cancers

MDPI

Article

Germline Variants of *CYBA* and *TRPM4* Predispose to Familial Colorectal Cancer

Lizhen Zhu [1,2,†], Beiping Miao [1,3,4,†], Dagmara Dymerska [5], Magdalena Kuswik [5], Elena Bueno-Martínez [6], Lara Sanoguera-Miralles [6], Eladio A. Velasco [6], Nagarajan Paramasivam [7], Matthias Schlesner [8], Abhishek Kumar [1,9,10], Ying Yuan [2], Jan Lubinski [5], Obul Reddy Bandapalli [1,3,4,11,*,‡], Kari Hemminki [1,12,13,*,‡] and Asta Försti [1,3,4]

1 Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany; azhulijane@163.com (L.Z.); b.miao@kitz-heidelberg.de (B.M.); abhishek.abhishekkumar@gmail.com (A.K.); a.foersti@kitz-heidelberg.de (A.F.)
2 Department of Medical Oncology, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou 310009, China; yuanying1999@zju.edu.cn
3 Hopp Children's Cancer Center (KiTZ), D-69120 Heidelberg, Germany
4 Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), D-69120 Heidelberg, Germany
5 Department of Genetics and Pathology, Hereditary Cancer Center, Pomeranian Medical University, Unii Lubelskiej 1, 71-252 Szczecin, Poland; dagmara.dymerska@gmail.com (D.D.); magdalenakuswik@gmail.com (M.K.); lubinski@pum.edu.pl (J.L.)
6 Splicing and Genetic Susceptibility to Cancer, Instituto de Biología y Genética Molecular (CSIC-UVa), 47003 Valladolid, Spain; buenome@usal.es (E.B.-M.); lara.sanoguera@alumnos.uva.es (L.S.-M.); eavelsam@ibgm.uva.es (E.A.V.)
7 Computational Oncology, Molecular Diagnostics Program, National Center for Tumor Diseases (NCT), D-69120 Heidelberg, Germany; n.paramasivam@dkfz-heidelberg.de
8 Bioinformatics and Omics Data Analytics, German Cancer Research Center (DKFZ), D-69120 Heidelberg, Germany; matthias.schlesner@informatik.uni-augsburg.de
9 Institute of Bioinformatics, International Technology Park, Bengaluru 560066, India
10 Manipal Academy of Higher Education (MAHE), Manipal 576104, India
11 Medical Faculty Heidelberg, Heidelberg University, D-69120 Heidelberg, Germany
12 Faculty of Medicine and Biomedical Center in Pilsen, Charles University in Prague, 30605 Pilsen, Czech Republic
13 Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), D-69120 Heidelberg, Germany
* Correspondence: bandapalli@gmail.com (O.R.B.); k.hemminki@dkfz.de (K.H.)
† These authors contributed equally to this work.
‡ Shared senior authorship.

check for updates

Citation: Zhu, L.; Miao, B.; Dymerska, D.; Kuswik, M.; Bueno-Martínez, E.; Sanoguera-Miralles, L.; Velasco, E.A.; Paramasivam, N.; Schlesner, M.; Kumar, A.; et al. Germline Variants of *CYBA* and *TRPM4* Predispose to Familial Colorectal Cancer. *Cancers* **2022**, *14*, 670. https://doi.org/10.3390/cancers14030670

Academic Editor: Alessandro Cama

Received: 22 December 2021
Accepted: 26 January 2022
Published: 28 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Simple Summary: Whole-genome sequencing and bioinformatics analysis on unique colorectal cancer families revealed two attractive candidate predisposition genes, *CYBA* and *TRPM4*, each with a loss-of-function variant. Supported by our functional studies, we suggest that the two gene defects mechanistically involve intestinal barrier integrity through reactive oxygen species and mucus biology, which converges in chronic bowel inflammation, a known risk factor for colorectal cancer.

Abstract: Familial colorectal cancer (CRC) is only partially explained by known germline predisposing genes. We performed whole-genome sequencing in 15 Polish families of many affected individuals, without mutations in known CRC predisposing genes. We focused on loss-of-function variants and functionally characterized them. We identified a frameshift variant in the *CYBA* gene (c.246delC) in one family and a splice site variant in the *TRPM4* gene (c.25–1 G > T) in another family. While both variants were absent or extremely rare in gene variant databases, we identified four additional Polish familial CRC cases and two healthy elderly individuals with the *CYBA* variant (odds ratio 2.46, 95% confidence interval 0.48–12.69). Both variants led to a premature stop codon and to a truncated protein. Functional characterization of the variants showed that knockdown of *CYBA* or *TRPM4* depressed generation of reactive oxygen species (ROS) in LS174T and HT-29 cell lines. Knockdown of *TRPM4* resulted in decreased MUC2 protein production. *CYBA* encodes a component in the NADPH oxidase system which generates ROS and controls, e.g., bacterial colonization in

the gut. Germline *CYBA* variants are associated with early onset inflammatory bowel disease, supported with experimental evidence on loss of intestinal mucus barrier function due to ROS deficiency. *TRPM4* encodes a calcium-activated ion channel, which, in a human colonic cancer cell line, controls calcium-mediated secretion of MUC2, a major component of intestinal mucus barrier. We suggest that the gene defects in *CYBA* and *TRPM4* mechanistically involve intestinal barrier integrity through ROS and mucus biology, which converges in chronic bowel inflammation.

Keywords: whole-genome sequencing; cancer predisposition; mucin; reactive oxygen species

1. Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and the second leading cause of death from cancer [1]. Familial CRC accounts for some 15% of all cases, and the twin estimate of heritability amounts up to 30%, but only 2–5% of CRCs are confirmed to be caused by inherited syndromes related to CRC [2–6]. Among these, Lynch-syndrome-related CRC, caused by germline mutations in mismatch repair (MMR) genes *MLH1*, *MSH2*, *MSH6*, *PMS2* and *EPCAM*, accounts for the majority of hereditary CRC. These tumors are characterized by deficient MMR (dMMR). Rarer syndromes are related to mutations in genes *APC*, *MUTYH*, *STK11*, *PTEN*, *BMPR1A* and *SMAD4* [6].

With the advancement of next-generation sequencing technologies, increasing number of genes have been reported to be possible candidates for CRC predisposition. These include *POLE*, *POLD1*, *NTHL1*, *GERM1*, *GALNT12*, *RNF43*, *RPS20*, *MLH3* and *MSH3* [3,6–9]. According to the National Comprehensive Cancer Network clinical practice guidelines, most of these new candidates do not have well-established evidence of increased risk for CRC [3]; however, with more data, some of these genes will be proved to be CRC predisposing genes [3,8,10]. Epidemiological evidence, such as lacking correlation of CRC risk between spouses, suggests that most of familial aggregation in this cancer is genetic [11]. Thus, it is likely that novel predisposing genes will be identified in families of affected individuals.

Gut homeostasis is maintained by a physical separation of the microbial community from the gut epithelium by a mucus barrier. MUC2 mucin, expressed in the goblet cells of the intestine, is the main component of the intestinal mucus barrier; its expression is commonly lost in CRC, and its loss predicts adverse outcome [12]. MUC5AC and MUC6 are rarely observed in the normal colon, but their expression in CRC is associated with favorable outcome. MUC1 is expressed on the apical surface of most epithelial cells, and it does not seem to affect the outcome.

Most 'classical' cancer-predisposing genes were found in linkage studies in families with multiple affected patients [13]. Following this paradigm, we performed germline whole-genome sequencing (WGS) in 15 Mendelian type of CRC families unrelated to known CRC-predisposing genes, focusing on loss-of function variants. We identified a frameshift variant in *CYBA* in one family and a splice site variant in *TRPM4* in another family; while the genes encode proteins in diverse pathways, their functions appeared to converge in intestinal-barrier integrity and mucus-biology-targeting inflammatory bowel disease, a known risk factor of CRC [14].

2. Materials and Methods

2.1. Population Recruitment

In several regions of Poland, population screening was performed mainly in the years 2000–2014, in which questionnaires about cancer family history were collected. Persons with positive CRC family history were invited to genetic outpatient clinics all over Poland, and their more detailed family histories were collected through face-to-face detailed interviews. Similarly, persons with negative cancer family history were interviewed. An average review

took 20–30 min. Eligible individuals were asked to participate in the study, and they signed an informed consent form.

2.2. Colorectal Cancer Families

Fifteen families with strong CRC aggregation compatible with an autosomal dominant pattern of inheritance were recruited to the study. Each family had at least three pathologically confirmed CRC cases; 13 families had at least one case diagnosed below the age of 55 years. In Poland, colonoscopy is offered for the family members, so many families also had members diagnosed with polyps. Twelve families had CRC in at least two generations. In the other three families, one had 5 siblings diagnosed with CRC and 5 of their children had polyps at the age of 41–57 years. In another family, 3 siblings were diagnosed with CRC at the age of 53, 54 and 63 years; their father had been diagnosed with prostate cancer and had died at the age of 72 years and their children were born in 1980s and thus uninformative regarding the CRC status. The third family had 3 siblings diagnosed with CRC at the age of 50, 55 and 62 years; their father had suffered from cancer and died at the age of 55 years. All families were screened for alterations in *APC*; the mismatch repair genes *MLH1, MSH2* and *MSH3*; and large deletions in *EPCAM* and *POLE* p.Leu424Val, *POLD1* p.Ser478Asn and *NTHL1* p.Gln90* mutations and were found to be negative.

2.3. Validation Cohort

Altogether, 1705 unrelated familial CRC cases and 1674 healthy elderly individuals without a family history of cancer were included in the validation cohort.

2.4. DNA Isolation

Peripheral blood samples were collected from affected and unaffected family members who agreed to participate in the study, as well as from the validation cohort. Genomic DNA was isolated by using a modified Lahiri and Schnabel method [15].

