Prediagnostic selenium status, selenoprotein gene variants and association with breast cancer risk in a European cohort study

David J. Hughes a,*, Lutz Schomburg b, Mazda Jenab c, Carine Biessy c, Catherine Méplan d, Aurelie Moskal e, f, Qian Sun b, Kamil Demircan b, Veronika Fedirko f, Elisabete Weiderpass c, Maryam Mukhtar a, Anja Olsen g, h, Anne Tjonneland d, i, Kim Overvad j, k, Matthias Schulze l, Therese Haugdahl Nøst m, Guri Skeie m, Karina Standahl Olsen m, Fulvio Ricceri b, o, Sara Grioni p, Domenico Palli q, Giovanna Masala q, Rosario Tumino f, Fabrizio Pasanisi q, Pilar Amiano t, u, v, Sandra M. Colorado Yohar w, x, y, Antonio Agudo z, a, Maria-Jose Sánchez z, a, b, ac, ad, Eva Ardanaz z, a, c, Malin Sund g, a, Anne Andersson a, Aurora Perez-Cornago aj, Ruth Travis aj, Alicia K. Heath ak, Laure Dossus c

a Cancer Biology and Therapeutics Group, School of Biomolecular and Biomedical Science, UCD Conway Institute, University College Dublin, Dublin, Ireland
b Institute for Experimental Endocrinology, Charité – Medical University, Berlin, Germany
c School of Biomedical, Nutritional and Sport Sciences, Newcastle University, Newcastle Upon Tyne, UK
d Research on Healthcare Performance (RESHAPE), INSERM U1290, Université Claude Bernard Lyon 1, Lyon, France
e Department of Epidemiology, MD Anderson Cancer Centre, Houston, TX, USA
f Danish Cancer Society Research Center, Copenhagen, Denmark

Contents lists available at ScienceDirect

Prediagnostic selenium status, selenoprotein gene variants and association with breast cancer risk in a European cohort study

David J. Hughes a,*, Lutz Schomburg b, Mazda Jenab c, Carine Biessy c, Catherine Méplan d, Aurelie Moskal e, f, Qian Sun b, Kamil Demircan b, Veronika Fedirko f, Elisabete Weiderpass c, Maryam Mukhtar a, Anja Olsen g, h, Anne Tjonneland d, i, Kim Overvad j, k, Matthias Schulze l, Therese Haugdahl Nøst m, Guri Skeie m, Karina Standahl Olsen m, Fulvio Ricceri b, o, Sara Grioni p, Domenico Palli q, Giovanna Masala q, Rosario Tumino f, Fabrizio Pasanisi q, Pilar Amiano t, u, v, Sandra M. Colorado Yohar w, x, y, Antonio Agudo z, a, Maria-Jose Sánchez z, a, b, ac, ad, Eva Ardanaz z, a, c, Malin Sund g, a, Anne Andersson a, Aurora Perez-Cornago aj, Ruth Travis aj, Alicia K. Heath ak, Laure Dossus c

a Cancer Biology and Therapeutics Group, School of Biomolecular and Biomedical Science, UCD Conway Institute, University College Dublin, Dublin, Ireland
b Institute for Experimental Endocrinology, Charité – Medical University, Berlin, Germany
c School of Biomedical, Nutritional and Sport Sciences, Newcastle University, Newcastle Upon Tyne, UK
d Research on Healthcare Performance (RESHAPE), INSERM U1290, Université Claude Bernard Lyon 1, Lyon, France
e Department of Epidemiology, MD Anderson Cancer Centre, Houston, TX, USA
f Danish Cancer Society Research Center, Copenhagen, Denmark

Contents lists available at ScienceDirect

Free Radical Biology and Medicine 209 (2023) 381–393
Available online 1 November 2023
0891-5849/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Keywords:
Selenium status
Selenoprotein P
Glutathione peroxidase 3
Breast cancer risk
Gene-nutrient interaction

ABSTRACT
Selenium (Se) may help prevent breast cancer (BC) development. Owing to limited observational evidence, we investigated whether prediagnostic Se status and/or variants in the selenoprotein genes are associated with BC risk in a large European cohort. Se status was assessed by plasma measures of Se and its major circulating proteins, selenoprotein P (SELENOP) and glutathione peroxidase 3 (GPX3), in matched BC case-control pairs (2208 for SELENOP; 1785 for GPX3 and Se) nested within the European Prospective Investigation into Cancer and Nutrition (EPIC). Single nucleotide polymorphisms (SNPs, n = 452) in 55 selenoprotein and Se metabolic pathway genes and an additional 18 variants previously associated with Se concentrations were extracted from existing genotyping data within EPIC for 1564 case-control pairs. Multivariable-adjusted logistic regression models were used to calculate the odds ratios (ORs) and 95% confidence intervals (CIs) of the association between Se status markers, SNP variants and BC risk. Overall, there was no statistically significant association of Se status with BC risk. However, higher GPX3 activity was associated with lower risk of premenopausal BC (4th versus 1st quartile, OR = 0.54, 95% CI: 0.30–0.99, P trend = 0.013). While none of the genetic variant associations (P ≤ 0.05) retained significance after multiple testing correction, rs1004243 in the SELENOM selenoprotein gene and two SNPs in the related antioxidant TXN2 gene (rs4821494 and rs5750261) were associated with respective lower and higher risks of BC at a significance threshold of P < 0.01. Fourteen SNPs in twelve Se pathway genes (P ≤ 0.01) in interaction with Se status were also associated with BC risk. Higher Se status does not appear to be associated with BC risk, although activity of the selenoenzyme GPX3 may be inversely associated with premenopausal BC risk, and SNPs in the Se pathway alone or in combination with suboptimal Se status may influence BC risk.

List of abbreviations (in alphabetical order)

BC breast cancer
BH Benjamini–Hochberg
BMI body mass index
BPC3 Breast and Prostate Cancer Cohort Consortium
CIs confidence intervals
CRC colorectal cancer
CVs coefficients of variation
DOBS dietary oxidative balance score
E2 Oestradiol
EPIC European Prospective Investigation into Cancer and Nutrition
ER endoplasmic reticulum
ER oestrogen receptor negative
ER+ oestrogen receptor positive
FTP full-term pregnancy
GPX1/eGPX glutathione peroxidase 1/erythrocyte GPx
GPX3 glutathione peroxidase 3
GPXs glutathione peroxidases
GSH glutathione
GWAS genome-wide association study
HER2 human epidermal growth factor receptor 2
HT hormonal therapy
IARC International Agency for Research on Cancer
ICD-O-2 International Classification of Diseases for Oncology second revision
Leu Leucine
MAF Minor Allele Frequency
METS metabolic equivalents
NSCLC non-small cell lung cancer
OR Q4 vs. Q1 ORs for the fourth versus the first quintile
OR Q5 vs. Q1 ORs for the fifth versus the first quintile
ORs odds ratios
P trend P-values for tests of trend of ordinal variables
ref reference
ROS reactive oxygen species
SDs standard deviations
Se selenium
Sec selenocysteine
SELENOP selenoprotein P
SNPs Single nucleotide polymorphisms
SOD2 superoxide dismutase 2
STROBE Strengthening the Reporting of Observational studies in Epidemiology
tagSNPs haplotype tagging SNPs
TXN2 thioredoxin 2
TXNRD1-3 thioredoxin reductases
TXRF total reflection X-ray fluorescence
WHI Women’s Health Initiative

1. Introduction
In Europe, breast cancer (BC) is the most frequently diagnosed cancer in women and was recently assessed as the most common cause of cancer-related mortality [1]. Several risks factors have been shown to contribute to BC risk, including genetic susceptibility traits, sex-hormones, where oestrogen receptor positive (ER+) BC is strongly driven by exposure to oestrogens, and environmental, reproductive and lifestyle factors. However, understanding of the complexity of BC aetiology, especially ER negative (ER-) subtypes, remains challenging as BC exists in several molecular classes and is likely to result from the combination of multiple factors [2, 3]. Regarding diet, there is observational evidence that nutrients from consuming diets high in non-starchy vegetables, carotenoids, fibre, cereals, fruits, calcium, and phytoestrogens are associated with a lower risk of developing this neoplasm [4, 5].

