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

Hughes, David J.; Schomburg, Lutz; Jenab, Mazda; Biessy, Carine; Méplan, Catherine; Moskal, Aurelie; Sun, Qian; Demircan, Kamil; Fedirko, Veronika; Weiderpass, Elisabete; Mukhtar, Maryam; Olsen, Anja; Tjønneland, Anne; Overvad, Kim; Schulze, Matthias; Nøst, Therese Haugdahl; Skeie, Guri; Olsen, Karina Standahl; Ricceri, Fulvio; Grioni, Sara; Palli, Domenico; Masala, Giovanna; Tumino, Rosario; Pasanisi, Fabrizio; Amiano, Pilar; Colorado Yohar, Sandra M.; Agudo, Antonio; Sánchez, Maria-Jose; Ardanaz, Eva; Sund, Malin; Andersson, Anne; Perez-Cornago, Aurora; Travis, Ruth; Heath, Alicia K.; Dossus, Laure

Published in:
Free Radical Biology and Medicine

DOI (link to publication from Publisher):
10.1016/j.freeradbiomed.2023.10.401

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

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

Contents lists available at ScienceDirect

Free Radical Biology and Medicine

journal homepage: www.elsevier.com/locate/freeradbiomed

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 [c,e], Qian Sun [b], Kamil Demircan [b], Veronika Fedirko [f], Elisabete Weiderpass [c], Maryam Mukhtar [a], Anja Olsen [g,h], Anne Tjønneland [g,i], Kim Overvad [j,k], Matthias Schulze [l], Therese Haugdahl Nøst [m], Guri Skeie [m], Karina Standahl Olsen [m], Fulvio Ricceri [n,o], Sara Grioni [p], Domenico Palli [q], Giovanna Masala [q], Rosario Tumino [r], Fabrizio Pasanisi [s], Pilar Amiano [t,u,v], Sandra M. Colorado Yohar [w,x,y], Antonio Agudo [z,aa], Maria-Jose Sánchez [x,ab,ac,ad], Eva Ardanaz [x,ae,af], Malin Sund [ag,ah], Anne Andersson [ai], 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] International Agency for Research on Cancer (IARC-WHO), Lyon, France
[d] School of Biomedical, Nutritional and Sport Sciences, Newcastle University, Newcastle Upon Tyne, UK
[e] Research on Healthcare Performance (RESHAPE), INSERM U1290, Université Claude Bernard Lyon 1, Lyon, France
[f] Department of Epidemiology, MD Anderson Cancer Centre, Houston, TX, USA
[g] Diet, Genes, and Environment, Danish Cancer Society Research Center, Copenhagen, Denmark
[h] Institute of Public Health, Aarhus University, Aarhus, Denmark
[i] Institute of Public Health, University of Copenhagen, Copenhagen, Denmark
[j] Department of Cardiology, Aalborg University Hospital, Aalborg, Denmark
[k] Section for Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark
[l] Department of Molecular Epidemiology, German Institute of Human Nutrition, 14558, Nuthetal, Germany
[m] Department of Community Medicine, UiT the Arctic University of Norway, N-9037, Tromsø, Norway
[n] Department of Clinical and Biological Sciences, University of Turin, Turin, Italy
[o] Unit of Epidemiology, Regional Health Service ASL TO3, Grugliasco, TO, Italy
[p] Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale Dei Tumori di Milano, 20133, Milano, Italy
[q] Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network (ISPRO), Florence, Italy
[r] Hyblean Association for Epidemiological Research, AIRE ONLUS Ragusa, Italy
[s] Departiment Di Medicina Clinica E Chirurgia Federico Ii University, Naples, Italy
[t] Ministry of Health of the Basque Government, Sub Directorate for Public Health and Addictions of Gipuzkoa, San Sebastian, Spain
[u] Biodonostia Health Research Institute, Epidemiology of Chronic and Communicable Diseases Group, San Sebastián, Spain
[v] Spanish Consortium for Research on Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain
[w] Department of Epidemiology, Murcia Regional Health Council, IMIB, Murcia, Spain
[x] Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[y] Research Group on Demography and Health, National Faculty of Public Health, University of Antioquia, Medellín, Colombia
[z] Unit of Nutrition and Cancer, Catalan Institute of Oncology - ICO, L'Hospitalet de Llobregat, Spain
[aa] Nutrition and Cancer Group, Epidemiology, Public Health, Cancer Prevention and Palliative Care Program, Bellvitge Biomedical Research Institute - IDIBELL, L'Hospitalet de Llobregat, Spain
[ab] Escuela Andaluza de Salud Pública (EASP), 18011, Granada, Spain
[ac] Instituto de Investigación Biosanitaria Ibs.GRANADA, 18012, Granada, Spain
[ad] Department of Preventive Medicine and Public Health, University of Granada, 18071, Granada, Spain
[ae] Navarra Public Health Institute, Pamplona, Spain
[af] IdiSNA, Navarra Institute for Health Research, Pamplona, Spain
[ag] Department of Surgery and Perioperative Sciences, Umeå University, Umeå, Sweden
[ah] Department of Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[ai] Department of Radiation Sciences/Oncology, Umeå University, Umeå, Sweden
[aj] Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, OX3 7LF, UK

