Identification of Genetic Biomarkers for Colorectal Cancer Risk in Inflammatory Bowel Disease in Patients

Arav Shelat[1], Francesca Perrone, Ph.D.[2]

[1]Jericho High School, Jericho, New York, USA
[2] Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK

Abstract

Inflammatory bowel disease (IBD) is a chronic condition associated with an increased risk of colorectal cancer (CRC), driven by inflammation-induced genetic mutations. This study identifies genetic biomarkers linking IBD to CRC and explores their functional roles in cancer progression. Publicly available datasets were analyzed, including whole-genome sequencing data from IBD patients (Olafsson et al., 2020) and CRISPR-Cas9 knockout screens from the Cancer Dependency Map (DepMap). Somatic mutations from 446 colonic crypts of IBD patients and 412 non-IBD controls were assessed and cross-referenced with The Cancer Genome Atlas (TCGA) to determine oncogenic significance. Functional relevance was evaluated through dependency scores derived from CRISPR knockout screens of 61 CRC-relevant cell lines, with statistical analyses including Fisher's Exact Test and multivariate regression models. Seven genes—ARID2, FBXW7, IL17RA, MYH11, PIGR, ZC3H12A, and ZNF521—were identified as high-priority biomarkers. These genes exhibited higher mutation rates in IBD patients and played critical roles in CRC cell survival. ARID2 and FBXW7, involved in chromatin remodeling and cell cycle regulation, showed synergistic effects on cancer progression. IL17RA and ZC3H12A contributed to immune modulation and inflammation-driven oncogenesis. This study highlights

molecular mechanisms linking IBD to CRC and explores the influence of epigenetic and environmental factors. The identified biomarkers offer potential targets for therapeutic interventions, aiding in personalized approaches for CRC prevention and treatment in high-risk IBD populations.

Keywords: IBD, colorectal cancer, biomarkers, CRISPR

Introduction

Inflammatory Bowel Disease (IBD) is a chronic inflammatory condition of the gastrointestinal tract, comprising two primary subtypes: Crohn's disease (CD) and ulcerative colitis (UC). IBD affects millions of people worldwide, with the highest prevalence in North America and Europe and increasing incidence in newly industrialized regions due to changes in diet, lifestyle, and environmental factors (Kaplan, 2015; Ng et al., 2017). Although IBD can occur at any age, it is most commonly diagnosed between 15 and 35 years, with distinct clinical presentations observed in pediatric cases. Symptoms include abdominal pain, diarrhea, and nutrient malabsorption, alongside systemic complications such as an elevated risk of colorectal cancer (CRC), particularly in cases of long-standing, untreated, or severe inflammation. This condition significantly impacts the quality of life due to its cycles of relapse and remission. (Canavan et al., 2006).

Treatments are aimed at symptom relief and induction of remission and vary based on disease severity, location, and patient response. For mild to moderate UC or CD, aminosalicylates such as mesalamine are effective (Feuerstein et al., 2020). Severe cases often require corticosteroids like prednisone for flare control, though these are unsuitable for long-term use due to side effects (Harbord et al., 2017). Immunomodulators like azathioprine or biologic agents, including anti-TNF therapies (infliximab and adalimumab), are commonly employed for

maintenance of remission (Lichtenstein et al., 2018, Sandborn et al., 2017). Emerging treatments, including Janus kinase (JAK) inhibitors, such as tofacitinib, and sphingosine-1-phosphate receptor modulators, are expanding options for patients with refractory disease (Panés et al., 2020). Surgical intervention is reserved for patients with complications or refractory disease (Baumgart & Sandborn, 2012). The emergence of small-molecule drugs and personalized medicine has enabled tailored treatments based on individual profiles. However, challenges in managing refractory cases and preventing complications underscore the need for continued research in IBD therapeutics.

Colorectal cancer (CRC) is the third most diagnosed cancer globally and the second leading cause of cancer-related deaths (Sung et al., 2021). It often arises from polyps that progress to malignancy, driven by genetic mutations in genes such as APC, KRAS, TP53, and mismatch repair (MMR) genes. These mutations disrupt cell cycle regulation and DNA repair, fostering malignant tumor development (Markowitz & Bertagnolli, 2009). Treatment for CRC depends on the stage of the cancer and surgery is the primary treatment for localized CRC, aiming to remove tumors and affected lymph nodes, whereas chemotherapy is typically used in advanced stages to reduce recurrence risk (Van Cutsem et al., 2016). Targeted therapies, including inhibitors of epidermal growth factor receptor (EGFR) and vascular endothelial growth factor (VEGF), are effective for metastatic CRC. Immunotherapy, particularly immune checkpoint inhibitors, shows promise for CRC cases with high microsatellite instability (MSI-high) (Le et al., 2017).

IBD is associated with an increased risk of CRC by two to three times compared to the general population, particularly in patients with long-standing and extensive disease (Jess et al., 2012). Chronic inflammation in the intestinal epithelium promotes DNA damage, epigenetic

alterations, clonal expansion of mutated cells, and subsequent neoplastic transformation due to genetic mutation accumulation (Axelrad et al., 2016).

Advancements in genetic research, especially the development of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), have revolutionized the study of gene function and disease mechanisms. CRISPR employs a guide RNA (gRNA) to direct the Cas9 protein to a specific DNA sequence. Cas9 then creates a double-strand break, which can be repaired by cellular mechanisms in ways that allow for gene disruption, replacement, or insertion (Jinek et al., 2012; Cong et al., 2013). CRISPR has become a powerful tool in cancer research, enabling identification of genes involved in cancer progression, particularly in inflammation-driven cancers like CRC (Doudna & Charpentier, 2014).

Among CRISPR approaches, CRISPR knock-out (KO) is applied to disrupt specific genes to elucidate their functional roles in cellular behavior. In this context, the complex CRISPR-Cas9 makes a double-strand break in the target gene. The cell attempts to repair the break, but the repair process is error-prone, introducing insertions or deletions at the break site, which can lead to frameshift mutations that disrupt the normal function of the gene. These mutations typically result in a loss of function, effectively "knocking out" the gene. For example, CRISPR-KO studies targeting genes associated with cellular proliferation can identify key drivers of tumor growth that can be used as biomarkers (Shalem et al., 2014, Feero et al., 2010).

Integrating tools like CRISPR and bioinformatics we can enhance our understanding of the genetic factors underlying CRC risk in IBD. Therefore, we leveraged publicly available CRISPR-KO screen datasets, alongside with datasets highlighting genes frequently mutated in IBD and CRC patients, to identify potential genes linking the IBD condition to the risk of developing CRC. Specifically, we identified genes shared between IBD and CRC patients and used CRISPR-KO data to investigate their roles in cancer promotion. By identifying actionable

biomarkers, these approaches can inform personalized strategies to reduce CRC risk in this high-risk population, bridging the gap between scientific discovery and clinical application.

Methods

Datasets and sample information

IBD dataset

IBD-related mutations were extrapolated from the dataset published by Olafsson et al. (2020), where they analyzed somatic mutations in non-neoplastic colonic tissue from 46 IBD patients, including 28 diagnosed with UC and 18 with CD. This cohort included 52% males and 48% females, ranging in age from 21 to 73 years. For the study, a total of 446 colonic crypts from IBD patients and 412 crypts from non-IBD controls were isolated from routine endoscopic biopsies using laser capture microdissection (LCM), which ensured precise capture of crypts while minimizing contamination from stromal or immune cells. The isolated crypts underwent whole-genome sequencing to detect somatic mutations, clonal structures, and genes under positive selection. Data are accessible via the European Genome-phenome Archive under accession numbers EGAD00001006061 for IBD patients and EGAD00001004192 and EGAD00001004193 for controls.

Notably, this dataset does not include information on potential environmental exposures (e.g., diet, medications, or toxin exposure) or epigenetic modifications (e.g., DNA methylation), limiting our ability to account for these factors in our analyses. Furthermore, demographic metadata (beyond basic age and sex) and detailed clinical variables, such as disease duration or

severity, were not consistently available. As a result, any evaluation of non-genetic contributors to disease pathogenesis was necessarily excluded from this study.