Regarding mechanisms of tumorigenesis, experimental and
observational evidence suggests that oxidative stress and accumulation of free radicals within the breast tissue play a role in BC initiation and progression by causing DNA damage [6–8]. Accordingly, antioxidant mechanisms are likely critical to the maintenance of healthy breast tissue. Data from animal models, genomic studies and epidemiological settings suggest the essential trace micronutrient Selenium (Se) may help prevent BC [9–11], its recurrence, or mortality of BC patients [12]. Selenium’s anti-carcinogenic properties are mainly attributed to selenoproteins, a group of 25 proteins in humans that contain Se in the form of the amino acid selenocysteine (Sec). In particular, the presence of Sec in the active site of selenoenzymes such as the glutathione peroxidases (GPxs) and thioredoxin reductases (TXNRD1-3) is essential to their antioxidant properties and ability to counter oxidative and inflammatory stress-induced tumorigenic potential of malignant mammary cells [13–15]. The two major plasma selenoproteins, i.e., the transporter Selenoprotein P (SELENOP), critical for Se distribution from the liver to distal tissues, and the selenoenzyme GPX3 also exhibit antioxidant activity [14,16]. Indeed, besides SELENOP, the plasma GPX3 is the only known selenocysteine-containing extracellular antioxidant isoform capable of catalysing the reduction of peroxides and lipid hydroperoxides using glutathione (GSH) as a reducing co-factor [17].

Optimal Se status is defined as the plasma Se concentration required to support synthesis of selenoproteins [18]. However, in most European countries Se intakes and status are low/suboptimal [19,20]. Both suboptimal Se status and genetic risk variants in selenoprotein genes have been proposed to impair the response of breast epithelial cells to oxidative challenges via inadequate selenoprotein production and function and have been associated with a higher risk of developing BC [21–23], as supported by meta-analyses of observational studies of Se status measurements [10,11]. Conversely, other observational studies, or the limited studies that have examined Se concentration and genetic interactions with a few single nucleotide polymorphisms (SNPs), such as in the Malmö Diet and Cancer cohort [23], have not found associations between Se intake or status alone and BC development [24–26].

However, to date, no prospective studies have examined the combined impact of a range of Se status markers along with extensive data on selenoprotein gene variation. In the present study, we investigated if BC risk was associated with pre-diagnostic Se status and/or SNP variation in all 25 selenoprotein genes and functionally associated genes in the Se pathway, using samples taken from 2208 BC cases and 2208 matched controls nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Se status was robustly ascertained by circulating measures of Se and selenoproteins SELENOP and GPX3, as the two major functional markers of Se biology in blood [27]. This study was conducted in EPIC as we hypothesised that the contribution of Se levels to BC risk is most relevant for populations, like many in Europe, where Se intake is suboptimal [19,28].

2. Materials and methods

2.1. Study cohort and population

The present nested case-control study was conducted within the EPIC study [29,30]. This large prospective cohort study was designed with the objective of exploring the impact of diet, lifestyle, and environmental factors on the incidence of cancer and other chronic diseases. Briefly, between 1992 and 2000, 521,324 participants (approximately 2/3 female) aged between 25 and 70 years, were enrolled in 23 different sub-cohorts in centres throughout 10 Western European countries (Denmark, France, Greece, Germany, Italy, Netherlands, Norway, Spain, Sweden, United Kingdom). At recruitment, detailed information about standardized dietary, lifestyle and sociodemographic data was collected. Questionnaires included information on dietary intakes, physical activity, education, smoking and medical history. Anthropometric data and blood samples were also obtained from participants for biomarker measurements and genetic analyses. Collected blood samples were stored in Lyon, France at the International Agency for Research on Cancer (IARC) in liquid nitrogen at −196 °C for all countries except for Denmark (which are held in nitrogen vapour at −150 °C) and Sweden (where they are stored in −80 °C freezers). Standardisation of sample storage, including protocols for DNA extraction and quantification, were previously described [31].

The work described in this study was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. All study participants provided written informed consent, and ethical approval for the EPIC study was obtained from the review boards of IARC (IARC Ethics Committee) and the relevant local participating centres. 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 occurrence and mortality

All participants were followed over time for the occurrence of cancer and other diseases, as well as for overall and cause-specific mortality. Cancer incidence was determined through record linkage with population-based cancer registries (Denmark, all centres in Italy except Naples, Netherlands, Norway, Spain, Sweden, United Kingdom) or via the use of health insurance records, cancer and pathology registries, and active contact with study subjects or next-of-kin (France, Germany, Naples). For the nested case-control cohort used in this study, the last updates of complete endpoint data including alive, dead and cause of death information, occurred between 2005 and 2010, depending on the centre. There was no data on BC treatment or recurrence post diagnosis.

2.3. Selection of cases and controls and study design

Case subjects were women who developed first incident BC after recruitment and before the latest follow-up date. Invasive (primary and malignant) BC cases were classified as per the International Classification of Diseases for Oncology (Topography C50), second revision (ICD-O-2). Cases were selected to include those with existing SNP genotyping for rs1050450 in GPX1 and rs4880 in superoxide dismutase 2 (SOD2) and/or with genome-wide association study (GWAS) data, ER information (progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) status may have been missing). There was no exclusion of subjects for whom information on the use of exogenous hormones (hormonal contraceptive pills or hormonal therapy; HT) at blood collection was missing. The present study includes all participating countries within EPIC except for Greece, which was excluded due to current data restriction issues. All subjects with any prior cancer diagnoses (except non-melanoma skin cancer) at baseline were excluded from this study. Cases were matched to controls 1:1 by study centre of enrolment, age at blood collection, time and fasting status at blood collection, menopausal status and exogenous hormone use at blood collection. Premenopausal women (ascertained at baseline) were also matched according to their phase of menstrual cycle.

2.4. Plasma Se, SELENOP and GPX3 status determination

Se status was successfully assessed in 50 μl of plasma samples taken at study enrollment before disease diagnosis for 2204 matched BC case-control pairs for SELENOP, 1783 pairs for Se, and 1743 pairs for GPX3 (due to limited sample availability for 461 case-control pairs and the choice of SELENOP as the primary bioavailable Se status biomarker).