* Corresponding author. Cancer Biology and Therapeutics Group, School of Biomolecular and Biomedical Science, UCD Conway Institute, University College Dublin, D04 V1W8, Dublin, Ireland.
E-mail address: david.hughes@ucd.ie (D.J. Hughes).

https://doi.org/10.1016/j.freeradbiomed.2023.10.401
Received 2 August 2023; Received in revised form 10 October 2023; Accepted 19 October 2023
Available online 1 November 2023

[ak] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK

ARTICLE INFO

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: 0.30–0.98, $P_{trend} = 0.013$). While none of the genetic variant associations ($P \leq 0.05$) 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 $P \leq 0.01$. Fourteen SNPs in twelve Se pathway genes ($P \leq 0.01$) 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
$P_{heterogeneity}$	*P*-value for heterogeneity
PPARγ	peroxisome proliferator-activated receptor γ
PR	progesterone receptor
Pro	Proline
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
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, Billingstad,

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. Replicate 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 Fedirko et al (but restricted to the primary pathway 1 and 2 genes in this study) [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 Haploview 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 ($r^2 \geq 0.8$). 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 r^2 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 (FTP) (yes/no), number of FTPs and age at first FTP. 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 Benjamini–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 ($R_s = 0.60$, as expected in a non-Se replete population, and moderately correlated with GPX3 ($R_s = 0.38$), while GPX3 and SELENOP were moderately correlated with each other ($R_s = 0.33$). 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
Selected characteristics of the study participants, EPIC study.

Baseline characteristic[a]	Cases	Controls
N	N = 2208	N = 2208
Age at blood collection, years	54.6 (7.7)	54.6 (7.6)
Educational attainment, % (157 missing)		
None/Primary	34	35
Technical/professional school	25	25
Secondary	19	19
University degree	18	18
Smoking status, % (88 missing)		
Never smoker	55	55
Former smoker	23	23
Current smoker	20	20
Physical activity, % (59 missing)		
Inactive	22	20
Moderately inactive	37	36
Moderately active	23	24
Active	17	20
METS recreational and household activity (309 missing)	92.9 (52.7)	95.4 (52.5)
Menopausal status, %		
Premenopausal	21	21
Postmenopausal, aged 55+	50	50
Postmenopausal, aged 45-54	28	28
Postmenopausal, aged <45	1	1
Hormonal therapy use (ever), % (250 missing)	58	58
Age at first menstrual period, yrs (191 missing)	13.1 (1.6)	13.2 (1.6)
Number of full-term pregnancies, (5 missing)	1.9 (1.2)	2.0 (1.2)
Age at first full-term pregnancy, yrs (17 missing)	25.3 (4.2)	24.8 (4.2)
BMI, kg/m^2	25.3 (4.5)	25.0 (3.9)
Height, cm	162.4 (6.5)	162.0 (6.4)
Baseline dietary intakes		
Total energy, kcal/d	1960.5 (536.0)	1935.5 (543.6)
Alcohol, g/d	10.1 (13.3)	9.6 (12.9)
Calcium, mg/d	1012.1 (390.2)	999.5 (385.4)
Fibre, g/d	22.2 (7.6)	22.1 (7.2)
Folate, g/d	294.5 (109.2)	294.4 (109.3)
Fruit, nuts, seeds and vegetables, g/d	442.3 (244.9)	450.0 (245.7)
Fruit and vegetables, g/d	435.6 (243.1)	443.5 (243.3)
Fish and shellfish, g/d	32.6 (26.7)	32.9 (27.4)
Red and processed meat, g/d	71.1 (39.7)	68.5 (39.6)
Baseline plasma biomarkers, geometric mean (SD)		
Glutathione peroxidase 3 (GPX3; U/L) (924 missing)	244.9 (58.0)	246.8 (56.4)
Selenium, μg/L (846 missing)	65.3 (21.6)	64.7 (20.8)
Selenoprotein P, mg/L (4 missing)	5.0 (1.2)	5.0 (1.1)
Tumour characteristics, % missing		
ER+/− (0 missing)	78.80/21.20	
PR+/− (635 missing)	46.60/24.60	
HER2+/− (1510 missing)	6.80/24.80	
Ductal/Lobular (314 missing)	71.4/14.4	

