Epistatic effect of TLR3 and cGAS-STING-IKKε-TBK1-IFN signaling variants on colorectal cancer risk

Calogerina Catalano1,2 | Miguel Inacio da Silva Filho1 | Christoph Frank1 | Shun Lu3 | Katerina Jiraskova4,5 | Veronika Vymetalkova4,5,6 | Miroslav Levy7 | Vaclav Liska6,8 | Ondrej Vycital6,8 | Alessio Naccarati4,9 | Ludmila Vodickova4,5,6 | Kari Hemminki1,10 | Pavel Vodicka4,5,6 | Alexander N. R. Weber11 | Asta Försti1,10,12,13

1Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
2Department of Internal Medicine V, University of Heidelberg, Heidelberg, Germany
3Sichuan Cancer Center, School of Medicine, Sichuan Cancer Hospital & Institute, University of Electronic Science and Technology of China, Chengdu, China
4Department of Molecular Biology of Cancer, Institute of Experimental Medicine, the Czech Academy of Sciences, Prague, Czech Republic
51st Medical Faculty, Institute of Biology and Medical Genetics, Charles University, Prague, Czech Republic
6Faculty of Medicine in Pilsen, Biomedical Center, Charles University Prague, Pilsen, Czech Republic
7First Medical Faculty, Department of Surgery, Charles University and Thomayer Hospital, Prague, Czech Republic
8Department of Surgery, Teaching Hospital and Medical School of Charles University, Pilsen, Czech Republic
9Molecular and Genetic Epidemiology, Italian Institute for Genomic Medicine (IIGM), Turin, Italy
10Center for Primary Health Care Research, Clinical Research Center, Lund University, Malmö, Sweden
11Department of Immunology, Interfaculty Institute for Cell Biology, University of Tübingen, Baden-Württemberg, Tübingen, Germany
12Hopp Children's Cancer Center (KiTZ), Heidelberg, Germany
13Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), Heidelberg, Germany

Correspondence
Calogerina Catalano and Asta Försti, Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, Heidelberg, Baden-Württemberg, 69120, Germany.
Email: cal.catalano88@gmail.com (C. C.); a.foersti@kitz-heidelberg.de (A. F.)

Funding information
The work was supported by Internal Grant Agency of the Ministry of Health of the Czech Republic under Grant AZV 15-27580A and AZV 15-26535A; the Grant Agency of the Czech Republic under Grant 17-16857S, and the Center Of Clinical And Experimental Liver Surgery under Grant UNCE/MED/006. AW was supported by the University of Tübingen, the Federal

Abstract
Objective: The TLR3/cGAS-STING-IFN signaling has recently been reported to be disturbed in colorectal cancer due to deregulated expression of the genes involved. Our study aimed to investigate the influence of potential regulatory variants in these genes on the risk of sporadic colorectal cancer (CRC) in a Czech cohort of 1424 CRC patients and 1114 healthy controls.

Methods: The variants in the TLR3, CGAS, TMEM173, IKBKE, and TBK1 genes were selected using various online bioinformatic tools, such as UCSC browser, HaploReg, Regulome DB, Gtex Portal, SIFT, PolyPhen2, and miRNA prediction tools.

Results: Logistic regression analysis adjusted for age and sex detected a nominal association between CRC risk and three variants, CGAS rs72960018 (OR: 1.68, 95% CI: 1.11-2.53, P-value = .01), CGAS rs9352000 (OR: 2.02, 95% CI: 1.07-3.84, P-value = .03) and TMEM173 rs13153461 (OR: 1.53, 95% CI: 1.03-2.27, P-value = .03). Their cumulative effect revealed a threefold increased CRC risk in...
1 | INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer and the fourth leading cause of cancer mortality worldwide, with an estimate of 1.8 million new cases and close to 1 million deaths in 2018. It originates from the epithelial cells lining the colorectal tract, as a consequence of the gradual accumulation of epigenetic and genetic alterations that lead to the transformation of physiological colonic mucosa to adenocarcinoma. About 85% of CRCs are sporadic and occur in people that have no family history of CRC. So far, genome-wide association studies have reported ~100 risk loci for CRC highlighting new genes and pathways contributing to CRC susceptibility and suggesting roles for Hedgehog signaling, Krüppel-like factors, Hippo-YAP signaling, and immune function. Hua et al have also suggested that polymorphisms within xeroderma pigmentosum group C (XPC) and G (XPG) genes may affect CRC susceptibility through impairment of the nucleotide excision repair (NER) pathway. Moreover, chronic intestinal inflammation has long been recognized as a prominent CRC driver. Patients affected by inflammatory bowel diseases (IBD), such as Crohn’s disease or ulcerative colitis, have been reported to have an increased risk of CRC development. Another factor modulating CRC risk appears to be the intestinal microbiota, the plethora of microorganisms populating the human intestine. The immune system plays an important role in keeping the balance between commensalism, harmful pathogen elimination and self-tolerant maintenance; a disruption of this balance greatly contributes to chronic inflammation. In this scenario, a pivotal role is played by the pattern recognition receptors (PRRs), among which Toll-like receptors (TLRs) are able to recognize different microbe-associated molecular patterns (MAMPs) and/or damage-associated molecular patterns (DAMPs), induce expression of several cytokines, and stimulate activation and differentiation of dendritic cells (DCs). Especially, TLR3, TLR7, and TLR9 are able to stimulate both interferon α (IFNα) and IFNβ. Focusing on TLR3, it is able to recognize viral dsRNA and to activate mitogen-activated protein kinase 1 (MAPK), nuclear factor-κB (NF-κB) and type I IFN signaling pathways through TIR domain-containing adaptor-inducing interferon-β (TRIF), leading to the production of chemokines and cytokines, such as IL-1β, IL-6, and TNFα. Particularly, TLR3 uses the TRIF - TNF Receptor Associated Factor 3 (TRAF3) -TANK-binding kinase 1 (TBK1) + Inhibitor of nuclear factor kappa-B kinase subunit epsilon (IKKε)—Interferon regulatory factor 3 (IRF3) axis to trigger IFNβ and antiviral responses. Moreover, many studies have reported that TLR3 signaling is not only able to induce type I IFN pathways, but indirectly also a strong CD8 T cell response. Indeed, TLR3 induces a cross-presentation of cell-associated antigens, pivotal for cytotoxic T lymphocyte induction, implying an important role in starting adaptive immune responses. The same IKKε-TBK1-IRF3 axis is used by cyclic-GMP-AMP synthase (cGAS), which can be activated by the recognition of cytosolic DNA, derived either from pathogens or self-DNA. Once activated, cGAS activates stimulator of interferon genes protein (STING) (encoded by (transmembrane protein 173 (TMEM173)) via the cyclic-adenosine-guanosine-monophosphate (cGAMP) second messenger to activate the TBK1-IRF3-dependent signaling. IRF3 phosphorylation and nuclear translocation then triggers the type I IFN response (Figure S1). Recently, a deregulation of these pathways in CRC has been reported, mainly caused by an imbalanced expression of the coding genes. Impaired expression of STING has been revealed to favor persistent inflammation and allow the tumor to evade immunosurveillance, thus laying the foundation for tumor initiation and progression. TLR3 expression in carriers of 5-6 risk alleles compared to those with 0-2 risk alleles. Epistatic interactions between these genes and the previously genotyped IFNAR1, IFNAR2, IFNA, IFNB, IFNK, IFNW, IRF3, and IRF7 genes, were computed to test their effect on CRC risk. Overall, we obtained nine pair-wise interactions within and between the CGAS, TMEM173, IKBKE, and TBKI genes. Two of them remained statistically significant after Bonferroni correction. Additional 52 interactions were observed when IFN variants were added to the analysis. Conclusions: Our data suggest that epistatic interactions and a high number of risk alleles may play an important role in CRC carcinogenesis, offering novel biological understanding for the CRC management.

