



**EMORY**  
LIBRARIES &  
INFORMATION  
TECHNOLOGY

**OpenEmory**

## **Association of Selenoprotein and Selenium Pathway Genotypes with Risk of Colorectal Cancer and Interaction with Selenium Status**

[Veronika Fedirko](#), *Emory University*

[Mazda Jenab](#), *International Agency for Research on Cancer*

[Catherine Meplan](#), *Newcastle University*

[Jeb Jones](#), *Emory University*

[Wanzhe Zhu](#), *Emory University*

[Lutz Schomburg](#), *University Medical School Berlin*

[Afshan Siddiq](#), *Imperial College London*

[Sandra Hybsier](#), *University Medical School Berlin*

[Kim Overvad](#), *Aarhus University*

[Anne Tjonneland](#), *Danish Cancer Society*

*Only first 10 authors above; see publication for full author list.*

---

**Proceedings Title:** Nutrients

**Publisher:** MDPI

**Conference Place:** Switzerland

**Volume/Issue:** Volume 11 | Issue 4

**Publication Date:** 2019-04-01

**Type of Work:** Conference | Final Publisher PDF

**Publisher DOI:** 10.3390/nu11040935

**Permanent URL:** <https://pid.emory.edu/ark:/25593/tqv6p>

---

Final published version: <http://dx.doi.org/10.3390/nu11040935>

### **Copyright information:**

© 2019 by the authors.

This is an Open Access work distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>).

*Accessed October 29, 2020 11:36 AM EDT*



Article

# Association of Selenoprotein and Selenium Pathway Genotypes with Risk of Colorectal Cancer and Interaction with Selenium Status

Veronika Fedirko <sup>1</sup>, Mazda Jenab <sup>2</sup>, Catherine Méplan <sup>3</sup>, Jeb S. Jones <sup>1</sup>, Wanzhe Zhu <sup>1</sup>, Lutz Schomburg <sup>4</sup>, Afshan Siddiq <sup>5</sup>, Sandra Hybsier <sup>4</sup>, Kim Overvad <sup>6</sup>, Anne Tjønneland <sup>7</sup>, Hanane Omichessan <sup>8,9</sup>, Vittorio Perduca <sup>8,9,10</sup>, Marie-Christine Boutron-Ruault <sup>8,9</sup>, Tilman Kühn <sup>11</sup>, Verena Katzke <sup>11</sup>, Krasimira Aleksandrova <sup>12</sup>, Antonia Trichopoulou <sup>13</sup>, Anna Karakatsani <sup>13,14</sup>, Anastasia Kotanidou <sup>13,15</sup>, Rosario Tumino <sup>16</sup>, Salvatore Panico <sup>17</sup>, Giovanna Masala <sup>18</sup>, Claudia Agnoli <sup>19</sup>, Alessio Naccarati <sup>20</sup>, Bas Bueno-de-Mesquita <sup>5,21,22,23</sup>, Roel C.H. Vermeulen <sup>24</sup>, Elisabete Weiderpass <sup>25,26,27,28</sup>, Guri Skeie <sup>28</sup>, Therese Haugdahl Nøst <sup>28</sup>, Leila Lujan-Barroso <sup>29</sup>, J. Ramón Quirós <sup>30</sup>, José María Huerta <sup>31,32</sup>, Miguel Rodríguez-Barranco <sup>32,33</sup>, Aurelio Barricarte <sup>32,34,35</sup>, Björn Gylling <sup>36</sup>, Sophia Harlid <sup>37</sup>, Kathryn E. Bradbury <sup>38</sup>, Nick Wareham <sup>39</sup>, Kay-Tee Khaw <sup>40</sup>, Marc Gunter <sup>2</sup>, Neil Murphy <sup>2</sup>, Heinz Freisling <sup>2</sup>, Kostas Tsilidis <sup>5,41</sup>, Dagfinn Aune <sup>5,42,43</sup>, Elio Riboli <sup>5</sup>, John E. Hesketh <sup>3</sup> and David J. Hughes <sup>44,\*</sup>

<sup>1</sup> Department of Epidemiology, Rollins School of Public Health & Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA; veronika.fedirko@emory.edu (V.F.); jeb.jones@emory.edu (J.S.J.); WZHU4@emory.edu (W.Z.)

<sup>2</sup> Section of Nutrition and Metabolism, International Agency for Research on Cancer, 69372 Lyon, France; jenabm@iarc.fr (M.J.); GunterM@iarc.fr (M.G.); MurphyN@iarc.fr (N.M.); FreislingH@iarc.fr (H.F.)

<sup>3</sup> School of Biomedical Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK; Catherine.Meplan@newcastle.ac.uk (C.M.); j.hesketh@rgu.ac.uk (J.E.H.)

<sup>4</sup> Institute for Experimental Endocrinology, University Medical School, D-13353 Berlin, Germany; lutz.schomburg@charite.de (L.S.); sandra.hybsier@charite.de (S.H.)

<sup>5</sup> Department of Epidemiology and Biostatistics, The School of Public Health, Imperial College London, London W2 1PG, UK; afshan.siddiq@genomicsengland.co.uk (A.S.); bas.bueno.de.mesquita@rivm.nl (B.B.-d.-M.); ktsilidis@gmail.com (K.T.); d.aune@imperial.ac.uk (D.A.); e.riboli@imperial.ac.uk (E.R.)

<sup>6</sup> Department of Public Health, Section for Epidemiology, Aarhus University, 8000 Aarhus, Denmark; ko@dce.au.dk

<sup>7</sup> Diet, Genes and Environment Unit, Danish Cancer Society Research Center, DK 2100 Copenhagen, Denmark; annet@cancer.dk

<sup>8</sup> Faculty of Medicine, CESP, University of Paris-Sud, Faculty of Medicine UVSQ, INSERM, University of Paris-Saclay, 94805 Villejuif, France; HANANE.OMICHESSAN@gustaveroussy.fr (H.O.); vittorio.perduca@gmail.com (V.P.); boutron@igr.fr (M.-C.B.-R.)

<sup>9</sup> Centre for Research in Epidemiology and Population Health (CESP), F-94805 Gustave Roussy, Villejuif, France

<sup>10</sup> Laboratory of Applied Mathematics, MAP5 (UMR CNRS 8145), University of Paris Descartes, 75270 Paris, France

<sup>11</sup> Division of Cancer Epidemiology, German Cancer Research Centre (DKFZ), 69120 Heidelberg, Germany; t.kuehn@dkfz-Heidelberg.de (T.K.); v.katzke@dkfz-Heidelberg.de (V.K.)

<sup>12</sup> Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbrücke, 14558 Nuthetal, Germany; Krasimira.Aleksandrova@dife.de

<sup>13</sup> Hellenic Health Foundation, 115 27 Athens, Greece; atrichopoulou@hhf-greece.gr (A.T.); a.karakatsani@hhf-greece.gr (A.K.); a.kotanidou@hhf-greece.gr (A.K.)

<sup>14</sup> 2nd Pulmonary Medicine Department, School of Medicine, National and Kapodistrian University of Athens, "ATTIKON" University Hospital, 106 79 Haidari, Greece

<sup>15</sup> 1st Department of Critical Care Medicine and Pulmonary Services, University of Athens Medical School, Evangelismos Hospital, 106 76 Athens, Greece

- <sup>16</sup> Cancer Registry and Histopathology Department, Civic M.P. Arezzo Hospital, 97100 Ragusa, Italy; rtumino@tin.it
- <sup>17</sup> Department of Clinical Medicine and Surgery, Federico II University, 80138 Naples, Italy; spanico@unina.it
- <sup>18</sup> Cancer Risk Factors and Life-Style Epidemiology Unit, Cancer Research and Prevention Institute—ISPO, 50141 Florence, Italy; g.masala@ispo.toscana.it
- <sup>19</sup> Epidemiology and Prevention Unit, IRCCS Foundation National Cancer Institute, 20133 Milan, Italy; claudia.agnoli@istitutotumori.mi.it
- <sup>20</sup> Molecular and Genetic Epidemiology Unit, Italian Institute for Genomic Medicine (IIGM) Torino, 10126 Torino, Italy; alessio.naccarati@hugef-torino.org
- <sup>21</sup> Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), 3720 Bilthoven, The Netherlands
- <sup>22</sup> Department of Gastroenterology and Hepatology, University Medical Centre, 3584 CX Utrecht, The Netherlands
- <sup>23</sup> Department of Social and Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur 50603, Malaysia
- <sup>24</sup> Institute of Risk Assessment Sciences, Utrecht University, 3512 JE Utrecht, The Netherlands; R.C.H.Vermeulen@uu.nl
- <sup>25</sup> Department of Research, Cancer Registry of Norway, Institute of Population-Based Cancer Research, N-0304 Oslo, Norway; WeiderpassE@iarc.fr
- <sup>26</sup> Department of Medical Epidemiology and Biostatistics, Karolinska Institute, SE-171 77 Stockholm, Sweden
- <sup>27</sup> Genetic Epidemiology Group, Folkhälsan Research Center, and Faculty of Medicine, Helsinki University, 00014 Helsinki, Finland
- <sup>28</sup> Department of Community Medicine, University of Tromsø, The Arctic University of Norway, 9019 Tromsø, Norway; Guri.Skeie@ism.uit.no (G.S.); therese.h.nost@uit.no (T.H.N.)
- <sup>29</sup> Unit of Nutrition and Cancer, Catalan Institute of Oncology (ICO-IDIBELL), L'Hospitalet de Llobregat, 08908 Barcelona, Spain; llujan@iconcologia.net
- <sup>30</sup> EPIC Asturias, Public Health Directorate, 33006 Oviedo, Asturias, Spain; joseamon.quirosgarcia@asturias.org
- <sup>31</sup> Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, 30008 Murcia, Spain; jmhuerta.carm@gmail.com
- <sup>32</sup> CIBER Epidemiology and Public Health (CIBERESP), 28029 Madrid, Spain; miguel.rodriguez.barranco.easp@juntadeandalucia.es (M.R.-B.); aurelio.barricarte.gurrea@cfnavarra.es (A.B.)
- <sup>33</sup> Andalusia School of Public Health, Institute for Biosanitary Research, University Hospital of Granada, University of Granada, 18011 Granada, Spain
- <sup>34</sup> Epidemiology, Prevention and Promotion Health Service, Navarra Public Health Institute, 31003 Pamplona, Spain
- <sup>35</sup> Navarra Institute for Health Research (IdiSNA), 31008 Pamplona, Spain
- <sup>36</sup> Department of Medical Biosciences, Pathology, Umea University, 901 87 Umea, Sweden; bjorn.gyilling@umu.se
- <sup>37</sup> Department of Radiation Sciences, Oncology, Umea University, 901 87 Umea, Sweden; sophia.harlid@umu.se
- <sup>38</sup> Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK; kathryn.bradbury@ceu.ox.ac.uk
- <sup>39</sup> MRC Epidemiology Unit, University of Cambridge, CB2 0QQ Cambridge, UK; nick.wareham@mrc-epid.cam.ac.uk
- <sup>40</sup> School of Clinical Medicine, University of Cambridge, Clinical Gerontology Unit, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK; kk101@medschl.cam.ac.uk
- <sup>41</sup> Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, 45110 Ioannina, Greece
- <sup>42</sup> Department of Nutrition, Bjørknes University College, 0456 Oslo, Norway
- <sup>43</sup> Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, 0372 Oslo, Norway
- <sup>44</sup> Cancer Biology and Therapeutics Group, UCD Conway Institute, School of Biomolecular and Biomedical Science, University College Dublin, D04 V1W8 Dublin, Ireland
- \* Correspondence: david.hughes@ucd.ie; Tel.: +353-1-716-6988