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

[a] 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 ($OR_{Q5\ vs.\ Q1}$) of Se, SELENOP, and GPX3 levels with BC risk were 1.12 (95 % CI: 0.84–1.48, $P_{trend} = 0.59$), 0.89 (95 % CI: 0.72–1.10, $P_{trend} = 0.42$), and 0.79 (95 % CI: 0.60–1.04, $P_{trend} = 0.071$). However, GPX3 activity was statistically significantly inversely associated with BC risk in the analysis including matching factors only ($OR_{Q5\ vs.\ Q1} = 0.75$; 95 % CI: 0.57–0.98; $P_{trend} = 0.028$). While the point estimates remained similar, the *P*-values lost significance in the partially adjusted ($OR_{Q5\ vs.\ Q1} = 0.78$; 95 % CI: 0.59–1.03; $P_{trend} = 0.058$) and fully adjusted ($OR_{Q5\ vs.\ Q1}$ given above) multivariable analyses. No statistical significance was observed for this biomarker in the continuous analysis either, where the risk of BC was assessed on the continuous log-transformed GPX3 activity ($OR_{fully\ adjusted} = 0.79$; 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 subsite (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: $OR_{Q4\ vs.\ Q1} = 0.54$ (0.30–0.98), $P_{trend} = 0.013$, and OR = 0.25 (95 % CI: 0.09–0.69) $P = 0.007$, respectively (Table 3). The *P*-values for heterogeneity between pre- and post-menopausal 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 ($OR_{Q4\ vs\ Q1} = 0.72$; 95 % CI: 0.53–0.98; $P_{trend} = 0.043$). GPX3 activity appeared to have no association with BC risk in women using exogenous hormones ($P_{heterogeneity} = 0.60$).

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+ ($OR_{Q4\ vs\ Q1} = 0.63$; 95 % CI: 0.42–0.95; $P_{trend} = 0.022$) and PR– BC ($OR_{Q4\ vs\ Q1} = 0.55$; 95 % CI: 0.32–0.94; $P_{trend} = 0.044$), with no evidence for heterogeneity ($P_{heterogeneity} = 0.53$).

Sensitivity analyses were conducted by examining subgroups at $</\geq$ 2 years and $</\geq$ 5 years between blood collection and diagnosis (see Supplementary Table S8). There were no substantive changes observed for the results in subgroups $</\geq$ 2 years. However, higher Se concentrations were associated with higher BC risk in women diagnosed more than 5 years after baseline ($OR_{Q4\ vs\ Q1} = 1.58$; 95 % CI: 1.06–2.37; $P_{trend} = 0.044$), but not in those diagnosed within 5 years ($P_{heterogeneity} = 0.087$). As menopausal status was defined at blood collection (and not at diagnosis), we performed a further sensitivity analysis using age of

Table 2
ORs (95 % CIs) for breast cancer risk in association with concentrations of circulating selenium and selenoprotein P, and with glutathione peroxidase 3 activity, EPIC study.