2.5. Germline Whole-Genome Sequencing (WGS)

WGS was performed in the Illumina X10 platform, using DNA extracted from the blood samples as paired-end sequencing with a read length of 150 bp. Mapping of reads to the human reference genome (GRCh37 assembly version hs37d5) was performed by using BWA mem (version 0.7.8), and duplicates were marked by using Picard (version 1.125). Small variants, single nucleotide variants (SNVs) and indels were called by using Platypus (version 0.8.1). Variants were annotated by using ANNOVAR [16], dbNSFP v2.9 [17], 1000 Genomes phase III [18], dbSNP [19] and ExAC [20], as described previously [21]. On the variants that passed all the internal Platypus filters, we performed a further filtering while considering the QUAL score >20 and coverage of minimum 5 reads. The sequencing coverage and quality statistics for each sample are summarized in Supplementary Table S1. Minor allele frequency (MAF) of 0.1% was used with respect to population databases (the 1000 Genomes phase III [18] and non-TCGA ExAC [20] data), and the variant frequency of 5% from the local datasets was used to remove technical artefacts. Pairwise comparison of variants among the cohort was performed to check for sample swaps and family relatedness.

2.6. Variant Identification

Variants were filtered based on the pedigree information, considering the members diagnosed with CRC as cases. Family members who were diagnosed with polyps were considered possible variant carriers, as were individuals whose parents were diagnosed with CRC and who had not yet reached the age of diagnosis of the youngest CRC case in the family. Unaffected family members were considered as controls. All variants segregating with the disease were filtered for loss-of-function (stop gain, frameshift and splicing) variants. All variants with a MAF of <0.1% in the gnomAD (https://gnomad.broadinstitute.org/ (accessed on 17 January 2022)) non-Finnish European population were further screened for their location in the gene. Variants located in the last exon or in the

non-protein coding transcript of the gene were excluded. Sequencing data were visually inspected by using the Integrative Genomic Viewer (IGV [22]) to exclude false positive variants. Combined Annotation-Dependent Depletion (CADD [23]) score was used to evaluate the deleteriousness of the variants; the scores >20 and >30 are indicative of the top 1% and top 0.1% of deleterious variants, respectively. The final variant selection was based on a search of the literature.

2.7. Copy Number Variants

Structural variants were analyzed by using the SmallPedigree-WGS workflow of Canvas (version 1.40.0.1613 [24]) separately to detect larger copy number variants as described earlier [25]. Variants that affected known cancer-predisposing genes were manually inspected by using IGV.

2.8. Variant Confirmation

Candidate variants and their segregation with the disease in the families were confirmed in all available family members by Sanger sequencing with the primers *CYBA* F 5′-GGAGCTTGGTTTCTCACTTGG-3′, R 5′-GGAGCTCCTCGGATTTGGA-3′; and *TRPM4* F 5′-GTGGCTCTGTGTCCCATAGG-3′, R 5′-TCTACACAGACCCAAACGCT-3′. The variants were checked for frequency in the 1705 familial CRC cases and 1674 healthy elderly individuals, using custom-made Taqman assays, and the existence of the heterozygous variants was confirmed by using the Sanger sequencing.

2.9. TRPM4 Minigene Assay

2.9.1. Splicing Predictions

In silico analysis of the wild-type (WT) and mutant sequences was made with the algorithm Max Ent Scan [26] that is integrated into the splicing software of Human Splicing Finder version 3.1 (HSF, http://www.umd.be/HSF3/ (accessed on 17 January 2022)) [27]. Variant and transcript descriptions were according to the Human Genome Variation Society (HGVS) guidelines on the basis of the *TRPM4* GenBank sequence NM_017636.4.

2.9.2. Construction of the Minigene mgTRPM4_ex1-2

A single fragment with exons 1 and 2 of the *TRPM4* gene (728 bp) was amplified from a patient (III-5) DNA with the variant with Kapa High Fidelity polymerase (Kapa Biosystems, Wilmington, MA) and the primers mgTRPM4_ex1-fw 5′ AGTCACCTGGACAACCTCAAAGGCACCTTT GAAGCAGAGCCGGCGGAGGG 3′ and mgTRPM4_ex2-rv 5′ ATAAGCTTGATATCGAATTCCTGCAGCCCG AGCACATAGAAGAGACATCGG 3′ (cloning tails are underlined). This insert has the following structure: EX1 [90 bp]–ivs1 [301 bp]–EX2 [68 bp]–ivs2 [269bp]. This fragment was cloned into the splicing vector pSAD v9.0 (Patent P201231427, CSIC) [28] by using Overlap Extension PCR [29] to generate a chimeric fusion of vector exon V1-*TRPM4* exon 1. Several colonies were sequenced at the Macrogen facility (Macrogen, Madrid, Spain) to confirm the presence of both WT and mutant minigenes (c.25–1G > T).

2.9.3. Transfection of Eukaryotic Cells

Transfections were carried out as previously described [30]. Cells were transfected with 1 μg of each construct, using low-toxicity Lipofectamine (Life Technologies, Carlsbad, CA, USA). To inhibit nonsense mediated decay (NMD), a 4 h incubation with cycloheximide 300 μg/mL (Sigma-Aldrich, St. Louis, MO, USA) was carried out.

2.9.4. Reverse-Transcription PCR of Minigene RNA

RNA was purified with the Genematrix Universal RNA Purification Kit (EURx), including on-column DNAse I digestion. Reverse transcription was carried out with 400 ng of RNA and the RevertAid First Strand cDNA Synthesis Kit (Life Technologies), using the vector-specific primer RTPSPL3-RV (5′-TGAGGAGTGAATTGGTCGAA-3′). Samples were

incubated at 42 °C for 1 h, followed by 5 min at 70 °C. Then, 40 ng of cDNA (final volume of 50 µL) was amplified with RT-pMAD502-FW (5′-GAGGTTCTTCGAGTCCTTT-3′) and RTpSAD-RV (Patent P201231427) (full-length transcript 394 nt), using Platinum-Taq DNA polymerase (Life Technologies). Samples were denatured at 94 °C for 2 min, followed by 35 cycles x (94 °C, 30 s; 60 °C, 30 s; 72 °C, 30 s), and a final extension step at 72 °C for 5 min. RT-PCR products were run on 1.2% agarose gels and sequenced at the Macrogen facility.

2.9.5. Capillary Electrophoresis of Fluorescent RT-PCR

In order to relatively quantify all transcripts, semi-quantitative fluorescent RT-PCRs were performed in triplicate with primers RT-pMAD502-FW and FAM-RTpSAD-RV, as described earlier [30].

2.10. Cell Culture

Human colorectal adenocarcinoma cell line HT-29 (RRID:CVCL_0320) (a kind gift from Peter Krammer's lab, DKFZ) was cultured in RPMI 1640 media (21875091, Life Technologies, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (FBS) (10500064, Life Technologies, Carlsbad, CA, USA). Mucin-secreting colorectal adenocarcinoma cell line LS174T (RRID:CVCL_1384) (C0009013, AddexBio, San Diego, CA, USA) was cultured in RPMI 1640 media supplemented with 10% FBS supplemented with 2 mM L-glutamine (51411C, Sigma-Aldrich, Saint Louis, MO, USA). HEK293T (RRID: CVCL_0063) cells were a gift from Andreas Trump (DKFZ, Heidelberg). The cells were maintained in DMEM high glucose supplemented with 10% FBS (Gibco), penicillin (50 U/mL, Life Technologies, Carlsbad, CA, USA) and streptomycin (50 µg/mL, Life Technologies, Carlsbad, CA, USA). The cells were authenticated by using SNP or STR profiling within the last 3 years, and all experiments were performed with mycoplasma-free cells.

2.11. Site-Directed Mutagenesis

pcDNA4TO-HA-*TRPM4* was a generous gift from Hugues Abriel [31]. Gateway clone of pENTR221-*CYBA* was obtained from DKFZ's genomics and proteomics core facility (GPCF). We used LR Clonase Enzyme Mix (11791-019, Invitrogen, Waltham, MA, USA) to make pENTR221-*CYBA* into pDEST26-*CYBA*. Variants identified in the WGS were introduced into respective plasmids by using QuikChange II XL Site-Directed Mutagenesis Kit (200521, Agilent Technologies, Santa Clara, CA, USA). Mutated cDNA clones were confirmed by Sanger sequencing before using them in further experiments. Plasmids were scaled up by transforming into bacteria and harvested by HiPure Plasmid Midiprep Kit (K210005, Invitrogen, Waltham, MA, USA), and the sequences were checked by Sanger sequencing. Since the *TRPM4* variant was a splice site variant not affecting an exon, we created the pcDNATO-HA-*TRPM4* c.25delAG according to the minigene results of the original *TRPM4* variant. Primers for site directed mutagenesis to get pcDNATO-HA-*TRPM4* c.25delAG and pDEST26-*CYBA* c.246delC were *TRPM4* F 5′-CGGAGAAGGAGCAGCTGGATCCCCAAGA-3′, R 5′-TCTTGGGGATCCAGCTGCTCCTTCTCCG-3′; *CYBA* F 5′-TAGTAATTCCTGGTAAAGGCCCGAACAGCTTCAC-3′, R 5′-GTGAAGCTGTTCGGGCCTTTACCAGGAATTACTA-3′. Sanger sequencing primers for mutated plasmids were *TRPM4* F 5′-CACGCTGTTTTGACCTCCAT-3′, R 5′-CGGAGGAAATTGCTGTGCTT-3′; *CYBA* F 5′-CATGTGGGCCAACGAACAG-3′, R 5′-TCAGTAGGTAGATGCCGCTC-3′.