KEYWORDS CRC, interaction, polygenic-risk-score, risk
CRC is quite controversial; indeed, while Nojiri et al reported a similar expression pattern between non-malignant epithelial and colon carcinoma cells,21 Niedzielska et al reported an inversely proportional relation between TLR3 expression level and malignancy stage.22,23 On the other hand, germline variation on TLR3 has been associated with poor prognosis.24

To shed light on the potential role of the single nucleotide polymorphisms (SNPs) within the TLR3, CGAS, TMEM173, TBK1, and IKBKE genes, we genotyped a set of 11 potential regulatory SNPs in a case-control study of 1424 CRC patients and 1114 healthy controls from the Czech Republic and evaluated their association with CRC risk. Moreover, we investigated whether their combined effect and/or pair-wise interactions between all the evaluated SNPs and the previously genotyped SNPs in the IFNA, IFNB, IFNK, IFNW1, IRF3, IRF7, and IFNAR1/2 genes may influence CRC risk.25

2 | MATERIALS AND METHODS

2.1 | Ethics statement

The ethical approval for this study design was obtained from the Institute of Experimental Medicine Academy of Sciences of the Czech Republic (Prague, Czech Republic) and the Institute for Clinical and Experimental Medicine and Faculty Thomayer Hospital (Prague, Czech Republic). Written informed consent was signed by each participant in accordance with the Helsinki declaration.

2.2 | Study population

Details of the studied populations are described elsewhere.26 Briefly, the case group contained 1424 CRC patients recruited between the years 2004 and 2013 by several oncological departments in the Czech Republic (Table 1). Their mean age was 62.7 years, and 61.8% of them were men. The patients showed positive colonoscopic results for malignancy, histologically confirmed as colon or rectal carcinomas. Patients with any previous history of cancer or those who met the Amsterdam criteria I or II for hereditary nonpolyposis colorectal cancer were not included in the study. The control group contained 1114 healthy individuals recruited by the blood-donor centers in Kralovske Vinohrady Hospital and Vojkov hospital in Prague.27 Their mean age was 47.1 years, and 53.4% of them were men. Other characteristics, such as smoking, drinking status and body mass index were not available for the vast majority of the individuals, therefore none of them was taken into consideration in the analysis.

| TABLE 1 | Characteristics of the study population |
|------------------|------------------|------------------|------------------|
| **CRC risk analysis** | **Cases** | **Controls** | **P-value** |
| All patients | 1424 | 1114 | |
| Age at diagnosis | | | |
| Mean (range) | 62.7 (24-90) | 47.1 (18-94) | <.0001* |
| Median | 63 | 47 | |
| Sex | | | |
| Male | 880 (61.8%) | 595 (53.4%) | 2.6e-05b |
| Female | 544 (38.2%) | 519 (46.6%) | |
| Tumor location | | | |
| Colon | 889 (62.4%) | | |
| Rectum | 398 (27.9%) | | |
| Missing information | 137 (9.6%) | | |

Note: Significant results are in bold.
*Z statistics: Wilcoxon Rank-Summ-Test
bChi-square.

2.3 | SNP selection

A total of 11 common SNPs (minor allele frequency, MAF ≥ 0.10 in the CEU population), with a pairwise linkage disequilibrium (LD) $r^2 \leq 0.80$, were selected for genotyping within five genes, namely TLR3, CGAS, TMEM173, IKBKE, and TBK1. Candidate SNPs were non-coding SNPs located within the 5' flanking region, 5' and/or 3' untranslated regions (UTRs) or they were expression quantitative trait loci (eQTL) SNPs for the selected genes or non-synonymous SNPs, validated by 1000 Genomes in the CEU population (Table S1).

Additionally, a total of 24 potentially functional SNPs within promoter, or 5'UTR or 3'UTR of the genes involved in the IFN signaling pathway, including IFNA (1, 2, 4, 5, 7, 8, 13, 16, 17, and 21), IFNB1, IFNK, IFNW1, IRF3, IRF7, and IFNAR1/2, were selected from our previous study25 (Table S2). Notably, IFNAS, IFNB1, IFNK, IFNW1 genes are all located at the same chromosome location (9p21.3), and capture many other SNPs in linkage, supplying further information on other genes at the given locus.