		Matching Factors[a]	Fully adjusted[b]
Plasma Biomarker	Cases/ Controls	OR (95 % CI)	OR (95 % CI)
Selenium, μg/L			
<49.02	348/357	1.00 (ref.)	1.00 (ref.)
49.02–57.76	382/358	1.11 (0.88–1.40)	1.12 (0.88–1.42)
57.77–65.78	329/356	0.96 (0.75–1.22)	0.97 (0.75–1.24)
65.79–77.66	363/356	1.06 (0.82–1.37)	1.09 (0.85–1.42)
≥77.67	361/356	1.06 (0.81–1.40)	1.12 (0.84–1.48)
P_{trend}		*0.86*	*0.59*
Continuous log-transformed variable		1.19 (0.89–1.59)	1.23 (0.92–1.66)
Selenoprotein P, mg/L			
<4.10	523/497	1.00 (ref.)	1.00 (ref.)
4.10–4.60	422/432	0.92 (0.77–1.11)	0.92 (0.76–1.11)
4.61–5.10	430/448	0.91 (0.75–1.09)	0.90 (0.75–1.09)
5.11–5.80	434/421	0.97 (0.80–1.17)	0.96 (0.79–1.16)
≥5.81	395/406	0.91 (0.74–1.13)	0.89 (0.72–1.10)
P_{trend}		*0.539*	*0.416*
Continuous log-transformed variable		1.02 (0.76–1.38)	0.98 (0.72–1.34)
Glutathione peroxidase 3, U/L			
<202.40	377/350	1.00 (ref.)	1.00 (ref.)
202.41–229.80	390/349	1.00 (0.80–1.25)	1.02 (0.81–1.27)
229.81–252.60	294/347	0.74 (0.58–0.94)	0.77 (0.60–0.98)
252.61–289.10	364/349	0.89 (0.70–1.13)	0.90 (0.70–1.15)
≥289.11	318/348	0.75 (0.57–0.98)	0.79 (0.60–1.04)
P_{trend}		*0.028*	*0.071*
Continuous log-transformed variable		0.73 (0.50–1.07)	0.79 (0.54–1.18)

Abbreviations: BMI = body mass index; OR = odds ratio; CI = confidence interval; ref = reference; FTP = full-term pregnancy; SD = standard deviation.

[a] Model based on matching factors only (study centre, age, time at blood collection, fasting status, menopausal status and exogenous hormone use at blood collection. Premenopausal women were also matched according to their phase of menstrual cycle.

[b] Model based on matching factors plus additional adjustments for smoking status, BMI, physical activity, education level, energy intake, calcium intake, fruit & vegetable intake, red & processed meat intake, alcohol intake, FTP (yes/no), number of FTP, age at first FTP

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 these 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 *TXN2* 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 *TXNRD1* 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 ≤0.0001 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*), where both analyses showed inverse associations.

Table 3
Odds ratios (ORs) and 95 % confidence intervals (95 % CI) for breast cancer risk by quartiles of plasma selenium, selenoprotein P and glutathione peroxidase 3 concentrations by menopausal status, EPIC study.