Received: 27 March 2019; Accepted: 22 April 2019; Published: 25 April 2019



**Abstract:** Selenoprotein genetic variations and suboptimal selenium (Se) levels may contribute to the risk of colorectal cancer (CRC) development. We examined the association between CRC risk and genotype for single nucleotide polymorphisms (SNPs) in selenoprotein and Se metabolic pathway genes. *Illumina Goldengate* assays were designed and resulted in the genotyping of 1040 variants in 154 genes from 1420 cases and 1421 controls within the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Multivariable logistic regression revealed an association of 144 individual SNPs from 63 Se pathway genes with CRC risk. However, regarding the selenoprotein genes, only *TXNRD1* rs11111979 retained borderline statistical significance after adjustment for correlated tests ( $P_{ACT} = 0.10$ ;  $P_{ACT}$  significance threshold was  $P < 0.1$ ). SNPs in Wingless/Integrated (Wnt) and Transforming growth factor (TGF) beta-signaling genes (*FRZB*, *SMAD3*, *SMAD7*) from pathways affected by Se intake were also associated with CRC risk after multiple testing adjustments. Interactions with Se status (using existing serum Se and Selenoprotein P data) were tested at the SNP, gene, and pathway levels. Pathway analyses using the modified Adaptive Rank Truncated Product method suggested that genes and gene  $\times$  Se status interactions in antioxidant, apoptosis, and TGF-beta signaling pathways may be associated with CRC risk. This study suggests that SNPs in the Se pathway alone or in combination with suboptimal Se status may contribute to CRC development.

**Keywords:** selenium; selenium status; selenoprotein gene variation; selenium pathway; colorectal neoplasms; selenoprotein P; prospective cohort; colorectal cancer risk; genetic epidemiology; biomarkers

## 1. Introduction

In Europe, colorectal cancer (CRC) is the cancer type with both the second highest incidence and mortality rate [1]. Substantial CRC risk may derive from dietary factors, genetic variants, and their interactions [2,3].

Experimental and observational evidence suggests that suboptimal dietary intakes of the micronutrient selenium (Se) contribute to greater risk for the development of cancers at several anatomical sites, including the colorectum [4–6]. In humans, Se exerts its potential anti-carcinogenic properties through incorporation into 25 selenoproteins by the amino acid selenocysteine [7,8]. Several selenoproteins protect cells from damaging oxidative radicals including the glutathione peroxidases (notably GPX1 and GPX4), components of the thioredoxin reductase system (*TXNRD1-3*) and selenoprotein P (SELENOP; please note the modified selenoprotein nomenclature [9]) which is also critical for Se transport [8,10,11].

The major mechanism through which Se is thought to influence the risk of CRC development is variation in gene expression and biosynthesis of protective selenoproteins [12,13]. In rodent models, adequate Se intake and selenoprotein expression have been shown to prevent colon cancer while selenoprotein dysregulation may increase colon cancer risk [14–16]. Data from nutritional intervention trials and epidemiological studies suggest implications for Se intake regarding CRC risk could potentially be more important in individuals with particular selenoprotein genotypes and/or in populations with low Se status, such as in Western Europe where the present study was conducted [4,5,17,18]. Risk modification by sex has also been observed for CRC risk associations with selenoprotein genotypes [19,20] and Se status [5,17].

Genetic variations in approximately half of the 25-known human selenoprotein genes have been associated with susceptibility to CRC and/or colorectal adenoma (CRA) risk in at least seven populations from Asia, North America and Europe; in addition some of these variants have been shown to impact survival outcomes (reviewed in [4,21]). Although some of these studies have been performed in suboptimal Se intake areas, large studies have more generally been conducted in Se-replete

environments in North America and these have reported evidence both for [22,23] and against [17] an association of selenoprotein genes with CRC risk. However, only a limited number of single nucleotide polymorphisms (SNPs) in selected selenoprotein genes have been analyzed, while in several of these reports the Se status of the analyzed cohort was not assessed.

To our knowledge, it is unknown which selenoproteins are critical in maintaining colonic health and no study has comprehensively evaluated variation in all selenoprotein genes for association with CRC risk. Moreover, interactions of selenoprotein genetic variations according to robust Se status biomarkers have not been explored. As both genetic factors and dietary Se intake can influence the pattern of selenoprotein expression and biosynthesis, we hypothesized that variation in selenoprotein genes, and in related signaling pathway genes influenced by Se intake (together comprising the 'Se pathway'), affect CRC development risk, while Se status may modify this risk.

In this study, we have examined for the first time the association of detailed Se pathway gene variation with cancer risk in 1420 CRC cases and 1421 controls within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. We previously reported in a subset of this nested cohort with 966 case-control pairs that a higher Se status (ascertained by serum levels of Se and SELENOP) was associated with a lower CRC risk [5]. In these Western European subjects, the mean Se and SELENOP circulating levels were 84.0 µg/L and 4.3 mg/L in cases and 85.6 µg/L and 4.4 mg/L in controls, respectively. Thus, our present study was conducted in a generally suboptimal Se status population, as these Se concentrations are insufficient for optimal GPX3 expression and SELENOP saturation [5,6]. We now report the interaction between these genes and their corresponding pathways with Se status biomarkers and CRC risk.

## 2. Materials and Methods

### 2.1. Study Population and Design

EPIC is a multicenter prospective cohort study designed to investigate the association between diet, lifestyle, genetic and environmental factors and the incidence of cancers. The rationale and methods of the EPIC design have been described previously [24,25]. Briefly, 521,448 men and women mostly aged 25–70 years were enrolled between 1992–2000 in 23 sub-cohorts in 10 European countries (Denmark, France, Germany, Greece, Italy, The Netherlands, Norway, Spain, Sweden, and United Kingdom). The present analysis is based on participant data from all sub-cohorts except for Norway. At recruitment, standardized dietary, lifestyle and socio-demographic questionnaires including information on physical activity, education, smoking and medical history; anthropometric data, and blood samples were collected from participants. Blood and DNA samples are stored at the International Agency for Research on Cancer (IARC, Lyon, France) at  $-196\text{ }^{\circ}\text{C}$  under liquid nitrogen for all countries except Denmark ( $-150\text{ }^{\circ}\text{C}$ , nitrogen vapor) and Sweden ( $-80\text{ }^{\circ}\text{C}$  freezers). Sample storage standardization including DNA extraction and quantification protocols were previously described in [26].

All study participants provided written informed consent. Ethical approval for the EPIC study was obtained from the review boards of the IARC (IARC Ethics Committee) and local participating centers. Study design methods were performed in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (<https://www.strobe-statement.org/index.php?id=strobe-home>).

### 2.2. Follow-Up for Cancer Incidence

Cancer incidence was determined through record linkage with population-based cancer registries (Denmark, Italy, Netherlands, Spain, Sweden, United Kingdom) or via a combination of methods, including the use of health insurance records, cancer and pathology registries, and active contact of study subjects or next-of-kin (France, Germany, Greece). Complete follow-up censoring dates for this study varied among centers, ranging between June 2002 and June 2003.

### 2.3. Selection of Cases and Controls and Study Design

Case subjects were men and women who developed first incident CRC after recruitment and before the latest follow-up date. Cancer incidence data were coded using the 10th Revision of the International Classification of Diseases (ICD-10) and the second revision of the International Classification of Disease for Oncology (ICDO-2). Colon cancers were defined as tumors in the cecum, appendix, ascending colon, hepatic flexure, transverse colon, splenic flexure, descending and sigmoid colon (C18.0-C18.7), and overlapping or unspecified origin tumors (C18.8 and C18.9). Rectal cancers were defined as tumors occurring at the recto-sigmoid junction (C19) or rectum (C20). Anal canal cancers (C21) were excluded. Colorectal cancer is the combination of the colon and rectal cancer cases.

All subjects with prior cancer diagnosis at any site (except non-melanoma skin cancer) were excluded. Cases were matched 1:1 by study center of enrollment, sex, age at blood collection, time of blood collection and fasting status, and menopausal status among women. Premenopausal women were matched on phase of menstrual cycle and postmenopausal women were matched on current hormonal therapy (HT) use. The matching was done as part of a previously published study on Se status [5], except for cases from Denmark for which new control subjects were identified due to problems with accessing the biobank. Furthermore, additional newly identified cases with their matching controls were also included for the genotyping from all participating countries but did not have biomarkers of Se status. Sweden was the only country of the nine participating in the genetic analysis for which we had no Se status data. Hence, there were 1478 cases and 1478 controls available for genotyping but the Se status information was only available for 966 of the cases and for 966 of the controls.

### 2.4. Gene Selection and Rationale

To examine selenoprotein gene and wider Se pathway gene variations in relation to CRC risk, we selected 1264 functional and haplotype tagging SNPs (tagSNPs) to comprehensively analyze common SNP variation in 164 Se pathway genes, which we assigned into eight functional pathways (listed in Supplementary Table S1). These included 42 genes in the primary selenoprotein pathway 1 (25 selenoprotein genes and 17 genes involved in Se transport and metabolism), and 122 genes in pathways 2–8 from (i) pathways affected by Se intake (Wnt, mTOR, Nrf2 and NF- $\kappa$ B signaling, endoplasmic reticulum and oxidative stress responses), and (ii) associated pathways of inflammatory response, apoptosis, DNA repair, Transforming growth factor (TGF) beta-signaling, and cell-cycle control [12] as detailed in Méplan and Hesketh, 2012 [13]. Variants in several genes from these affiliated pathways have been associated with CRC risk including regions of the Wnt signaling gene *C-MYC* in CRC genome-wide association studies (GWAS) [27]. Our SNP analysis set was substantially enlarged from and included the 384 Se pathway SNPs (in 72 Se related genes) Se ‘SNP-Chip’ devised for a similar study of gene-Se interaction in a prostate cancer study within EPIC [28].

### 2.5. Tagging Single Nucleotide Polymorphism (tagSNP) Selection Protocol

A list of SNPs in all gene regions was compiled using the data from HapMap (release 27, based on dbSNP version b126 and NCBI genome build 36). TagSNPs were selected by use of the Tagger algorithm as implemented in the Haploview 3.2 software (Broad Institute, Cambridge, MA, USA). Parameters used for SNP selection were a Minor Allele Frequency (MAF)  $\geq 5\%$  in Caucasians and pairwise tagging ( $r^2 \geq 0.8$ ). To include SNPs in promoter and potential regulatory regions,  $\pm 2$  to 5 kilo base-pairs beyond the 5' and 3' ends were included. Additionally, known functional variants in our selected genes were added to the tagSNP list, e.g., for the selenoproteins these included rs7579, rs297299, and rs3877899 in *SELENOP* [4]. Selected SNPs were then assessed for suitability for the *Illumina GoldenGate*<sup>TM</sup> (Saffron Walden, Essex, UK) genotyping platform using Illumina's custom assay building platform ([https://www.illumina.com/Documents/products/technotes/technote\\_goldengate\\_design.pdf](https://www.illumina.com/Documents/products/technotes/technote_goldengate_design.pdf)).