2.4 | In-silico analysis

Online bioinformatic tools were used to explore and select the SNPs of interest, including UCSC browser (https://genome-euro.ucsc.edu/), HaploReg (http://www.broadinstitute.org/mammals/haplogreg/haplogreg.php), Regulome DB (http://www.regulomedb.org/), Gtex Portal (https://gtexportal.org/home/), MicroSNiPer (http://epicenter.ifeFreiburg.mp.de/services/microsniper/), SIFT (Sorting Intolerant from Tolerant) (http://sift.jcvi.org/) and PolyPhen2 (Polymorphism Phenotyping v2) (http://genetics.bwh.harvard.edu/pph2/).
LD and haplotype blocks within the genes were examined based on pairwise LD $r^2$ (Table S1).

2.5 | Genotyping

SNP genotyping was performed on genomic DNA from peripheral blood leukocytes using KASP (LGC genomics, Hoddesdon, Hertfordshire, UK) and TaqMan (Thermo Fisher Scientific) as allelic discrimination methods. DNA amplification was carried out in accordance with the LGC genomics’ and TaqMan’s PCR cycling conditions. Genotype detection was performed using the ViiA 7 Real-Time PCR System (Thermo Fisher Scientific), and setting the range 94.0%-100% as a threshold for the genotype call rate. The genotype correlation between the 142 duplicated samples, used as quality controls, was higher than 95%. Samples with <50% call rate over all assays were excluded from the study, leaving 1396 cases and 1111 controls for the association analysis.

2.6 | Statistical analysis

The chi-square test was performed to test the deviation of genotype frequencies in the controls from Hardy-Weinberg equilibrium (HWE). Logistic regression analysis adjusted for age and sex was used to calculate odds ratios (ORs) and 95% confidence intervals (CIs) for associations between genotypes and CRC risk (SAS Version 9.3; SAS Institute).

In the combined analysis of the three SNPs that showed a nominal association with CRC, the allelic model was calculated for each SNP whereby the genotypes were converted into 0, 1, and 2 risk alleles. On the basis of the number of risk alleles, a genotype score ranging from 0 to 6 was constructed. Samples with one or more missing genotypes were not included.

To evaluate the significant findings, the false-positive report probability (FPRP) was calculated. A prior probability of 0.1 and an FPRP threshold of 0.2 were assigned to detect the OR of 0.67/1.50 (protective/risk effects) for the association with genotypes and alleles numbers under investigation. Only the associations with an FPRP value less than 0.2 were considered noteworthy findings.

Binary interactions for all different SNP combinations were evaluated to investigate whether the non-additive effect can improve the prediction of the disease risk. The newly genotyped SNPs were complemented and analyzed with the SNPs in the IFNA, IFNB, IFNK, IFNW1, IRF3, IRF7, and IFNARI/2 genes previously genotyped in a subset of 1327 CRC patients and 758 controls from the same Czech cohort. Details of the pair-wise interaction analyses are described elsewhere. Briefly, four different modes of inheritance were calculated and tested for each pair: the three genotypes model, the log-additive model, the dominant model, and the recessive model. To assess whether the SNP-SNP interaction term led to a considerably better fit of the data, likelihood ratio tests were performed. The best model for SNPs that showed significant interactions with each other by more than one model was selected on the basis of their Akaike information criterion (AIC). The smaller the value of AIC, the better the model data fit. To evaluate the benefit of all genetic components (including SNPs and the interaction term) to the model, likelihood ratio test-based P-values were calculated. The corresponding ORs and the Wald estimate for their 95% CIs and P-values were computed for the best model of each SNP pair. As the reference genotype combination we used the major allele genotype combination based on the best model of each interaction. Altogether, 55 (11 SNPs*(11-1)/2) independent tests were performed between the TLR3, CGAS, TMEM173, IKBKE, and TBK1 genes, giving a Bonferroni corrected p-value of 0.05/55 = 0.0009. Inclusion of the IFN pathway genes to the study increased the number of independent tests to 275, giving a Bonferroni corrected P-value of 0.05/275 = 0.0002.

3 | RESULTS

3.1 | Single SNP analysis

The minor allele frequencies of the genotyped SNPs were similar to the ones reported in the European population in the 1000 Genomes Project (http://www.internationalgenome.org/) and in the non-Finnish European population in the Genome Aggregation Database (https://gnomad.broadinstitute.org/).

The genotype distribution of all SNPs was consistent with HWE in the control group ($P > .05$). Three SNPs, two located within CGAS, rs72960018 (OR: 1.68, 95% CI: 1.11-2.53, P-value = .01, under dominant model) and rs9352000 (OR: 2.02, 95% CI: 1.07-3.84, P-value = .03, under recessive model), and one within TMEM173, rs13153461 (OR: 1.53, 95% CI: 1.03-2.27, P-value = .03, under recessive model), exhibited moderate associations with CRC risk (Table 2). However, when considering an FPRP threshold of 0.2, none of them was considered a noteworthy finding (Table S3).

3.2 | Combined analysis

Since CGAS and TMEM173 encode proteins that are interacting with each other through a second messenger, cGAMP, we further estimated the cumulative effect of the three SNPs reporting a nominal association with CRC susceptibility.
Patients in the highest risk score group (5-6 risk alleles) had about threefold augmented risk of developing CRC compared to those in the lowest risk score group (0-2 risk alleles) (adjusted OR = 2.98, 95% CI: 1.35-6.56, \( P \) for trend: \( 6 \times 10^{-4} \)) (Table 3). An FPRP value less than 0.2 was observed for the score group containing individuals carrying 3-4 risk alleles,