CRISPR screen dataset

Validation of the interesting genes was performed by using datasets from the Broad Institute's Cancer Dependency Map (DepMap, release 24Q2), which provides results from high-throughput CRISPR-Cas9 gene-editing screens. A genome-wide single-guide RNA (sgRNA) library targeting all protein-coding genes was introduced into various cancer cell lines (see Table 1) via transduction. After allowing 2–3 days for recovery and expansion, samples were collected to capture the initial distribution of sgRNAs. Final samples were then collected after a 14–21 day culture period. Genomic DNA was extracted from the collected cells at each time point using phenol-chloroform extraction or commercial DNA purification kits (Broad Institute, 2019).

The extracted DNA underwent PCR amplification to enrich sgRNA sequences, followed by the addition of sequencing adapters and barcodes for multiplexing. High-throughput sequencing, typically performed on Illumina platforms, determines the abundance of each sgRNA in the population, with sequencing depth optimized to provide comprehensive coverage of the sgRNA library and ensure accuracy in representing the effects of gene knockouts (Broad Institute, 2019).

This study analyzed colorectal cancer (CRC)-relevant cell lines available on DepMap to evaluate the functional relevance of seven genes of interest. Out of 90 CRC-relevant cell lines, 61 were selected because they contained knockout data for the genes of interest isolated in the first part of the analysis. Dependency scores obtained from DepMap were used to identify genes critical for cancer cell survival. Dependency scores are a quantitative measure of how essential a gene is for the survival of a given cell line. The scores are calculated by comparing the

abundance of single-guide RNAs (sgRNAs) at the final time point to their initial baseline abundance. This calculation involves several steps, including normalizing raw sequencing reads to account for library composition and sequencing depth, calculating the log-fold change in sgRNA abundance, and aggregating sgRNA data to generate gene-level scores. Statistical models, such as the CERES algorithm, are used to adjust for confounding factors like copy number variations, ensuring accuracy. Dependency scores range from -1 to +1: scores near -1 indicate high dependency, meaning that the gene is essential for cell survival; scores around 0 suggest the gene is non-essential; and scores closer to +1 may indicate negligible impact or even a slight growth advantage upon gene knockout (Broad Institute, n.d.-a; Broad Institute, n.d.-b).

To ensure robustness and reproducibility, multiple replicates were included for each cell line (Table 1). Each gene was analyzed individually to assess its impact on cell viability across the selected cell lines. Genes with dependency scores near -1, indicating high dependency and essentiality for cell survival, were prioritized for further study. Additionally, interaction models were developed to investigate potential additive or synergistic effects of gene knockouts. These models revealed gene networks influencing CRC progression in the context of IBD, uncovering complex biological interactions.

Table 1

DepMap Cell Line ID with Corresponding Cell Line Name Selected for the 7 Genes of Interest

Cell Line ID	Cell Line
ACH-000007	LS513
ACH-000009	C2BBE1
ACH-000202	COLO320
ACH-000249	CL11
ACH-000252	LS1034
ACH-000253	COLO201
ACH-000286	SNU1033
ACH-000296	OUMS23
ACH-000350	COLO678
ACH-000381	T84
ACH-000403	NCIH747
ACH-000421	SW837
ACH-000467	HCC56
ACH-000470	SW1463
ACH-000489	SW1116
ACH-000491	NCIH716
ACH-000501	LS123
ACH-000532	SNU61
ACH-000552	HT29
ACH-000565	RCM1
ACH-000651	SW620
ACH-000683	SNU503
ACH-000722	SNUC1
ACH-000798	CL40
ACH-000820	SW403
ACH-000926	HT55
ACH-000935	MDST8
ACH-000943	RKO
ACH-000950	LOVO

Cell Line ID	Cell Line
ACH-000957	LS180
ACH-000958	SW48
ACH-000959	SNUC4
ACH-000963	CCK81
ACH-000969	KM12
ACH-000970	SNUC5
ACH-000971	HCT116
ACH-000985	LS411N
ACH-000986	HT115
ACH-000991	SNU81
ACH-000997	HCT15
ACH-001039	CO205
ACH-001061	DLD1
ACH-001081	HCC299
ACH-001345	GP5D
ACH-001399	SW626
ACH-001454	C10
ACH-001458	C75
ACH-001459	C80
ACH-001460	C84
ACH-001461	C99
ACH-001786	SNU1544
ACH-002024	ECC4
ACH-002025	TT1TKB
ACH-002233	DIFI
ACH-002535	SNU254
ACH-002654	JVE015
ACH-002659	JVE126
ACH-002660	JVE187
ACH-002662	SC002662
ACH-002664	JVE253
ACH-002669	KP363T

Pathway analysis

To gain further insights into the biological roles of these genes, pathway analysis was conducted using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases.

Statistical analysis

Statistical analyses were conducted using Python (version 3.9). Data processing and organization were performed using pandas, facilitating the preparation of dependency scores, cell line metadata, and gene annotations for analysis. Colorectal cancer cell lines were selected and analyzed using data from the DepMap database (Broad Institute, n.d.-a). Small bowel cancer cell lines were initially considered; however, as no data were available for the genes of interest, they were automatically excluded from the analysis. Dependency scores, mutation rates, and gene interactions were the primary metrics used to evaluate gene function and pathway relevance. Statistical tests were executed using *scipy*, including Fisher's Exact Test to analyze mutation rates between the control and test groups. This test was selected due to its suitability for categorical data with small sample sizes or low mutation frequencies, enabling precise evaluation of the null hypothesis that mutation rates were independent of group membership.

Correlation analyses (e.g., Pearson and Spearman rank correlations) and regression models were applied to assess relationships between dependency scores and pathway involvement. Multivariate interaction models were employed to evaluate the effects of multiple gene knockouts on cell viability. A one-sample t-test was used to analyze dependency scores for gene combinations, as advised by a consulting statistician. Statistical significance was defined as $p < 0.05$, with adjustments such as the Benjamini-Hochberg method applied to control for

multiple comparisons where necessary. For the gene interaction analysis, combinations of up to six genes were tested, resulting in 119 possible interactions. Only gene combinations with a mean dependency score less than or equal to -0.1 and a p-value below 0.05 were retained for further analysis.

Data visualizations were created using matplotlib and seaborn. These included histograms for dependency score distributions, box plots comparing scores across pathways, bar charts summarizing pathway enrichment, and network diagrams illustrating gene-gene interactions. These tools facilitated a clear presentation of the data, ensuring reproducibility and comprehensive interpretations of the relationships among dependency scores, pathway involvement, and gene interactions.

Processed interaction data are provided in Supplementary Tables S3–S5

Ethical considerations

This study exclusively relied on publicly available datasets from Olafsson et al. (2020), TCGA, and DepMap, complying with their respective terms of use. No experimental data involving human or animal subjects were generated, and all analyses were conducted on de-identified data. Consequently, ethical approval was not required.

