Identifying and Characterizing Genetic Variants Associated with Colorectal Cancer

Medha Kaul1,2, Yao Yu, PhD3, Ryan Bohlender, PhD2, Chad Huff, PhD2

Johns Hopkins University Bloomberg School of Public Health1, Baltimore, MD Department of Epidemiology, The University of Texas MD Anderson Cancer Center2, Houston, TX

Background

- Colorectal cancer (CRC) is the third most common cancer diagnosed in both men and women in the US1.
- Several genes are known to affect CRC risk, but they only explain a small proportion of the disease heritability2.
- Limited representation of diverse groups in MDA and UKB may exacerbate disparities in understanding genetic basis of CRC.
- These disparities will be explored through African, Asian, and Hispanic representation in AOS.

Methods

- Whole-exome and whole-genome sequencing of CRC cases and controls
- Generate variant-based odds ratios
- Meta-analyze ACAT-combined p values from MDA, UKB, and AOS
- Meta-analyze p values from MDA, UKB, and AOS
- Meta-analyze variant-based odds ratios
- Generate ORs for each study using logistic regression
- Meta-analyze ORs from MDA, UKB, and AOS using METAL (weighted by square root of effective sample size)
- Calculate percent of familial relative risk explained using logistic regression
- Use OR estimates and carrier counts to calculate the percent of log familial relative risk explained by the identified variants in each gene
- Combine VAAST and CMC p values

Results

Table 1. Results of gene-based association analyses: VAAST and CMC p values from MDA, UKB, and AOS. Based on the meta-analysis results, the most significant established CRC genes and nominally significant a priori candidates with previous germline evidence for CRC were included in the table. Significant p values are bolded (genome-wide or nominal). P values with a ≤ sign are the smallest obtainable value (Medha Kaul, Shine Chang, Ph.D., Principal Investigator).

Table 2. ACAT-combined p values from MDA, UKB, and AOS and meta-analysis p values. Based on the meta-analysis results, the most significant established CRC genes and nominally significant a priori candidates with previous germline evidence for CRC were included in the table. Significant p values are bolded (genome-wide or nominal). Genes with an asterisk reached genome-wide significance. P values with a ≤ sign are the smallest obtainable value (Medha Kaul, Shine Chang, Ph.D., Principal Investigator).

Table 3. Meta-analyzed variant-based odds ratios and percent of log familial relative risk explained. Based on the meta-analysis results, the most significant established CRC genes and nominally significant a priori candidates with previous germline evidence were included in the table. Significant ORs are bolded. (Medha Kaul, Shine Chang, Ph.D., Principal Investigator).

References


Acknowledgements: This work was supported by NIH/NCI R25CA056452 (Medha Kaul, Shine Chang, Ph.D., Principal Investigator). Thank you to my PI, Dr. Chad Huff, and my lab team members Yao and Ryan for all the guidance and support on this project.