Total Se concentrations were measured in plasma samples diluted 1:2 with a standard solution containing 1000 μg/L gallium and analysed with total reflection X-ray fluorescence (TXRF) spectroscopy using an ultra-trace element analysis system (S4 T-star; Bruker Nano GmbH, Berlin, Germany), as previously described [32]. To ensure accuracy, a serum standard (Seronorm serum standard, SERO AS, Billingsstad,
Norway) was measured alongside each sample. To determine SELENOP concentrations, a colorimetric enzyme-linked immunoassay (selenOtest; selenOmed GmbH, Berlin, Germany) was used [28]. Per sample, a total of 5 µL of blood plasma was diluted 1:33 and the assay carried out according to the manufacturer’s instructions. Coefficients of variation (CVs) were determined by establishing 3 controls which covered the upper, middle and lower part of the assay’s working range (13.5–484.8 µg/L). These controls were included in the approximately 40 separate batches needed to assay all samples. GPX3 activity was quantified in triplicates (using 5 µL of plasma for each with the average value taken as the sample enzyme activity) via a coupled enzyme reaction method using hydrogen peroxide as substrate, as previously described [33,34]. For quality-control of intra- and inter-assay variability, case-control status was blinded for analysis and one sample of known Se concentration and GPX3 activity, and 3 controls for SELENOP concentrations (upper, middle, and lower range values) were used in each analysis run. The samples were measured in single measurements for Se, duplicates for SELENOP and triplicates for GPX3, and mean concentration values, SDs (standard deviations), and CVs were calculated. Duplicate samples with differences in CVs >10 % were measured again to corroborate the results. Intra- and inter-assay CVs were <10 % for Se and SELENOP, and <15 % for GPX3. The evaluation was performed with GraphPad Prism 6.01 (La Jolla, CA, USA) by using a 4-parameter logistic function.

2.5. Gene and tagging single nucleotide polymorphism (tagSNP) selection

To examine selenoprotein gene and wider Se pathway gene variations in relation to BC risk, we selected 452 functional and haplotype tagging SNPs (tagSNPs) to comprehensively analyse common SNP variation in 55 Se pathway genes. These genes were considered as belonging to two main most relevant functional pathways to Se biology: 41 in the primary selenoprotein pathway 1, relating to Se transport and biosynthesis, and 14 in the second pathway relating to antioxidant and redox reactions. The selection of these SNPs was described in our previous study of colorectal cancer (CRC) by Pedirko et al. (restricted conditions) [35]. Briefly, HapMap data (release 27, based on dbSNP version b126 and NCBI genome build 36) was used to compile a list of SNPs in all gene regions. The tagger algorithm as implemented in the haplovie 3.2 software (Broad Institute, Cambridge, MA, USA) was used to identify the tagSNPs. SNP selection was based on a Minor Allele Frequency (MAF) >5 % in Caucasians and on pairwise tagging (r² > 0.8). SNPs in promoter and potential regulatory regions (i.e., those within 2–5 kilo-base-pairs of the 5’ and 3’ ends) and known functional variants, such as rs7579, rs297299, and rs3877899 in SELENOP and rs713041 in GPX4, were included in the analysis [14,21]. Additionally in this study, 18 variants were also selected which had previously been associated with Se levels in GWAS reports [36–39]. All the included relevant results tables and supplementary tables clarify the pathway designation.

2.6. Imputation of existing SNP genotyping data

The SNP data for this study was imputed and harmonised from two BC GWAS projects in EPIC that were performed with different chips: (1) The BPC3 ER-GWAS (Breast and Prostate Cancer Cohort Consortium), performed in 2010 on 1011 subjects from the EPIC cohort; Chip: Hg18, Human 660 W-Quad, (Illumina, San Diego, CA, USA) [40], and (2) The OncoArray GWAS, carried out in 2013 on 7489 subjects from the EPIC cohort; Chip: Hg19, Infinium OncoArray-500 k (Illumina, San Diego, CA, USA) [41].

There were 452 SNPs available for our analysis. Of these, 450 SNPs were imputed and assigned continuous values between 0 and 2 and analysed using an additive genetic model. For the other two SNPs (rs1050450 in GPX1 and rs4880 in SOD2), the data was taken from their direct genotyping in a previous BPC3 study examining the association of these SNPs with BC risk [42], and thus they were coded as 0,1,2 according to genotype. There was 74 % overlap for the subjects with GWAS data and genotyping data for these two variants, where only <1 % did not have data for these SNPs, while approximately 75 % had GWAS data.

Prior to SNP data imputation, a bioinformatics pipeline was created for the data from each project (nested-case control data) to assign the EPIC IDs, complete the chromosomal position information, flip and annotate the SNPs, and check for duplicates (based on SNP call rates/ close relative tests). The Michigan Imputation platform (Reference panel: 1000 Genomes Phase 3 v5; Population: European) was used to impute the SNPs from human genome build hg37. An imputation quality metric r² filter of 0.3 was used, so that the different projects did not yield the exact same number of SNPs. Respectively, 2 and 4 SNPs were absent from the imputed data for BPC3 (rs42828087 & rs30546778), and Oncoarray (rs26118299, rs42828087, rs78425188, rs30546778).

2.7. Statistical analyses

Analysis of covariance including values natural logarithm transformed to approximate a normal distribution was used to evaluate geometric mean differences in Se, SELENOP and GPX3 concentrations among the controls based on baseline characteristics, with adjustment for study centre and laboratory batches. P-values for tests of trend of ordinal variables (P_trend) or of heterogeneity (P_heterogeneity) were determined. To identify factors associated with Se status, we conducted a Pearson correlation analysis of the Plasma Se, SELENOP and GPX3 concentrations among controls.

Multivariable adjusted conditional logistic regression models were used to calculate the odds ratio (OR) and 95 % confidence interval (CI) of the association between Se status markers and BC risk. Se and SELENOP concentrations and GPX3 activity were analysed as categorical variables, with quintile cut-points based on the distribution in the control subjects (quartile cut-points were used for the stratified analyses), and continuous log-transformed variables. To test dose-response relationships, trend values were assigned to quintile or quartile categories for Se, SELENOP, and GPX3.

Three models were used in these analyses: 1) the crude analysis model for the case-control matching factors only (age, study centre, time of blood collection and fasting status at blood collection, menopausal status and exogenous hormone use at blood collection); 2) the partially adjusted model was based on these matching factors plus additional dietary and lifestyle factors including smoking status (never, former, current), body mass index (BMI), physical activity (inactive, moderately inactive, moderately active, active), education level, energy intake, and intake of calcium, fruit and vegetables, red and processed meat, and alcohol; and 3) the fully adjusted model was based on all these covariates along with full-term pregnancy (FPT) (yes/no), number of FPTs and age at first FPT. As there was no substantive difference in results for the partially adjusted and fully adjusted models, results are presented only for the fully adjusted model in the main tables. Sub-group analyses by receptor status (ER+/−; PR+/−; HER2+/−), by cancer subsite (ductal/lobular), by menopausal status (pre-/post-menopausal), and by hormonal therapy use at blood collection were conducted. Sensitivity analyses were performed by examining subgroups at ≤2 years and ≥5 years between blood collection and diagnosis, and age of diagnosis ≤55 years at diagnosis.