Results

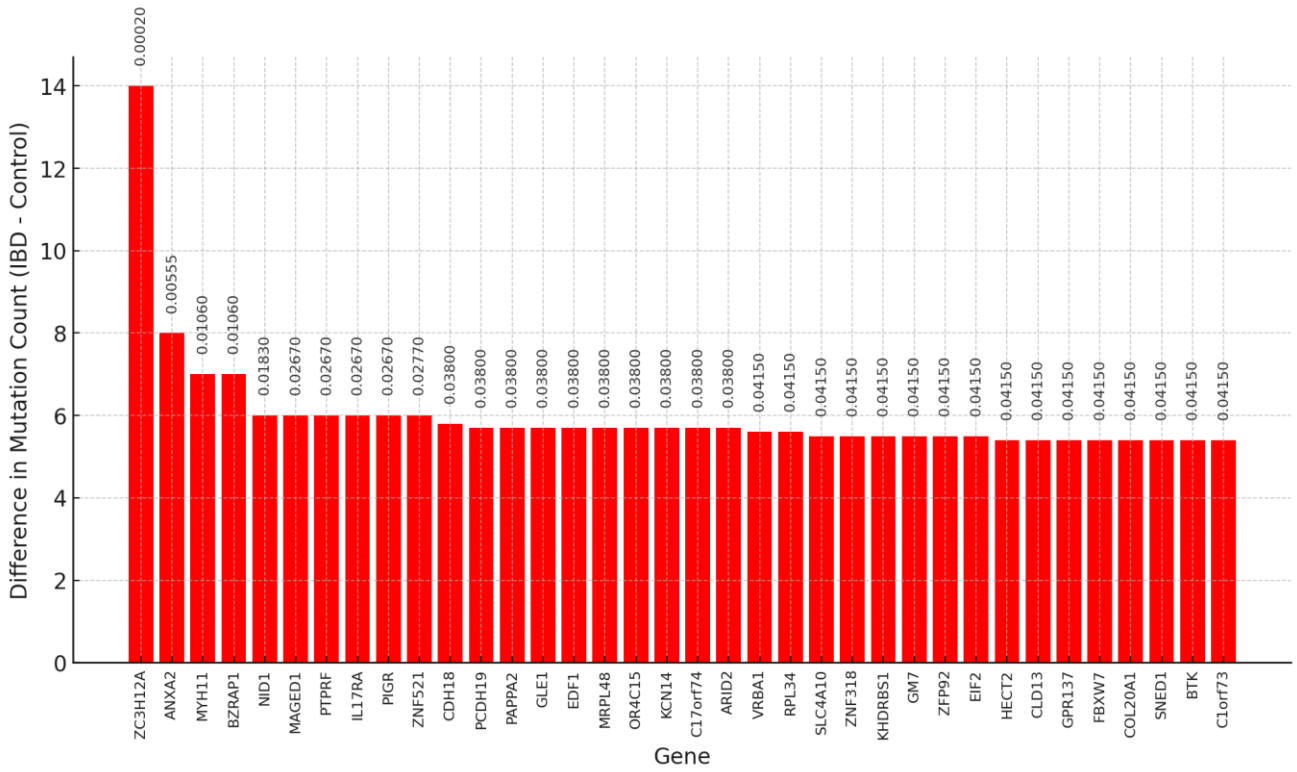
Identifying High-Priority Biomarkers Linking IBD to CRC Progression

This study employed a multi-faceted approach to identify and analyze genetic biomarkers linking inflammatory bowel disease (IBD) to colorectal cancer (CRC) progression. Using the

Olafsson dataset, 39 candidate genes were initially identified based on statistically significant differences in mutation rates between IBD patients and controls (Figure 1). Metrics such as dN/dS ratios, which measure the ratio of nonsynonymous to synonymous mutations, and mutation hotspots indicative of positive selection were employed. These 39 genes were further refined by cross-referencing with The Cancer Genome Atlas (TCGA) to identify those classified as cancer hotspots (frequently mutated) or drivers (functionally significant in cancer progression, Figure 2). This process narrowed the list to seven high-priority genes shared between IBD and CRC: ARID2, FBXW7, IL17RA, MYH11, PIGR, ZC3H12A, and ZNF521.

Figure 1

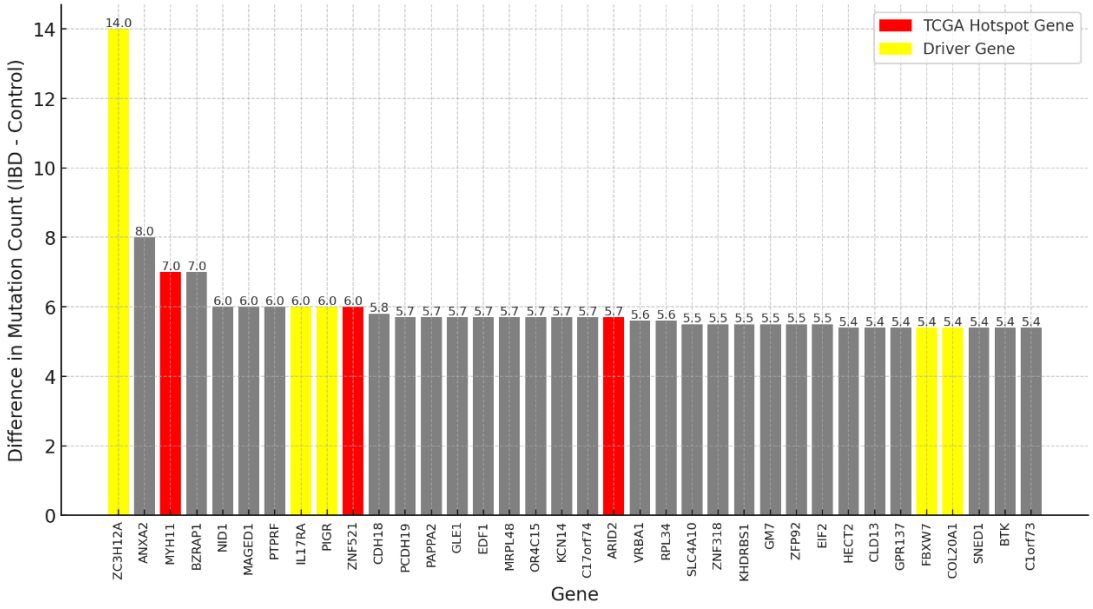
Genes with statistically significant ($p<0.05$) higher mutations rate in IBD group versus control group



Note: The bar plot illustrates the differences in mutation counts between Inflammatory Bowel Disease (IBD) cases and controls for genes exhibiting significantly higher mutation rates in IBD group. Each bar represents the difference in mutation count (IBD-Controls) for a gene, with statistical significance indicated by the respective p-values. The genes are listed in the x-axis, while the y-axis reports the mutation counts. Only genes with p-values below 0.05 are included (statistical method; report the number of samples (N= (total number of samples, IBD=n; controls=n).

Figure 2

Genes with higher mutation rates in IBD Group with TGCA Hotspots and Drivers



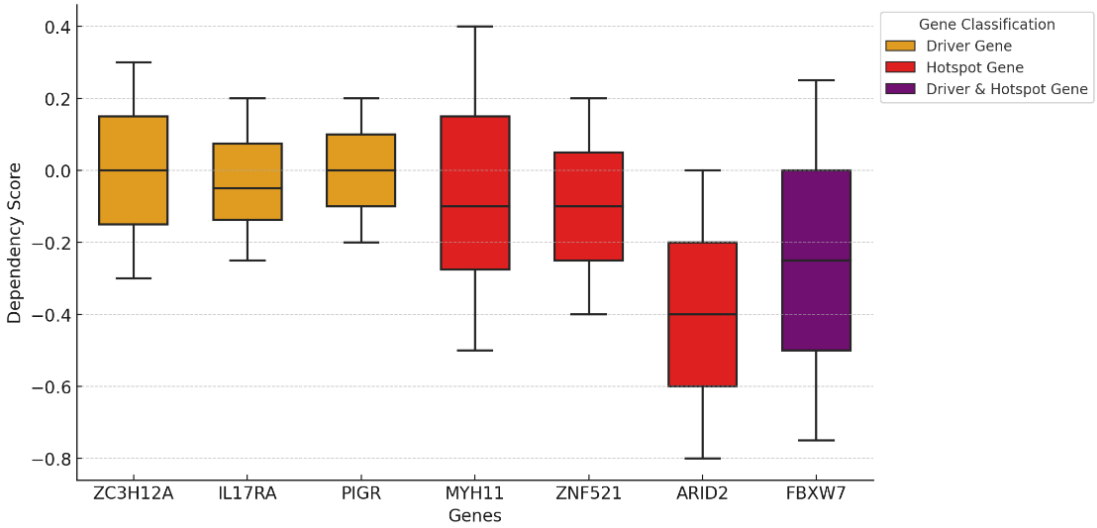
Note*:* Bar graph shows genes with statistically significant higher mutation rate in IBD patients with genes known to be TGCA hotspots and/or driver genes shown in color.