2.12. SiRNA Mediated Knockdown of CYBA and TRPM4

HT-29 and LS174T cells were transfected with human CYBA siRNA (sc-36149, Santa Cruz Biotechnology, INC) to knockdown CYBA or human TRPM4 siRNA (sc-45439, Santa Cruz Biotechnology, Heidelberg, Germany) to knockdown TRPM4, or scrambled siRNA (sc-37007, Santa Cruz Biotechnology, Heidelberg, Germany) as a control. All the siRNA products consist of pools of three to five target-specific 19–25 nt siRNAs designed to knockdown gene expression. HEK293T cells were transfected with wild-type (WT) or mutated (MUT) pDEST26-CYBA or pDEST26 and WT or MUT pcDNA4TO-HA-TRPM4

or pcDNA4TO. All the transfections were performed with Lipofectamine 2000 (11668027, Invitrogen, Waltham, MA, USA).

2.13. In Vitro Cell Proliferation Assay

Cell proliferation was estimated by using a cell counting kit 8 (96992-500TESTS-F, Sigma-Aldrich, Saint Louis, MO, USA). HT-29 cells with a density of 5×10^3 and LS174T cells with a density of 7×10^3 cells per well were seeded in 96-well plates. The absorbance optical (OD) density value was measured at 450 nm, using Multiskan FC (Thermo Scientific, Waltham, MA, USA), for 5 days after transfection.

2.14. Real-Time PCR

Total RNA was extracted by TRIzol (15596018, Invitrogen, Waltham, MA, USA), and 1 µg of RNA was reverse-transcribed into cDNA, using a ProtoScript First Strand cDNA Synthesis Kit (E6300L, NEW ENGLAND BioLabs, Ipswich, MA, USA). Quantitative expression analysis was performed with QuantiFast SYBR Green (204056, Qiagen, Hilden, Germany) with 20 ng of cDNA. *HPRT1* was used as a housekeeping gene. Primers of *MUC2* and *HPRT1* were from QuantiTect Primer Assay (QT00059066, QT01004675, Qiagen, Hilden, Germany). The detections were performed by Applied Biosystems 7300. The primer sequences for real-time PCR were *TRPM4* F 5′-TGGCTCTCACCTGCTTCCT-3′, R 5′-CCGCACCGTGAAAACCATG-3′; *CYBA* F 5′-ACCAGGAATTACTATGTTCGGGC-3′, R 5′-TAGGTAGATGCCGCTCGCAATG-3′; *MUC1* F 5′-CTGGTCTGTGTTCTGGTTGC-3′, R 5′-CCACTGCTGGGTTTGTGTAA-3′.

2.15. Western Blot

Protein lysates were prepared and quantified by using Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA, #23225). NuPAGE™ 4–12% Bis-Tris Protein Gels and the respective running buffer (Thermo Fisher Scientific; Waltham, MA, USA, #NP0321PK2, #NP0001) were used for separation of 20 µg of each protein sample. Blotted membranes were blocked with 5% milk for 1 h, incubated overnight at 4 °C with the primary antibodies (HA tag monoclonal antibody (26183, Invitrogen, 1:5,000), anti-CYBA antibody (sc-271968, Santa Cruz, Heidelberg, Germany 1:250), anti-TRPM4 antibody (ab123936, Abcam, Cambridge, UK, 1:1000) and loading control alpha tubulin antibody (ab4074, Abcam, Cambridge, UK, 0.25 µ/mL) and subsequently for 1 h at room temperature with anti-rabbit IgG antibody (7074S, Cell Signaling Technology, Frankfurt, Germany, 1:10,000) and anti-mouse IgG antibody (7076S, Cell Signaling Technology, Frankfurt, Germany, 1:2500) diluted in 5% milk. The membranes were washed, and signals were detected by Immobilon Western Chemiluminescent HRP substrate (WBKLS0500, Millipore, Burlington, MA, USA) and scanned by INTAS Chemostar.

2.16. Immunofluorescence

Cells were seeded on Chamber slide (C6932, Sigma-Aldrich) with the seeding density of 30×10^3 per well. The cells were allowed to attach to the wells overnight before transfecting with siRNA. After transfection with the siRNA for 72 h, slides were incubated with 4% paraformaldehyde (sc-281692, Santa Cruz, Heidelberg, Germany) 20 min at room temperature; PBS 5 min 3 times, 0.1% Triton X-100 (11332481001, Sigma-Aldrich, Saint Louis, MO, USA) 10 min at room temperature, PBS 5 min 3 times, 3% BSA block 1 h at 37 °C. After that, the slides were incubated with anti-MUC2 antibody (ab11197, Abcam, Cambridge, UK, 1:500) for 3 h at 37 °C, PBS 5 min 3 times, and with anti-mouse IgG (ab150113, Abcam, Cambridge, UK, 1:1000) for 1 h at 37 °C, PBS 5 min 3 times. The nucleus was dyed with 1 ug/mL DAPI (62248, Thermo Scientific) for 2 min at room temperature, PBS 2 min 3 times. The slides were mounted with Limonene Mounting Medium (ab104141, Abcam, Cambridge, UK,) and covered by coverslips (7695031, Th. Geyer, Renningen, Germany). Confocal microscopy was performed with Leica SP8, and images were saved as

TIFF file formats. For each group in every experiment, we randomly selected three images with similar numbers of cells and uniform background.

2.17. Reactive Oxygen Species (ROS) Detection

Cellular ROS detection was performed with DCFDA Cellular ROS Detection Assay Kit (ab113851, Abcam, Cambridge, UK) in both HT-29 and LS174T cell lines, according to the manufacturer's instructions.

2.18. Statistical Analysis

The comparison of proliferation, real-time PCR and ROS between the knockdown group and the control group were analyzed by non-paired Student's *t*-test. Image J was used to calculate the fluorescence intensity and area in every group, and quantitative statistical comparison of fluorescence intensity of MUC2 was performed with non-paired Student's *t*-test [32,33]. Statistical analysis was performed in GraphPad Prism 5.0. Significant differences were considered when $p < 0.05$. Means and standard errors of the means of each assay were presented on the graphs. For real-time PCR assays and ROS detection, we performed three independent experiments in triplicate for each analysis, and two for Western blot and immunofluorescence.

3. Results

3.1. Whole-Genome Sequencing Revealed No Variants in Known Colorectal Cancer Predisposing Genes

None of the 15 families had copy number variants or pathogenic or likely pathogenic variants or variants of unknown significance in any of the genes attributed to the hereditary CRC syndromes including *MLH1*, *MSH2*, *MSH6*, *PMS2*, *EPCAM*, *APC*, *MUTYH*, *STK11*, *PTEN*, *BMPR1A*, *SMAD4* and *TP53* nor in the genes suggested to be related with familial CRC including *POLE*, *POLD1*, *NTHL1*, *GERM1*, *GALNT12*, *CHEK2*, *BLM*, *AXIN2*, *ATM*, *BMPR1A*, *RECQL*, *RPS20*, *MLH3* and *MSH3*.

3.2. Identification of Loss-of-Function Variants by Whole-Genome Sequencing

We identified, altogether, 14 rare loss-of-function variants (nonsense variants, indels leading to a frameshift and variants in canonical splice sites) that segregated with the disease in 9 of the 15 whole-genome sequenced families (Supplementary Table S2). CADD scores of the variants were between 24.7 and 44, supporting their deleterious nature. The literature search indicated a role in inflammation and mucin homeostasis of the intestine for two genes, *CYBA* and *TRPM4*, that were mutated in two unrelated families (Figure 1); these were investigated further.

3.3. CYBA c.246delC and TRPM4 c.25-1 G > T Variant Segregation in CRC Families

In Family 8, we found a heterozygous germline *CYBA* c.246delC (ENST00000261623) variant, which segregated with the disease in the pedigree (Figures 1A and S1, upper panel). In this family, four family members were diagnosed with CRC at ages from 42 to 67 years (II-5, III-9, III-12 and IV-11) and two with cancer in the abdominal cavity (II-3 and III-5). Several family members had also polyps in the colon. The two sequenced family members diagnosed with CRC (III-12 and IV-11) carried the variant, as did III-3, who was diagnosed with polyps in the colon and his two sons (IV-4 and IV-6) who were also diagnosed with polyps in their 40s. Of note, the mother of IV-11 (III-9) was diagnosed with both colon and ovarian cancer. In contrast, III-7 and both of his sons (IV-8 and IV-9) who did not carry the variant had no cancer, nor did III-13 or III-15, although they had polyps diagnosed at the age of 61 and 56 years.

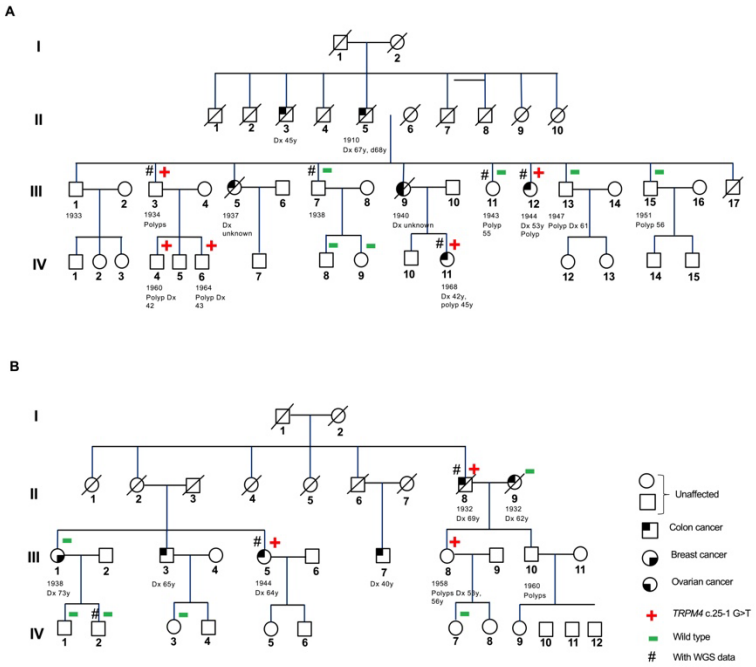

Figure 1. Pedigrees of colon cancer families. (**A**) Family 8 with CYBA c.246delC variant. (**B**) Family 11 with TRPM4 c.25-1 G > T variant.