Fifty-five SNPs which failed assay development criteria were replaced by proxy SNPs, i.e., those within the same genic region in high LD ( $r^2 > 0.8$ ) to the original SNP. Proxy SNP replacements for functional selenoprotein SNPs which failed assay design included rs1800668 for *GPX1*-rs1050450, and rs5845 plus rs540049 for *SELENOP*-rs5859. However, there were no adequate proxies for *SELENOS*-rs34713741 or *GPX4*-rs713041.

## 2.6. Genotyping

A total of 1264 SNPs from 164 Se pathway genes were genotyped by *Illumina Goldengate*<sup>TM</sup> in DNA samples available for 1478 case-control pairs matched within EPIC. Genotyping was performed simultaneously for cases and controls, blinded to case-control status (but with matched pairs analyzed in the same batch). A total of 62 replicate samples were genotyped to test for internal quality control, approximately 2 per genotyping plate, with the lowest reproducibility frequency for each of the replicates of 0.98. Samples with unclear or failed genotype calls were excluded from the analysis, leaving 1420 cases and 1421 controls for subsequent analyses.

From the 1264 initially selected, 96 SNPs failed genotyping, 27 failed Hardy–Weinberg Equilibrium (HWE), and 101 had less than 80% successfully genotyped samples. Thus, 1040 SNPs in 154 Genes (24 selenoprotein genes analyzed of 25, and 130 other Se pathways genes) with at least 80% genotypes across all genotyped samples were included in the final dataset (with a final genotyping call rate of 0.97, excluding zero call rate and those removed). Supplementary Table S2 provides the full gene and SNP list successfully analyzed in the current study.

## 2.7. Selenium Status Assays

Measurements of serum Se and SELENOP were previously done for a subset (966 cases and 966 controls) of the current analyzed cohort. The methods used were described in Hughes et al., 2015 and Hybsier et al., 2017 [5,29]. Briefly, total Se levels were measured in 4  $\mu$ L of each serum sample using a bench-top total reflection X-ray fluorescence (TXRF) spectrometer (Picofox<sup>TM</sup> S2, Bruker Nano GmbH, Berlin, Germany). SELENOP protein concentrations were ascertained from 20  $\mu$ L of each serum sample by a colorimetric enzyme-linked immunoassay (Selenotest<sup>TM</sup>, ICI GmbH, Berlin, Germany). For quality-control, the sample type (case or control) was blinded and two serum samples of known Se and SELENOP concentrations for intra-assay variability were included in each analysis plate. The samples were measured in duplicate and the mean concentration values, standard deviation (SD), and coefficient of variation (CV) were calculated. Duplicate samples with variances in concentration over 10% were re-measured. The evaluation was performed using GraphPad Prism 6.01 (GraphPad Software, La Jolla, CA, USA) using a four-parameter logistic function. The CV was 7.3% and 7.2% for controls 1 (SELENOP: 1.5 mg/L) and 2 (SELENOP: 8.6 mg/L), respectively.

## 2.8. Statistical Analysis

Both unconditional and conditional logistic regression analysis were carried out to assess the association of individual SNPs with CRC risk, adjusting for age (as a continuous variable), sex, and study center and provided similar results. We present the data for the unconditional logistic regression. Four standard genetic analysis models were tested for disease penetrance: multiplicative, additive, common recessive, and common dominant models [30]. Sub-group analyses by sex and anatomical sub-site of the colorectum (colon and rectum) were conducted. The associations between Se and SELENOP concentrations and genetic variants (coded as 0, 1, 2 corresponding to the number of minor alleles) were assessed among controls using linear regression models adjusted for age, sex, and center. Further adjustment by body mass index (BMI), smoking status, and physical activity did not change the results substantially.

Multiple testing corrections were performed by the Benjamini–Hochberg (BH) procedure [31]. *P*-values were also adjusted for correlated tests ( $P_{ACT}$ ) to take account of the correlated nature of the

SNP data in biologically relevant and related pathways [32]. BH was performed for all SNPs, followed by  $P_{ACT}$  for the genes that had SNPs with  $P < 0.01$ .

We further employed exploratory gene- and pathway-based testing based on overall SNP variation to help identify possible important Se related biological pathways and genes with multiple risk variants that may be discounted in multiple testing corrections for the large number of SNPs with small effect sizes in a SNP by SNP approach. Genes were classified a priori into a primary best-known functional pathway based on the literature (listed in Supplementary Table S1). Gene- and pathway-based  $P$ -values were computed using the PIGE (Self-Contained Gene Set Analysis for Gene- and Pathway-Environment Interaction Analysis) R package which implements the modified Adaptive Rank Truncated Product (ARTP) test using a permutation algorithm [33] to accommodate gene-environment interactions (<https://cran.rproject.org/web/packages/PIGE/index.html>). Prior to this analysis, SNPs in high linkage disequilibrium (LD) were removed using AdaJoint [34] and the online tool SNPsnap (<https://data.broadinstitute.org/mpg/snpnap/about.html>) so that all SNP pairs had LD  $r^2 < 0.8$ . Gene  $\times$  Se status interactions were also examined using the PIGE R package. Although these methods do not identify individual susceptibility loci, they may help to identify a pathway that could modify the association between Se status and CRC risk. An advantage is that they do not require a priori knowledge of directionality for the variants.

All statistical tests were two-sided, and  $P$ -values  $< 0.05$  were considered statistically significant (except  $P < 0.1$  for  $P_{ACT}$ ). Analyses were conducted using SAS version 9.2 (SAS Institute, Cary, NC, USA) and R (R Foundation for Statistical Computing, Vienna, Austria; <http://www.R-project.org/>) statistical packages.

### 3. Results

#### 3.1. Baseline Characteristics of Participants

The baseline characteristics of participants are presented in Table 1. Colon and rectal cancer cases were diagnosed, on average, 4.1 and 4.2 years after blood collection, respectively. CRC cases were overall less likely to be physically active compared to controls. There were no data on Se supplement use for our study participants.

**Table 1.** Selected baseline characteristics of incident colon and rectal cancer cases and controls, the European Prospective Investigation into Cancer and Nutrition (EPIC) study, 1992–2003.

Characteristic	Colon Cancer Cases		Rectal Cancer Cases		Controls	
N	900		520		1419	
Women, N (%)	475	(52.8)	230	(44.2)	701	(49.4)
Mean age at blood collection, (SD) yrs	58.8	(7.5)	58.0	(6.9)	58.6	(7.4)
Mean years of follow-up (SD) yrs	4.1	(2.3)	4.2	(2.2)		
Smoking status, N (%) *						
Never	385	(42.8)	195	(37.5)	594	(41.9)
Former	299	(33.2)	177	(34)	460	(32.4)
Smoker	204	(22.7)	142	(27.3)	349	(24.6)
Physical activity, N (%) *						
Inactive	129	(14.3)	73	(14)	183	(12.9)
Moderately inactive	257	(28.6)	145	(27.9)	367	(25.9)
Moderately active	374	(41.6)	209	(40.2)	612	(43.1)
Active	75	(8.3)	55	(10.6)	148	(10.4)
BMI, kg/m <sup>2</sup> , (SD)	26.9	(4.36)	26.6	(3.92)	26.3	(3.84)



Table 1. Cont.

Characteristic	Colon Cancer Cases		Rectal Cancer Cases		Controls	
Country, N (%)						
Sweden	55	(6.1)	33	(6.3)	86	(6.1)
Denmark	174	(19.3)	164	(31.5)	340	(24)
The Netherlands	99	(11)	54	(10.4)	158	(11.1)
United Kingdom	166	(18.4)	74	(14.2)	250	(17.6)
Germany	110	(12.2)	69	(13.3)	169	(11.9)
France	22	(2.4)	6	(1.2)	29	(2)
Italy	144	(16)	58	(11.2)	198	(14)
Spain	101	(11.2)	45	(8.7)	141	(9.9)
Greece	29	(3.2)	17	(3.3)	48	(3.4)

\* Percentages do not add up to 100% due to missing values. Abbreviations: BMI, body mass index; N, sample size; SD, standard deviation; yrs, years.

### 3.2. Se Pathway Genetic Variation and Colorectal Cancer (CRC) Risk Association

The 1040 tagging SNPs successfully analyzed from 154 genes and in HWE are shown in Supplementary Table S2 (which also provides all the genetic analysis results for CRC, plus stratified analyses for colon and rectal sub-site and by sex). These include 325 SNPs from 41 selenoprotein and Se transport/selenoprotein biosynthesis genes (designated as the primary Se pathway 1), and 715 variants from the other 113 wider Se metabolic pathway genes (pathways 2–8). A summary of the genetic associations before and after multiple testing corrections is provided in Supplementary Figure S1. Prior to adjustment for multiple comparisons, there were 144 SNPs in 63 genes nominally associated with CRC risk ( $P < 0.05$  in at least one of the disease penetrance models tested; listed in Supplementary Table S3). There were 28 unique SNPs in LD with other associated SNPs and these are listed and highlighted in Supplementary Table S3 (tab 'LD CEU'). Among the 40 SNPs significantly associated with CRC risk from pathway 1, approximately half (21) were in 12 selenoprotein genes (i.e., 50% of the 24 selenoprotein genes successfully genotyped out of 25) and have the potential to affect the function or expression of individual selenoproteins, although this remains to be investigated. These 12 selenoprotein genes include those previously found associated with CRC risk (*GPX1*, *GPX4*, *SELENOF*, *TXNRD1*, *TXNRD2*, *TXNRD3*; for reviews, see [4,21]) and several with limited prior or no previous evidence of association with CRC risk (*DI01*, *GPX6*, *SELENOM*, *SELENON*, *SELENOT*, *SELENOV*). The other 19 SNPs associated with CRC in pathway 1 are in 8 of the 17 (47%) other Se transport/selenoprotein biosynthesis genes. Therefore, they have the potential to affect the synthesis of most selenoproteins (which also needs to be examined). Notably, 31% of the genes harboring SNPs associated with CRC risk (20 of 63) were related to selenoprotein biosynthesis and function implicated in protection from cancer development [4,21] with pathway 1 and 2 proteins involved in (1) Se homeostasis (*SELENOP*, *SEPHS1*, *SEPSEC*, *EFSEC*, *SCLY*), (2) antioxidant enzymes (*GPXs*, *TXNRDs*, *SELENON*), and (3) endoplasmic reticulum (ER) function or stress (*SELENOF*, *SELENOM*, *SELENOT*, and again *SELENON*). Additionally, several of these genes (e.g., *GPX1*, *GPX5*, *LRP2*, *SEPHS1*, *SELENOM*, *SELENON*, *TXNRD1*, and *TXNRD2*) had multiple SNPs and/or SNPs with raw  $P$ -values  $< 0.01$  associated with CRC risk further supporting a role of selenoproteins, selenoprotein metabolism, ER stress, and oxidative stress in CRC development. Table 2 lists the SNPs in the primary Se pathway 1 with raw  $P$ -values  $< 0.01$  associated with CRC risk.