Mutation frequency analysis highlighted significant differences between control and IBD groups, as shown in Figures 1 and 2. In the control group, mutations predominantly occurred in genes such as TNFRSF10C, MSH6, and NID1, which are associated with DNA damage response and repair pathways. In contrast, the IBD group displayed significantly higher mutation frequencies in these genes, as well as additional mutations in IL17RA and ZC3H12A. Chronic inflammation in IBD patients likely creates a mutagenic microenvironment, amplifying genetic

instability and increasing the risk of oncogenic transformation. Further stratification of the mutations reveals additional differences between IBD and Controls patients (Figures 3 and 4) identifying genes such as ZC3H12A, MYH11, and IL17RA as disproportionately burdened by mutations ($p < 0.01$). Notably, genes identified as hotspots and drivers in TCGA, including FBXW7 and ARID2, exhibited similar trends, further underscoring the involvement of these mutated genes in both genomic stability and CRC progression.

Figure 3

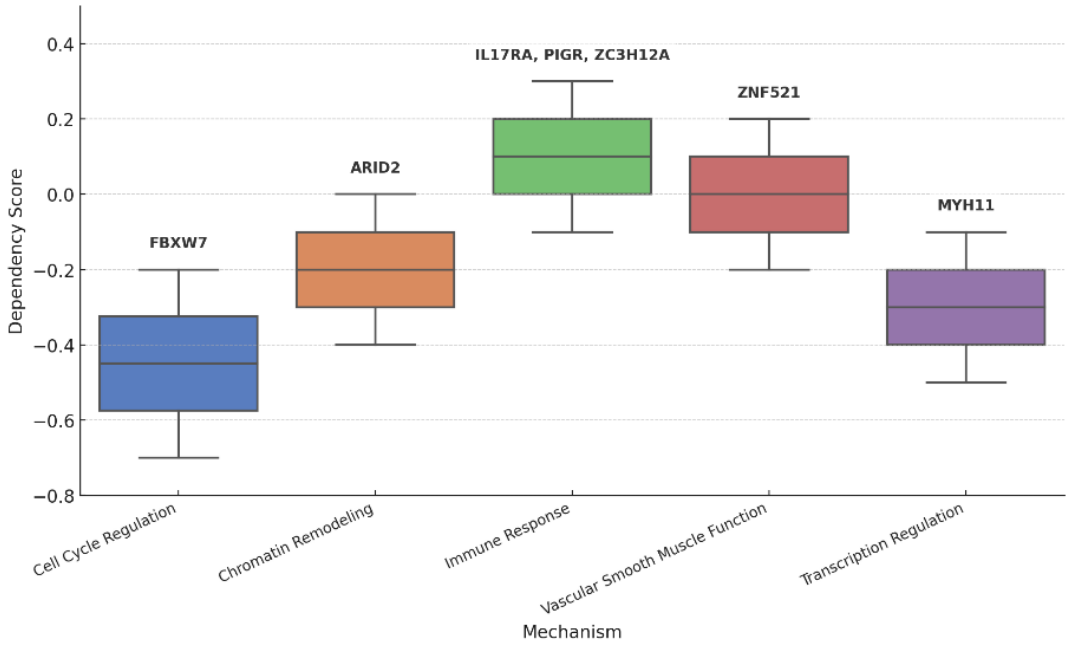
Gene Dependency Whisker Chart by Classification



Note: Whisker chart demonstrating dependency scores based on the primary pathway supported by each gene. Genes involved in Chromatin Remodeling or Cell Cycle Regulation are demonstrated to have dependency scores that suggest a critical role in cancer progression.

Figure 4

Dependency score range based on primary gene mechanism



Note: Whisker chart demonstrating the range of dependency scores in genes with higher mutation rates in IBD patients and are also TCGA Hotspot and/or Drive genes. The closer the dependency score is to -1.00, the more critical the mutation is toward cancer progression.

Linking Chronic Inflammation to CRC Progression in IBD

To gain insights into the functional significance of these genes, dependency scores derived from DepMap were analyzed (Broad Institute, n.d.-a). Of the 90 CRC-relevant cell lines available on DepMap, 61 were selected for this study because they contained knockout data for the seven genes of interest (Listed in Table 1). Dependency scores revealed that genes like FBXW7 and ARID2, which are classified as both drivers and hotspots in TCGA, exhibited strong negative scores near -0.5, indicating that their knockout significantly impairs cancer cell survival. This is likely due to their critical roles in chromatin remodeling and cell cycle progression, processes critical for maintaining tumor viability under inflammatory stress.

Conversely, genes like IL17RA, PIGR, and ZNF521, which are involved in immune modulation and transcriptional regulation processes, showed moderate dependency scores from +0.25 to -0.35. This suggests that their functional significance may be more pronounced in the context of IBD compared to the tumor microenvironment. In IBD, these genes primarily contribute to immune regulation, epithelial integrity, and the resolution of inflammation, which are central to disease pathology. In contrast, their roles in CRC extend to processes such as tumorigenesis and progression, where their contributions may be more context-dependent or secondary. The distinction likely reflects the divergent nature of the diseases: IBD arises from chronic inflammation driven by immune dysregulation, whereas the immune response in cancer is typically more selective and shaped by specific tumor-immune interactions. This difference underscores the importance of these genes in maintaining tissue homeostasis and regulating inflammation in IBD, which contrasts with their broader and less consistent involvement in CRC progression. MYH11 displayed a narrower range of dependency scores, likely reflecting its involvement in tumor microenvironment regulation rather than direct cancer cell survival.

Pathways Linking Inflammation to CRC in IBD

To understand how inflammation-linked genes cooperate to drive colorectal cancer (CRC) in inflammatory bowel disease (IBD), we analyzed every possible CRISPR knock-out combination of seven candidates—ARID2, FBXW7, IL17RA, ZC3H12A, ZNF521, MYH11 and PIGR—previously flagged for their high mutation frequency in IBD colonic crypts and for their strong single-gene dependency scores in 61 CRC cell lines. The 21 resulting double knock-outs are plotted in Figure 5. In this figure, the horizontal axis reports the average CRISPR dependency score, so points that sit further to the left correspond to combinations that impair cell

viability more severely, whereas the vertical axis displays $-\log_{10} p$, meaning that higher points are supported by stronger statistical evidence. A horizontal dashed line at $p = 0.05$ separates significant from non-significant effects, and a vertical dashed line at a mean CRISPR dependency value of roughly –0.19 (the fifth percentile of all pairwise scores) marks a stringent synthetic-lethal boundary. Only four double knock-outs occupy the upper-left quadrant delineated by these two guides, and therefore satisfy both statistical and biological cut-offs: ARID2 plus FBXW7, ARID2 plus IL17RA, ARID2 plus ZC3H12A, and ARID2 plus ZNF521. The most potent of these is the ARID2–FBXW7 combination, which shows an average CRISPR dependency score of about –0.23 with a p-value below 0.01 and thus links the failure of SWI/SNF chromatin remodeling (ARID2) with loss of an SCF-ubiquitin-ligase checkpoint that normally restrains MYC, Cyclin-E, and other proliferative drivers. The additional synergies between ARID2 and the immune-signaling genes IL17RA or ZC3H12A, as well as with the stem-cell transcriptional repressor ZNF521, suggest that chronic cytokine stress and an undifferentiated transcriptional state intensify the genomic instability initiated by chromatin-remodeling defects.