The association of individual SNPs (coded as 0, 1, 2 based on the number of minor alleles) with BC risk, and associations between Se, SELENOP concentrations, GPX3 activity, and genetic variants with BC risk, were assessed using conditional logistic regression analyses (no adjustments subsequent to the crude matching model). Multiple testing corrections were carried out on all SNPs using the Benjamin-Hochberg (BH) method [43]. All statistical tests were two-sided, and P-values ≤0.05 were considered statistically significant. Analyses were performed using the SAS version 9.2 (SAS Institute, Cary, NC, USA) statistical package.
3. Results

3.1. Baseline characteristics

Demographic and lifestyle characteristics of the study participants are shown in Table 1. On average, cases were less physically active than controls (P = 0.03), had higher BMI (P = 0.005), were older at first full-term pregnancy (P = 0.003) and had higher daily intakes of red and processed meats (P = 0.02). Biomarker measures by country are available in Supplementary Table S1. The respective mean and median follow-up times to cancer diagnosis were 4.8 (SD = 2.7) and 4.7 years, while the average follow-up time to BC diagnosis for the pre- and post-menopausal groups was 5.2 and 4.7 years respectively. Se concentrations were indicative of an overall low/suboptimal Se status in this population [19,44]. In the controls, Spearman correlation analysis of the Se status biomarkers (adjusted by age, BMI and batch) with corresponding scatterplots (computed on the raw values), see Supplementary Table S2, indicated that Se was highly correlated with SELENOP (R$^2$ = 0.60, as expected in a non-Se replete population, and moderately correlated with GPX3 (R$^2$ = 0.38), while GPX3 and SELENOP were moderately correlated with each other (R$^2$ = 0.33). There was no marked correlation of the Se status biomarkers with BMI (slightly positive with Se and SELENOP concentrations and slightly negative with GPX3 activity; Supplementary Table S2).

Table 1

<table>
<thead>
<tr>
<th>Baseline characteristic</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>N = 2208</td>
<td>N = 2208</td>
</tr>
<tr>
<td>Age at blood collection, years</td>
<td>54.6 (7.7)</td>
<td>54.6 (7.6)</td>
</tr>
<tr>
<td>Educational attainment, % (157 missing)</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>None/Primary</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Technical/professional school</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Secondary</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>University degree</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Smoking status, % (88 missing)</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>Never smoker</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Former smoker</td>
<td>37</td>
<td>36</td>
</tr>
<tr>
<td>Current smoker</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Physical activity, % (59 missing)</td>
<td>92.9 (52.7)</td>
<td>95.4 (52.5)</td>
</tr>
<tr>
<td>Menopausal status, %</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Postmenopausal, aged 55+</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Postmenopausal, aged 45-54</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Postmenopausal, aged &lt;45</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hormonal therapy use (ever), % (250 missing)</td>
<td>13.1 (1.6)</td>
<td>13.2 (1.6)</td>
</tr>
<tr>
<td>Age at first menstrual period, yrs (191 missing)</td>
<td>1.9 (1.2)</td>
<td>2.0 (1.2)</td>
</tr>
<tr>
<td>Number of full-term pregnancies, % (missing)</td>
<td>25.3 (4.2)</td>
<td>24.8 (4.2)</td>
</tr>
<tr>
<td>BMI, kg/m$^2$</td>
<td>25.3 (4.5)</td>
<td>25.0 (3.9)</td>
</tr>
<tr>
<td>Height, cm</td>
<td>162.4 (6.5)</td>
<td>162.0 (6.4)</td>
</tr>
<tr>
<td>Total energy intake, kcal/d</td>
<td>1960.5</td>
<td>1935.5</td>
</tr>
<tr>
<td>Alcohol, g/d</td>
<td>10.1 (13.3)</td>
<td>9.6 (12.9)</td>
</tr>
<tr>
<td>Calcium, mg/d</td>
<td>1012.1</td>
<td>999.5 (385.4)</td>
</tr>
<tr>
<td>fibre, g/d</td>
<td>22.2 (7.6)</td>
<td>22.1 (7.2)</td>
</tr>
<tr>
<td>Folate, g/d</td>
<td>294.5 (109.2)</td>
<td>294.4 (109.3)</td>
</tr>
<tr>
<td>Fruit, nuts, seeds and vegetables, g/d</td>
<td>442.3 (244.9)</td>
<td>450.0 (245.7)</td>
</tr>
<tr>
<td>Fruit and vegetables, g/d</td>
<td>435.6 (243.1)</td>
<td>443.5 (243.3)</td>
</tr>
<tr>
<td>Fish and shellfish, g/d</td>
<td>32.6 (26.7)</td>
<td>32.9 (27.4)</td>
</tr>
<tr>
<td>Red and processed meat, g/d</td>
<td>71.1 (39.7)</td>
<td>68.5 (39.6)</td>
</tr>
<tr>
<td>Baseline plasma biomarkers, geometric mean (SD)</td>
<td>244.9 (58.0)</td>
<td>246.8 (56.4)</td>
</tr>
<tr>
<td>Glutathione peroxidase 3 (GPX3; U/L) (924 missing)</td>
<td>65.3 (21.6)</td>
<td>64.7 (20.8)</td>
</tr>
<tr>
<td>Selenium, μg/L (846 missing)</td>
<td>0.5 (1.2)</td>
<td>5.0 (1.1)</td>
</tr>
<tr>
<td>Tumour characteristics, % missing</td>
<td>1-2% (0 missing)</td>
<td>78.80/21.20</td>
</tr>
<tr>
<td>ER+ (901 missing)</td>
<td>46.60/24.60</td>
<td></td>
</tr>
<tr>
<td>HER2+ (1510 missing)</td>
<td>6.80/24.80</td>
<td></td>
</tr>
<tr>
<td>Baseline plasma biomarkers, geometric mean (SD)</td>
<td>71.4/14.4</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: BMI = Body Mass Index; ER = oestrogen receptor; HER2 = human epidermal growth factor receptor 2; METS = metabolic equivalents; PR = progesterone receptor.

* Data are given as means (SD) unless otherwise specified. Missing values were not excluded from percentage calculations; therefore, the sum of percentages across subgroups may not add up to 100 %.

3.2. Associations of Se, SELENOP, and GPX3 levels with breast cancer risk

There was no significant association of any of the three assessed Se status biomarkers with risk of BC in the fully adjusted models (Table 2). Respective multivariable adjusted ORs for the fifth versus the first quintile (ORQ5 vs Q1) of Se, SELENOP, and GPX3 levels with BC risk were 1.12 (95 % CI: 0.84–1.48, P$^\text{trend}$ = 0.59), 0.89 (95 % CI: 0.72–1.10, P$^\text{trend}$ = 0.42), and 0.79 (95 % CI: 0.60–1.04, P$^\text{trend}$ = 0.071). However, GPX3 activity was statistically significantly inversely associated with BC risk in the analysis including matching factors only (ORQ5 vs Q1 = 0.75; 95 % CI: 0.57–0.98; P$^\text{trend}$ = 0.028). While the point estimates remained similar, the P-values lost significance in the partially adjusted (ORQ5 vs Q1 = 0.78; 95 % CI: 0.59–1.03; P$^\text{trend}$ = 0.058) and fully adjusted (ORQ5 vs Q1 given above) multivariable analyses. No statistical significance was observed for this biomarker in the continuous log-transformed GPX3 activity (ORfully adjusted = 0.79; 95 % CI: 0.54–1.18).

3.3. Association of breast cancer risk with tumour subtype, tumour receptor status, menopausal status, and exogenous hormone use

Stratified analyses were carried out by menopausal status, hormonal therapy use, tumour subtype (ductal or lobular), and ER, PR and HER2 hormone receptor status.

Higher GPX3 activity was associated with lower BC risk in premenopausal women, but not in postmenopausal women, in both the categorical and continuous analyses: ORQ4 vs Q1 = 0.54 (0.30–0.98), P$^\text{trend}$ = 0.013, and OR = 0.25 (95 % CI: 0.09–0.69) = 0.007, respectively (Table 3). The P-values for heterogeneity between pre- and postmenopausal cohorts were 0.40, 0.67 and 0.19 for Se and SELENOP concentrations, and GPX3 activity, respectively. The other stratified analyses are summarised in Supplementary Tables S3–S7, respectively.