In pathways 3–8, considering genes with multiple SNPs associated with CRC risk or SNPs with raw  $P$ -values  $< 0.01$  for at least one genetic model, there were several notable and some novel associations with CRC risk for genes in pathways 3 (*C-MYC*, *FRZB*), 4 (*APAF1*, *BAX*, *FOXO3*), 5 (*IL12B*, *RPS6KA2*, *TRL4*), 6 (*MSH2*, *MSH3*) and 7 (*BMP2*, *BMPR2*, *SMAD3*, *SMAD7*, *TFGB1*).

**Table 2.** Single nucleotide polymorphisms (SNPs) associated with colorectal cancer (CRC) risk in primary selenium pathway 1 (selenium and selenoprotein transport, biosynthesis and metabolism) with raw *P*-values < 0.01 in at least one genetic model prior to multiple testing adjustment, the EPIC study, 1992–2003.

Gene/SNP/Genotype	CRC	Control	OR (95% CI)	<i>P</i>	<i>P</i> <sub>BH</sub> <sup>+</sup>
<i>GPX1</i> /rs17080528					
GG	700	620	1.00 (ref)	0.010	0.703
GA	580	636	0.81 (0.69,0.95)		
AA	131	154	0.75 (0.58,0.97)		
Additive *	1411	1410	0.84 (0.75,0.95)	0.003	0.554
Dominant (GA + AA vs. GG)	1411	1410	0.80 (0.69,0.92)	0.003	0.534
Recessive (AA vs. GG + GA)	1411	1410	0.83 (0.65,1.06)	0.137	0.854
<i>SELENOM</i> /rs11705137					
AA	367	346	1.00 (ref)	0.024	0.753
AG	631	648	0.91 (0.75,1.09)		
GG	288	359	0.74 (0.60,0.92)		
Additive *	1286	1353	0.86 (0.78,0.96)	0.008	0.684
Dominant (AG + GG vs. AA)	1286	1353	0.85 (0.71,1.01)	0.064	0.815
Recessive (GG vs. AA + AG)	1286	1353	0.79 (0.66,0.95)	0.012	0.710
<i>SELENON</i> /rs4659382					
GG	783	713	1.00 (ref)	0.019	0.747
GC	509	573	0.80 (0.69,0.94)		
CC	96	107	0.82 (0.61,1.10)		
Additive *	1388	1393	0.86 (0.76,0.97)	0.011	0.710
Dominant (GC + CC vs. GG)	1388	1393	0.81 (0.69,0.94)	0.005	0.625
Recessive (CC vs. GG + GC)	1388	1393	0.90 (0.67,1.20)	0.455	0.954
<i>SEPHS1</i> /rs2275129					
GG	361	423	1.00 (ref)	0.032	0.780
GC	726	690	1.23 (1.03,1.47)		
CC	321	295	1.28 (1.04,1.58)		
Additive *	1408	1408	1.14 (1.02,1.26)	0.017	0.747
Dominant (GC + CC vs. GG)	1408	1408	1.25 (1.05,1.47)	0.010	0.697
Recessive (CC vs. GG + GC)	1408	1408	1.12 (0.94,1.34)	0.217	0.885
<i>TXNRD1</i> /rs11111979 <sup>^</sup>					
GG	395	429	1.00 (ref)	0.015	0.745
GC	627	680	1.00 (0.84,1.20)		
CC	279	230	1.34 (1.07,1.67)		
Additive *	1301	1339	1.14 (1.02,1.27)	0.022	0.749
Dominant (GC + CC vs. GG)	1301	1339	1.09 (0.92,1.28)	0.315	0.932
Recessive (CC vs. GG + GC)	1301	1339	1.33 (1.10,1.62)	0.004	0.566

<sup>+</sup> = After Benjamini–Hochberg (BH) multiple testing correction; \* = Additive models impose a structure in which each additional copy of the variant allele increases the response (log odds ratio) by the same amount; <sup>^</sup> = *TXNRD1* rs11111979 was borderline significant after adjustment for correlated tests (*P*<sub>ACT</sub> = 0.10). EPIC = European Prospective Investigation into Cancer and Nutrition.

None of the SNPs in the primary Se pathway 1 remained significant after multiple testing corrections by the BH procedure. Overall, only 6 SNPs harbored by more distantly related genes in cell-signaling pathways retained significance (*FRZB*, *SMAD3*, and *SMAD7*; see Table 3). Genes harboring SNPs with raw *P*-values < 0.01 with CRC risk for at least one genetic model (21 genes/34 SNPs) were further considered for gene-wide variance significance by the *P*<sub>ACT</sub> method. For pathway 1, the *TXNRD1* selenoprotein variant rs11111979, an intron 3′–5′UTR SNP previously associated with healthy aging [35], remained borderline significant for an association with CRC risk (*P*<sub>ACT</sub> = 0.100; *P*<sub>ACT</sub> significance threshold was *P* < 0.1) in the recessive genetic model. Including *FRZB*, *SMAD3* and

*SMAD7*, the other wider pathway genes retaining significance were *C-MYC* ( $P_{ACT} = 0.032$ ), *BMP2* ( $P_{ACT} = 0.012$ ) and *BAX* ( $P_{ACT} = 0.035$ ).

**Table 3.** Single Nucleotide Polymorphisms (SNPs) statistically significantly associated with colorectal cancer (CRC) risk after Benjamini–Hochberg (BH) multiple testing correction, the EPIC study, 1992–2003.

Gene/SNP/Genotype	CRC	Control	OR (95% CI)	<i>P</i>	<i>P<sub>BH</sub></i>
<i>FRZB/rs17265803</i> ^					
AA	844	976	1.00 (ref)	3.04E-06	0.003
AG	315	240	1.56 (1.28,1.89)		
Additive *	1163	1232	1.35 (1.13,1.62)	0.001	0.372
Dominant (AG + GG vs. AA)	1163	1232	1.48 (1.22,1.79)	6.77E-05	0.034
Recessive (GG vs. AA + AG)	1163	1232	0.26 (0.09,0.80)	0.018	0.747
<i>SMAD3/rs7180244</i> ^					
GG	994	1183	1.00 (ref)	2.22E-16	1.12E-12
GC	372	198	2.33 (1.92,2.83)		
Additive *	1372	1388	2.16 (1.79,2.60)	1.11E-15	3.74E-12
Dominant (GC + CC vs. GG)	1372	1388	2.28 (1.88,2.77)	1.11E-16	1.12E-12
Recessive (CC vs. GG + GC)	1372	1388	0.87 (0.29,2.61)	0.804	0.997
<i>SMAD7/rs11874392</i>					
AA	478	400	1.00 (ref)	1.39E-07	2.82E-04
AT	704	671	0.88 (0.74,1.04)		
TT	222	337	0.55 (0.44,0.68)		
Additive *	1404	1408	0.75 (0.68,0.84)	1.99E-07	3.36E-04
Dominant (AT + TT vs. AA)	1404	1408	0.77 (0.65,0.90)	0.001	0.372
Recessive (TT vs. AA + AT)	1404	1408	0.59 (0.49,0.72)	6.26E-08	1.58E-04
<i>SMAD7/rs12953717</i>					
GG	366	470	1.00 (ref)	3.47E-05	0.019
GA	671	643	1.36 (1.14,1.62)		
AA	335	273	1.60 (1.29,1.98)		
Additive *	1372	1386	1.27 (1.14,1.41)	9.07E-06	0.006
Dominant (GA + AA vs. GG)	1372	1386	1.43 (1.21,1.68)	2.38E-05	0.014
Recessive (AA vs. GG + GA)	1372	1386	1.32 (1.10,1.59)	0.003	0.534
<i>SMAD7/rs4939827</i>					
AA	433	378	1.00 (ref)	6.46E-06	0.005
AG	664	634	0.92 (0.77,1.09)		
GG	248	357	0.60 (0.49,0.75)		
Additive *	1345	1369	0.79 (0.71,0.87)	9.46E-06	0.006
Dominant (AG + GG vs. AA)	1345	1369	0.80 (0.68,0.95)	0.009	0.697
Recessive (GG vs. AA + AG)	1345	1369	0.64 (0.53,0.77)	1.66E-06	0.002
<i>SMAD7/rs6507874</i>					
AA	467	389	1.00 (ref)	1.93E-06	0.002
AG	705	677	0.87 (0.73,1.03)		
GG	234	336	0.57 (0.46,0.71)		
Additive *	1406	1402	0.77 (0.69,0.86)	1.26E-06	0.002
Dominant (AG + GG vs. AA)	1406	1402	0.77 (0.65,0.90)	0.001	0.420
Recessive (GG vs. AA + AG)	1406	1402	0.63 (0.52,0.76)	1.16E-06	0.002

^ = Results for the rare homozygous genotypes are omitted for these SNPs due to the small sample numbers with these genotypes; \* = Additive models impose a structure in which each additional copy of the variant allele increases the response (log odds ratio) by the same amount. EPIC = European Prospective Investigation into Cancer and Nutrition.

Supplementary Table S3 also catalogs the SNPs with raw significant *P*-values and the BH corrections stratified by cancer sub-site (comprising 138 SNPs in 65 genes for colon in tab ‘colon cancer’

and 123 SNPs in 54 genes for rectum listed in tab 'rectal cancer'). Additionally, the tab 'All' lists all the SNPs showing an association for CRC, colon only, rectal only, plus the analyses stratified by sex. Generally, there was predominate overlap in the genes associated with CRC and sub-site risks. Genes containing variants uniquely associated only with sub-site risk plus raw  $P$ -values  $< 0.01$  comprised rs12124257 in *PTGS2* for colon cancer and 4 SNPs in *IL10* for rectal cancer.

### 3.3. Associations Between Se Pathway Genetic Variation and Se Status

Among controls, 99 different SNPs in 55 genes were nominally associated with Se status (raw  $P$ -values  $< 0.05$ ). From these 99 variants, 87 SNPs from 45 genes were associated with either Se or SELENOP levels (55 SNPs in 33 genes and 56 SNPs in 37 genes, respectively) while the other 12 variants from 10 genes were associated with both Se status measures, including 2 each in *HIF1A* and *SMAC*. Eight pathway 1 genes harbored 14 SNPs significant for Se level changes (including 7 SNPs in 5 selenoprotein genes) while 9 pathway 1 genes carried 15 variants significant for SELENOP status association (11 SNPs in 6 selenoprotein genes). However, none of the associations retained significance after BH multiple testing adjustments. These SNP IDs together with the beta coefficients for change in Se ( $\mu\text{g/L}$ ) or SELENOP ( $\text{mg/L}$ ) are listed according to gene pathway in Supplementary Table S4.