We next asked whether combinations involving three, four, or five genes could expose still deeper vulnerabilities. Accordingly, we screened all thirty-five possible triplets, thirty-five quadruplets, and twenty-one quintuplets. The strongest triplet, which couples ARID2 and FBXW7 with IL17RA, reached a mean CRISPR dependency score of roughly –0.18 with an exceptionally small p-value of 1.1×10^{-19}, yet it improved lethality by only about twenty percent relative to the ARID2–FBXW7 pair. The best quadruplets and quintuplets plateaued around –0.13 despite p-values in the 10^{-16} range, indicating diminishing biological returns beyond two lesions. Since these higher-order combinations add therapeutic complexity without a

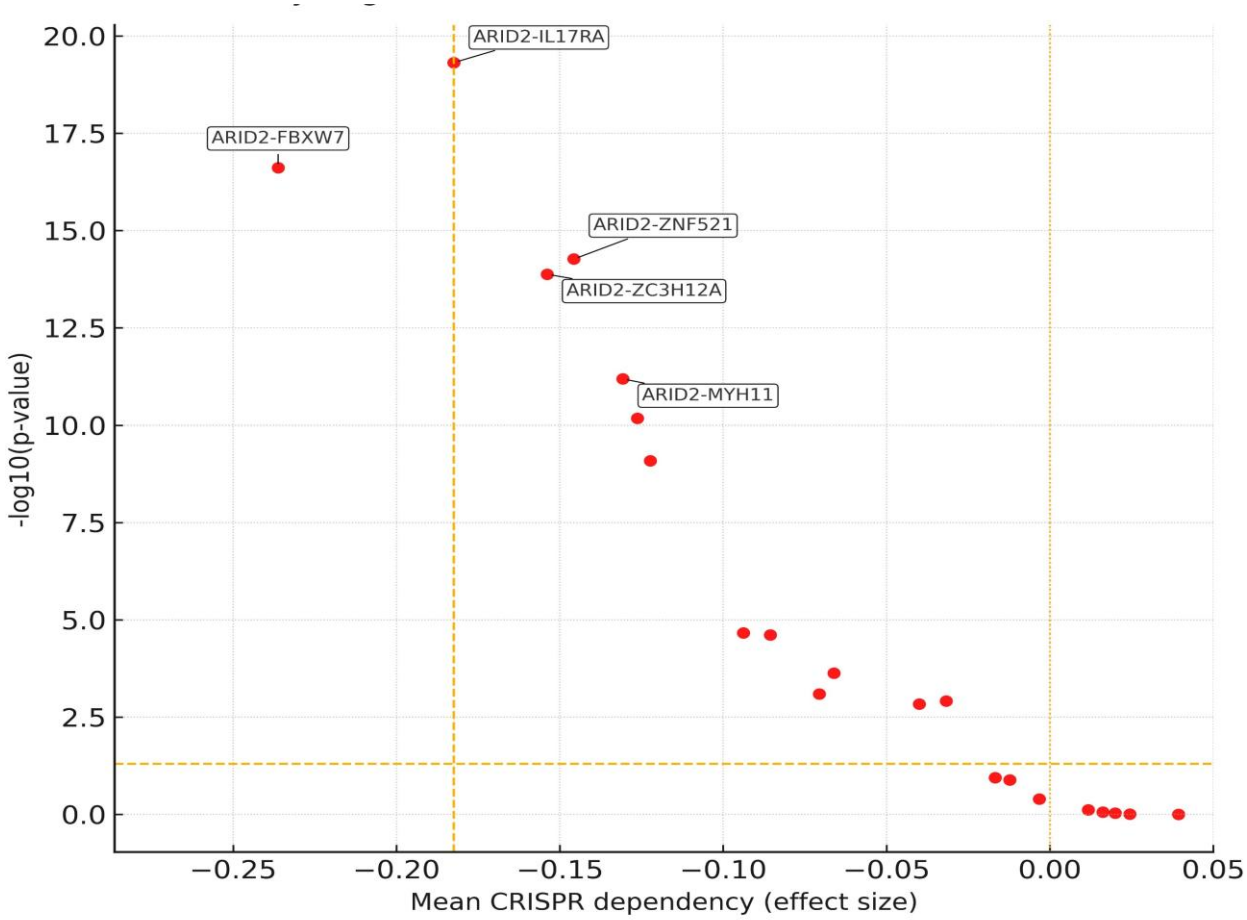
commensurate gain in effect size, we restricted the main discussion to the biologically and clinically tractable doubles shown in Figure 5 and present the complete pairs, triplet, quadruplet, and quintuplet data in Supplementary Tables S2–S5.

When mutation prevalence, single-gene dependency, and interaction results are considered together, five genes (FBXW7, ARID2, ZC3H12A, IL17RA, and MYH11) emerge repeatedly as high-priority drivers. ARID2 and FBXW7 form the core synthetic-lethal axis that disrupts chromatin integrity and cell-cycle surveillance. IL17RA and ZC3H12A inject persisting IL-17/NF-κB signaling into this vulnerability, while MYH11, which is enriched in IBD crypts and essential in DepMap, may remodel the tumor microenvironment through stromal or smooth-muscle interactions. Together, these observations support a model in which long-standing cytokine stress exacerbates chromatin-remodeling defects, collapses checkpoint control, and nudges epithelial cells toward an EMT-like progenitor state that favors malignant transformation.

Clinically, co-mutation or co-loss of ARID2 and FBXW7, alone or in combination with an IL-17 pathway signature, could help stratify IBD patients for intensified colonoscopic surveillance. Therapeutically, a synthetic-lethal strategy that pairs a SWI/SNF inhibitor with blockade of IL-17A or NF-κB may selectively eliminate premalignant colonic epithelium in high-risk individuals. Future studies in patient-derived organoids and murine colitis models will be needed to confirm whether interrupting this chromatin-immune axis delays or prevents tumor formation, but the present data position ARID2–FBXW7, modulated by IL-17 signaling, as a tractable target set for both risk stratification and drug development.

Figure 5

Synergistic Gene-Pair Knockout in CRC Cell Lines



Note: Graph shows how strongly each two-gene knock-out weakens colorectal-cancer cells (leftward shift on the x-axis) and how reliable that effect is (upward shift on the y-axis). The horizontal dashed line marks p = 0.05; any gene pair above it is statistically significant. The vertical dashed line at a mean CRISPR dependency score of about –0.19 separates mild from strong synthetic-lethal effects. Gene pairs that lie left of the vertical cut-off and above the horizontal line are our most promising synergistic targets.

Discussion

This study provides critical insights into the genetic underpinnings linking inflammatory bowel disease (IBD) to colorectal cancer (CRC), emphasizing the role of chronic inflammation in driving oncogenesis. By identifying seven key genetic biomarkers - ARID2, FBXW7, IL17RA, MYH11, PIGR, ZC3H12A, and ZNF521- the paper highlights their potential as critical markers for understanding and predicting the link between persistent inflammation and cancer risk. Through an innovative approach of combining mutation frequency analyses and CRISPR-Cas9 knockout data, this research bridges significant gaps in understanding inflammation-driven carcinogenesis. These findings not only expand the current scientific knowledge but also provide a foundation for improving diagnosis and developing targeted interventions that may mitigate CRC risk in IBD patients.

The study builds on prior research that established the mutagenic microenvironment in IBD tissues, such as the work by Olafsson et al. (2020) and the Cancer Genome Atlas (TCGA) Network (2012). While earlier studies focused on the genetic mutations associated with chronic inflammation, this research goes further by evaluating the functional importance of these mutations using dependency scores derived from CRISPR-Cas9 screens.

By integrating mutation data with CRISPR knockout screens, this study represents a methodological advancement that validates its findings while paving the way for future research. By combining genomic analyses with functional experiments, we were able to not only identify genes commonly mutated in IBD and CRC but also determine their importance for cancer cell survival. This approach strengthens the case for these genes as therapeutic targets and provides a roadmap for investigating other inflammation-driven cancers. For example, similar studies could be conducted in pancreatic, gastric, or liver cancers, where chronic inflammation is also a

significant risk factor. This methodology could uncover genetic drivers unique to these cancers, offering insights leading to more effective, disease-specific treatments.