### Table 2: Association of single SNPs with CRC risk

| Gene ID | SNP ID      | Genotype | Cases | Controls | OR    | 95% CI          | \( P \)  |
|---------|-------------|----------|-------|----------|-------|----------------|--------|
| CGAS    | rs72960018  | A/A      | 69    | 70       | 1.00  |                |        |
|         |             | A/G      | 505   | 408      | 1.6   | (1.04-2.46)    | .03    |
|         |             | G/G      | 789   | 614      | 1.73  | (1.13-2.63)    | .01    |
|         |             | A/A      | 69    | 70       | 1.00  |                |        |
|         |             | A/G + G/G| 1294 | 1022     | 1.68  | (1.11-2.53)    | .01    |
|         | rs9352000   | T/T      | 935   | 761      | 1.00  |                |        |
|         |             | G/T      | 323   | 237      | 1.03  | (0.82-1.30)    | .77    |
|         |             | G/G      | 44    | 19       | 2.04  | (1.07-3.88)    | .03    |
|         |             | T/T + G/T| 1258 | 998      | 1.00  |                |        |
|         |             | G/G      | 44    | 19       | 2.02  | (1.07-3.84)    | .03    |
|         | rs34413328  | A/A      | 854   | 666      | 1.00  |                |        |
|         |             | A/-      | 452   | 377      | 0.97  | (0.79-1.19)    | .75    |
|         |             | -/-      | 56    | 49       | 1.1   | (0.68-1.80)    | .69    |
|         | rs610913    | T/T      | 499   | 387      | 1.00  |                |        |
|         |             | G/T      | 617   | 521      | 0.91  | (0.73-1.12)    | .37    |
|         |             | G/G      | 228   | 171      | 1.13  | (0.85-1.51)    | .41    |
| TMEM173 | rs7380272   | C/C      | 1053  | 887      | 1.00  |                |        |
|         |             | C/T      | 283   | 199      | 1.21  | (0.94-1.55)    | .13    |
|         |             | T/T      | 28    | 15       | 1.57  | (0.75-3.26)    | .23    |
|         | rs13153461  | A/A      | 751   | 646      | 1.00  |                |        |
|         |             | A/G      | 509   | 396      | 1.12  | (0.91-1.38)    | .28    |
|         |             | G/G      | 106   | 60       | 1.6   | (1.07-2.39)    | .02    |
|         |             | A/A + A/G| 1260 | 1042     | 1.00  |                |        |
|         |             | GG       | 106   | 60       | 1.53  | (1.03-2.27)    | .03    |
| IKBKE   | rs2297549   | T/T      | 816   | 663      | 1.00  |                |        |
|         |             | C/T      | 481   | 387      | 0.97  | (0.79-1.19)    | .75    |
|         |             | C/C      | 71    | 42       | 1.35  | (0.84-2.19)    | .22    |
|         | rs2297548   | T/T      | 890   | 752      | 1.00  |                |        |
|         |             | C/T      | 419   | 303      | 1.2   | (0.97-1.49)    | .10    |
|         |             | C/C      | 56    | 39       | 1.43  | (0.86-2.38)    | .17    |
|         | rs15672     | G/G      | 378   | 302      | 1.00  |                |        |
|         |             | G/A      | 676   | 510      | 1.15  | (0.91-1.45)    | .23    |
|         |             | A/A      | 276   | 265      | 0.87  | (0.66-1.15)    | .33    |
| TBK1    | rs61933195  | C/C      | 1007  | 819      | 1.00  |                |        |
|         |             | A/C      | 319   | 262      | 0.9   | (0.72-1.13)    | .37    |
|         |             | A/A      | 35    | 19       | 1.22  | (0.61-2.42)    | .57    |
| TLR3    | rs3775291   | C/C      | 663   | 512      | 1.00  |                |        |
|         |             | C/T      | 573   | 479      | 0.92  | (0.75-1.13)    | .42    |
|         |             | T/T      | 115   | 113      | 0.73  | (0.52-1.04)    | .08    |

Note: Significant results are highlighted in bold.

Abbreviations: 95% CI: 95% Confidence Interval; OR, Odds Ratio; \( P \), \( P \)-value.
but not for the highest risk score group, which showed also a low statistical power (Table S3). This suggests some possible bias in the findings due to reduced sample size, which need to be further validated in larger studies. Interestingly, no synergistic interaction was observed between these SNPs (data not shown).

### 3.3 SNP-SNP interactions in CRC risk

We further evaluated whether a synergistic effect of the 11 SNPs within TLR3, CGAS, TMEM173, TBK1, and IKBKE genes may impact CRC risk. After setting our significance level of \( P \)-value < .05, nine interactions, counting interactions between SNPs both within a gene and between the five genes, were observed (Table S4). Two out of nine interactions, \textit{IKBKE} rs2297549 - \textit{TMEM173} rs13153461 and \textit{IKBKE} rs2297549 - \textit{TMEM173} rs7380272, passed the Bonferroni correction (\( P \)-value < .0009). The association with the risk of CRC was estimated for the best model of each SNP-SNP interaction (Table S4).

The two \textit{IKBKE} SNPs, rs2297549 and rs15672 (\( r^2 = .01 \)), showed an interesting and complex interaction with the same two \textit{TMEM173} SNPs, rs13153461 and rs7380272 (\( r^2 = .38 \)). An increased CRC risk was observed particularly between \textit{IKBKE} rs2297549 and the two \textit{TMEM173} SNPs when the minor allele homozygote genotypes of one gene interacted with the major allele containing genotypes of the other gene. On the other hand, an increased and decreased risk of CRC was observed when \textit{IKBKE} rs15672 (minor allele homozygote genotype) interacted with \textit{TMEM173} rs13153461 (minor allele homozygote genotype) and \textit{TMEM173} rs7380272 (major allele homozygote genotype), respectively (Table S4).