### 3.4. Pathway Analysis

An exploratory analysis of CRC risk with gene variation and gene  $\times$  Se status interaction within eight predefined pathways was performed using the PIGE package. A summary of the main PIGE results per pathway is presented in Table 4, while Supplementary Table S5 provides all the  $P$ -values for each gene per pathway designation. Considering nominal significance for association with disease risk by pathway of  $P < 0.05$ , then these analyses suggest that TGF-beta signaling ( $P < 0.001$ ) is the sole pathway highly associated with CRC risk independent of Se status interaction. Antioxidant/redox pathway genetic variation combined with Se status interactions was associated with a significant effect on CRC risk ( $P = 0.011$  and  $0.010$  for Se and SELENOP interactions, respectively), possibly driven by SNPs in *HIF1A*, *KEAP1*, *GPX7*, *CAT*, and *SOD2* (when considering the  $P$ -values for each individual gene regarding gene only variation and gene  $\times$  Se status interaction; see Supplementary Table S5). In contrast, the risk association with gene variation in the apoptosis pathway seems to depend more on interaction with Se levels ( $P = 0.003$ ) but not with SELENOP concentrations ( $P = 0.105$ ). For gene only analyses there were several genes across the pathways associated with CRC risk including the pathway 1 genes *GPX1*, *SELENOM*, *SELENON*, and *SEPHS1*, from which only *SELENON* was significant for both the gene only and gene  $\times$  Se interaction PIGE analyses (Supplementary Table S5). In agreement with previous large gene association and GWAS studies, overall genetic variation in the *SMAD3*, *SMAD7*, *BMP2*, and *BMP2* genes was associated with CRC risk [36–38]. Excluding individuals with no measurements of Se or SELENOP concentrations did not substantially change the main gene only PIGE results (both sets of results are provided in Supplementary Table S5).

**Table 4.**  $P$ -values for genetic pathways and pathway-selenium (Se) status interactions and colorectal cancer risk, the EPIC study, 1992–2003.

Pathway	$P_{\text{Pathway Only}}$	$P_{\text{Pathway Only (non-Missing Se Status) ***}}$	$P_{\text{Pathway} \times \text{Se Interaction}}$	$P_{\text{Pathway} \times \text{SELENOP Interaction}}$
Se and Selenoproteins *	0.217	0.098	0.615	0.726
Antioxidant and Redox	0.173	0.072	0.011	0.010
Cell signaling **	0.307	0.489	0.223	0.872
Apoptosis	0.361	0.097	0.003	0.105
Inflammation	0.822	0.262	0.199	0.607
DNA repair	0.739	0.432	0.175	0.088
TGF $\beta$ signaling	$<0.001$	0.001	0.061	0.764
Cell cycle control	0.398	0.475	0.097	0.449

\* Se and selenoprotein transport, biosynthesis & metabolism. \*\* Includes Wnt, mTOR, Nf $\kappa$ B, and Nrf2 signaling.

\*\*\* Includes only participants with non-missing blood Se or SELENOP concentrations. EPIC = European Prospective Investigation into Cancer and Nutrition.

#### 4. Discussion

The results of this prospective nested case-control study represent the largest reported analysis of both the association of Se pathway SNP variation and the interaction with Se status biomarkers (serum Se levels and SELENOP protein concentrations) with CRC risk. The analysis of 1,040 tagSNPs in 154 Se pathway genes in DNA samples from 1,420 CRC cases and 1,421 controls within EPIC indicated that 144 of these SNPs in 63 genes were nominally associated with CRC risk. However, for pathway 1 only the *TXNRD1* selenoprotein gene rs11111979 SNP retained borderline significance after correction for multiple testing. For pathways 2–8, variants in *BAX*, *BMP2*, *C-MYC*, *FRZB*, *SMAD3*, and *SMAD7* passed significance thresholds following these adjustments.

Selenoprotein genes nominally associated with CRC risk included several with limited or no prior evidence (*DIO1*, *GPX6*, *SELENOM*, *SELENON*, *SELENOT*, *SELENOV*) and those reported in several studies (*GPX1*, *GPX4*, *SELENOF*, *TXNRD1*, *TXNRD2*, *TXNRD3*) for an association with CRC (or specifically colon or rectal cancer) risk. This latter group of genes has been more extensively examined due to their putative roles related to cancer prevention in colonic tissue (for reviews, see [4,21]), while the former group of selenoprotein genes have generally less well-characterized function, especially regarding how they may affect colorectal function and CRC development. Overall, any functional consequences from genetic variations in these genes, together with Se status, may affect several oxidative stress, inflammatory, and signal translation pathways implicated in colorectal carcinogenesis [13,39]. Notably several of these genes are ER-resident selenoproteins (*SELENOF*, *SELENOM*, *SELENON*, *SELENOT*), thought to be involved in ER-stress response and calcium flux, comprising a potentially important mechanism of selenoprotein-related cancer prevention or promotion [40].

None of the 3 *GPX1* SNPs (rs17080528, rs3448, rs9818758) or rs2074451 in *GPX4* associated with CRC risk are in high LD (i.e.,  $r^2 \geq 0.8$ ) to the functional *GPX1* Pro/Leu rs1050450 and *GPX4* rs713041 SNPs (for which the *Illumina* assays failed) previously implicated in prostate, breast, lung (rs1050450), and CRC risk (rs713041) [4]. However, from these pathway 1 genes, only a *TXNRD1* selenoprotein variant (rs11111979), one of the three thioredoxin reductases which function in redox control [8], remained borderline significant for an increased CRC risk when applying gene-wide variance considerations by the  $P_{ACT}$  method. Interestingly, this SNP inducing a change in the 5' untranslated region of *TXNRD1*, among others in *TXNRD1*, was previously observed to be associated with age-related physical performance [35], and age is a primary risk factor for CRC development [41]. In the wider metabolic pathway, following adjustment for multiple testing, genotypes for SNPs in Wnt, TGF-beta signaling, and apoptosis pathway genes (*C-MYC*, *FRZB*, *SMAD3*, *SMAD7*, *BMP2*, and *BAX*) were also significantly associated with CRC risk.

Positive associations of selenoprotein gene variants with CRC risk have been more commonly reported in areas with suboptimal Se availability such as European populations, than regions with generally adequate Se intake (e.g., North America). However, tagSNPs in several selenoprotein genes (*GPX3*, *TXNRD3*, *SELENON*, *SELENOF*, and *SELENOX*) were also associated with colon or rectal cancer risk and/or survival outcomes in two separate studies of several large case-control USA cohorts drawn from populations with generally adequate dietary Se intakes [22,23]. Associations of multiple SNPs in the same selenoprotein gene with CRC risk were observed in this study for *GPX1*, *SELENON*, *TXNRD1*, and *TXNRD2* (3, 4, 3, and 3 SNPs, respectively), broadly comparable to previous reports [4]. As reported by Slattery et al. in 2012 many of the same selenoprotein genes were separately associated with colon and rectal cancer risk in sub-site analyses although risks often differed by SNP [22]. From the 4 *SELENON* variants that were associated with CRC risk (rs11247735, rs2072749, rs4659382, and rs11247710), the first 3 were previously associated with rectal cancer risk in this North American cohort [22]. In our sub-site analyses, rs11247735, rs2072749, and rs11247710 were associated with rectal cancer risk only, and rs4659382 with both colon and rectal cancer. We also found further modifications of gene only and gene-Se risk for CRC by sex, as indicated by our previous studies of selenoprotein genetic variation in a Czech population [19] and Se status in the EPIC study [5]. This reflects the

importance of interactions between Se intake, Se status and genotype, sex and CRC sub-type risks (reviewed in [4]).

Prior to this study there were few data available on the interaction of selenoprotein genotype and Se status regarding CRC risk, apart from a study in a Se replete population of North American women which reported that the null results for serum Se did not differ by selenoenzyme (*GPX1-4* and *SELENOP*) genetic variants [17]. The effect of Se pathway SNPs on the efficacy of Se utilization may be particularly relevant to CRC risk in populations with sub-optimal Se status, such as this study within EPIC [5]. We observed that numerous genetic variations were associated with Se status levels (as assessed by serum Se and *SELENOP* concentration), although these were not significant after adjustment for multiple tests. However, in the PIGE analysis overall gene and pathway genetic variation interacted with biomarkers of Se status to alter CRC risk. As expected, there were several variations in pathway 1 nominally associated with Se status levels. These included selenoprotein genes *SELENON*, *SELENOP*, *SELENOS*, and *TXNRD1* that are regulated by Se availability and whose genetic variations have been previously shown to affect blood and tissue Se levels [6,21,42,43]. Transgenic mouse studies underlined the critical function of *SELENOP* for Se organification and transport [11]. Thus, SNP interactions with *SELENOP* levels may be particularly important regarding CRC risk as serum *SELENOP* is a functional marker of Se status and is more associated with CRC risk than Se in this cohort [5]. Notably, all 4 of the variants associated with both *SELENOP* status and CRC risk in the SNP only analysis were from 2 selenoprotein genes; rs4659382, rs11247710, and rs2072749 in *SELENON* and the  $P_{ACT}$  borderline significant rs11111979 variant in *TXNRD1*. Selenoproteins *SELENOP*, *SELENON*, and *TXNRD1* are antioxidant enzymes and their genetic variations plus regulation by *SELENOP* levels may be important factors in relation to colorectal carcinogenesis. We observed an association of *SELENOP* levels with rs6413428 in *SELENOP*, which in a SNP-only analysis was previously observed to be associated with CRC risk in the USA [44], an area of generally high Se status, but not in our study. Another *SELENOP* variant, rs3877899, was also associated with *SELENOP* status, while the GG genotype for this SNP previously showed the highest significant correlation of all selenoprotein genotypes tested between serum Se and activity of the vital antioxidant enzyme thioredoxin reductase [45]. This latter study also showed a correlation between serum Se and increased DNA damage with *SELENOS*-rs4965373 under peroxide challenge. We selected rs12910524 in *SELENOS* as a proxy tagSNP for this variant (as rs4965373 failed *Illumina* assay development) and found that it was significantly associated with Se levels. The lipoprotein megalin receptor (LRP2) protein appears to mediate *SELENOP* uptake to various tissues and affect plasma Se status levels [46,47]. Here, the *LRP2* SNPs rs12614394, rs2229266, rs2389557, rs700552, and rs9789747 were associated with CRC risk alone while the rs3755166 promoter SNP was associated with Se levels. Previously, rs3755166 has been associated with Alzheimer's disease with the rare allele showing decreased transcriptional activity [48]. Intriguingly, this indicates a potential mechanism for the suggested link of sub-optimal Se status with neurodegenerative disease [49].