The identification of FBXW7 and ARID2 as critical players in chromatin remodeling and cell cycle regulation underscores their essential roles in CRC survival. Similarly, the interaction between IL17RA and ZC3H12A highlights the influence of immune signaling pathways on tumor progression under conditions of inflammation-driven genetic instability. These findings demonstrate how chronic inflammation not only accelerates mutation accumulation but also creates a microenvironment conducive to tumorigenesis.

One of the most striking findings is the synergy observed between ARID2 and FBXW7, two genes deeply involved in chromatin remodeling and cell cycle regulation. These processes are critical for maintaining genomic stability and preventing the uncontrolled cell proliferation characteristic of cancer. The interconnected roles of these genes suggest that targeting their pathways simultaneously could significantly disrupt cancer cell survival, particularly in the inflammatory microenvironment unique to IBD. Therapies designed to inhibit these pathways could offer a twofold benefit: directly impairing cancer cell growth while also stabilizing the genetic environment of inflamed tissues and reducing the likelihood of further oncogenic mutations.

In addition to these central players, the study highlights genes like IL17RA and ZC3H12A, which appear to have more specialized roles in immune modulation. These genes contribute to the inflammatory cascade that links chronic IBD to an elevated risk of CRC. IL17RA, for instance, is a key component of the interleukin-17 signaling pathway, which is known to exacerbate inflammation and promote tumor progression under certain conditions. ZC3H12A, on the other hand, regulates the immune response by controlling the degradation of

inflammatory mRNA transcripts. The moderate dependency scores observed for these genes suggest that they are not strictly essential for cancer cell survival but play significant roles in shaping the tumor microenvironment. Therapeutics targeting these pathways could help "cool down" the inflammation associated with IBD, effectively reducing the tumor-promoting effects of the immune system in these patients.

Beyond individual gene identification, the study's exploration of clinically significant gene interactions provides an even deeper understanding of cancer progression. Indeed, this research highlights the interconnected nature of tumor biology and inflammation, providing a roadmap for developing innovative therapies that address both aspects simultaneously. The synergy between ARID2 and FBXW7 offers a promising target for dual-inhibition strategies, while the roles of IL17RA and ZC3H12A suggest new opportunities to modulate the inflammatory microenvironment. By focusing on how genes like ARID2 and FBXW7 interact, the research sheds light on the broader biological networks that govern tumor behavior. These networks often involve multiple, interconnected pathways, and targeting them collectively could yield better therapeutic outcomes than addressing single genes in isolation. The development of combination therapies that inhibit multiple pathways simultaneously could prove particularly effective for high-risk patients, such as those with long-standing IBD, by attacking the cancer on multiple fronts.

The methodology employed in this study could be applied to other cancers, enabling the discovery of novel genetic drivers and therapeutic targets. Ultimately, the insights gained from this work have the potential to transform cancer treatment, moving the field closer to a future where personalized, effective, and less invasive therapies are available for even the most complex diseases.

With the availability of genomic screening technologies, it may soon become possible to identify patients whose genetic profiles place them at a higher risk of developing CRC in the context of IBD. These patients could then receive tailored interventions targeting the specific pathways identified in this study, potentially preventing the onset of cancer altogether. For patients already diagnosed with CRC, personalized treatments based on their tumor's genetic makeup could improve survival rates and reduce treatment-related side effects.

The implications extend beyond clinical applications to the broader field of cancer biology. By demonstrating the value of integrating genomic data with CRISPR-based functional screens, this study sets a new standard for investigating complex diseases. The ability to not only identify mutations but also understand their functional consequences represents a significant leap forward. This approach could accelerate the pace of discovery across a wide range of diseases, enabling researchers to uncover actionable insights with unprecedented efficiency.

As the field of oncology continues to evolve, the findings from this study exemplify how precision medicine can benefit from computational tools and large-scale genetic screening. The ability to stratify patients based on their genetic profiles offers the potential for early intervention and personalized treatment strategies. In the context of CRC prevention, these findings could guide the design of diagnostic tools and therapeutic approaches that reduce the burden of cancer in patients with chronic inflammatory diseases.

Despite its notable strengths, the study has several limitations that must be acknowledged, particularly regarding the scope and depth of its data and analyses. A primary limitation is its reliance on publicly available datasets, such as those from Olafsson et al. (2020) and the Cancer Dependency Map (DepMap). While these datasets are robust and widely regarded as reliable, they inherently carry restrictions. They focus on specific aspects of genetic

mutations and functional dependency, which may not fully represent the complex interplay of biological factors involved in IBD and CRC pathogenesis. For instance, these datasets do not include information about epigenetic modifications- chemical changes to DNA and histones that influence gene expression without altering the underlying genetic sequence- thus leaving the role of these modifications in IBD-related carcinogenesis underexplored.

Furthermore, environmental factors, such as diet**,** smoking**,** medication use, and exposure to toxins**,** are not accounted for in the datasets used here. These factors can influence both IBD severity and CRC risk, often through mechanisms like oxidative stress or DNA damage. The gut microbiota, closely tied to lifestyle and dietary choices, can also modulate inflammation and the activity of critical biomarkers. Future investigations could integrate metagenomic analyses or dietary assessments to determine how specific microbial communities or dietary components interact with the identified genetic pathways, potentially altering the expression of these biomarkers.

Additionally, while this study identifies crucial genetic biomarkers, it does not fully address genetic and clinical heterogeneity among IBD patients. Variables such as age, ethnicity, and disease subtype (Crohn’s disease vs. ulcerative colitis) may influence the frequency or impact of particular mutations. Personalized treatment plans that consider both genetic susceptibility and environmental exposures could lead to more effective interventions, especially in patients at elevated risk. Moreover, patient-derived organoids or animal models would help validate the in vivo relevance of these biomarkers, as CRISPR-Cas9 knockout data from cell lines alone may not capture the full complexity of the tumor microenvironment.

By embracing these additional dimensions that account or epigenetic modifications**,** environmental influences**,** gut microbiota, and individual variability, future research could yield a

more holistic understanding of IBD-associated CRC. Such efforts would likely enhance the robustness of the findings, allowing for more accurate risk stratification and facilitating the design of longitudinal studies to monitor disease progression over time. In turn, these advances would enable truly personalized treatment plans that integrate genetic, environmental, and clinical factors, bringing the field closer to targeted interventions that address the multifactorial nature of inflammation-driven cancers.

This research has significant implications for medical science, providing a critical foundation for understanding the genetic mechanisms linking inflammatory bowel disease (IBD) to colorectal cancer (CRC). By identifying actionable genetic targets such as ARID2, FBXW7, IL17RA, and ZC3H12A, and exploring their roles in cancer progression, the study offers a roadmap for translational research aimed at reducing cancer risk in high-risk IBD patients. These findings not only highlight the potential for therapies that address both tumor biology and the inflammatory microenvironment but also set the stage for integrating genetic insights into clinical practice—potentially including experimental validation in living systems, broader population studies, and environmental assessments to capture the full array of factors influencing CRC development in IBD populations.