### 3.4 SNP-SNP interactions in CRC risk including the IFN variants

The two \textit{TMEM173} SNPs were not only shown to be the main interaction partners within our candidate genes but...
also exhibited an interaction with many of the previously genotyped IFN variants (Table 5). Especially, TMEM173 rs13153461 showed four more interactions with IRF3 rs2304204, IRF7 rs1061502, IFNB1 rs1424855, and IFNK rs700782, which are not in LD with each other. Compared to the reference genotype pair, an increased risk was observed for specific genotype pairs when TMEM173 rs13153461 interacted with the IRF3, IFNB1, and IFNK SNPs (Table S5). No significant ORs were detected for the IRF7 interaction. On the other hand, TMEM173 rs7380272 showed interactions with another set of three IFN SNPs, IFNA7/14 rs6475526, IFNA16 rs10964912, and IFNA21 rs12376071, which were in moderate LD with each other \( (r^2 = .40-.50) \). A strong interaction was reported between TMEM173 rs7380272 (major homozygote genotype) and IFN\(\gamma\)/IFNK rs6475526 (minor allele genotypes). The other interactions were more complex and depended on the genotype combinations (Table S5, Figure S4).

The highest number of interactions was represented by the four CGAS SNPs, rs72960018 \( (n = 8) \), rs9352000 \( (n = 9) \), rs34413328 \( (n = 3) \), and rs610913 \( n = 10 \) when analyzed in interplay with the previously genotyped IFN variants (Table 5, Figure S5). The unlinked SNPs, rs72960018, rs9352000, rs34413328 \( (r^2 < .08) \) shared several interactions with rs610913, which was in a moderate LD with the other SNPs \( (r^2 = .20-.38) \). Especially, we observed a decreased risk of CRC development when CGAS rs610913 and rs72960018 interacted with IFNA4 rs2383183 and IFNA13 rs641734 \( (r^2 = .43) \). Many genotype combinations of the CGAS SNPs rs610913 and rs34413328 with IFNA7/14 rs6475526 and IFNK rs700782 were associated to an increased risk of CRC. Similarly, many genotype combinations in the shared interactions of CGAS rs610913 and rs9352000 with IFNA2 rs10120977, IFNA16 rs10964912 \( (r^2 = .43) \), and IFN\(R\)2 rs1131668 seemed to increase CRC risk (Table S5). Furthermore, the IKBKE SNPs, rs15672, rs2297549, rs2297548, the T"BK1 SNP rs61933195, and the TLR3 SNP rs3775291 showed a few interactions with the IFN genes (Table 5, Figure S3). There was no overlap between the IKBKE-IFNs and T"BK1-IFNs interactions. Interestingly, the IKBKE interactions led to increased risk of CRC, while the TLR3 interactions decreased the risk. IKBKE rs2297549 shared two interactions with IKBKE rs2297548, comprising IFNK rs700782 and IFN\(AR\)1 rs2834202, while only one with IKBKE rs15672, with IFN\(AR\)1 rs2856968 (Table S5).

It is interesting to note that the interactions and \( r^2 \) values do not seem to correlate; indeed, most of the previously genotyped IFN SNPs, located at the same locus on the chromosome 5 and involved in interactions with the same SNP, do not show a high LD (Figure S6).

In summary, all these regulatory SNPs could affect the expression of the corresponding genes leading to protective/harmful effects when interacting with each other.

4 | DISCUSSION

Balance is the key to everything, especially when it concerns the immune system, which can highly contribute to both suppression and promotion of cancer. Recent studies have shown that the cGAS-STING and TLR3 pathways, which through the TBK1-IKKe phosphorylation induce the type I IFNs production, are disturbed in CRC, mainly because of an imbalanced expression of their coding genes.\(^{20,29}\) cGAS produces cGAMP in response to cytosolic DNA, which in turn can bind and activate STING.\(^{30}\) It has been described that the levels of 2', 3'-cGAMP, or its analogs are important for the immune system to decide which direction to follow. Indeed, high levels of STING activators have been shown to lead the immune system toward sustained inflammation and consequent tumor initiation and progression.\(^{20}\) Furthermore, cGAS plays an important role in controlling cellular senescence a delicate cellular state vital for the elimination of pre-cancerous state but also a reservoir of potentially harmful tumorigenic progenitors.\(^{31}\) Impaired STING expression may also allow the cancer cells to escape the immunosurveillance system. Here, we showed that inherited genetic variation potentially affecting gene expression of the cGAS-STING-IFN pathway may contribute to CRC susceptibility. Individually, the studied SNPs showed only nominal if any associations with CRC risk, however, they seem to interact and by that affect the risk.

So far, about 100 CRC susceptibility loci have been identified through genome-wide association studies.\(^{5}\) Polygenic risk scores derived from these studies have evaluated that some 5% of the study populations have over twofold increased risk of CRC.\(^{32,33}\) In our study also, we observed an increased risk for individuals with increasing number of alleles causing a moderately increased CRC risk. However, polygenic risk scores do not take into account epistatic interactions, which may by far cause a more pronounced risk compared to single variants, as shown in our study.

In this research, the two-way interaction, as well as the cumulative risk analyses, uncovered associations, which were substantial compared to individual SNP associations. Our results suggested that studying the interplay and/or the cumulative effects instead of the single effect of SNPs within genes involved in the immunity could be of interest to help our understanding of the mechanisms underlying the CRC development.