	Pre-menopause				Post-menopause			
		Matching factors[a]	Multivariate adjusted[b]			Matching factors[a]	Multivariate adjusted[b]	
Plasma Biomarker	Ca/Co	OR (95 % CI)	OR (95 % CI)	Plasma Biomarker	Ca/Co	OR (95 % CI)	OR (95 % CI)	$P_{heterogeneity}$
Selenium, μg/L				Selenium, μg/L				
23.86–48.88	83/83	1.00 (ref.)	1.00 (ref.)	6.50–52.30	372/364	1.00 (ref.)	1.00 (ref.)	*0.40*
48.89–58.50	81/82	0.98 (0.58–1.67)	0.91 (0.51–1.62)	52.31–62.01	343/363	0.92 (0.74–1.15)	0.95 (0.76–1.20)	
58.51–69.10	87/83	1.05 (0.60–1.84)	0.97 (0.53–1.77)	62.02–74.92	359/364	0.97 (0.76–1.24)	0.99 (0.77–1.26)	
≥69.11	78/81	0.95 (0.53–1.72)	0.90 (0.48–1.71)	≥74.93	380/363	1.06 (0.81–1.39)	1.12 (0.85–1.49)	
P_{trend}		*0.926*	*0.831*	P_{trend}		*0.638*	*0.424*	
Continuous log-transformed variable		1.19 (0.57–2.50)	1.27 (0.57–2.83)	Continuous log-transformed variable		1.19 (0.86–1.63)	1.23 (0.89–1.70)	
Selenoprotein P, mg/L				Selenoprotein P, mg/L				
2.10–3.90	114/122	1.00 (ref.)	1.00 (ref.)	1.10–4.30	471/474	1.00 (ref.)	1.00 (ref.)	*0.67*
3.91–4.55	125/110	1.23 (0.85–1.79)	1.51 (1.00–2.28)	4.31–4.90	428/427	1.01 (0.84–1.22)	1.02 (0.84–1.23)	
4.56–5.15	119/116	1.09 (0.76–1.57)	1.16 (0.78–1.72)	4.91–5.70	438/440	1.00 (0.83–1.21)	0.99 (0.81–1.20)	
≥5.16	106/116	0.97 (0.65–1.43)	1.04 (0.68–1.59)	≥5.71	403/399	1.02 (0.82–1.27)	1.00 (0.80–1.25)	
P_{trend}		*0.777*	*0.834*	P_{trend}		*0.877*	*0.946*	
Continuous log-transformed variable		0.85 (0.43–1.67)	0.84 (0.41–1.74)	Continuous log-transformed variable		1.07 (0.76–1.49)	1.03 (0.73–1.45)	
Glutathione peroxidase 3, U/L				Glutathione peroxidase 3, U/L				
47.70–210.55	100/79	1.00 (ref.)	1.00 (ref.)	75.10–209.10	373/357	1.00 (ref.)	1.00 (ref.)	*0.19*
210.56–238.65	83/79	0.76 (0.48–1.23)	0.73 (0.44–1.22)	209.11–241.10	354/357	0.93 (0.74–1.17)	0.97 (0.77–1.23)	
238.66–269.05	64/79	0.59 (0.36–0.95)	0.50 (0.29–0.86)	241.11–279.60	366/357	0.94 (0.74–1.20)	0.98 (0.77–1.26)	
≥269.06	69/79	0.58 (0.34–0.99)	0.54 (0.30–0.98)	≥279.61	334/356	0.84 (0.64–1.10)	0.90 (0.68–1.19)	
P_{trend}		*0.02*	*0.013*	P_{trend}		*0.263*	*0.516*	
Continuous log-transformed variable		0.30 (0.12–0.76)	0.25 (0.09–0.69)	Continuous log-transformed variable		0.90 (0.59–1.38)	0.98 (0.63–1.52)	

Abbreviations: Ca = cases; Co = controls; OR = odds ratio; CI = confidence interval; ref = reference; FTP = full-term pregnancy.
[a] Model based on matching factors only.
[b] Model based on matching factors plus additional adjustments for BMI, height, smoking status, physical activity, energy intake, calcium intake, alcohol intake, consumption of red & processed meat, consumption of fruit and veg, education level, FTP, number of FTP, age at first FTP

4. Discussion

As far as we are aware, we have conducted the largest reported prospective investigation 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 higher GPX3 activity and lower BC risk was observed, especially in premenopausal women. There was some limited evidence that genetic variation in Se pathway genes alone or in interaction with Se status may affect BC development, although none of the findings retained significance 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 with BC recurrence and mortality [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

Table 4
All selenium pathway SNPs (by gene and pathway) significantly associated with overall breast cancer risk before multiple testing corrections, EPIC study.

