Rho GTPase Activating Protein 15 (ARHGAP15) increases cancer risk in patients with diverticular disease

Steven Lehrer, MD
Department of Radiation Oncology
Icahn School of Medicine at Mount Sinai
New York

Peter H. Rheinstein, MD, JD, MS
Severn Health Solutions
Severna Park, Maryland

correspondence to Dr. Steven Lehrer, Box 1236 Radiation Oncology, Mount Sinai Medical Center, 1 Gustave L. Levy Place, New York 10029 or steven.lehrer@mssm.edu

Keywords: diverticulosis, colon, cancer

Conflicts of interest: none

Author contributions: Dr. Lehrer and Dr. Rheinstein contributed equally to the conception, writing, and data analysis of this study.

Funding: This work was supported in part through the computational resources and staff expertise provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai. Research reported in this paper was also supported by the Office of Research Infrastructure of the National Institutes of Health under award numbers S10OD018522 and S10OD026880. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data availability: Data in this study available from UK Biobank https://www.ukbiobank.ac.uk/ after their approval of a formal application.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Introduction: Patients with diverticular disease who have colorectal histopathology, normal or abnormal, have increased risk of overall incident cancer. To determine whether the cancer risk might have a genetic basis, we performed a Genome Wide Association Study (GWAS) with data from UK Biobank.

Methods: We used PLINK 2, a whole-genome association analysis toolset, to analyze the UKB chromosome files. To identify cases, we used ICD10 code K57.3, diverticular disease of large intestine without perforation or abscess. We selected subjects younger than 45, since genetic forms of disease tend to manifest in younger patients.

Results: GWAS summary (Manhattan) plot of the meta-analysis association statistics, highlighting one susceptibility locus, Rho GTPase Activating Protein 15 (ARHGAP15), with genome wide significance for diverticulosis.

Conclusion: We conclude that ARHGAP15 could be responsible for the association of diverticular disease and cancer.
Ma et al. studied cancer risk in diverticular disease and reported that patients with diverticular disease who have colorectal histopathology, normal or abnormal, have increased risk of overall incident cancer [1]. The cancers included colorectal cancer, lung cancer, pancreatic cancer, and liver cancer. To determine whether the cancer risk might have a genetic basis, we performed a Genome Wide Association Study (GWAS) with data from UK Biobank [2].

Methods

We used PLINK 2, a whole-genome association analysis toolset, to analyze the UKB chromosome files. We adhered to recommended quality control procedures [3] that consisted of the following:

1) Missingness of SNPs 0.05: This command excluded SNPs that are missing in a large proportion of the subjects. In this step, SNPs with low genotype calls were removed.
2) Missingness of individuals 0.05: This command excluded individuals who had high rates of genotype missingness. In this step, individuals with low genotype calls were removed.
3) Hardy Weinberg equilibrium 1e-6: This command excluded markers which deviate from Hardy–Weinberg equilibrium.
4) Minor allele frequency (MAF) threshold 0.01: This command included only SNPs above the set MAF threshold.

To identify cases, we used ICD10 code K57.3, diverticular disease of large intestine without perforation or abscess. We selected subjects younger than 45, since genetic forms of disease tend to manifest in younger patients [4]. We analyzed data from 1828 subjects, 50.5% women and 49.5% men.

Results

GWAS summary (Manhattan) plot of the meta-analysis association statistics, highlighting one susceptibility locus, Rho GTPase Activating Protein 15 (ARHGAP15), with genome wide significance for diverticulosis, is shown in figure 1A. The red line indicates the genome wide significance threshold of a p value less than 5×10^-8.

LocusZoom plot of ARHGAP15 association is in figure 1B. Genomic position is depicted on the x-axis. The left y-axis shows the -log10 of the p-value. SNPs are colored based on their correlation (r^2) with the labeled top SNP, rs7607879 (purple diamond), which has the smallest p value in the region. The fine-scale recombination rates estimated from 1000 Genomes (EUR) data (right y axis) are indicated by the fluctuating blue line. The position of ARHGAP15 relative to rs7607879 is displayed.

The SNP rs7607879 is an intron variant, position chr2:143532194, alleles C>T, reference allele (C) frequency 0.2.

Discussion

ARHGAP15 inhibits PAK kinase. The regulation of cell motility, survival, metabolism, cell cycle, proliferation, transformation, stress, inflammation, and gene expression are just a few of the critical
activities that PAK kinases affect. PAK kinase dysregulation interferes with cellular homeostasis and negatively affects essential cell activities, many of which are linked to a variety of human disorders, including cancer.

ARHGAP15 is associated with diverticular disease and expressed in the colon [5]. ARHGAP15 may be a tumor suppressor during colorectal cancer progression [6] and regulates lung cancer cell proliferation and metastasis via the STAT3 pathway [7].

We conclude that ARHGAP15 could be responsible for the association of diverticular disease and cancer that Ma et al reported [1].
Figure 1 (A) GWAS Summary (Manhattan) Plot of the meta-analysis association statistics highlighting one susceptibility locus (ARHGAP15) on chromosome 2q22 with genome wide significance for colon diverticula. The red line indicates the genome wide significance threshold of a p value less than $5 \times 10^{-8}$.  
(B) LocusZoom plot of ARHGAP15 association. Genomic position is depicted on the x-axis. The left y-axis shows the $-\log_{10}$ of the p-value. SNPs are colored based on their correlation ($r^2$) with the labeled top SNP, rs7607879 (purple diamond), which has the smallest p value in the region. The fine-scale recombination rates estimated from 1000 Genomes (EUR) data (right y axis) are indicated by the fluctuating blue line. The position of ARHGAP15 relative to rs7607879 is displayed.
References