Supporting the data presented here, *GPX1* and *GPX4* selenoprotein gene loci have been implicated in GWAS of inflammatory bowel disease (IBD), which is a risk factor for CRC development [50,51]. Additionally, the rs7901303 variant from the selenophosphate synthetase 1 (*SEPHS1*) gene, which plays a major role in selenoprotein synthesis, was associated in this study with CRC risk (before multiple testing corrections). rs7901303 was previously associated with risk of Crohn's disease in interaction with serum Se levels in a sub-optimal Se population of New Zealand [52]. The genetic associations identified in these studies suggest therefore a key role of the corresponding proteins in colorectal function and/or the carcinogenic process.

Genomic studies and animal models have shown Se intake to not only affect expression of selenoprotein genes but also pathways key to colorectal carcinogenesis such as the antioxidant response, immune and inflammatory pathways (including NFkB and Nrf2 signaling) and the Wnt signaling pathway [4,13,53]. Furthermore, expression of constituents of these metabolic pathways has been shown to be affected by Se level in human rectal biopsies [54]. Therefore, in addition to a focus on Se metabolism and selenoprotein genes, the present analysis also encompassed a substantial

examination of genetic variations in these selenium-relevant pathways. Associations of multiple SNPs in the same gene (several of which are novel) with CRC risk were observed in genes from pathways 2–8, e.g., *BAX*, *GPX5*, *FOXO3*, *IL12B*, *TLR4*, *MSH2*, *MSH3*, *TGFB1*, as well as *IL10* with rectal cancer risk. Polymorphisms in several of these genes have previously been associated with CRC risk [55–62]. After BH multiple testing corrections, SNPs in cell-signaling pathways retained significance (*FRZB*, *SMAD3*, and *SMAD7*). The variants in both *SMAD* genes were previously linked to CRC by GWAS, suggesting a role of these variants in CRC development [38]. The association of rs17265803 in *FRZB* appears to be novel and it is not in LD with the functional *FRZB* genetic variant Arg324Gly (rs7775) previously reported to be associated with an increased CRC risk [63], although this was not replicated in a nested case-control study [64]. Additional genetic variants retaining significance by  $P_{ACT}$  were rs6983267 in *C-MYC* previously identified in a meta-GWAS [65], rs235770 in *BMP2* previously associated with colon cancer risk [37], and a novel association of rs4645887 in *BAX*.

In the pathway analysis, all the Se pathway genes were grouped into a primary best-known functional pathway and were analyzed for the association of whole gene and whole pathway genetic variation with CRC risk, and in interaction with Se status. Neither gene only variation or interactions with Se status in the core pathway 1 selenoprotein and biosynthesis pathway were associated with CRC risk by pathway, although gene only variation for *GPX1*, *SELENOM*, *SELENON*, and *SEPHS1* plus gene x Se status interactions for *SELENON* (with *SELENOP*) and *PSTK* (with Se) were associated with CRC risk. In the gene x Se analysis, only pathway 2 (antioxidant/redox) was significant for an association with CRC risk for both Se and *SELENOP*. Alternatively, it also remains possible that the genetic ‘noise’ from any irrelevant selenoproteins masked the overall risk associations for pathway 1 (based on the rationale that most genes in pathways like oxidative stress are important in cancer prevention but that some of the selenoproteins may be irrelevant to colorectal carcinogenesis, as they are included solely because they share selenocysteine motifs). This is partly supported by the strongest association (by PIGE) in pathway 1 for gene variance in *GPX1*, which has been previously implicated in risk of various cancers [4]. However, these pathway divisions cannot reflect, for example, the biological overlaps with the antioxidant selenoprotein genes in pathway 1 and their non-Se containing counterparts in pathway 2. Several aspects of our data suggest a potentially under-appreciated focus on variation in apoptosis genes (pathway 4) and CRC risk that may also be modified by Se status. These comprise the association of several SNPs in both the *FOXO3* and *BAX* genes (including the  $P_{ACT}$  significant rs4645887 variant in *BAX*) with CRC risk, significance of *FOXO3* for overall gene variation, and several significant findings of SNP x Se and gene x Se status interactions for genes in this pathway (e.g., *SMAC*, *CASP8*, *MAPK8*, and *MAPK9*). Overall, our analyses suggest that genetic variation in TGF beta signaling (pathway 7), which includes members such as *BMP2*, *BMPR2*, *SMAD3*, and *SMAD7* implicated in CRC risk by previous large case-control and GWAS reports [36–38], is sufficient to alter CRC risk, independent of Se status interaction, while SNP risk associations attributed to the antioxidant and apoptosis pathways may be significantly modified by Se status interactions.

Strengths and weaknesses of our study design for the Se status analyses have been discussed earlier [5]. The hypothesis-driven approach and appreciable sample size within a large, prospective study allowed an extensive examination of Se pathway genetic variation (including gene pathway analyses) and the interaction with robust markers of Se status regarding CRC risk. Despite the large sample size, gene pathway, gene–Se interaction analysis and some stratified analyses had limited power, particularly analyses by sex and anatomical sub-sites. The pathway designations were assigned based on known function from the literature, and there will be interactions between these pathways that we were not able to model. Finally, as most of the reported associations involve tagSNPs of no known functionality (or the actual contributing functional variant(s) they tag) additional genetic mapping and lab-based studies will be needed to explore these aspects.

## 5. Conclusions

In summary, the present study indicates that genetic variation in selenoprotein genes and genes in antioxidant/redox, Wnt, apoptotic, and TGF-beta signaling pathways may modify risk of CRC development. Furthermore, for genes in antioxidant/redox and apoptotic pathways the influence of SNPs on the disease risk is also dependent on interaction with Se status. Overall, these results taken together with our previous study [5] suggest that risk of CRC may be modified by genotype, Se status, sex, and gene variation interactions within biological pathways. Thus, will individuals harboring these genotypes benefit from increased Se intake, including consideration of ‘Se adequate’ environments, such as the US, where Se intervention trials have not shown a significant benefit in the general population [66]? Before such a recommendation can be defined, further examination of these findings in other populations and investigation of Se metabolism is needed to clarify the relevance of the Se pathway and signaling genotypes for CRC etio-pathogenesis, especially for individuals with suboptimal Se status.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-6643/11/4/935/s1>: Figure S1: Selenium pathway genetic analysis flowchart, Table S1: Se pathway SNP study: List of genes and the 8 pathway designations, the EPIC study, 1992–2003, Table S2: All Se pathway SNPs by gene and pathway and association with CRC risk, the EPIC study, 1992–2003, Table S3: All Se pathway SNPs (by gene and pathway) significantly associated with CRC risk before multiple testing corrections, the EPIC study, 1992–2003, Table S4: Se pathway SNPs from pathways 1–8 associated with Se or SELENOP Concentrations Among Controls, the EPIC study, 1992–2003, Table S5: Se pathway PIGE P values for each gene and CRC risk association per pathway designation, the EPIC study, 1992–2003.

**Author Contributions:** Conceptualization, D.J.H., V.F., C.M., J.E.H., L.S., M.J. and E.R.; Data curation, D.J.H., J.E.H., V.F., L.S. and M.J.; Formal analysis, V.F., D.J.H., J.S.J., W.Z., C.M. and M.J.; Funding acquisition, D.J.H., V.F., J.E.H., C.M., L.S. and M.J. Investigation, D.J.H., V.F., C.M., J.E.H., L.S., S.H. (Sandra Hybsier), A.S. and M.J.; Methodology, V.F., D.J.H., J.E.H., C.M., M.J., L.S. and E.R.; Project administration, D.J.H., V.F., C.M., J.E.H. and M.J. Resources, D.J.H., V.F., C.M., J.E.H., M.J., and all other EPIC co-authors (K.O., A.T. (Anne Tjønneland), H.O., V.P., M.-C.B.-R., T.K., V.K., K.A., A.T., A.K. (Anna Karakatsani), A.K. (Anastasia Kotanidou), R.T., S.P., G.M., C.A., A.N., B.B.-d.-M., R.C.H.V., E.W., G.S., T.H.N., L.L.-B., J.R.Q., J.M.H., M.R.-B., A.B., B.G., S.H. (Sophia Harlid), K.E.B., N.W., K.-T.K., M.G., N.M., H.F., K.T., D.A., and E.R.); Software, V.F., D.J.H., A.S., J.S.J., W.Z., L.S. and M.J.; Supervision, D.J.H., V.F., L.S. and M.J.; Validation, D.J.H., V.F., C.M., A.S., S.H. (Sandra Hybsier) and L.S.; Writing—original draft, D.J.H., V.F., M.J., C.M., and J.E.H.; Writing—review and editing, all other EPIC co-authors (K.O., A.T. (Anne Tjønneland), H.O., V.P., M.-C.B.-R., T.K., V.K., K.A., A.T. (Antonia Trichopoulou), A.K. (Anna Karakatsani), A.K. (Anastasia Kotanidou), R.T., S.P., G.M., C.A., A.N., B.B.-d.-M., R.C.H.V., E.W., G.S., T.H.N., L.L.-B., J.R.Q., J.M.H., M.R.-B., A.B., B.G., S.H. (Sophia Harlid), K.E.B., N.W., K.-T.K., M.G., N.M., H.F., K.T., D.A. and E.R.) reviewed and approved the manuscript and commented on the analysis and interpretation of the findings. The lead authors (D.J.H., V.F., C.M., J.E.H. and M.J.) affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

**Funding:** Funding for this study was provided by the Health Research Board of Ireland health research awards HRA\_PHS/2013/397 and HRA\_PHS/2015/1142 (principal investigator: D.J.H.). We thank all the members of EPIC for their work on the EPIC cohort study. The EPIC study was supported by “Europe Against Cancer” Programme of the European Commission (SANCO); Ligue contre le Cancer; Institut Gustave Roussy; Mutuelle Générale de l’Education Nationale; Institut National de la Santé et de la Recherche Médicale (INSERM); German Cancer Aid; German Cancer Research Center; German Federal Ministry of Education and Research; Danish Cancer Society; Health Research Fund (FIS) of the Spanish Ministry of Health; the CIBER en Epidemiología y Salud Pública (CIBERESP), Spain; ISCIII RETIC (RD06/0020); Spanish Regional Governments of Andalusia, Asturias, Basque Country, Murcia (No 6236) and Navarra and the Catalan Institute of Oncology; Cancer Research UK; Medical Research Council, UK; the Hellenic Health Foundation; Italian Association for Research on Cancer; Italian National Research Council; Compagnia di San Paolo; Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (ZorgOnderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); Swedish Cancer Society; Swedish Scientific Council; Regional Governments of Skane and Vasterbotten, Sweden; and Nordforsk center of excellence programme HELGA. L.S. was supported by Deutsche Forschungsgemeinschaft (DFG Research Unit 2558 TraceAge, Scho 849/6-1). D.P. was supported by the Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy.