In the no exogenous hormone use group, there was an inverse association between GPX3 activity and BC risk, but it was only statistically significant in the crude analysis (ORQ4 vs Q1 = 0.72; 95 % CI: 0.53–0.98; P$^\text{trend}$ = 0.043). GPX3 activity appeared to have no association with BC risk in women using exogenous hormones (P$^\text{heterogeneity}$ = 0.60).

In nearly all the analyses stratified by receptor status (ER, PR, HER2), and considering the multivariable adjusted results, there was no marked association with BC risk observed. The only exception was the association of higher GPX3 activity with a lower risk of both PR+ (ORQ4 vs Q1 = 0.63; 95 % CI: 0.42–0.95; P$^\text{trend}$ = 0.022) and PR– BC (ORQ4 vs Q1 = 0.55; 95 % CI: 0.32–0.94; P$^\text{trend}$ = 0.044), with no evidence for heterogeneity (P$^\text{heterogeneity}$ = 0.53).

Sensitivity analyses were conducted by examining subgroups at $<\approx$ 2 years and $\leq 5$ years between blood collection and diagnosis (see Supplementary Table S8). There were no substantive changes observed for the results in subgroups $<\approx 2$ years. However, higher Se concentrations were associated with higher BC risk in women diagnosed more than 5 years after baseline (ORQ4 vs Q1 = 1.58; 95 % CI: 1.06–2.37; P$^\text{trend}$ = 0.044), but not in those diagnosed within 5 years (P$^\text{trend}$ = 0.087). As menopausal status was defined at blood collection (and not at diagnosis), we performed a further sensitivity analysis using age of
diagnosis at 55 years as a proxy for menopause onset (results presented in Supplementary Table S9). There were no notable differences in the results for Se and SELENOP, while in individuals diagnosed with BC before 55 years, GPX3 activity was non-statistically significantly inversely associated with BC risk in the fully adjusted model (OR_{Q4} vs Q1 = 0.65; 95 % CI: 0.40–1.06; P_{trend} = 0.091), which was broadly equivalent to the point estimates for the association in premenopausal women.

### 3.4. Association between breast cancer risk, genetic variants in the selenium pathway, and interaction between SNPs and selenium status

The 55 genes from the Se metabolic pathway were allocated to 2 sub-pathways, where pathway 1 comprised genes coding for selenoproteins (including those with antioxidant function) and Se biosynthesis and transport (14 genes), while pathway 2 comprised genes involved in antioxidant function and redox control (41 genes). A total of 450 tagSNPs across these genes were assessed for their associations with BC risk, along with 18 further SNPs which had previously shown significant associations with circulating Se concentrations in GWAS reports (Supplementary Table S10).

Significant associations with BC risk were observed for 18 tagSNPs, including 7 genetic variants in 7 genes within pathway 1 and 10 genetic variants in 7 genes from pathway 2, and 1 of the SNPs (located at the CBS gene region) identified by GWAS to influence Se concentrations. In stratified analyses, 110 of these variants in 46 genes, comprising 69 SNPs in 30 pathway 1 genes, 35 SNPs in 12 pathway 2 genes and 6 GWAS SNPs at 4 genomic loci, were associated with BC risk for at least one hormone receptor subtype (ER+/−, PR+/−, HER2+/−), with 16 of these SNPs significant for more than one receptor type. Furthermore, 20 of the 110 SNPs (18 %) were previously found to be nominally associated with CRC risk (or for colon or rectal subsites) in our previous EPIC study [35]. Table 4 shows the significant SNP associations for overall BC risk, while Supplementary Table S10 contains all the individual SNPs analysed, their variant alleles and frequencies and the corresponding gene/locus and pathway designations. Supplementary Table S11 lists all significant SNPs for overall BC and by hormone receptor subtype, and whether they were also reported as significantly associated with CRC risk in EPIC [35].

While none of the 110 SNPs retained significance following multiple testing correction by the BH procedure, 27 SNPs did show nominal P-values ≤0.01 for association with risk of overall BC and/or at least one subtype. Of these 27, 15 were in 9 pathway 1 genes and 12 were in 7 pathway 2 genes. Only TNR2 rs4821494 had a nominal P-value ≤0.01 for overall BC risk and for more than one sub-analysis (ER+ and PR+), while also being nominally significant (P = 0.03) for HER2+ BC. However, several genes harboured >1 SNP with nominal P-values ≤0.01 for association with BC risk in various sub-analyses; for example, GSR contained 4 SNPs associated with HER2- BC, GPX2 harboured 3 SNPs associated with PR+ BC, and GPX7 and TNR11 included SNPs each associated with HER2- and PR+ BC, respectively. Furthermore, GPX7, GSR and SECISBP2 all harboured at least one SNP which had a nominal P-value ≤0.01 in this study and which was also associated with CRC risk in EPIC [35].

Among the control subjects, we detected 88 SNPs in 37 genes nominally associated with levels of at least one of the three Se status biomarkers. None of these associations retained significance following multiple testing correction. However, 19 of these SNPs had nominal P-values ≤0.01 for association with at least one Se status marker. The SNPs rs3877899 and rs6413428 in SELENOP had nominal P-values ≤0.01 for association with both Se and SELENOP levels (and both SNPs had nominal P-values ≤0.001 for the SELENOP association); the other 17 SNPs were nominally associated with only one of the three biomarkers (Supplementary Tables S12–S14).

Alterations to the functional efficiency of selenoproteins that might result from genetic variation in the Se pathway may be compensated by a higher Se intake; thus, these variants may have a stronger association with cancer risk in suboptimal Se status populations. Therefore, we also assessed if any interactions between individual SNPs and Se status measures, at levels both above and below their respective medians, were evident in the associations with BC risk. Supplementary Table S15 shows the 34 SNPs nominally associated with BC above and/or below the median level of at least one of the three Se status biomarkers (25 SNPs in 14 pathway 1 genes, 8 SNPs in 7 pathway 2 genes and 1 GWAS-associated SNP). While none of the SNP associations retained significance following adjustment for multiple testing, 14 SNPs (in 12 genes) had a nominal P-value ≤0.01 (Table 5). Only one of these was also associated with BC risk alone (rs451774, GPX5), while both analyses showed inverse associations.
4. Discussion

As far as we are aware, we have conducted the largest reported prospective investigative study of the associations between Se status (as assessed by measuring circulating Se and two functional selenoprotein biomarkers of Se activity), common genetic variations in all selenoprotein genes (and related redox pathway genes), and their association with risk of developing BC. Although none of the plasma biomarkers were associated with modification of BC risk in the fully adjusted models, an association between genetic variation in Se pathway genes alone or in interaction with Se status may affect BC development, although none of the findings remained significant after multiple testing adjustment.

Considering associations at a nominal P-value cut-off for significance of ≤0.01, in the SNP-alone analyses among the selenoprotein genes only rs1004243 in SELENOM was associated with BC risk (inversely), while there were several associations for interactions of Se status and genetic variation with BC risk.