In Family 11, a heterozygous germline *TRPM4* c.25-1 G > T (ENST00000252826) variant segregated with the disease in the pedigree (Figures 1B and S1, lower panel). In this family, two siblings (III-3 and III-5), their cousin (III-7) and their uncle (II-8) were diagnosed with CRC at the age ranging from 40 to 69 years. Several family members had also polyps. The two sequenced family members (II-8 and III-5) diagnosed with CRC carried the variant, as did the daughter of II-8 (III-8), who had multiple polyps. Of note, the wife of II-8 (II-9), who was also diagnosed with CRC, did not carry the variant. No *TRPM4* c.25-1 G > T variant was found in a female member (III-1) diagnosed with breast cancer at the age of 73 years, whose colonoscopy was negative until the age of 74 years, as with her two sons (IV-1 and IV-2).

3.4. CYBA c.246delC and TRPM4 c.25-1 G > T Are Rare Variants

CYBA c.246delC is absent from the population-based variant databases gnomAD, including data from 76156 individuals from the main world populations (Available online: https://gnomad.broadinstitute.org/ (accessed on 17 January 2022)) in the Exome Variant Server, including data from 6503 individuals from European and African American populations (Available online: http://evs.gs.washington.edu/EVS (accessed on 17 January 2022)) and Leiden Open Variation Database (LOVD; Available online: https://www.lovd.nl/ (accessed on 17 January 2022)), but it was reported in dbSNP (hg38) (Available online: http://www.ncbi.nlm.nih.gov/snp (accessed on 17 January 2022)), showing an allele frequency of 0.00004 (5/125,568, Available online: https://www.ncbi.nlm.nih.gov/snp/rs1439134665 (accessed on 17 January 2022)) and Clinvar (Available online: https://www.ncbi.nlm.nih.gov/clinvar/ (accessed on 17 January 2022)), in which it was suggested by a single submitter to be an apathogenic variant of granulomatous disease (Available online: https://www.ncbi.nlm.nih.gov/clinvar/variation/619030/ (accessed on 17 January 2022)). *TRPM4* c.25-1 G > T is absent in all of the above population databases.

3.5. Screening of a Large Cohort of Familial CRC Patients

Screening of the *CYBA* c.246delC and *TRPM4* c.25-1 G > T variants in 1705 familial CRC cases and 1674 healthy elderly individuals, both from Poland, using custom-made Taqman assays, confirmed the presence of the CYBA variant in Family 8 and identified the

CYBA c.246delC variant in four additional familial CRC cases and two healthy individuals (odds ratio 2.46, 95% confidence interval 0.48–12.69). Two of the CRC patients, diagnosed at ages of 59 years and 64 years, had a family history of CRC. One CRC patient, diagnosed at the age of 44 years, had no contact to family members, and another one, diagnosed at the age of 63 years, had a family history of female genital tract and larynx cancers. The sampling ages of the two healthy individuals were 94 years and 69 years, and they had no family history of any cancer. The existence of the heterozygous *CYBA* c.246delC variant was confirmed by Sanger sequencing in all positive samples. In addition to the *TRPM4* variant in Family 11, no other families were found to carry the variant.

3.6. CYBA c.246delC Led to Loss of CYBA Protein

As shown in Figure 2A, HEK293T cells transfected with pDEST26-*CYBA* c.246delC expressed less CYBA protein than those transfected with wild-type pDEST26-*CYBA*. Moreover, pDEST26 only expressed a similar amount of CYBA to that of the mutant, suggesting that *CYBA* c.246delC led to a loss of the CYBA protein.

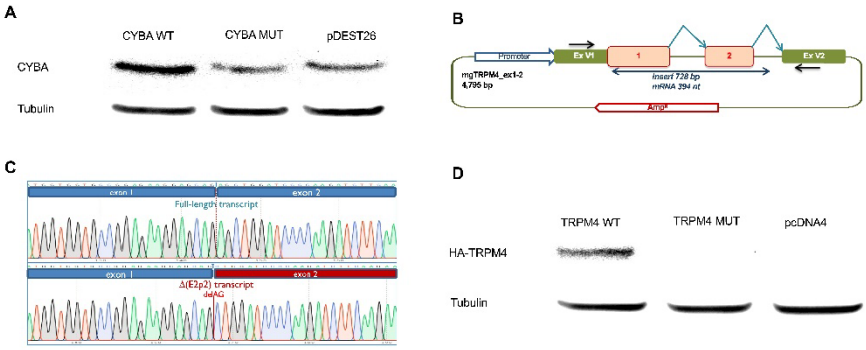

Figure 2. *CYBA* c.246delC and *TRPM4* c.25-1G > T variants led to loss of the corresponding proteins. (**A**) Western blot detection of the CYBA protein in HEK293T transfected with wild-type (CYBA WT), mutated (CYBA MUT) or control plasmids (pDEST26). HEK293T cells transfected with the mutant CYBA plasmid expressed less protein compared to cells transfected with the wild-type CYBA plasmid. CYBA c.246delC-transfected cells expressed a similar amount of protein to those transfected with pDEST26 only. (**B**) Outline of the *TRPM4* minigene construct. The black arrows in vector exons V1 and V2 indicate specific RT-PCR minigene primers; broken arrows represent the expected splicing reactions. (**C**) Sequencing traces of the transcripts generated by the wild-type (above) and mutant (c.25-1G > T) minigenes, suggesting that TRPM4 c.25-1G > T led to a frameshift transcript r.25_26del, that would be equivalent to c.25delAG. (**D**) *TRPM4* c.25delAG led to the loss of TRPM4 protein. Western blot of TRPM3-HA Tag antibody in HEK293T cells transfected with wild-type, mutated or control plasmids. HEK293T cells transfected with pcDNA4TO-HA-TRPM4 c.25delAG (TRPM4 MUT) and pcDNA4 did not express any HA-TRPM4; only HEK293T cells transfected with wild-type pcDNA4TO-HA-TRPM4 (TRPM4 WT)-expressed HA-TRPM4.

3.7. TRPM4 c.25-1G > T Led to a Frameshift Transcript r.25_26del and Loss of TRPM4 Protein

The in silico analysis of the *TRPM4* wild type and c.25-1G > T sequences with MaxEnt Scan (MES) showed that this variant disrupted the canonical acceptor site of TRPM4 exon 2 (MES = 11.38 → 2.78) but simultaneously created a strong de novo acceptor site two nucleotides downstream (MES = 9.9).

The minigene mgTRPM4_ex1-2 was constructed by using DNA from a patient (III-5). The wild-type and mutant minigenes were introduced into MCF-7 cells, and RNA was isolated and reverse transcribed (Figure 2B). The wild-type minigene produced a single transcript of the expected size (394 nt) and structure (V1- TRPM4 ex1-2 –V2). The c.25-1G>T-minigene also generated a single transcript whose sequence revealed the loss of the first two nucleotides of exon 2 (r.25_26del), using the new alternative acceptor site predicted by MES. This transcript was called Δ(E2p2) by following previously suggested

rules [34]. Moreover, Δ(E2p2) would introduce a frameshift and a premature protein truncation 17 codons downstream (p. Ser9Leufs*17) that would inactivate the TRPM4 protein. High-resolution fragment analysis of fluorescent RT-PCR products confirmed the presence of this aberrant transcript (Figures 2C and S2).

As *TRPM4* c.25-1G > T was not in the cDNA, and the minigene assay showed that *TRPM4* c.25-1G>T led to r.25_26del., i.e., equivalent of c.25delAG, we created pcDNATO-HA-*TRPM4* c.25delAG. We transfected the pcDNATO-HA-*TRPM4* c.25delAG into HEK293T cells to see if the variant affects the expression of TRPM4. As shown in Figure 2D, only HEK293T cells transfected with wild-type pcDNA4TO-HA-*TRPM4*-expressed HA-TRPM4, while HEK293T cells transfected with pcDNA4TO-HA-*TRPM4* c.25delAG or pcDNA4 did not express HA-TRPM4, confirming that *TRPM4* c.25delAG resulted from *TRPM4* c.25-1G > T led to a loss of TRPM4 protein.

3.8. Knockdown of CYBA or TRPM4 Promoted Proliferation in HT-29 Cells

To study the effect of *CYBA* and *TRPM4* on cell proliferation, we knocked down *CYBA* and *TRPM4* with respective siRNAs, separately, in two colorectal adenocarcinoma cell lines, LS174T and HT-29; LS174T is a mucin secreting cell line. Moreover, qPCR and Western blot analysis confirmed the absence of CYBA or TRPM4 mRNA and protein expression in both LS174T and HT-29 cells (Supplementary Figure S3A–D). While knockdown of *CYBA* or *TRPM4* did not seem to affect cell proliferation of LS174T cells in in vitro cell proliferation assay (Figure 3A,B), knockdown of *CYBA* or *TRPM4* in the HT-29 cells led to an increase in cell proliferation throughout the time up to 96 h (Figure 3C,D; $p < 0.05$).