**Conflicts of Interest:** The authors declare no competing interests. Funding support for the EPIC study is described in the acknowledgements; there were no financial relationships with any organizations that might have an interest in the submitted work in the previous three years, and no other relationships or activities that could appear to have influenced the submitted work. While L.S. is the founder of selenOmed GmbH, a



company involved in improving Se diagnostics, the data were analyzed before the founding of selenOmed and the analyses were conducted completely blinded to any clinical finding. For information on how to apply for gaining access to EPIC data and/or biospecimens, please follow the instructions at <http://epic.iarc.fr/access/index.php>. **IARC Disclosure Statement:** Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

## References

1. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D.M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in globocan 2012. *Int. J. Cancer* **2015**, *136*, E359–E386. [[CrossRef](#)] [[PubMed](#)]
2. Cappellani, A.; Zanghi, A.; Di Vita, M.; Cavallaro, A.; Piccolo, G.; Veroux, P.; Lo Menzo, E.; Cavallaro, V.; de Paoli, P.; Veroux, M.; et al. Strong correlation between diet and development of colorectal cancer. *Front. Biosci.* **2013**, *18*, 190–198.
3. Slattery, M.L.; Lundgreen, A.; Herrick, J.S.; Caan, B.J.; Potter, J.D.; Wolff, R.K. Diet and colorectal cancer: Analysis of a candidate pathway using SNPs, haplotypes, and multi-gene assessment. *Nutr. Cancer* **2011**, *63*, 1226–1234. [[CrossRef](#)] [[PubMed](#)]
4. Meplan, C.; Hesketh, J. Selenium and Cancer: A Story that Should not be Forgotten—Insights from Genomics. *Cancer Treat. Res.* **2014**, *159*, 145–166. [[PubMed](#)]
5. Hughes, D.J.; Fedirko, V.; Jenab, M.; Schomburg, L.; Meplan, C.; Freisling, H.; Bueno-de-Mesquita, H.B.; Hybsier, S.; Becker, N.P.; Czuban, M.; et al. Selenium status is associated with colorectal cancer risk in the European prospective investigation of cancer and nutrition cohort. *Int. J. Cancer* **2015**, *136*, 1149–1161. [[CrossRef](#)] [[PubMed](#)]
6. Combs, G.F., Jr. Biomarkers of selenium status. *Nutrients* **2015**, *7*, 2209–2236. [[CrossRef](#)]
7. Fairweather-Tait, S.J.; Bao, Y.; Broadley, M.R.; Collings, R.; Ford, D.; Hesketh, J.E.; Hurst, R. Selenium in human health and disease. *Antioxid. Redox Signal.* **2011**, *14*, 1337–1383. [[CrossRef](#)] [[PubMed](#)]
8. Labunskyy, V.M.; Hatfield, D.L.; Gladyshev, V.N. Selenoproteins: Molecular pathways and physiological roles. *Physiol. Rev.* **2014**, *94*, 739–777. [[CrossRef](#)]
9. Gladyshev, V.N.; Arnér, E.S.; Berry, M.J.; Brigelius-Flohé, R.; Bruford, E.A.; Burk, R.F.; Carlson, B.A.; Castellano, S.; Chavatte, L.; Conrad, M.; et al. Selenoprotein Gene Nomenclature. *J. Biol. Chem.* **2016**, *291*, 24036–24040.
10. Hatfield, D.L.; Tsuji, P.A.; Carlson, B.A.; Gladyshev, V.N. Selenium and selenocysteine: Roles in cancer, health, and development. *Trends Biochem. Sci.* **2014**, *39*, 112–120. [[CrossRef](#)] [[PubMed](#)]
11. Burk, R.F.; Hill, K.E. Regulation of Selenium Metabolism and Transport. *Annu. Rev. Nutr.* **2015**, *35*, 109–134. [[CrossRef](#)] [[PubMed](#)]
12. Hesketh, J.; Meplan, C. Transcriptomics and functional genetic polymorphisms as biomarkers of micronutrient function: Focus on selenium as an exemplar. *Proc. Nutr. Soc.* **2011**, *70*, 365–373. [[CrossRef](#)]
13. Meplan, C.; Hesketh, J. The influence of selenium and selenoprotein gene variants on colorectal cancer risk. *Mutagenesis* **2012**, *27*, 177–186. [[CrossRef](#)] [[PubMed](#)]
14. Irons, R.; Carlson, B.A.; Hatfield, D.L.; Davis, C.D. Both selenoproteins and low molecular weight selenocompounds reduce colon cancer risk in mice with genetically impaired selenoprotein expression. *J. Nutr.* **2006**, *136*, 1311–1317. [[CrossRef](#)]
15. Irons, R.; Tsuji, P.A.; Carlson, B.A.; Ouyang, P.; Yoo, M.H.; Xu, X.M.; Hatfield, D.L.; Gladyshev, V.N.; Davis, C.D. Deficiency in the 15-kDa selenoprotein inhibits tumorigenicity and metastasis of colon cancer cells. *Cancer Prev. Res.* **2010**, *3*, 630–639. [[CrossRef](#)]
16. Hu, Y.; McIntosh, G.H.; Le Leu, R.K.; Young, G.P. Selenium-enriched milk proteins and selenium yeast affect selenoprotein activity and expression differently in mouse colon. *Br. J. Nutr.* **2010**, *104*, 17–23. [[CrossRef](#)] [[PubMed](#)]
17. Takata, Y.; Kristal, A.R.; King, I.B.; Song, X.; Diamond, A.M.; Foster, C.B.; Hutter, C.M.; Hsu, L.; Duggan, D.J.; Langer, R.D.; et al. Serum selenium, genetic variation in selenoenzymes, and risk of colorectal cancer: Primary analysis from the Women’s Health Initiative Observational Study and meta-analysis. *Cancer Epidemiol. Biomark. Prev.* **2011**, *20*, 1822–1830. [[CrossRef](#)] [[PubMed](#)]

18. Steinbrenner, H.; Speckmann, B.; Sies, H. Toward understanding success and failures in the use of selenium for cancer prevention. *Antioxid. Redox Signal.* **2013**, *19*, 181–191. [[CrossRef](#)] [[PubMed](#)]
19. Meplan, C.; Hughes, D.J.; Pardini, B.; Naccarati, A.; Soucek, P.; Vodickova, L.; Hlavata, I.; Vrana, D.; Vodicka, P.; Hesketh, J.E.; et al. Genetic variants in selenoprotein genes increase risk of colorectal cancer. *Carcinogenesis* **2010**, *31*, 1074–1079. [[CrossRef](#)] [[PubMed](#)]
20. Sutherland, A.; Kim, D.H.; Relton, C.; Ahn, Y.O.; Hesketh, J. Polymorphisms in the selenoprotein S and 15-kDa selenoprotein genes are associated with altered susceptibility to colorectal cancer. *Genes Nutr.* **2010**, *5*, 215–223. [[CrossRef](#)]
21. Meplan, C. Selenium and chronic diseases: A nutritional genomics perspective. *Nutrients* **2015**, *7*, 3621–3651. [[CrossRef](#)]
22. Slattery, M.L.; Lundgreen, A.; Welbourn, B.; Corcoran, C.; Wolff, R.K. Genetic variation in selenoprotein genes, lifestyle, and risk of colon and rectal cancer. *PLoS ONE* **2012**, *7*, e37312. [[CrossRef](#)]
23. Haug, U.; Poole, E.M.; Xiao, L.; Curtin, K.; Duggan, D.; Hsu, L.; Makar, K.W.; Peters, U.; Kulmacz, R.J.; Potter, J.D.; et al. Glutathione peroxidase tagSNPs: Associations with rectal cancer but not with colon cancer. *Genes Chromosomes Cancer* **2012**, *51*, 598–605. [[CrossRef](#)]
24. Riboli, E.; Kaaks, R. The EPIC Project: Rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int. J. Epidemiol.* **1997**, *26*, S6–S14. [[CrossRef](#)]
25. Riboli, E.; Hunt, K.J.; Slimani, N.; Ferrari, P.; Norat, T.; Fahey, M.; Charrondière, U.R.; Hémon, B.; Casagrande, C.; Vignat, J.; et al. European Prospective Investigation into Cancer and Nutrition (EPIC): Study populations and data collection. *Public Health Nutr.* **2002**, *5*, 1113–1124. [[CrossRef](#)]
26. Caboux, E.; Lallemand, C.; Ferro, G.; Hémon, B.; Mendy, M.; Biessy, C.; Sims, M.; Wareham, N.; Britten, A.; Boland, A.; et al. Sources of pre-analytical variations in yield of DNA extracted from blood samples: Analysis of 50,000 DNA samples in EPIC. *PLoS ONE* **2012**, *7*, e39821. [[CrossRef](#)] [[PubMed](#)]
27. Hutter, C.M.; Chang-Claude, J.; Slattery, M.L.; Pflugeisen, B.M.; Lin, Y.; Duggan, D.; Nan, H.; Lemire, M.; Rangrej, J.; Figueiredo, J.C.; et al. Characterization of gene-environment interactions for colorectal cancer susceptibility loci. *Cancer Res.* **2012**, *72*, 2036–2044. [[CrossRef](#)]
28. Meplan, C.; Rohrmann, S.; Steinbrecher, A.; Schomburg, L.; Jansen, E.; Linseisen, J.; Hesketh, J. Polymorphisms in thioredoxin reductase and selenoprotein K genes and selenium status modulate risk of prostate cancer. *PLoS ONE* **2012**, *7*, e48709. [[CrossRef](#)]
29. Hybsier, S.; Schulz, T.; Wu, Z.; Demuth, I.; Minich, W.B.; Renko, K.; Rijntjes, E.; Köhrle, J.; Strasburger, C.J.; Steinhagen-Thiessen, E.; et al. Sex-specific and inter-individual differences in biomarkers of selenium status identified by a calibrated ELISA for selenoprotein P. *Redox Biol.* **2017**, *11*, 403–414. [[CrossRef](#)]
30. Clarke, G.M.; Anderson, C.A.; Pettersson, F.H.; Cardon, L.R.; Morris, A.P.; Zondervan, K.T. Basic statistical analysis in genetic case-control studies. *Nat. Protoc.* **2011**, *6*, 121–133. [[CrossRef](#)]
31. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **1995**, *57*, 289–300. [[CrossRef](#)]
32. Conneely, K.N.; Boehnke, M. So many correlated tests, so little time! Rapid adjustment of P values for multiple correlated tests. *Am. J. Hum. Genet.* **2007**, *81*, 1158–1168. [[CrossRef](#)] [[PubMed](#)]
33. Yu, K.; Li, Q.; Bergen, A.W.; Pfeiffer, R.M.; Rosenberg, P.S.; Caporaso, N.; Kraft, P.; Chatterjee, N. Pathway analysis by adaptive combination of P-values. *Genet. Epidemiol.* **2009**, *33*, 700–709. [[CrossRef](#)] [[PubMed](#)]
34. Zhang, H.; Shi, J.; Liang, F.; Wheeler, W.; Stolzenberg-Solomon, R.; Yu, K. A fast multilocus test with adaptive SNP selection for large-scale genetic-association studies. *Eur. J. Hum. Genet.* **2014**, *22*, 696–702. [[CrossRef](#)]
35. Dato, S.; De Rango, F. Antioxidants and Quality of Aging: Further Evidences for a Major Role of TXNRD1 Gene Variability on Physical Performance at Old Age. *Oxid. Med. Cell Longev.* **2015**, *2015*, 926067. [[CrossRef](#)]
36. Peters, U.; Hutter, C.M.; Hsu, L.; Schumacher, F.R.; Conti, D.V.; Carlson, C.S.; Edlund, C.K.; Haile, R.W.; Gallinger, S.; Zanke, B.W.; et al. Meta-analysis of new genome-wide association studies of colorectal cancer risk. *Hum. Genet.* **2012**, *131*, 217–234. [[CrossRef](#)]
37. Slattery, M.L.; Lundgreen, A.; Herrick, J.S.; Kadlubar, S.; Caan, B.J.; Potter, J.D.; Wolff, R.K. Genetic variation in bone morphogenetic protein and colon and rectal cancer. *Int. J. Cancer* **2012**, *130*, 653–664. [[CrossRef](#)] [[PubMed](#)]
38. Peters, U.; Bien, S.; Zubair, N. Genetic architecture of colorectal cancer. *Gut* **2015**, *64*, 1623–1636. [[CrossRef](#)] [[PubMed](#)]