The Danish, “Diet, Cancer, and Health” study observed that erythrocyte GPx (eGPx, also known as cGPx or GPX1 activity), a signifier of GPX1 expression, was lower in BC cases than in controls [45]. Subsequently, it was shown in the same cohort that T/T (Leucine; Leu) homozygous genotype for the rs1050450 functional SNP in the GPX1 gene was associated with a higher risk of developing non-ductal BC [13]. The Leu variant has been shown to reduce GPX1 activity compared with the Proline (Pro) counterpart [15]. Furthermore, an association was reported between the Leu carriers for rs1050450, increased GPX1 activity in HT current users (at time of GPX1 measurement) compared to never or former users, and later BC development. Here, the authors suggested that postmenopausal women harbouring the rs1050450 Leu allele and who have low GPX1 activity when using HT may be more susceptible to BC development [13]. There was some overlap in the postmenopausal cases used in this study and in the Danish contribution to the present report (while the controls were sampled differently). In our study, higher levels of GPX3 were statistically significantly associated with a lower risk of BC only for premenopausal women across multiple centres (Table 3). Measurement of GPX1 activity was beyond the scope of this study but would be warranted in future investigations to better elucidate the potential biological interaction between various GPXs, menopausal status, HT use, oestrogen, and BC risk.

A few extensive observational studies have indicated that a higher Se status improves survival outcomes from BC [46–48]. This observation has recently been substantiated by a linear and dose-dependent associations of autoimmunity to SELENOP that negatively affects Se transport [12]. However, many previous epidemiological studies of Se status and BC risk are limited and equivocal including issues such as possible reverse causality, “inactive” (not bioavailable) Se measures from toenail clippings, and low statistical power (<150 cases) [11,49–51]. There has not yet been a robust intervention trial published to indicate whether Se supplementation may help prevent BC in women with suboptimal Se status. Meta-analyses and systematic reviews of perspective and case-control studies indicate that a higher dietary Se intake and status levels are associated with BC prevention [10,11], which is more apparent for circulating Se concentrations [11]. However, this is not supported by the evidence from intervention studies [11]. An updated Cochrane review of observational studies concluded that there was no marked association between baseline Se levels and BC risk, while there were no intervention studies primarily addressing BC assessed [52]. The authors did note that
nutritional status and the interaction with inherited genetic variation should be further explored, as included in this study. More recent and higher-powered observational studies have further disputed this evidence. The Malmö Diet and Cancer Study of 1186 case-control pairs found no association between prediagnostic serum Se levels and BC risk [24,26], while an investigation conducted in the Women’s Health Initiative (WHI) found no link between Se intake and BC incidence in a US prospective cohort with 9487 accrued postmenopausal cases [25]. However, Se intake was ascertained from a self-assessed baseline food frequency and supplemental intake questionnaire, and the study was conducted in postmenopausal women.

All these studies only measured total Se (usually circulating or in toenails) or extrapolated Se intake from dietary information. The effects of Se could depend on different biological interactions in pre- or postmenopausal women or with different Se bioavailability and functionally relevant selenoproteins, as suggested by the interaction between GPX3 and BC risk in this study and by the association between HT use, GPX1 activity and BC risk identified by Mèplan and colleagues [13]. Our finding of the possible link of GPX3 activity with BC is bolstered by evidence that the GPX3 gene is downregulated in BC cells and that its overexpression in vitro suppresses BC cell proliferation and migration [53]. Furthermore, gene expression profiling in a mouse model revealed that Gpx3 could act as a key mediator of oestrogen in relation to fat mass in white adipose tissue [54]. In stratifying the association of increasing GPX3 activity by existing measures of circulating oestradiol (E2) concentration in EPIC in the crude model, we observed a significant inverse association with BC risk only at or above the median E2 value compared to below the median (OR = 0.22, 95% CI: 0.05–0.94, P = 0.04 and OR = 0.65, 95% CI: 0.19–2.17, P = 0.48, respectively). Overall, our findings support a potential link between GPX3 and changes in female sex-hormones in premenopausal women that could impact BC development.

Beside Se transportation, GPX3 is also regulated at the transcriptional level such as Se dependent translation by under-regulating the peroxisome proliferator-activated receptor γ (PPARY), as demonstrated by knock-down or activation/overexpression of PPARY, respectively decreasing or increasing the cellular expression of GPX3 [55]. Similar to SELENOP, GPX3 may also act as a major scavenger receptor of reactive oxygen species (ROS) and function by acting at ROS generation hubs that increase lipid hydroperoxide generation within lipoxigenases vicinity at the cell surface; however, the exact mechanism is still unclear [56,57]. Furthermore, in silico analysis of protein and gene expression databases suggests that GPX3 is downregulated in BC [58], while modestly sized patient cohort studies have reported decreased GPX3 expression in breast carcinoma vs normal breast tissue [58,59], that was associated with no HT use in subjects with BC [60].

No significant associations were observed following multiple testing correction between individual SNPs in the selenoprotein, Se biosynthesis and transport, and related antioxidant pathway genes with overall BC development risk. We note that some of the studied genes have no major known expression or that for their corresponding proteins in breast tumours, based on databases such as the Cancer Genome Atlas (TCGA; https://gdc.cancer.gov/) and a recent study of the selenoprotein transcriptome by Demircan and colleagues [61] showing that GPX6 was the only selenoprotein gene not expressed in breast tumours. Thus, the SNP associations for these genes (such as rs974343 in GPX6, with also a P-value of just 0.04) may simply reflect false positives due to multiple testing. However, as we do not have BC tissue available in this cohort, we cannot measure the expression levels of these genes or proteins in the study samples. Although in some cases genetic variants affect gene/protein expression levels, they can also affect enzyme activity, protein and mRNA stability, post-translational modifications, interaction with other factors, and subcellular locations. These changes would not necessarily result in observable changes in gene or protein expression levels. As we mainly used a tagSNP approach to cover common variation in the studied genes, we therefore do not know if the SNPs found to be associated with a modified BC risk were directly functional (i.e., affecting the corresponding protein function, regulation, or activity). However, one SNP (rs10449982) in the SELENOM selenoprotein gene and 2 variants (rs4821494 and rs5750261) in TXN2 showed nominal P-values ≤0.01 with a decreased and increased BC risk, respectively. SELENOM is an endoplasmic reticulum (ER) resident selenoprotein reported to contribute to ER-stress response and calcium signalling, which may confer selenoprotein-related effects in counteracting or promoting carcinogenic processes [62]. SELENOM gene expression with increasing Se concentrations has also recently been observed to be associated with lower mortality in BC patients in a large, prospective Swedish study [61]. TXN2, a member of the thioredoxin family, encodes a mitochondrial redox protein, with a vital role in the maintenance of mitochondrial reactive oxygen species homeostasis, alongside functions in the regulation of apoptosis and cell viability [63]. TXN2 rs4821494 and rs5750261 were the only SNPs in our analysis with nominal P-values ≤