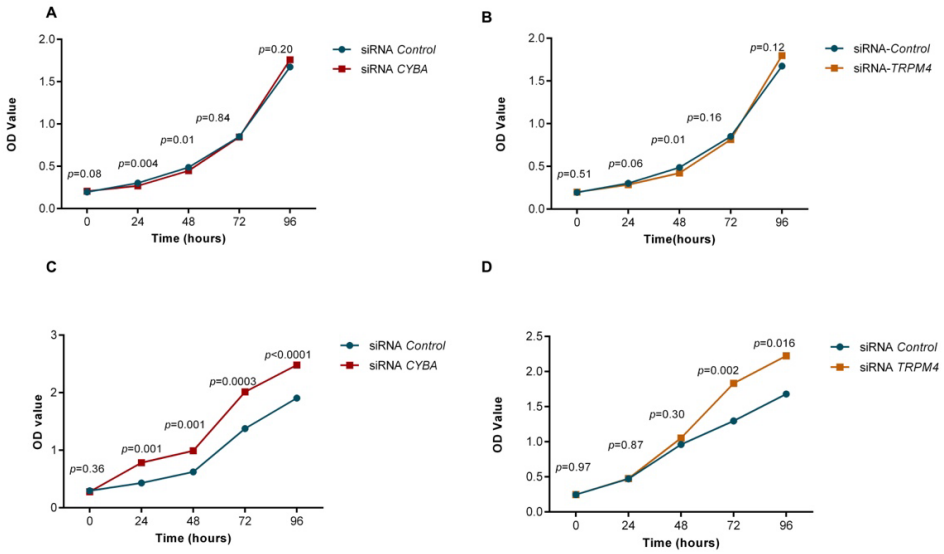

Figure 3. Effect of CYBA or TRPM4 knockdown on cell proliferation of LS174T and HT29 cells measured by CCK-8 cell assay. (A) LS174T cells transfected with siRNA-CYBA (**A**) or siRNA-TRPM4 (**B**) showed similar proliferation compared to siRNA-control transfected cells. HT29 cells transfected with siRNA-CYBA (**C**) or siRNA-TRPM4 (**D**) showed a significant increase in proliferation compared to siRNA-control transfected cells.

3.9. Knockdown of TRPM4 Decreased MUC1 and MUC2 in Both LS174T and HT-29 Cells

In order to test the hypothesis that loss of *TRPM4* disrupts the mucus barrier by inhibiting mucin secretion, we measured the expression of *MUC1* and *MUC2* by qPCR after downregulating *TRPM4* in LS174T and HT-29 cell lines. The mRNA levels of *MUC1* and *MUC2* were decreased after knockdown *TRPM4* in both LS174T and HT-29 cells ($p \leq 0.0001$; Figure 4A,B). Moreover, the protein expression of MUC2, measured by immunofluorescence was decreased in both LS174T and HT-29 cells after the knockdown of *TRPM4* compared to those treated with control siRNA (LS174T: $p = 0.002$; HT-29: $p = 0.006$; Figure 4C,D).

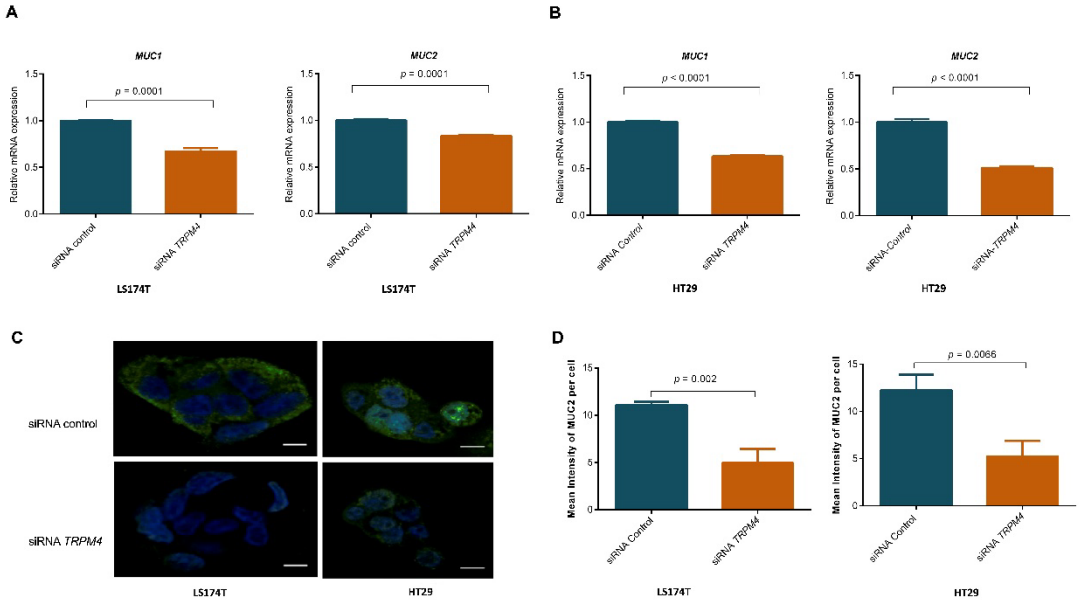

Figure 4. RT-PCR results showed that mRNA levels of *MUC1* and *MUC2* decreased after siRNA knockdown of *TRPM4* in LS174T (**A**) and HT-29 (**B**) cells. Three independent experiments in triplicate were performed, and means and standard errors of the means are presented on the graphs. A clear decrease of MUC2 expression is seen in immunofluorescence after knockdown of *TRPM4*, especially in mucin-secreting cell line LS174T (**C**), but also in HT-29 (**D**) cells. Two independent experiments were performed, and means and standard errors of the means are presented on the graphs.

3.10. Knockdown of CYBA or TRPM4 Led to ROS Deficiency

As *CYBA* is involved in ROS modulation, we measured the ROS activity in both LS174T and HT-29 cells after siRNA-mediated knockdown of *CYBA* or *TRPM4*. There was a small, but significant decrease in ROS activity in LS174T cells after the knockdown of *CYBA* or *TRPM4* ($p = 0.009$ and 0.014; Figure 5A), but no decrease in HT-29 cells (Figure 5B).

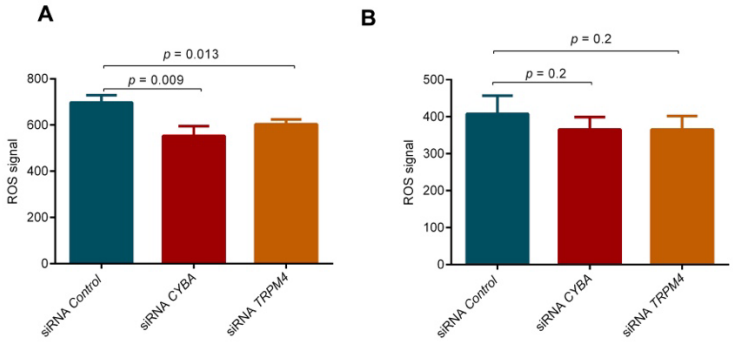

Figure 5. ROS activity upon *CYBA* or *TRPM4* knockdown. *CYBA* or *TRPM4* depletion led to significant reduction of ROS in LS174T cells (**A**), but no decrease in HT-29 cells (**B**). Three independent experiments in triplicate were performed, and means and standard errors of the means are presented on the graphs.

4. Discussion

We carried out germline WGS analyses of 15 Polish CRC families, each of which presented with several CRCs. Loss-of-function variant analysis in two separate families pointed out two rare variants in the genes *CYBA* and *TRPM4* that segregated with the disease in these families. The variant in *CYBA* (*CYBA* c.246delC) leads to a frameshift and a premature protein truncation 41 codons downstream (p.Phe83Asnfs*41), while the

splice site variant in *TRPM4* (*TRPM4* c.25-1 G > T) leads to a frameshift and a premature protein truncation 17 codons downstream (p. Ser9Leufs*17). We showed that both variants lead to a loss of protein expression. *CYBA* encodes the alpha chain of cytochrome B-245, which is part of the NADPH oxidase (NOX) system, to generate superoxide [35]. It is a component in NOX1–4, of which NOX1 is expressed in the gastrointestinal tract and NOX2 in phagocytes [36]. *TRPM4* encodes a nonselective monovalent cation channel, the upregulation of which enhances sodium entry, which, in turn, leads to depolarization of the membrane potential [37]. Both genes are highly expressed in the colon and are active in mucin-secreting goblet cells [38,39]. Although the functions of these two genes appeared initially to be quite diverse, our literature search resulted in a challenging combination of pathways targeting inflammatory bowel disease, a risk factor of CRC [14]. The present study gives further insight into the complementary function of the *CYBA* and *TRPM4* pathways, which converge in the protection of the colonic epithelium. The colon is covered by an inner colon mucus layer that is proteolytically converted into an outer mucus layer, both consisting of the MUC2 mucin as the main protein component [40]. The inner mucus layer acts as a barrier against bacteria; only when it is damaged are bacteria able to reach the epithelial surface, which may lead to severe inflammation [40].

Superoxide, generated by NOX analogues, is a member of ROS for which the optimal concentration is critical; overproduction will lead to oxidative stress and development of disease, and, likewise, insufficient ROS production may be detrimental to health [38]. Germline variants in *CYBA* are associated with autosomal recessive chronic granulomatous disease, characterized by the failure of activated phagocytes (neutrophils and macrophages), to generate enough superoxide to accomplice intracellular killing of pathogens. Patients suffer from life-threatening infections and from excessive inflammatory reactions [41]. About half of chronic granulomatous disease colon patients present with granuloma formation and acute or chronic inflammation mimicking inflammatory bowel disease [42]. Germline *CYBA* variants have been associated with early onset inflammatory bowel disease [43,44]. Bacterial penetration of colonic epithelium, the normally restricted zone, is observed in many colitis models and also in patients with inflammatory bowel disease, such as ulcerative colitis [45]. Mucus defects that allow bacteria to reach the epithelium and to stimulate an immune response can lead to the development of intestinal inflammation [46]. The comparison of inflamed ulcerative colitis patients and $MUC2^{-/-}$ mice revealed that bacteria in both models had been able to penetrate the colon epithelium, thus causing inflammation [46]. $MUC2^{-/-}$ mice have also been shown to develop adenomas in the intestine, the majority of which even progressed into adenocarcinomas [47].