In our analyses, nine interactions between CGAS, TMEM173, T"BK1, and IKBKE and further 52 interactions together with IFN\(As\), IFNB, IFNW1, IFNK, IRF3, IRF7, and IFN\(AR\)1/2 in the smaller sample set were observed. For all interactions, the global null hypothesis test was highly significant \( (P\text{-value} < .0001) \). Two out of the nine interactions, TMEM173 rs13153461-IBKKE rs2297549, and TMEM173 rs7380272-IBKKE rs2297549, passed the Bonferroni multiple testing corrections \( (P\text{-value} < .0009) \).
| Variable 1 | Mode of Inheritance | Variable 2 | Mode of Inheritance | Interaction | SNP total |
|------------|---------------------|------------|---------------------|-------------|-----------|
| Gene       | SNP                 | Gene       | SNP                 | LRT Statistic | DG | P-value |
| TMEM173    | rs13153461          | IRE7       | rs1061502           | 7.08        | 2 | .029    |
|            | Recessive           |            | dominant            |             |   |         |
| TMEM173    | rs13153461          | IRE7       | rs1424855           | 7.19        | 2 | .028    |
|            | Three Genotypes     |            | three genotypes     |             |   |         |
| TMEM173    | rs13153461          | IFNB1      | rs700782            | 3.99        | 1 | .046    |
|            | Recessive           |            | dominant            |             |   |         |
| IFN7/      | rs6475526           | TMEM173    | rs7380272           | 4.13        | 1 | .042    |
| IFN14      | Allele Number       |            | recessive           |             |   |         |
| IFNA6      | rs10964912          | TMEM173    | rs7380272           | 5.46        | 1 | .019    |
|            | dominant            |            | recessive           |             |   |         |
| IFNA21     | rs12376071          | TMEM173    | rs7380272           | 7.45        | 1 | .006    |
|            | Allele Number       |            | recessive           |             |   |         |
| IFNA4      | rs2383183           | CGAS       | rs72960018          | 6.32        | 1 | .012    |
|            | Dominant            |            | recessive           |             |   |         |
| IFNA4      | rs2383183           | CGAS       | rs610913            | 4.83        | 1 | .028    |
|            | Dominant            |            | recessive           |             |   |         |
| IFNA13     | rs641734            | CGAS       | rs72960018          | 7.40        | 2 | .025    |
|            | dominant            |            | recessive           |             |   |         |
| IFN13      | rs641734            | CGAS       | rs610913            | 3.90        | 1 | .048    |
|            | dominant            |            | recessive           |             |   |         |
| CGAS       | rs34413328          | IFNK       | rs700782            | 6.59        | 1 | .010    |
|            | dominant            |            | recessive           |             |   |         |
| CGAS       | rs610913            | IFNK       | rs700782            | 4.92        | 1 | .027    |
|            | Allele Number       |            | recessive           |             |   |         |
| CGAS       | rs34413328          | IFN7/IFN14 | rs6475526           | 4.53        | 1 | .033    |
|            | Allele Number       |            | recessive           |             |   |         |
| IFNA5      | rs12156640          | TBK1       | rs61933195          | 11.86       | 1 | .0006   |

(Continues)
| Gene   | SNP           | Mode of Inheritance | Gene          | SNP           | Mode of Inheritance | Interaction | SNP total |
|--------|---------------|---------------------|---------------|---------------|---------------------|-------------|-----------|
|        |               |                     |               |               |                     | LRT Statistic | DG | P-value | LRT Statistic | DG | P-value |
| IRF3   | rs2304204     | Dominant            | CGAS          | rs9352000     | Recessive           | 6.22        | 1  | .013   | 11.32        | 3  | .010    |
| IFNAR2 | rs1131668     | Allele Number       | TLR3          | rs3775291     | Recessive           | 11.10       | 1  | .0009  | 13.44        | 3  | .004    |
| IFNAR1 | rs2257167     | Dominant            | IKBKE         | rs15672       | Allele Number      | 9.88        | 1  | .002   | 11.84        | 3  | .008    |
| CGAS   | rs72960018    | Dominant            | IFNA17        | rs7873404     | Three Genotypes    | 8.17        | 2  | .017   | 13.54        | 5  | .019    |
| IFNA21 | rs2939        | Dominant            | CGAS          | rs610913      | Allele Number      | 5.38        | 1  | .020   | 9.49         | 3  | .023    |
| IFNAR1 | rs2850015     | Allele Number       | TLR3          | rs3775291     | Recessive           | 5.31        | 1  | .021   | 8.31         | 3  | .040    |
| IFNB1  | rs1424855     | Allele Number       | TBL1          | rs61933195    | Dominant            | 8.87        | 1  | .003   | 8.90         | 3  | .031    |
| IFNA21 | rs2939        | Dominant            | IKBKE         | rs2297548     | Allele Number      | 5.03        | 1  | .025   | 9.83         | 3  | .020    |
| IFNAR1 | rs2856968     | Allele Number       | CGAS          | rs72960018    | Dominant            | 7.68        | 1  | .006   | 16.93        | 3  | .001    |
| IFNA7/ IFN14 | rs6475526 | Dominant            | TBL1          | rs61933195    | Allele Number      | 4.84        | 1  | .028   | 11.07        | 3  | .011    |
| IFNAR1 | rs2834202     | Three Genotypes     | CGAS          | rs610913      | Recessive           | 10.31       | 2  | .006   | 15.25        | 5  | .009    |
| IFNA8  | rs10811536    | Three Genotypes     | IKBKE         | rs2297549     | Recessive           | 7.12        | 2  | .028   | 12.15        | 5  | .033    |
| IFNW1  | rs10757189    | Dominant            | TBL1          | rs61933195    | Dominant            | 7.55        | 1  | .006   | 8.74         | 3  | .033    |
| IFNA21 | rs2939        | Dominant            | CGAS          | rs72960018    | Dominant            | 7.54        | 1  | .006   | 12.79        | 3  | .005    |
| IFNAR1 | rs2856968     | Recessive           | CGAS          | rs9352000     | Recessive           | 4.44        | 1  | .035   | 8.63         | 3  | .035    |
| IFNAR1 | rs2856968     | Three Genotypes     | CGAS          | rs610913      | Recessive           | 10.12       | 2  | .006   | 20.42        | 5  | .001    |
| IFNA4  | rs2383183     | Dominant            | IKBKE         | rs2297548     | Allele Number      | 4.33        | 1  | .037   | 8.04         | 3  | .045    |
| IFNAR1 | rs2850015     | Dominant            | CGAS          | rs72960018    | Three Genotypes     | 6.57        | 2  | .038   | 11.34        | 5  | .045    |
| IRF7   | rs1061502     | Recessive           | TLR3          | rs3775291     | Three Genotypes     | 9.77        | 2  | .008   | 13.95        | 5  | .016    |
| IFNAR2 | rs12376071    | Dominant            | CGAS          | rs9352000     | Recessive           | 7.03        | 1  | .008   | 13.06        | 3  | .005    |
| IFNW1  | rs10757189    | Recessive           | CGAS          | rs9352000     | Recessive           | 6.84        | 1  | .009   | 11.69        | 3  | .009    |
| IFNAR2 | rs1131668     | Recessive           | IKBKE         | rs2297548     | Recessive           | 6.78        | 1  | .009   | 9.25         | 3  | .026    |
| TLR3   | rs3775291     | Dominant            | IFNA13        | rs641734      | Dominant            | 3.90        | 1  | .048   | 8.62         | 3  | .035    |
| IFNAR1 | rs2257167     | Allele Number       | CGAS          | rs72960018    | Dominant            | 6.60        | 1  | .010   | 11.75        | 3  | .008    |