---

**Table 4**

All selenium pathway SNPs (by gene and pathway) significantly associated with overall breast cancer risk before multiple testing corrections, EPIC study.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP Reference</th>
<th>Variant Allele</th>
<th>OR (95% CI&lt;sup&gt;a&lt;/sup&gt;)</th>
<th>P-value&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pathway 1</strong></td>
<td>Selenoproteins and Selenium biosynthesis and transport</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPX6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>rs974334&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>0.88 (0.77–1.00)</td>
<td>0.94 (0.79–0.98)</td>
</tr>
<tr>
<td>LRP2</td>
<td>rs700552</td>
<td>G</td>
<td>1.14 (1.01–1.29)</td>
<td>0.01 (0.80–0.97)</td>
</tr>
<tr>
<td>SELENOM</td>
<td>rs1004243</td>
<td>A</td>
<td>1.16 (1.00–1.34)</td>
<td>0.05 (1.00–1.23)</td>
</tr>
<tr>
<td>SELENOM&lt;sup&gt;c&lt;/sup&gt;</td>
<td>rs13329318</td>
<td>C</td>
<td>1.15 (1.01–1.29)</td>
<td>0.01 (0.77–1.25)</td>
</tr>
<tr>
<td>SEPHS1&lt;sup&gt;d&lt;/sup&gt;</td>
<td>rs2275129</td>
<td>G</td>
<td>1.18 (1.01–1.34)</td>
<td>0.03 (1.00–1.23)</td>
</tr>
<tr>
<td>TXNRD2&lt;sup&gt;d&lt;/sup&gt;</td>
<td>rs7777226</td>
<td>G</td>
<td>0.90 (0.80–1.00)</td>
<td>0.02 (0.79–0.98)</td>
</tr>
<tr>
<td><strong>Pathway 2</strong></td>
<td>Antioxidants and Redox function genes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>rs10449982</td>
<td>A</td>
<td>1.16 (1.00–1.35)</td>
<td>0.03 (0.89–0.99)</td>
</tr>
<tr>
<td>CAT</td>
<td>rs11027700</td>
<td>A</td>
<td>1.15 (1.01–1.32)</td>
<td>0.03 (1.01–1.28)</td>
</tr>
<tr>
<td>GPX3</td>
<td>rs451774</td>
<td>G</td>
<td>1.11 (1.00–1.23)</td>
<td>0.03 (1.00–1.23)</td>
</tr>
<tr>
<td>GSR</td>
<td>rs8109096</td>
<td>G</td>
<td>1.12 (1.01–1.32)</td>
<td>0.01 (1.00–1.23)</td>
</tr>
<tr>
<td>HIF1A</td>
<td>rs11549465</td>
<td>G</td>
<td>1.12 (1.01–1.32)</td>
<td>0.02 (1.00–1.23)</td>
</tr>
<tr>
<td>RPS6KB1</td>
<td>rs143652</td>
<td>A</td>
<td>1.13 (1.01–1.32)</td>
<td>0.01 (1.00–1.32)</td>
</tr>
<tr>
<td>SOD2</td>
<td>rs4880</td>
<td>G</td>
<td>1.11 (1.00–1.23)</td>
<td>0.03 (1.00–1.23)</td>
</tr>
<tr>
<td>TXN2</td>
<td>rs4821494</td>
<td>A</td>
<td>1.18 (1.01–1.23)</td>
<td>0.01 (1.00–1.23)</td>
</tr>
<tr>
<td>TXN2</td>
<td>rs5750261</td>
<td>G</td>
<td>1.18 (1.00–1.36)</td>
<td>0.03 (1.00–1.36)</td>
</tr>
<tr>
<td><strong>GWAS Selenium concentration-associated SNPs</strong></td>
<td>CBS</td>
<td>rs234709</td>
<td>C</td>
<td>1.12 (1.01–1.25)</td>
</tr>
</tbody>
</table>
pairs included in this study, found no significant interaction between previous 2006 study) [45] that rs1050450 was associated with an impacting BC development [13]. Subsequently, the large, international may be an interaction effect between the variant and GPX1 activity [13].

Many previously reported a lower BC risk associated with rs4821494 -0.01 for risks of overall BC, ER- BC and PR- BC. A large population-based case-control study of postmenopausal women in Germany previously reported a lower BC risk associated with rs4821494 [64]. Additionally, Table 5 summarizes the observed interactions between SNPs in TXNRD2 and TXNRD3 and Se biomarkers, supporting a role of the TXN system, redox control and mitochondria-mediated apoptosis signalling in BC development [65].

Other case-control studies have provided some evidence, albeit conflicting, that other SNPs in the selenoprotein-coding and related oxidative-stress response genes are associated with BC risk. Several studies have shown diverging evidence for an association of the GPX1 P198L-rs1050450 variant with BC risk. In the 2022 Swedish Malmö Diet and Cancer study, it was observed that T/T genotype carriers for rs1050450 were associated with a lower BC risk, while interaction with a higher Se intake increased the estimated protection from developing BC [23]. The Nurses’ Health Study in the US observed an association with BC [23]. The Nurses Health Study in the US observed an association with increased BC risk [66]. M

Table 5

<table>
<thead>
<tr>
<th>Pathway 1</th>
<th>Selenoproteins and Se biosynthesis and transport</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
<td>SNP</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>GPX3</td>
<td>rs10463312</td>
</tr>
<tr>
<td></td>
<td>rs2070593</td>
</tr>
<tr>
<td></td>
<td>rs434112</td>
</tr>
<tr>
<td>SCLY</td>
<td>rs1562337</td>
</tr>
<tr>
<td></td>
<td>rs4876979</td>
</tr>
<tr>
<td></td>
<td>rs17564336</td>
</tr>
<tr>
<td></td>
<td>rs2864148</td>
</tr>
<tr>
<td></td>
<td>rs75436938</td>
</tr>
<tr>
<td></td>
<td>rs3788314</td>
</tr>
<tr>
<td></td>
<td>rs9637365</td>
</tr>
<tr>
<td>GPX5</td>
<td>rs451774</td>
</tr>
<tr>
<td></td>
<td>rs7217337</td>
</tr>
<tr>
<td></td>
<td>rs4135212</td>
</tr>
<tr>
<td></td>
<td>rs11089790</td>
</tr>
</tbody>
</table>

Associations with breast cancer risk (conditional logistic regression). Results given as Odds Ratio (95% Confidence Interval); P-value (those in italics are at \( \leq 0.05 \)). Abbreviations: Se, Selenium; SELENOP, Selenoprotein P; GPX3, Glutathione peroxidase 3.

Median columns show odds ratios, P-values for association between each individual SNP and breast cancer risk when levels of each Se status marker are below their respective medians.

Pathway 2 = Antioxidants and Redox

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Se</th>
<th>Median</th>
<th>Median</th>
<th>SELENOP</th>
<th>Median</th>
<th>Median</th>
<th>GPX3</th>
<th>Median</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPX5</td>
<td>rs451774</td>
<td>0.79</td>
<td>1.11</td>
<td>0.005</td>
<td>0.85</td>
<td>0.80</td>
<td>0.82</td>
<td>0.22</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs7217337</td>
<td>0.90</td>
<td>0.10</td>
<td>0.13</td>
<td>0.89</td>
<td>0.93</td>
<td>1.17</td>
<td>0.57</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs4135212</td>
<td>1.02</td>
<td>0.96</td>
<td>0.84</td>
<td>1.00</td>
<td>0.93</td>
<td>1.17</td>
<td>0.57</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>TXN2</td>
<td>rs11089790</td>
<td>0.92</td>
<td>0.82</td>
<td>0.43</td>
<td>0.99</td>
<td>0.93</td>
<td>1.17</td>
<td>0.57</td>
<td>0.41</td>
<td></td>
</tr>
</tbody>
</table>

0.01 for risks of overall BC, ER- BC and PR- BC. A large population-based case-control study of postmenopausal women in Germany previously reported a lower BC risk associated with rs4821494 [64]. Additionally, Table 5 summarizes the observed interactions between SNPs in TXNRD2 and TXNRD3 and Se biomarkers, supporting a role of the TXN system, redox control and mitochondria-mediated apoptosis signalling in BC development [65].