An important clue about the mechanism of action of *CYBA* in inflammatory bowel disease was recently shown in *CYBA* mutant mice which generated low intestinal ROS [48]. The mice suffered from profound mucus layer disruption, with bacterial penetration into crypts, and from a compromised innate immune response to invading microbes, leading to mortality. The results implicated the loss of the mucus barrier and innate immune defense in the pathogenesis of intestinal inflammation, due to ROS deficiency [48]. In another study, using a dextran sodium sulfate–induced colitis mouse model, ROS deficient Ncf1-mutant mice developed well-differentiated adenocarcinomas, while only high-grade dysplasia without malignant invasion was observed in ROS-proficient mice with Ncf1 wild type, suggesting that ROS deficiency may cause CRC in response to environmental risk factors [49]. Our experimental results and data from the literature suggest that the loss-of function variant in *CYBA* promotes CRC by at least two mechanisms because of depressed ROS production, first by faltering defense against intestinal bacteria at colonic epithelium and second by suppressing bacterial killing by intestinal phagocytes, such as neutrophils and macrophages (Figure 6A) [38].

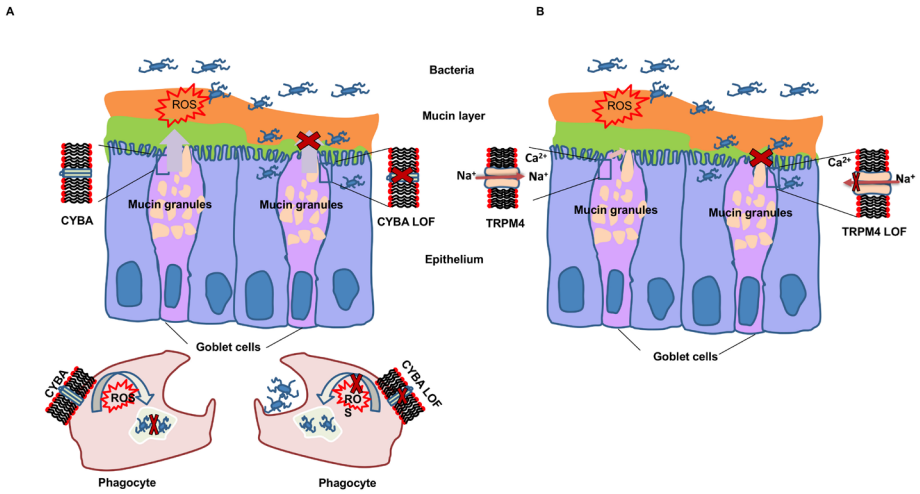

Figure 6. Schematic presentation of the suggested consequences of *CYBA* and *TRPM4* variants on colonic mucin layer integrity. (**A**) The loss-of-function variant in *CYBA* leads to decreased ROS production and promotes CRC by faltering defense against intestinal bacteria at colonic epithelium and by suppressing bacterial killing by intestinal phagocytes. (**B**) The loss-of-function variant in *TRPM4* leads to decreased mucus secretion due to inactivation of the TRPM4 channel, potentially leading to mucus-layer disruption with bacterial penetration into the mucin-protected epithelium, inflammation and colorectal cancer. The function of the wild-type CYBA and TRPM4 is shown in the left goblet cell; the function of CYBA and TRPM4 after the loss-of-function (LOF) variant is shown in the right goblet cell.

TRPM4 encodes a calcium-activated nonselective ion channel, the activity of which increases with increasing intracellular calcium concentration; however, this channel does not transport calcium. Disturbance of the membrane potential is deleterious to calcium homeostasis, and this is suggested to contribute to carcinogenesis through uncontrolled cell migration and invasion [37]. The *TRPM4* gene is highly expressed in CRC tumor buds, contributing to the proliferation and invasion of tumor cells [50]. TRPM4 cooperates in the control of mucin secretion from goblet cells in response to extracellular stimuli [39]. The results in the human colonic cancer goblet cell line (HT-29-18N2) showed that TRPM4 protein controls calcium-mediated secretion of MUC2 and MUC5AC in conjunction with a Na^+/Ca^{2+} exchanger NCX [39]. Knockdown of TRPM4 in the HT29-18N2 cells blocked MUC2 secretion [39]. In support of this finding, we showed for the first time that *TRPM4* knockdown also affects *MUC2* mRNA and protein expression. These findings suggest that the loss-of-function variant in *TRPM4* leads to disruption of the mucus layer, allowing bacterial penetration into the mucin-protected epithelium, and thus resulting in inflammation and risk of CRC (Figure 6B). In their publication, Cantero-Recasens et. al. speculated about two distinct modes of mucin secretion: In baseline mucin secretion, TRPM4 is inactive and mucin secretion is regulated by KChIP3 and intracellular Ca^{2+} oscillation [39]. Stimulated mucin secretion is caused by exogenous stimulation of the cells by, for example, ATP or IL-13, which leads to a rapid Ca^{2+} release from the endoplasmic reticulum, activation of TRPM4 and increase of intracellular Na^+. This, in turn, triggers NCX to export Na^+ and import Ca^{2+}, leading to a rapid burst in mucin secretion. The above authors showed similar results in goblet cells earlier with TRPM5, a homologue of TRPM4 [51]. *TRPM4* and *TRPM5* share high sequence homology and similar biophysical properties, yet they are not able to fully compensate each other [51]. In addition, by affecting mucin expression and secretion, TRPM4 may influence inflammation through the modulation of immune responses [31,37]. As an alternative mechanism, TRPM4 has been implicated in the regulation of the Wnt signaling pathway, which is often dysregulated in CRC [37].

The most commonly mutated gene in CRC is *KRAS*. When mutated, oncogenic RAS-mediated signaling cascades not only drive tumor initiation, maintenance and progression,

but also affect cellular metabolism, including ROS signaling [52]. While mutated *KRAS* reportedly is leading to increased ROS generation and tumor promotion, elevated ROS levels may also trigger tumor cell death. In our study, we observed a cell-line-dependent decrease in ROS generation and cell proliferation after the knockdown of *CYBA* or *TRPM4*. Thus, the effect on ROS generation and downstream signaling may be context dependent and requires further research [37]. On the other hand, oncogenic RAS signaling may promote tumorigenesis, both when the MUC2 production is suppressed or enhanced and depending on the specific subtype of CRC [53]. Unfortunately, no tumor sample or tumor pathology data were available from the CRC families to analyze the type of tumors of the patients.

The present variants were not present in the gnomAD database, which listed eight other loss-of-function variants for *CYBA* and 54 for *TRPM4*; the database covers 125,748 exome sequences. Of note, exactly the same *CYBA* c.246delC variant was submitted to the Clinvar database as a pathogenic variant of granulomatous disease, which is a recessive condition (Available online: https://www.ncbi.nlm.nih.gov/clinvar/variation/619030/ (accessed on 17 January 2022)). In the CRC family, the variant was present in the heterozygous form, and the phenotype would be expected to be milder with a later onset than in the case of granulomatous disease. Unfortunately, tumor samples from the patients were not available for testing this. Although the *CYBA* variant was absent from 125,748 exomes of the gnomAD database, it was found in four additional familial CRC patients among 1705 tested Polish patients and in 2/1674 controls, of whom only one was past the common diagnostic age of CRC. As a DNA sample was only available for the index case of the four families, we could not evaluate the segregation of the variant with the disease in these families, and this was a limitation of our study.

The Polish National Colorectal Cancer Screening program offers colonoscopy to all inhabitants between 50 and 66 years; individuals with family history of CRC are eligible starting at the age of 40 years. Among 236,089 individuals screened from 2000 to 2011, 17.7% were diagnosed with adenomas; among individuals with a family history, 18.9% [54]. As individuals with adenomas are at an increased risk of CRC, we considered family members diagnosed with polyps as potential carriers of the predisposing variant. Since we did not have access to any clinical or pathological features of the patients, their tumors and polyps, we were not able to evaluate the risk category of these individuals. Thus, we can only speculate about the role of inflammation in the patients carrying the variant.

In the present study, we focused on the loss-of-function variants, as these are considered to be the most deleterious types of sequence variants. We plan to assess the less deleterious variants, such as missense variants in the future, but they require an increased focus on functional characterization. Many biological pathways are cell and tissue specific, thus calling for the application of different cell lines. In the present study, MUC1 and MUC2 secretion was affected in both LS174T and HT-29 cell lines, while proliferation was affected only in the HT-29 cells, and the effect on ROS activity was more prominent in the LS174T cells.

5. Conclusions

Our germline sequencing efforts of familial CRC led to the identification of two possible pathogenic variants of two likely CRC-predisposing genes *CYBA* and *TRPM4*. The function of these two genes appeared to provide complementary pathways of protecting the colonic epithelium, as shown in Figure 6 [40]. Although our data did not allow us to demonstrate a direct mechanism to cancer formation, we showed associations of the *CYBA* and *TRPM4* functions with ROS and mucin production with a likely link to tumorigenesis of CRC.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers14030670/s1, Figure S1. Sanger sequencing results of the CYBA (upper panel) and TRPM4 (lower panel) wild type and mutant sequences; Figure S2. Splicing functional assays of TRPM4 variant c.25-1G>T. (A) Fluorescent fragment analysis by capillary

electrophoresis of RT-PCR products generated by the wild type and mutant minigenes. The canonical and the aberrant Δ(E2p2) transcripts are shown as blue and red peaks, respectively, while the LIZ500 size standard is displayed as orange peaks. (B) Schematic representation of the canonical (above) and anomalous (below) transcripts. Variant c.25-1G>T is shown in red; acceptor sites are underlined; splicing events are shown as broken arrows; Figure S3. SiRNA knockdown of CYBA or TRPM4 lead to reduced mRNA and protein of the respective genes: mRNA (A) and protein (B) levels of CYBA and TRPM4 decreased after siRNA knockdown of CYBA or TRPM4 in LS174T cells; mRNA (C) and protein (D) levels of CYBA and TRPM4, decreased after siRNA knockdown of CYBA or TRPM4 in HT-29 cells. For real-time PCR assays, three independent experiments in triplicate were performed and means and standard errors of the means are presented on the graphs. For western blot two independent experiments were performed;Table S1. Summary of the sequencing coverage and quality statistics for each sample; Table S2. Loss-of-function variants segregated with the disease in 15 colorectal cancer families.