39. Peters, K.M.; Carlson, B.A.; Gladyshev, V.N.; Tsuji, P.A. Selenoproteins in colon cancer. *Free Radic. Biol. Med.* **2018**, *127*, 14–25. [[CrossRef](#)]
40. Pitts, M.W.; Hoffmann, P.R. Endoplasmic reticulum-resident selenoproteins as regulators of calcium signaling and homeostasis. *Cell Calcium* **2018**, *70*, 76–86. [[CrossRef](#)]
41. Hagggar, F.A.; Boushey, R.P. Colorectal cancer epidemiology: Incidence, mortality, survival, and risk factors. *Clin. Colon Rectal Surg.* **2009**, *22*, 191–197. [[CrossRef](#)] [[PubMed](#)]
42. Ekoue, D.N.; Zaichick, S.; Valyi-Nagy, K.; Picklo, M.; Lacher, C.; Hoskins, K.; Warso, M.A.; Bonini, M.G.; Diamond, A.M. Selenium levels in human breast carcinoma tissue are associated with a common polymorphism in the gene for SELENOP (Selenoprotein P). *J. Trace Elem. Med. Biol.* **2017**, *39*, 227–233. [[CrossRef](#)]
43. Evans, D.M.; Zhu, G.; Dy, V.; Heath, A.C.; Madden, P.A.; Kemp, J.P.; McMahon, G.; St Pourcain, B.; Timpson, N.J.; Golding, J.; et al. Genome-wide association study identifies loci affecting blood copper, selenium and zinc. *Hum. Mol. Genet.* **2013**, *22*, 3998–4006. [[CrossRef](#)]
44. Hazra, A.; Chanock, S.; Giovannucci, E.; Cox, D.G.; Niu, T.; Fuchs, C.; Willett, W.C.; Hunter, D.J. Large-scale evaluation of genetic variants in candidate genes for colorectal cancer risk in the Nurses' Health Study and the Health Professionals' Follow-up Study. *Cancer Epidemiol. Biomark. Prev.* **2008**, *17*, 311–319. [[CrossRef](#)] [[PubMed](#)]
45. Karunasinghe, N.; Han, D.Y.; Zhu, S.; Yu, J.; Lange, K.; Duan, H.; Medhora, R.; Singh, N.; Kan, J.; Alzaher, W.; et al. Serum selenium and single-nucleotide polymorphisms in genes for selenoproteins: Relationship to markers of oxidative stress in men from Auckland, New Zealand. *Genes Nutr.* **2012**, *7*, 179–190. [[CrossRef](#)] [[PubMed](#)]
46. Chiu-Ugalde, J.; Theilig, F.; Behrends, T.; Drebes, J.; Sieland, C.; Subbarayal, P.; Köhrle, J.; Hammes, A.; Schomburg, L.; Schweizer, U. Mutation of megalin leads to urinary loss of selenoprotein P and selenium deficiency in serum, liver, kidneys and brain. *Biochem. J.* **2010**, *431*, 103–111. [[CrossRef](#)] [[PubMed](#)]
47. Krol, M.B.; Gromadzinska, J.; Wasowicz, W. SeP, ApoER2 and megalin as necessary factors to maintain Se homeostasis in mammals. *J. Trace Elem. Med. Biol.* **2012**, *26*, 262–266. [[CrossRef](#)]
48. Vargas, T.; Bullido, M.J.; Martinez-Garcia, A.; Antequera, D.; Clarimon, J.; Rosich-Estrago, M.; Martin-Requero, A.; Mateo, I.; Rodriguez-Rodriguez, E.; Vilella-Cuadrada, E.; et al. A megalin polymorphism associated with promoter activity and Alzheimer's disease risk. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **2010**, *153B*, 895–902. [[CrossRef](#)]
49. Burk, R.F.; Hill, K.E.; Motley, A.K.; Winfrey, V.P.; Kurokawa, S.; Mitchell, S.L.; Zhang, W. Selenoprotein P and apolipoprotein E receptor-2 interact at the blood-brain barrier and also within the brain to maintain an essential selenium pool that protects against neurodegeneration. *FASEB J.* **2014**, *28*, 3579–3588. [[CrossRef](#)]
50. Franke, A.; McGovern, D.P.; Barrett, J.C.; Wang, K.; Radford-Smith, G.L.; Ahmad, T.; Lees, C.W.; Balschun, T.; Lee, J.; Roberts, R.; et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat. Genet.* **2010**, *42*, 1118–1125. [[CrossRef](#)]
51. Jostins, L.; Ripke, S.; Weersma, R.K.; Duerr, R.H.; McGovern, D.P.; Hui, K.Y.; Lee, J.C.; Schumm, L.P.; Sharma, Y.; Anderson, C.A.; et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **2012**, *491*, 119–124. [[CrossRef](#)] [[PubMed](#)]
52. Gentschew, L.; Bishop, K.S.; Han, D.Y.; Morgan, A.R.; Fraser, A.G.; Lam, W.J.; Karunasinghe, N.; Campbell, B.; Ferguson, L.R. Selenium, selenoprotein genes and Crohn's disease in a case-control population from Auckland, New Zealand. *Nutrients* **2012**, *4*, 1247–1259. [[CrossRef](#)] [[PubMed](#)]
53. Kipp, A.; Banning, A.; van Schothorst, E.M.; Méplan, C.; Schomburg, L.; Evelo, C.; Coort, S.; Gaj, S.; Keijer, J.; Hesketh, J.; et al. Four selenoproteins, protein biosynthesis, and Wnt signalling are particularly sensitive to limited selenium intake in mouse colon. *Mol. Nutr. Food Res.* **2009**, *53*, 1561–1572. [[CrossRef](#)] [[PubMed](#)]
54. Meplan, C.; Johnson, I.T.; Polley, A.C.; Cockell, S.; Bradburn, D.M.; Commane, D.M.; Arasaradnam, R.P.; Mulholland, F.; Zupanic, A.; Mathers, J.C.; et al. Transcriptomics and proteomics show that selenium affects inflammation, cytoskeleton, and cancer pathways in human rectal biopsies. *FASEB J.* **2016**, *30*, 2812–2825. [[CrossRef](#)] [[PubMed](#)]
55. Sun, R.; Jia, F.; Liang, Y.; Li, L.; Bai, P.; Yuan, F.; Gao, L.; Zhang, L. Interaction analysis of IL-12A and IL-12B polymorphisms with the risk of colorectal cancer. *Tumour Biol.* **2015**, *36*, 9295–9301. [[CrossRef](#)]
56. Sheng, W.Y.; Yong, Z.; Yun, Z.; Hong, H.; Hai, L.L. Toll-like receptor 4 gene polymorphisms and susceptibility to colorectal cancer: A meta-analysis and review. *Arch. Med. Sci.* **2015**, *11*, 699–707. [[CrossRef](#)] [[PubMed](#)]

57. Wang, P.; An, J.; Zhu, Y.; Wan, X.; Zhang, H.; Xi, S.; Li, S. Association of three promoter polymorphisms in interleukin-10 gene with cancer susceptibility in the Chinese population: A meta-analysis. *Oncotarget* **2017**, *8*, 62382–62399. [[CrossRef](#)] [[PubMed](#)]
58. Tsilidis, K.K.; Helzlsouer, K.J.; Smith, M.W.; Grinberg, V.; Hoffman-Bolton, J.; Clipp, S.L.; Visvanathan, K.; Platz, E.A. Association of common polymorphisms in IL10, and in other genes related to inflammatory response and obesity with colorectal cancer. *Cancer Causes Control* **2009**, *20*, 1739–1751. [[CrossRef](#)] [[PubMed](#)]
59. Miao, H.K.; Chen, L.P.; Cai, D.P.; Kong, W.J.; Xiao, L.; Lin, J. MSH3 rs26279 polymorphism increases cancer risk: A meta-analysis. *Int. J. Clin. Exp. Pathol.* **2015**, *8*, 11060–11067. [[PubMed](#)]
60. Naccarati, A.; Pardini, B.; Hemminki, K.; Vodicka, P. Sporadic colorectal cancer and individual susceptibility: A review of the association studies investigating the role of DNA repair genetic polymorphisms. *Mutat. Res.* **2007**, *635*, 118–145. [[CrossRef](#)]
61. Zhang, B.; Jia, W.H.; Matsuda, K. Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. *Nat. Genet.* **2014**, *46*, 533–542. [[CrossRef](#)] [[PubMed](#)]
62. Wang, Y.; Yang, H.; Li, L.; Xia, X. An updated meta-analysis on the association of TGF-beta1 gene promoter -509C/T polymorphism with colorectal cancer risk. *Cytokine* **2013**, *61*, 181–187. [[CrossRef](#)] [[PubMed](#)]
63. Shanmugam, K.S.; Brenner, H.; Hoffmeister, M.; Chang-Claude, J.; Burwinkel, B. The functional genetic variant Arg324Gly of frizzled-related protein is associated with colorectal cancer risk. *Carcinogenesis* **2007**, *28*, 1914–1917. [[CrossRef](#)] [[PubMed](#)]
64. Berndt, S.I.; Huang, W.Y.; Yeager, M.; Weissfeld, J.L.; Chanock, S.J.; Hayes, R.B. Genetic variants in frizzled-related protein (FRZB) and the risk of colorectal neoplasia. *Cancer Causes Control* **2009**, *20*, 487–490. [[CrossRef](#)]
65. Hong, Y.; Wu, G.; Li, W.; Liu, D.; He, K. A comprehensive meta-analysis of genetic associations between five key SNPs and colorectal cancer risk. *Oncotarget* **2016**, *7*, 73945–73959. [[CrossRef](#)]
66. Lippman, S.M.; Klein, E.A.; Goodman, P.J.; Lucia, M.S.; Thompson, I.M.; Ford, L.G.; Parnes, H.L.; Minasian, L.M.; Gaziano, J.M.; Hartline, J.A.; et al. Effect of selenium and vitamin E on risk of prostate cancer and other cancers: The Selenium and Vitamin E Cancer Prevention Trial (SELECT). *JAMA* **2009**, *301*, 39–51. [[CrossRef](#)] [[PubMed](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).