The conflicting results for Se measures and selenoprotein gene variants with BC risk likely reflect study setting, design, and power differences with limited investigation of the interactions between Se intake, Se status, Se metabolism and genotype in association with BC and BC sub-type risks [21]. Moreover, there have been risk differences observed in sub-groups associated with health disparities and lower Se levels, even in generally Se replete geographic areas, as shown by the reported association between two SELENOP variants (rs230812 and rs6865453) and an increased BC risk in women of higher Native American ancestry [68].

Considering the other pathway 1 selenoprotein gene, TXNRD1 encodes one of three thioredoxin reductase enzymes (TXNRD1-3) which play a key role in redox homeostasis and Se metabolism and have been reported to be overexpressed in aggressive BC tumours [69]. TXNRD1 is expressed ubiquitously in the cytosol [70], and likely contributes
SNPs in the GPX family associated with overall BC risk, all members found to be significant for at least 2 of the 3 biomarkers, it is worth SNPs displayed opposite risk directions at Se status marker levels above median; all other SNPs were associated with BC risk either above or

tion of selenoprotein genotype and Se status regarding BC risk. Of the 34 cancer (NSCLC) among women in the US [76]. Finally, 13 of the 18 SNPs levels was also observed in controls from the EPIC CRC study [35], and

ation between another

2 of these SNPs, including rs1002149, were also associated with CRC risk in EPIC [35].

Numerous genetic variations were observed to be associated with Se status levels (as assessed by plasma Se and SELENOP concentration, and GPX3 activity), although none retained significance after multiple test corrections. Of the 19 SNPs with unadjusted P-values ≤0.01 for associations with Se status, only two SNPs, rs3877899 and rs6413428 in SELENOP, were associated with variance in both Se and SELENOP levels (but not with GPX3 activity). This result could be expected, since SELENOP is the primary Se transport protein in plasma, carrying approximately one half to two thirds of circulating Se [18], and rs3877899, which lies within the coding region of SELENOP, has previously been shown to influence plasma concentrations of Se and selenoproteins [73,74], and GPX3 enzyme activity in pregnant women [39]. The GG genotype of rs3877899 has also shown significant associations with TXNRD activity, an important antioxidant enzyme [75]. An association between another SELENOP variant (rs6413428) with SELENOP levels was also observed in controls from the EPIC CRC study [35], and the SNP has been previously associated with risk of non-small cell lung cancer (NSCLC) among women in the US [76]. Finally, 13 of the 18 SNPs previously linked with Se biomarker variance in GWAS studies were also associated with Se status levels in this study; 11 with Se concentrations (only one of which, rs685966, was also associated with SELENOP concentration), one with SELENOP alone (rs234709) and one with GPX3 activity alone (rs6586282).

Prior to this study there were limited data available on the interaction of selenoprotein genotype and Se status regarding BC risk. Of the 34 SNP-Se interactions which displayed nominal significance for BC risk, 3 SNPs (rs10463312 and rs5085435 in GPX3 and rs2695234 in SOD3) were significant when assessed at Se levels both above and below the median; of other SNPs were associated with BC risk either above or below the median levels of Se status biomarkers. Interestingly, these 3 SNPs displayed opposite risk directions at Se status marker levels above vs below the median. Although the SNP-Se interactions were not always found to be significant for at least 2 of the 3 biomarkers, it is worth noting that in most cases the estimates for the risks of developing BC were following the same direction below and above the median, suggesting that the potential impact of risk or protective alleles are modified by the various Se biomarkers.

Higher Se concentrations have been observed in breast tumour tissue in association with the AA homozygote of the missense functional rs3877899 (Alanine234Threonine) variant in SELENOP [77]. Additionally, rs3877899 has been associated with BC risk [13,78] or survival [79] and risk of prostate cancer [80]. Although there were opposite BC risk directions observed for the Danish report [13] and the modestly sized Iranian study [78]. Other investigations have observed no association for this SNP and risk of these two hormonal-related cancer types, either alone or with Se status [81–84]. In previous studies of interaction of selenoprotein SNPs with other biomarker assessments, Pellatt et al. identified four variants (GPX3 rs2070593, GPX4 rs2074451, SELENOS rs9874, and TXNRD1 rs17202060) that demonstrated significant interactions with dietary oxidative balance score (DOBS) to modify BC risk [68], further supporting the hypothesis of the roles of both antioxidative stress response selenoproteins and alterations in oxidative balance generating DNA-damaging free radicals in the prevention or development and progression of BC [6,16].

This study reports the most extensive observational data so far on the association of Se status and Se pathway genotypes with BC risk. The hypothesis-driven approach and large sample size within a prospective study allowed for the detailed examination of Se pathway genetic variation and interaction with the three, primary Se status biomarkers. The collection of blood samples prior to BC diagnoses minimised reverse causality bias, and our analysis models adjusted for several potential covariates to reduce the possibility of confounding.

Blood collection at a single time point is a limitation in this study, potentially giving rise to random error. Secondly, we have no data on use of mineral supplements containing Se to investigate if their use may confound the risk estimates. Thirdly, it is important to note that despite the large sample size, SNP-Se interaction analysis and some stratified analyses had limited power due to modest sizes for those sub-analyses, particularly analyses by HER2+ receptor status and anatomical sub-sites. Thus, such results are only tentative and need repeating with larger cohorts and/or those with multiple time-point measures. Finally, since most of the included variants in the genetic analyses were based on tagSNPs and not function (or the actual contributing functional variant (s) they tag), further genetic mapping and validation studies are required to explore these putative associations more comprehensively.

In summary, the present study indicates that higher prediagnostic Se status levels do not appear to be associated with BC overall, although GPX3 activity may be important for BC prevention in premenopausal women. There was limited evidence that genetic variation in selenoprotein genes, Se metabolism genes, and other antioxidant genes may be associated with BC risk, either alone or in interaction with Se status. While none of the genetic-based findings retained significance following multiple testing correction, some of them showed nominal P-values ≤0.01 for risk with overall BC or different tumour subtypes and/or interactions with Se status, with some also showing associations with CRC risk in a similar Se pathway study in EPIC [35], warranting further investigation.

Author contributions

(Ruth Travis), A.K.H.) reviewed and approved the manuscript and commented on the analysis and interpretation of the findings.

IARC disclaimer 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.

Declaration of competing interest

Lutz Schomburg is the founder of selenOmed GmbH, a company involved in improving Se diagnostics. The other 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. 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.

Acknowledgements

Funding for this study was provided by the Health Research Board of Ireland health research award HRA-PHS/2015/1142 (principal investigator: D.J.H.). L.S. was supported by Deutsche Forschungsgemeinschaft (DFG Research Unit 2558 TraceAge, Sch 849/6-2). We thank the members of EPIC and the study participants for their contributions to the EPIC cohort study. The coordination of EPIC is financially supported by International Agency for Research on Cancer (IARC) and by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London which has additional infrastructure support provided by the NIHR Imperial Biomedical Research Centre (BRC). The national cohorts are supported by: Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l’Éducation Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), German Institute of Human Nutrition Potsdam-Rehbrueck (DIfE), Federal Ministry of Education and Research (BMBF) (Germany); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy, Compagnia di SanPaolo and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); Health Research Fund (FIS) - Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology- ICO (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C8221/A29017 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk; MR/M012910/1 to EPIC-Oxford), (United Kingdom). We thank the National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands, for their contribution and ongoing support to the EPIC Study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.freeradbiomed.2023.10.401.

References

D.J. Hughes et al.

Free Radical Biology and Medicine 209 (2023) 381–393

392