Abbreviations: DG, Degrees of Freedom; LRT, Likelihood Ratio Test.
The three SNPs involved in the most significant interactions were \textit{TMEM173} rs13153461, which also associated with CRC risk as a single SNP, the \textit{TMEM173} eQTL SNP rs7380272, and the \textit{IKBKE} 5’UTR SNP rs2297549. Interestingly, \textit{TMEM173} rs13153461 and \textit{TMEM173} rs7380272 show a moderate LD ($r^2 = .38$), indicating that some of the interactions may be due to a modest LD between the SNPs.

As all the selected SNPs were potentially functional variants they are all located within enhancer/promoter histone marks, DNase hypersensitivity sites in different tissues, including gastrointestinal tract (GI) and whole blood, and are also predicted to affect several transcription factor-binding sites (TFBSs) (Table S1). Some of them have also an eQTL nature, such as \textit{TMEM173} rs7380272, whose T allele correlates with a decreased expression of \textit{TMEM173} in blood (P-values: $3.18 \times 10^{-31}$; Z-score: $-11.62$) (https://molgenis58.target.rug.nl/bloodeqtlbrowser/). Additionally, the selected SNPs capture many other SNPs, which can give us further information on additional SNPs or genes located at the same locus, for example \textit{TMEM173} rs7380272 is in LD with rs7380824, which is not only a missense variant mapping in a highly conserved region, predicted to be deleterious and probably damaging by SIFT and PolyPhen, respectively; it also acts as a \textit{TMEM173} eQTL in blood tissue (P-values: $2.73 \times 10^{-31}$; Z-score: $-11.64$) (https://molgenis58.target.rug.nl/bloodeqtlbrowser/). Hence, it could affect not only the expression of the gene, but also the function of the encoded protein.

When we included the previously genotyped \textit{IFN} variants to our analyses, further synergistic effects became evident. The main interactions were exhibited by the four \textit{CGAS} SNPs, rs72960018, rs9352000, rs34413328, and rs610913, among which a few were toward the same \textit{IFN} SNPs. Particularly, rs72960018, rs9352000, and rs610913 shared an interaction with the same \textit{IFNAR1} SNP, rs2856968, which additionally interplayed with the \textit{IKBKE} SNPs, rs15672, and rs2297549. A persistent increased risk was particularly exerted when their minor alleles interacted with each other. A possible explanation could be the potential involvement of \textit{IFNAR1} rs2856968 in altering protein binding regions, as predicted by Regulome DB, such as those of FOXM1, MXII, MAZ, MAX, and CHD1. Furthermore, it is in LD with many SNPs lying within regulatory regions, which map within TFBSs such as those of the polymerase epsilon catalytic subunit (POLE) or of AP-2. These transcription factors (TFs) have been shown to be associated with the risk of CRC development and its progression, respectively.\textsuperscript{35,36}

On the other hand, the four \textit{CGAS} SNPs were located within the binding sites of several TFs, among which NF-kB. Aberrant regulation of NF-kB and consequently of the downstream signaling pathways are involved in CRC initiation and progression, senescence regulation\textsuperscript{37,38} as well as in resistance to chemotherapy and in the immune response.\textsuperscript{39,40} Additionally, they were predicted to affect binding of several other TFs, such as Egr-1 (early growth response-1), YY1 (Yin Yang 1), BATF (Basic Leucine Zipper ATF-Like Transcription Factor), that have already been reported to be associated with apoptosis and tumor cell proliferation\textsuperscript{41} or with tumorigenesis in CRC\textsuperscript{42} or to be over-expressed in ulcerative colitis and CRC.\textsuperscript{43}

In this study, we included only five members of the TLR3/cGAS-STING-IKKe-TBK1 signaling cascade, which has recently been reported to be disturbed in CRC due to deregulated expression of the genes involved,\textsuperscript{19} in addition to nine \textit{IFN} genes from our previous studies to evaluate their genetic interactions. Inclusion of a large network of genes would have led to a higher number of multiple tests, increasing the likelihood of chance findings. This kind of genetic interaction study needs full genotyping data of all SNPs of interest, which lead to another limitation of our study, which is the lack of replication in another population. However, because these genes play a key role in the signaling cascade and there are emerging data about their importance in CRC, our study serves as a starting point for further studies including not only the genes and SNPs studied by us but also other genes important in the mucosal immune system.

Our data suggest that epistatic interactions and a high number of risk alleles may play an important role in explaining the CRC onset, offering novel biological understanding for the management of CRC patients. Our data warrant the exploration of these genetic variants for patient risk stratification and therapeutic decision making, including immune checkpoint inhibitors. Additionally, functional SNPs within these genes may represent potential biomarkers to be used to identify high-CRC-risk individuals and therefore direct them to colonoscopy. Indeed, their relative frequency within the European population (> 10%) makes them suitable for a widespread use. However, replication of these results in independent cohorts is needed, together with functional experimental studies in order to confirm the in silico-predicted effects of the identified variants and their combinations on CRC susceptibility.

ACKNOWLEDGEMENTS
This article is based upon work from COST action CA17118, supported by COST (European Cooperation in Science and Technology, www.cost.eu).

CONFLICT OF INTEREST
The authors and planners have disclosed no potential conflicts of interest, financial or otherwise.

ORCID
Calogerina Catalano  https://orcid.org/0000-0002-4047-8485
Shun Lu  https://orcid.org/0000-0002-1833-475X
REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, 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(6):394-424.