Author Contributions: Conceptualization, O.R.B., K.H. and A.F.; data curation, N.P. and M.S.; formal analysis, L.Z., N.P., M.S. and A.K.; funding acquisition, E.A.V. and K.H.; investigation, L.Z., B.M., E.B.-M., L.S.-M., E.A.V. and O.R.B.; methodology, L.Z., B.M., N.P., M.S. and O.R.B.; project administration, K.H. Y.Y. and A.F.; resources, D.D., M.K., N.P., M.S., Y.Y. and J.L.; software, N.P., M.S. and A.K.; supervision, O.R.B., K.H. and A.F.; validation, D.D. and M.K.; visualization, L.Z., B.M., E.B.-M., L.S.-M., E.A.V. and O.R.B.; writing—original draft, L.Z., O.R.B., K.H. and A.F.; writing—review and editing, all authors. All authors have read and agree to the published version of the manuscript.

Funding: L.Z. was supported by Zhejiang University Lu's Graduate Education International Exchange Fund. E.A.V.'s lab was supported by grants from the Spanish Ministry of Science, Innovation and Universities, Plan Nacional de I+D+I 2013-2016, ISCIII (PI17/00227) co-funded by FEDER from Regional Development European Funds (European Union), and grant CSI242P18 (actuación cofinanciada P.O. FEDER 2014-2020 de Castilla y León) from the Consejería de Educación, Junta de Castilla y León. L.S.-M. was supported by a predoctoral fellowship from the Spanish Association Against Cancer (AECC), Junta Provincial de Valladolid (2019-2023). K.H. was supported by the EU Horizon 2020 program, grant No. 856620. A.K. is a recipient of the Ramalingaswami Re-Retry Faculty Fellowship (Grant; BT/RLF/Re-entry/38/2017). This article is based upon work from COST Action CA17118, supported by COST (European Cooperation in Science and Technology).

Institutional Review Board Statement: The study was approved by the Bioethics Committee of the Pomeranian Medical Academy in Szczecin No. BN-001/174/05.

Informed Consent Statement: All participating individuals signed an informed consent.

Data Availability Statement: The WGS data generated in this study are available in EGA under accession number EGAS00001005118. Other data that support the findings of this study are available from the corresponding author upon request.

Acknowledgments: The authors thank the members of the families for participating in this study. We further thank the Genomics and Proteomics Core Facility (GPCF) of the German Cancer Research Center (DKFZ) for providing excellent library preparation and sequencing services and the Omics IT and Data Management Core Facility (ODCF) of the DKFZ for the whole-genome sequencing data management.

Conflicts of Interest: The authors disclose no conflict. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [CrossRef] [PubMed]
2. Frank, C.; Fallah, M.; Sundquist, J.; Hemminki, A.; Hemminki, K. Population Landscape of Familial Cancer. *Sci. Rep.* **2015**, *5*, 12891. [CrossRef]
3. Gupta, S.; Provenzale, D.; Regenbogen, S.E.; Hampel, H.; Slavin, T.P.; Hall, M.J.; Llor, X.; Chung, D.C.; Ahnen, D.J.; Bray, T.; et al. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Colorectal, Version 3.2017. *J. Natl. Compr. Cancer Netw.* **2017**, *15*, 1465–1475. [CrossRef] [PubMed]

4. Lichtenstein, P.; Holm, N.V.; Verkasalo, P.K.; Iliadou, A.; Kaprio, J.; Koskenvuo, M.; Pukkala, E.; Skytthe, A.; Hemminki, K. Environmental and heritable factors in the causation of cancer—Analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **2000**, *343*, 78–85. [CrossRef] [PubMed]
5. Mucci, L.A.; Hjelmborg, J.B.; Harris, J.R.; Czene, K.; Havelick, D.J.; Scheike, T.; Graff, R.E.; Holst, K.; Moller, S.; Unger, R.H.; et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *JAMA* **2016**, *315*, 68–76. [CrossRef] [PubMed]
6. Valle, L.; de Voer, R.M.; Goldberg, Y.; Sjursen, W.; Forsti, A.; Ruiz-Ponte, C.; Caldes, T.; Garre, P.; Olsen, M.F.; Nordling, M.; et al. Update on genetic predisposition to colorectal cancer and polyposis. *Mol. Asp. Med.* **2019**, *69*, 10–26. [CrossRef]
7. Bellido, F.; Pineda, M.; Aiza, G.; Valdes-Mas, R.; Navarro, M.; Puente, D.A.; Pons, T.; Gonzalez, S.; Iglesias, S.; Darder, E.; et al. POLE and POLD1 mutations in 529 kindred with familial colorectal cancer and/or polyposis: Review of reported cases and recommendations for genetic testing and surveillance. *Genet. Med.* **2016**, *18*, 325–332. [CrossRef]
8. Grolleman, J.E.; de Voer, R.M.; Elsayed, F.A.; Nielsen, M.; Weren, R.D.A.; Palles, C.; Ligtenberg, M.J.L.; Vos, J.R.; Ten Broeke, S.W.; de Miranda, N.; et al. Mutational Signature Analysis Reveals NTHL1 Deficiency to Cause a Multi-tumor Phenotype. *Cancer Cell* **2019**, *35*, 256–266. [CrossRef]
9. Jaeger, E.; Leedham, S.; Lewis, A.; Segditsas, S.; Becker, M.; Cuadrado, P.R.; Davis, H.; Kaur, K.; Heinimann, K.; Howarth, K.; et al. Hereditary mixed polyposis syndrome is caused by a 40-kb upstream duplication that leads to increased and ectopic expression of the BMP antagonist GREM1. *Nat. Genet.* **2012**, *44*, 699–703. [CrossRef]
10. Buchanan, D.D.; Stewart, J.R.; Clendenning, M.; Rosty, C.; Mahmood, K.; Pope, B.J.; Jenkins, M.A.; Hopper, J.L.; Southey, M.C.; Macrae, F.A.; et al. Risk of colorectal cancer for carriers of a germ-line mutation in POLE or POLD1. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **2018**, *20*, 890–895. [CrossRef]
11. Frank, C.; Fallah, M.; Ji, J.; Sundquist, J.; Hemminki, K. The population impact of familial cancer, a major cause of cancer. *Int. J. Cancer* **2014**, *134*, 1899–1906. [CrossRef]
12. Betge, J.; Schneider, N.I.; Harbaum, L.; Pollheimer, M.J.; Lindtner, R.A.; Kornprat, P.; Ebert, M.P.; Langner, C. MUC1, MUC2, MUC5AC, and MUC6 in colorectal cancer: Expression profiles and clinical significance. *Virchows Arch.* **2016**, *469*, 255–265. [CrossRef] [PubMed]
13. Rahman, N. Realizing the promise of cancer predisposition genes. *Nature* **2014**, *505*, 302–308. [CrossRef] [PubMed]
14. Keller, D.S.; Windsor, A.; Cohen, R.; Chand, M. Colorectal cancer in inflammatory bowel disease: Review of the evidence. *Tech. Coloproctol.* **2019**, *23*, 3–13. [CrossRef] [PubMed]
15. Lahiri, D.K.; Schnabel, B. DNA isolation by a rapid method from human blood samples: Effects of MgCl2, EDTA, storage time, and temperature on DNA yield and quality. *Biochem. Genet.* **1993**, *31*, 321–328. [CrossRef]
16. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [CrossRef]
17. Liu, X.; Wu, C.; Li, C.; Boerwinkle, E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum. Mutat.* **2016**, *37*, 235–241. [CrossRef]
18. Genomes Project, C.; Auton, A.; Brooks, L.D.; Durbin, R.M.; Garrison, E.P.; Kang, H.M.; Korbel, J.O.; Marchini, J.L.; McCarthy, S.; McVean, G.A.; et al. A global reference for human genetic variation. *Nature* **2015**, *526*, 68–74. [CrossRef]
19. Smigielski, E.M.; Sirotkin, K.; Ward, M.; Sherry, S.T. dbSNP: A database of single nucleotide polymorphisms. *Nucleic Acids Res.* **2000**, *28*, 352–355. [CrossRef]
20. Lek, M.; Karczewski, K.J.; Minikel, E.V.; Samocha, K.E.; Banks, E.; Fennell, T.; O'Donnell-Luria, A.H.; Ware, J.S.; Hill, A.J.; Cummings, B.B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [CrossRef]
21. Kumar, A.; Bandapalli, O.R.; Paramasivam, N.; Giangiobbe, S.; Diquigiovanni, C.; Bonora, E.; Eils, R.; Schlesner, M.; Hemminki, K.; Forsti, A. Familial Cancer Variant Prioritization Pipeline version 2 (FCVPPv2) applied to a papillary thyroid cancer family. *Sci. Rep.* **2018**, *8*, 11635. [CrossRef]
22. Thorvaldsdottir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192. [CrossRef] [PubMed]
23. Kircher, M.; Witten, D.M.; Jain, P.; O'Roak, B.J.; Cooper, G.M. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **2014**, *46*, 310–315. [CrossRef] [PubMed]
24. Roller, E.; Ivakhno, S.; Lee, S.; Royce, T.; Tanner, S. Canvas: Versatile and scalable detection of copy number variants. *Bioinformatics* **2016**, *32*, 2375–2377. [CrossRef] [PubMed]
25. Srivastava, A.; Giangiobbe, S.; Kumar, A.; Paramasivam, N.; Dymerska, D.; Behnisch, W.; Witzens-Harig, M.; Lubinski, J.; Hemminki, K.; Forsti, A.; et al. Identification of Familial Hodgkin Lymphoma Predisposing Genes Using Whole Genome Sequencing. *Front. Bioeng. Biotechnol.* **2020**, *8*, 179. [CrossRef]
26. Yeo, G.; Burge, C.B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* **2004**, *11*, 377–394. [CrossRef]
27. Desmet, F.O.; Hamroun, D.; Lalande, M.; Collod-Beroud, G.; Claustres, M.; Beroud, C. Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **2009**, *37*, e67. [CrossRef]
28. Acedo, A.; Hernandez-Moro, C.; Curiel-Garcia, A.; Diez-Gomez, B.; Velasco, E.A. Functional classification of BRCA2 DNA variants by splicing assays in a large minigene with 9 exons. *Hum. Mutat.* **2015**, *36*, 210–221. [CrossRef]
29. Bryksin, A.V.; Matsumura, I. Overlap extension PCR cloning: A simple and reliable way to create recombinant plasmids. *BioTechniques* **2010**, *48*, 463–465. [CrossRef]