References

Axelrad, J. E., Lichtiger, S., & Yajnik, V. (2016). Inflammatory bowel disease and cancer: The role of inflammation, immunosuppression, and cancer treatment. World Journal of Gastroenterology, 22(20), 4794–4801. https://doi.org/10.3748/wjg.v22.i20.4794

Broad Institute. (2019). Genome-wide CRISPR/Cas9 screening protocol for identifying cancer genes in cell lines. https://www.broadinstitute.org/files/publications/2019/01/Adelmann2019_Protool_Genome-WideCRISPRCas9Screening.pdf

Broad Institute. (n.d.-a). Cancer Dependency Map (DepMap) project: Genome-scale CRISPR-Cas9 genetic perturbation reagents. https://depmap.org/portal/achilles/

Broad Institute. (n.d.-b). CERES: A computational model to account for copy-number effects in CRISPR-Cas9 essentiality screens. https://depmap.org/ceres/

Baumgart, D. C., & Sandborn, W. J. (2012). Crohn's disease. The Lancet, 380(9853), 1590–1605. https://doi.org/10.1016/S0140-6736(12)60026-9

Canavan, C., Abrams, K. R., & Mayberry, J. (2006). Meta-analysis: Colorectal and small bowel cancer risk in patients with Crohn's disease. Alimentary Pharmacology & Therapeutics, 23(8), 1097–1104. https://doi.org/10.1111/j.13652036.2006.02854.x

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., ... & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science, 339(6121), 819-823. https://doi.org/10.1126/science.1231143

DepMap Consortium. (2020). DepMap 20Q2 public. Broad Institute. https://depmap.org/portal/

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. Science, 346(6213), 1258096. https://doi.org/10.1126/science.1258096

Feuerstein, J. D., Isaacs, K. L., Schneider, Y., & Siddique, S. M. (2020). AGA clinical practice guidelines on the management of mild-to-moderate ulcerative colitis. Gastroenterology, 158(4), 932–946. https://doi.org/10.1053/j.gastro.2020.01.020

Harbord, M., Eliakim, R., Bettenworth, D., Karmiris, K., Katsanos, K., Kopylov, U., ... & Sturm, A. (2017). Third European evidence-based consensus on diagnosis and management of ulcerative colitis. Journal of Crohn's and Colitis, 11(7), 769–784. https://doi.org/10.1093/ecco-jcc/jjx009

Jess, T., Rungoe, C., & Peyrin-Biroulet, L. (2012). Risk of colorectal cancer in patients with ulcerative colitis: A meta-analysis of population-based cohort studies. Clinical Gastroenterology and Hepatology, 10(6), 639–645. https://doi.org/10.1016/j.cgh.2011.12.031

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science, 337(6096), 816–821. https://doi.org/10.1126/science.1225829

Kaplan, G. G. (2015). The global burden of IBD: From 2015 to 2025. Nature Reviews Gastroenterology & Hepatology, 12(12), 720–727. https://doi.org/10.1038/nrgastro.2015.150

Le, D. T., Durham, J. N., Smith, K. N., Wang, H., Bartlett, B. R., Aulakh, L. K., ... & Diaz, L. A. (2017). Mismatch repair deficiency predicts the response of solid tumors to PD-1 blockade. Science, 357(6349), 409–413. https://doi.org/10.1126/science.aan6733

Lichtenstein, G. R., Loftus, E. V., Isaacs, K. L., Regueiro, M. D., Gerson, L. B., & Sands, B. E. (2018). ACG clinical guideline: Management of Crohn's disease in adults. American Journal of Gastroenterology, 113(4), 481–517. https://doi.org/10.1038/ajg.2018.27

Markowitz, S. D., & Bertagnolli, M. M. (2009). Molecular origins of cancer: Molecular basis of colorectal cancer. New England Journal of Medicine, 361(25), 2449-2460. https://doi.org/10.1056/NEJMra0804588

Ng, S. C., Shi, H. Y., Hamidi, N., Underwood, F. E., Tang, W., Benchimol, E. I., & Kaplan, G.G. (2017). Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: A systematic review of population-based studies. The Lancet, 390(10114), 2769–2778. https://doi.org/10.1016/S01406736(17)32448-0

Olafsson, S., McIntyre, R. E., Coorens, T. H. H., Butler, T., Salguero, I., Wedge, D. C., ...& Campbell, P. J. (2020). Somatic evolution in non-neoplastic IBD-affected colon. Cell, 182(4), 672–684. https://doi.org/10.1016/j.cell.2020.06.036

Panés, J., Vermeire, S., Lindsay, J. O., Sands, B. E., & Su, C. (2020). Tofacitinib for induction and maintenance therapy in ulcerative colitis: Results from the OCTAVE clinical trials. Journal of Crohn's and Colitis, 14(4), 606–619. https://doi.org/10.1093/ecco-jcc/jjaa012

Sandborn, W. J., Feagan, B. G., Rutgeerts, P., Hanauer, S., Colombel, J. F., Sands, B. E., ... & Reinisch, W. (2017). Vedolizumab as induction and maintenance therapy for Crohn's disease. New England Journal of Medicine, 369(8), 711–721. https://doi.org/10.1056/NEJMoa1215739

Shalem, O., Sanjana, N. E., & Zhang, F. (2014). High-throughput functional genomics using CRISPR-Cas9. Nature Reviews Genetics, 16(5), 299–311. https://doi.org/10.1038/nrg3899

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians, 71(3), 209–249. https://doi.org/10.3322/caac.21660

The Cancer Genome Atlas Research Network. (2012). Comprehensive molecular characterization of human colon and rectal cancer. Nature, 487(7407), 330–337. https://doi.org/10.1038/nature11252

Ungaro, R., Mehandru, S., Allen, P. B., Peyrin-Biroulet, L., & Colombel, J. F. (2017). Ulcerative colitis. The Lancet, 389(10080), 1756–1770. https://doi.org/10.1016/S0140-6736(16)32126-2

Van Cutsem, E., Cervantes, A., Adam, R., Sobrero, A., Van Krieken, J. H., Aderka, D., ...& Arnold, D. (2016). ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. Annals of Oncology, 27(8), 1386–1422. https://doi.org/10.1093/annonc/mdw235

SUPPLEMENTARY TABLES

Supplementary Table S2:

Gene Combination	Mean Effect Size	P-value
ARID2-FBXW7	-0.236222186	2.42E-17
ARID2-IL17RA	-0.182411494	4.80E-20
ARID2-ZC3H12A	-0.153890971	1.32E-14
ARID2-ZNF521	-0.145672396	5.36E-15
ARID2-MYH11	-0.130698817	6.46E-12
ARID2-PIGR	-0.126240262	6.61E-11
FBXW7-IL17RA	-0.122291352	8.19E-10
FBXW7-ZC3H12A	-0.093770829	2.18E-05
FBXW7-ZNF521	-0.085552253	2.44E-05
FBXW7-MYH11	-0.070578674	0.00080273
FBXW7-PIGR	-0.06612012	0.000235979
IL17RA-ZC3H12A	-0.039960137	0.001464934
IL17RA-ZNF521	-0.031741562	0.001213216
IL17RA-MYH11	-0.016767983	0.114362847
IL17RA-PIGR	-0.012309428	0.129635436
ZC3H12A-ZNF521	-0.003221039	0.40167146
MYH11-ZC3H12A	0.01175254	0.764514759
PIGR-ZC3H12A	0.016211095	0.871547043
MYH11-ZNF521	0.019971115	0.920264301
PIGR-ZNF521	0.02442967	0.987449065
MYH11-PIGR	0.039403249	0.997323747

Supplementary Table S3:

Gene Combination	Mean Effect Size	p-value	False Discovery Rate
ARID2-FBXW7-IL17RA	-0.1803	0	0
ARID2-FBXW7-ZC3H12A	-0.1613	0	0
ARID2-FBXW7-ZNF521	-0.1558	0	0
ARID2-FBXW7-MYH11	-0.1458	0	0
ARID2-FBXW7-PIGR	-0.1429	0	0
ARID2-IL17RA-ZC3H12A	-0.1254	0	0
ARID2-IL17RA-ZNF521	-0.1199	0	0
ARID2-IL17RA-MYH11	-0.11	0	0
ARID2-IL17RA-PIGR	-0.107	0	0
ARID2-ZC3H12A-ZNF521	-0.1009	0	0
ARID2-MYH11-ZC3H12A	-0.0909	0	0
ARID2-PIGR-ZC3H12A	-0.088	0	0
ARID2-MYH11-ZNF521	-0.0855	0	0
FBXW7-IL17RA-ZC3H12A	-0.0853	0	0
ARID2-PIGR-ZNF521	-0.0825	0	0
FBXW7-IL17RA-ZNF521	-0.0799	0	0
ARID2-MYH11-PIGR	-0.0725	0	0
FBXW7-IL17RA-MYH11	-0.0699	0	0
FBXW7-IL17RA-PIGR	-0.0669	0	0
FBXW7-ZC3H12A-ZNF521	-0.0608	0.0001	0.0002
FBXW7-MYH11-ZC3H12A	-0.0509	0.002	0.0031
FBXW7-PIGR-ZC3H12A	-0.0479	0.0013	0.002
FBXW7-MYH11-ZNF521	-0.0454	0.0028	0.004
FBXW7-PIGR-ZNF521	-0.0424	0.0012	0.002
FBXW7-MYH11-PIGR	-0.0324	0.0145	0.0195
IL17RA-ZC3H12A-ZNF521	-0.025	0.0078	0.0109
IL17RA-MYH11-ZC3H12A	-0.015	0.1118	0.145
IL17RA-PIGR-ZC3H12A	-0.012	0.1356	0.1695
IL17RA-MYH11-ZNF521	-0.0095	0.1833	0.2213
IL17RA-PIGR-ZNF521	-0.0065	0.2157	0.2517
IL17RA-MYH11-PIGR	0.0034	0.6306	0.7119
MYH11-ZC3H12A-ZNF521	0.0095	0.7862	0.86
PIGR-ZC3H12A-ZNF521	0.0125	0.8855	0.9391
MYH11-PIGR-ZC3H12A	0.0225	0.9661	0.9945
MYH11-PIGR-ZNF521	0.0279	0.9969	0.9969

Note: False Discovery Rate (FDR) is the Benjamini-Hochberg–adjusted *p*-value, calculated separately for triplets, quadruplets, and quintuplets. It estimates the fraction of false positives you would expect if you treated that row (and all rows with equal or smaller FDR) as significant.

Supplementary Table S4:

Gene Combination	Mean Effect Size	p-value	False Discovery Rate
ARID2-FBXW7-IL17RA-ZC3H12A	-0.1381	0	0
ARID2-FBXW7-IL17RA-ZNF521	-0.134	0	0
ARID2-FBXW7-IL17RA-MYH11	-0.1265	0	0
ARID2-FBXW7-IL17RA-PIGR	-0.1243	0	0
ARID2-FBXW7-ZC3H12A-ZNF521	-0.1197	0	0
ARID2-FBXW7-MYH11-ZC3H12A	-0.1122	0	0
ARID2-FBXW7-PIGR-ZC3H12A	-0.11	0	0
ARID2-FBXW7-MYH11-ZNF521	-0.1081	0	0
ARID2-FBXW7-PIGR-ZNF521	-0.1059	0	0
ARID2-FBXW7-MYH11-PIGR	-0.0984	0	0
ARID2-IL17RA-ZC3H12A-ZNF521	-0.0928	0	0
ARID2-IL17RA-MYH11-ZC3H12A	-0.0853	0	0
ARID2-IL17RA-PIGR-ZC3H12A	-0.0831	0	0
ARID2-IL17RA-MYH11-ZNF521	-0.0812	0	0
ARID2-IL17RA-PIGR-ZNF521	-0.079	0	0
ARID2-IL17RA-MYH11-PIGR	-0.0715	0	0
ARID2-MYH11-ZC3H12A-ZNF521	-0.067	0	0
ARID2-PIGR-ZC3H12A-ZNF521	-0.0647	0	0
FBXW7-IL17RA-ZC3H12A-ZNF521	-0.0628	0	0
ARID2-MYH11-PIGR-ZC3H12A	-0.0572	0	0
FBXW7-IL17RA-MYH11-ZC3H12A	-0.0553	0.0001	0.0001
ARID2-MYH11-PIGR-ZNF521	-0.0531	0	0
FBXW7-IL17RA-PIGR-ZC3H12A	-0.053	0	0
FBXW7-IL17RA-MYH11-ZNF521	-0.0512	0	0.0001
FBXW7-IL17RA-PIGR-ZNF521	-0.0489	0	0
FBXW7-IL17RA-MYH11-PIGR	-0.0414	0.0002	0.0003
FBXW7-MYH11-ZC3H12A-ZNF521	-0.0369	0.0047	0.0059
FBXW7-PIGR-ZC3H12A-ZNF521	-0.0347	0.0032	0.0041
FBXW7-MYH11-PIGR-ZC3H12A	-0.0272	0.0214	0.0258
FBXW7-MYH11-PIGR-ZNF521	-0.0231	0.0271	0.0316
IL17RA-MYH11-ZC3H12A-ZNF521	-0.01	0.1577	0.1781
IL17RA-PIGR-ZC3H12A-ZNF521	-0.0078	0.1865	0.204
IL17RA-MYH11-PIGR-ZC3H12A	-0.0003	0.489	0.5186
IL17RA-MYH11-PIGR-ZNF521	0.0038	0.679	0.6989
MYH11-PIGR-ZC3H12A-ZNF521	0.0181	0.9695	0.9695

Note: False Discovery Rate (FDR) is the Benjamini-Hochberg–adjusted *p*-value, calculated separately for triplets, quadruplets, and quintuplets. It estimates the fraction of false positives you would expect if you treated that row (and all rows with equal or smaller FDR) as significant.

Supplementary Table S5:

Gene Combination	Mean Effect Size	p-value	False Discovery Rate
ARID2-FBXW7-IL17RA-ZC3H12A-ZNF521	-0.1095	0.0000	0
ARID2-FBXW7-IL17RA-MYH11-ZC3H12A	-0.1035	0.0000	0
ARID2-FBXW7-IL17RA-PIGR-ZC3H12A	-0.1017	0.0000	0
ARID2-FBXW7-IL17RA-MYH11-ZNF521	-0.1002	0.0000	0
ARID2-FBXW7-IL17RA-PIGR-ZNF521	-0.0984	0.0000	0
ARID2-FBXW7-IL17RA-MYH11-PIGR	-0.0924	0.0000	0
ARID2-FBXW7-MYH11-ZC3H12A-ZNF521	-0.0888	0.0000	0
ARID2-FBXW7-PIGR-ZC3H12A-ZNF521	-0.087	0.0000	0
ARID2-FBXW7-MYH11-PIGR-ZC3H12A	-0.081	0.0000	0
ARID2-FBXW7-MYH11-PIGR-ZNF521	-0.0777	0.0000	0
ARID2-IL17RA-MYH11-ZC3H12A-ZNF521	-0.0673	0.0000	0
ARID2-IL17RA-PIGR-ZC3H12A-ZNF521	-0.0655	0.0000	0
ARID2-IL17RA-MYH11-PIGR-ZC3H12A	-0.0595	0.0000	0
ARID2-IL17RA-MYH11-PIGR-ZNF521	-0.0562	0.0000	0
ARID2-MYH11-PIGR-ZC3H12A-ZNF521	-0.0448	0.0000	0
FBXW7-IL17RA-MYH11-ZC3H12A-ZNF521	-0.0432	0.0002	0.0002
FBXW7-IL17RA-PIGR-ZC3H12A-ZNF521	-0.0414	0.0001	0.0001
FBXW7-IL17RA-MYH11-PIGR-ZC3H12A	-0.0354	0.0009	0.001
FBXW7-IL17RA-MYH11-PIGR-ZNF521	-0.0322	0.0007	0.0009
FBXW7-MYH11-PIGR-ZC3H12A-ZNF521	-0.0207	0.0331	0.0347
IL17RA-MYH11-PIGR-ZC3H12A-ZNF521	0.0008	0.537	0.537

Note: False Discovery Rate (FDR) is the Benjamini-Hochberg–adjusted *p*-value, calculated separately for triplets, quadruplets, and quintuplets. It estimates the fraction of false positives you would expect if you treated that row (and all rows with equal or smaller FDR) as significant.