2. Hong SN. Genetic and epigenetic alterations of colorectal cancer. In: Int Rev Cell Mol Biol. 2018;16(3):327-337.

3. Frank C, Sundquist J, Yu H, Hemminki A, Hemminki K. Concordant and discordant familial cancer: familial risks, proportions and population impact. Int J Cancer. 2017;140(7):1510-1516.

4. Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. Nat Rev Cancer. 2017;17(11):692-704.

5. Huyghhe JR, Bien SA, Harrison TA, et al. Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet. 2019;51(1):76-87.

6. Hua RX, Zhuo ZJ, Zhu J, et al. XPG gene polymorphisms contribute to colorectal cancer susceptibility: a two-stage case-control study. J Cancer. 2016;7(12):1731-1739.

7. Hua R-X, Zhu J, Jiang D-H, et al. Association of XPC gene polymorphisms with colorectal cancer risk in a southern Chinese population: a case-control study and meta-analysis. Genes. 2016;7(10):73.

8. Sun J, Kato I. Gut microbiota, inflammation and colorectal cancer. Int J Mol Sci. 2018;19(5):1203-1210.

9. Lukas M. Inflammatory bowel disease as a risk factor for colorectal cancer. Cancer J Clin. 2018;68(6):394-424.

10. Sun J, Kato I. Gut microbiota, inflammation and colorectal cancer. Int J Mol Sci. 2018;19(5):1203-1210.

11. Nojiri K, Sugimoto K, Shiraki K, et al. Expression and function of Toll-like receptors 3 and 9 in human colon carcinoma. Oncol Rep. 2013;29(5):1737-1743.

12. Niedzielska I, Orawczyk T, Ziaja K, et al. Toll-like receptors and the tendency of normal mucous membrane to transform to polyp or colorectal cancer. J Physiol Pharmacol. 2009;1:65-71.

13. Slattery ML, Herrick JS, Bondurant KL, Wolff RK. Toll-like receptor genes and their association with colon and rectal cancer development and prognosis. Int J Cancer. 2012;130(12):2974-2980.

14. Castro FA, Försti A, Buch S, et al. TLR9 polymorphism is an independent prognostic marker for stage II colorectal cancer. Eur J Cancer. 2011;47(8):1203-1210.

15. Lu S, Pardini B, Cheng B, et al. Single nucleotide polymorphisms within interferon signaling pathway genes are associated with colorectal cancer susceptibility and survival. PLoS ONE. 2014;9(10):e11061.

16. Catalano C, da Silva Filho MI, Frank C, et al. Investigation of single and synergic effects of NLRC5 and PD-L1 variants on the risk of colorectal cancer. PLoS ONE. 2018;13(2):e0192385.

17. Carrai M, Campa D, Vodicka P, et al. Association between taste receptor (TAS) genes and the perception of wine characteristics. Sci Rep. 2017;7(1):9239.

18. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst. 2004;96(6):434-442.

19. Bednarczyk M, Muc-Wierzgon M, Walkiewicz K, Kokot T, Fatyga E, Mazurek U. Profile of gene expression of TLR-signaling pathways in colorectal cancer tissues. Int J Immunopathol Pharmacol. 2017;30(3):322-326.

20. Diner E, Burdette D, Wilson S, et al. The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. Cell Rep. 2013;3(5):1355-1361.

21. Yang H, Wang H, Ren J, Chen Q, Chen ZJ. cGAS is essential for cellular senescence. Proc Natl Acad Sci U S A. 2017;114(23):E4611-E4620.

22. Schmit SL, Edlund CK, Schumacher FR, et al. Novel common genetic susceptibility loci for colorectal cancer. Nat Genet. 2018;50(2):183-189.

23. Frampton MJ, Law P, Litchfield K, et al. Implications of polygenic risk for personalised colorectal cancer screening. Ann Oncol. 2016;27(3):429-434.

24. Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet. 2013;45(10):1238-1243.

25. Esteban-Jurado C, Garre P, Vila M, et al. New genes emerging for colorectal cancer predisposition. World J Gastroenterol. 2014;20(8):1961-1971.

26. Valle L, Hernández-Illán E, Bellido F, et al. New insights into POLE and POLE1 germline mutations in familial colorectal cancer and polyposis. Hum Mol Genet. 2014;23(13):3506-3512.

27. Jing H, Lee S. NF-kappaB in cellular senescence and cancer treatment. Mol Cells. 2014;37(3):189-195.

28. Basseres DS, Baldwin AS. Nuclear factor-kappaB and inhibitor of kappaB kinase pathways in oncogenic initiation and progression. Oncogene. 2006;25(51):6817-6830.

29. Hassanzadeh P. Colorectal cancer and NF-kappaB signaling pathway. Gastroenterol Hepatol Bed Bench. 2011;4(3):127-132.
40. Dev A, Iyer S, Razani B, Cheng G. NF-kappaB and innate immunity. *Curr Top Microbiol Immunol*. 2011;349:115-143.

41. Myung D-S, Park Y-L, Kim N, et al. Expression of early growth response-1 in colorectal cancer and its relation to tumor cell proliferation and apoptosis. *Oncol Rep*. 2014;31(2):788-794.

42. Chinnappan D, Xiao D, Ratnasari A, Andry C, King TC, Weber HC. Transcription factor YY1 expression in human gastrointestinal cancer cells. *Int J Oncol*. 2009;34(5):1417-1423.

43. Punkenburg E, Vogler T, Büttner M, et al. Batf-dependent Th17 cells critically regulate IL-23 driven colitis-associated colon cancer. *Gut*. 2016;65(7):1139-1150.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Catalano C, da Silva Filho MI, Frank C, et al. Epistatic effect of TLR3 and cGAS-STING-IKKε-TBK1-IFN signaling variants on colorectal cancer risk. *Cancer Med*. 2020;9:1473–1484. [https://doi.org/10.1002/cam4.2804](https://doi.org/10.1002/cam4.2804)