30. Fraile-Bethencourt, E.; Diez-Gomez, B.; Velasquez-Zapata, V.; Acedo, A.; Sanz, D.J.; Velasco, E.A. Functional classification of DNA variants by hybrid minigenes: Identification of 30 spliceogenic variants of BRCA2 exons 17 and 18. *PLoS Genet.* **2017**, *13*, e1006691. [CrossRef]
31. Bianchi, B.; Smith, P.A.; Abriel, H. The ion channel TRPM4 in murine experimental autoimmune encephalomyelitis and in a model of glutamate-induced neuronal degeneration. *Mol. Brain* **2018**, *11*, 41. [CrossRef] [PubMed]
32. Tawiah, A.; Cornick, S.; Moreau, F.; Gorman, H.; Kumar, M.; Tiwari, S.; Chadee, K. High MUC2 Mucin Expression and Misfolding Induce Cellular Stress, Reactive Oxygen Production, and Apoptosis in Goblet Cells. *Am. J. Pathol.* **2018**, *188*, 1354–1373. [CrossRef] [PubMed]
33. Cobo, E.R.; Kissoon-Singh, V.; Moreau, F.; Holani, R.; Chadee, K. MUC2 Mucin and Butyrate Contribute to the Synthesis of the Antimicrobial Peptide Cathelicidin in Response to Entamoeba histolytica- and Dextran Sodium Sulfate-Induced Colitis. *Infect. Immun.* **2017**, *85*, e00905–e00916. [CrossRef] [PubMed]
34. Fraile-Bethencourt, E.; Valenzuela-Palomo, A.; Diez-Gomez, B.; Caloca, M.J.; Gomez-Barrero, S.; Velasco, E.A. Minigene Splicing Assays Identify 12 Spliceogenic Variants of BRCA2 Exons 14 and 15. *Front. Genet.* **2019**, *10*, 503. [CrossRef] [PubMed]
35. O'Neill, S.; Brault, J.; Stasia, M.J.; Knaus, U.G. Genetic disorders coupled to ROS deficiency. *Redox Biol.* **2015**, *6*, 135–156. [CrossRef]
36. Stasia, M.J. CYBA encoding p22(phox), the cytochrome b558 alpha polypeptide: Gene structure, expression, role and physiopathology. *Gene* **2016**, *586*, 27–35. [CrossRef]
37. Gao, Y.; Liao, P. TRPM4 channel and cancer. *Cancer Lett.* **2019**, *454*, 66–69. [CrossRef]
38. Aviello, G.; Knaus, U.G. NADPH oxidases and ROS signaling in the gastrointestinal tract. *Mucosal Immunol.* **2018**, *11*, 1011–1023. [CrossRef]
39. Cantero-Recasens, G.; Butnaru, C.M.; Brouwers, N.; Mitrovic, S.; Valverde, M.A.; Malhotra, V. Sodium channel TRPM4 and sodium/calcium exchangers (NCX) cooperate in the control of Ca(2+)-induced mucin secretion from goblet cells. *J. Biol. Chem.* **2019**, *294*, 816–826. [CrossRef]
40. Johansson, M.E.; Hansson, G.C. Mucus and the goblet cell. *Dig. Dis.* **2013**, *31*, 305–309. [CrossRef]
41. Roos, D. Chronic granulomatous disease. *Br. Med. Bull.* **2016**, *118*, 50–63. [CrossRef] [PubMed]
42. Alimchandani, M.; Lai, J.P.; Aung, P.P.; Khangura, S.; Kamal, N.; Gallin, J.I.; Holland, S.M.; Malech, H.L.; Heller, T.; Miettinen, M.; et al. Gastrointestinal histopathology in chronic granulomatous disease: A study of 87 patients. *Am. J. Surg. Pathol.* **2013**, *37*, 1365–1372. [CrossRef] [PubMed]
43. Ashton, J.J.; Andreoletti, G.; Coelho, T.; Haggarty, R.; Batra, A.; Afzal, N.A.; Beattie, R.M.; Ennis, S. Identification of Variants in Genes Associated with Single-gene Inflammatory Bowel Disease by Whole-exome Sequencing. *Inflamm. Bowel Dis.* **2016**, *22*, 2317–2327. [CrossRef] [PubMed]
44. Denson, L.A.; Jurickova, I.; Karns, R.; Shaw, K.A.; Cutler, D.J.; Okou, D.T.; Dodd, A.; Quinn, K.; Mondal, K.; Aronow, B.J.; et al. Clinical and Genomic Correlates of Neutrophil Reactive Oxygen Species Production in Pediatric Patients with Crohn's Disease. *Gastroenterology* **2018**, *154*, 2097–2110. [CrossRef]
45. Johansson, M.E. Mucus layers in inflammatory bowel disease. *Inflamm. Bowel Dis.* **2014**, *20*, 2124–2131. [CrossRef]
46. Wenzel, U.A.; Magnusson, M.K.; Rydstrom, A.; Jonstrand, C.; Hengst, J.; Johansson, M.E.; Velcich, A.; Ohman, L.; Strid, H.; Sjovall, H.; et al. Spontaneous colitis in Muc2-deficient mice reflects clinical and cellular features of active ulcerative colitis. *PLoS ONE* **2014**, *9*, e100217. [CrossRef]
47. Velcich, A.; Yang, W.; Heyer, J.; Fragale, A.; Nicholas, C.; Viani, S.; Kucherlapati, R.; Lipkin, M.; Yang, K.; Augenlicht, L. Colorectal cancer in mice genetically deficient in the mucin Muc2. *Science* **2002**, *295*, 1726–1729. [CrossRef]
48. Aviello, G.; Singh, A.K.; O'Neill, S.; Conroy, E.; Gallagher, W.; D'Agostino, G.; Walker, A.W.; Bourke, B.; Scholz, D.; Knaus, U.G. Colitis susceptibility in mice with reactive oxygen species deficiency is mediated by mucus barrier and immune defense defects. *Mucosal Immunol.* **2019**, *12*, 1316–1326. [CrossRef]
49. Carvalho, L.; Gomes, J.R.M.; Tavares, L.C.; Xavier, A.R.; Klika, K.D.; Holmdahl, R.; Carvalho, R.A.; Souto-Carneiro, M.M. Reactive Oxygen Species Deficiency Due to Ncf1-Mutation Leads to Development of Adenocarcinoma and Metabolomic and Lipidomic Remodeling in a New Mouse Model of Dextran Sulfate Sodium-Induced Colitis. *Front. Immunol.* **2018**, *9*, 701. [CrossRef]
50. Kappel, S.; Stoklosa, P.; Hauert, B.; Ross-Kaschitza, D.; Borgstrom, A.; Baur, R.; Galvan, J.A.; Zlobec, I.; Peinelt, C. TRPM4 is highly expressed in human colorectal tumor buds and contributes to proliferation, cell cycle, and invasion of colorectal cancer cells. *Mol. Oncol.* **2019**, *13*, 2393–2405. [CrossRef]
51. Guo, J.; She, J.; Zeng, W.; Chen, Q.; Bai, X.C.; Jiang, Y. Structures of the calcium-activated, non-selective cation channel TRPM4. *Nature* **2017**, *552*, 205–209. [CrossRef]
52. Mukhopadhyay, S.; Vander Heiden, M.G.; McCormick, F. The Metabolic Landscape of RAS-Driven Cancers from biology to therapy. *Nat. Cancer* **2021**, *2*, 271–283. [CrossRef]
53. Gundamaraju, R.; Chong, W.C. Consequence of distinctive expression of MUC2 in colorectal cancers: How much is actually bad? *Biochim. Biophys. Acta Rev. Cancer* **2021**, *1876*, 188579. [CrossRef]
54. Wieszczy, P.; Kaminski, M.F.; Franczyk, R.; Loberg, M.; Kobiela, J.; Rupinska, M.; Kocot, B.; Rupinski, M.; Holme, O.; Wojciechowska, U.; et al. Colorectal Cancer Incidence and Mortality After Removal of Adenomas During Screening Colonoscopies. *Gastroenterology* **2020**, *158*, 875–883. e5. [CrossRef]