subtypes of Colorectal Cancer and Obesity

Peter Abraham, Michael Greene, Elizabeth Lipke, and Peyton Kuhlert

According to the World Health Organization, over 650 million adults age 18 and older were obese in 2016 [1]. Within the United States, the prevalence of obesity has increased from 30.5% in 1999-2000 to 42.4% in 2017-18[2]. Furthermore, obesity has been linked to many diseases such as cardiovascular disease, diabetes, and cancer. Thus, it is important to understand the role of obesity in developing these health outcomes. This project focused specifically on colorectal cancer (CRC), which has been shown to be associated with obesity. However, CRC is a heterogeneous disease that differs genetically from patient to patient, resulting in a variety of patient outcomes. Recently, CRC was subdivided into four molecular subtypes, formally called the consensus molecular subtypes (CMSs), based on distinguishing genomic characteristics [3]. The primary goal of this work was to examine the links between the consensus molecular subtypes of colorectal cancer and obesity utilizing transcriptomic data.

Transcriptomic data were acquired from The Cancer Genomic Atlas (TCGA), an online cancer genomics database generated by the TCGA Research Network that contains RNA-sequencing samples and associated patient information. For this project, CRC RNA-sequencing data were utilized along with patient height and weight data from which body mass index (BMI) was calculated. Patients were then categorized as normal, overweight, or obese. The RNA-sequencing samples were classified into the CMSs using the CMS caller package [4] available in the R programming environment.

With the samples categorized as normal, overweight, or obese and as CMS1-4, the data could be analyzed further. First, Gene Set Enrichment Analysis (GSEA), a software developed by the Broad Institute, was utilized to examine whether obesity was affecting known biological pathways differently between the four CMSs

[5]. It was observed that inflammatory pathways such as TNF-alpha signaling were enriched in obese samples compared to normal samples in CMS1, the immune-infiltrated subtype, while mesenchymal-associated processes such as epithelial mesenchymal transition were enriched in obese samples compared to normal samples in CMS4, the mesenchymal subtype. These results suggest that obesity has CMS-specific effects on CRC.

Using the R package DESeq2, obese and normal samples were compared for each CMS, and the top significantly upregulated differentially expressed genes were utilized to create CMS-specific obesity-associated gene signatures [6]. With PROGgene V2, an online prognostics software [7], these gene signatures were evaluated, and high expression was found to result in significantly decreased survival in certain cohorts, suggesting that these genes may have prognostic value for CRC patients in a clinical setting.

In conclusion, the results indicate that obesity enhances distinct CMS characteristics such as the immune-infiltration of CMS1 tumors and the mesenchymal nature of CMS4 tumors. Obesity-associated genes for each CMS were also identified and found to have prognostic value. Taken together, obesity has CMS-specific effects on the CRC tumor transcriptome. Further analyses are underway to uncover the effect of obesity on the CMSs with respect to drug sensitivity. Future work will aim to validate these results in another CRC cohort.

Statement of Research Advisor

Peter has performed key analyses demonstrating that obesity affects CRC in a consensus molecular subtype specific manner. Having planned an experimental project that had to be postponed due to COVID, Peter adapted his plans quickly and learned an entirely new skill set to take on this research. Focused and persistent, Peter developed his understanding of the field through

reading the literature and has strong problem-solving abilities. His contribution was critical to elucidating this novel understanding of how obesity is linked to CRC.

–*Michael Greene, Nutrition, Dietetics, and Hospitality Management, and Elizabeth Lipke, Chemical Engineering*

References

1. Obesity and overweight. 2020 [cited 2021 May 19]; Available from: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>.
2. Craig M. Hales, M.D.C., Cheryl D. Fryar, Cynthia L. Ogden. Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017-2018. 2020 February 27, 2020 [cited 2021 May 19]; Available from: <https://www.cdc.gov/nchs/products/databriefs/db360.htm>.
3. Guinney, J., et al., The consensus molecular subtypes of colorectal cancer. *Nature Medicine*, 2015. 21(11): p. 1350-1356.
4. Eide, P. W., et al., CMS caller: An R package for consensus molecular subtyping of colorectal cancer pre-clinical models. *Scientific Reports*, 2017. 7(1): p. 1-8.
5. Subramanian, A., et al., Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 2005. 102(43): p. 15545-15550.
6. Love, M. I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 2014. 15(12).
7. Goswami, C.P. and H. Nakshatri, PROGgene V2: enhancements on the existing database. *BMC Cancer*, 2014. 14(1): p. 970.