Gene	SNP	Reference Allele	Variant Allele	OR (95 % CI)[a]	*P-value*[b]
Pathway 1 = Selenoproteins and Selenium biosynthesis and transport					
GPX6[c]	rs974334	G	C	0.88 (0.77–1.00)	0.04
LRP2[d]	rs700552	G	C	0.88 (0.79–0.98)	0.02
SELENOH[c]	rs527140	A	G	1.14 (1.01–1.29)	0.04
SELENOM[c]	rs1004243	G	A	0.88 (0.80–0.97)	0.01
SELENOS[c]	rs13329318	A	C	1.16 (1.00–1.34)	0.05
SEPHS1[d]	rs2275129	G	C	1.11 (1.00–1.23)	0.05
TXNRD3[c]	rs777226	G	A	1.15 (0.01–1.29)	0.03
Pathway 2 = Antioxidants and Redox function genes					
CAT	rs1049982	G	A	0.90 (0.80–1.00)	0.05
CAT	rs11032700	A	C	0.88 (0.79–0.98)	0.02
CAT	rs4755374	A	C	1.16 (1.00–1.35)	0.05
GPX5	rs451774	A	G	0.89 (0.80–0.99)	0.03
GSR	rs8190996	G	A	1.13 (1.01–1.26)	0.03
HIF1A	rs11549465	G	A	1.18 (1.02–1.38)	0.03
RPS6KB1	rs143652	A	G	1.15 (1.01–1.32)	0.04
SOD2	rs4880	T	G	1.11 (1.00–1.23)	0.05
TXN2	rs4821494	A	C	1.19 (1.06–1.33)	0.003
TXN2	rs5750261	G	A	1.18 (1.04–1.36)	0.01
GWAS Selenium concentration-associated SNPs					
CBS	rs234709	C	T	1.12 (1.01–1.25)	0.03

SNP data was extracted from GWAS data on 1564 cases and 1564 controls. Except for rs4880 in *SOD2* (taken from previous genotyping of 1462 cases and 1462 controls [42]).

[a] Associations with breast cancer risk (conditional logistic regression). Model based on matching factors only (study centre, age, time at blood collection, fasting status, menopausal status and exogenous hormone use at blood collection. Premenopausal women were also matched according to their phase of menstrual cycle.

[b] None retained significance following adjustment by Benjamini-Hochberg multiple testing correction.

[c] Selenoprotein gene.

[d] Selenium biosynthesis gene.

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 post-menopausal 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 Méplan 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 (PPARγ), as demonstrated by knock-down or activation/overexpression of PPARγ, 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 lipoxygenases 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 rs974334 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 (rs1004243) in the *SELENOM* selenoprotein gene and 2 variants (rs4821494 & 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 countering 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 5
Se pathway SNPs (by gene and pathway) associated with breast cancer risk when Se status biomarkers are above and below their respective median levels (for those with at least one *P*-value ≤0.01, before multiple testing correction), EPIC study.

		Se		SELENOP		GPX3	
Gene	SNP	< Median	≥ Median	< Median	≥ Median	< Median	≥ Median
Pathway 1 = Selenoproteins and Se biosynthesis and transport							
GPX3[a]	rs10463312	*1.17 (1.01-1.36); 0.04*	*0.76 (0.62-0.94); 0.01*	1.11 (0.91–1.36); 0.30	0.85 (0.70–1.04); 0.11	1.06 (0.92–1.22); 0.45	1.02 (0.82–1.28); 0.84
GPX3[a]	rs2070593	*0.77 (0.64-0.93); 0.007*	1.26 (0.97–1.63); 0.09	*0.76 (0.59-0.97); 0.03*	1.22 (0.96–1.57); 0.11	0.85 (0.70–1.02); 0.08	0.92 (0.70–1.21); 0.55
GPX6[a]	rs434112	*0.81 (0.70-0.95); 0.007*	1.14 (0.91–1.42); 0.26	0.85 (0.70–1.04); 0.12	1.01 (0.83–1.22); 0.96	*0.85 (0.74-0.98); 0.03*	0.84 (0.68–1.06); 0.14
SCLY[b]	rs1562337	0.99 (0.85–1.15); 0.87	1.08 (0.89–1.31); 0.46	1.15 (0.95–1.41); 0.16	*0.78 (0.64-0.95); 0.01*	1.01 (0.87–1.17); 0.91	0.90 (0.72–1.11); 0.30
SECISBP2[b]	rs4876978	0.92 (0.78–1.08); 0.30	0.94 (0.76–1.17); 0.60	1.05 (0.86–1.29); 0.61	*0.75 (0.60-0.93); 0.008*	0.96 (0.83–1.12); 0.64	0.98 (0.79–1.21); 0.83
SELENBP1[b]	rs17564336	0.93 (0.79–1.08); 0.33	*1.34 (1.08-1.68); 0.009*	0.96 (0.79–1.17); 0.70	1.08 (0.88–1.33); 0.44	0.95 (0.82–1.11); 0.53	*1.38 (1.10-1.73); 0.005*
SELENBP1[b]	rs2864118	0.91 (0.79–1.06); 0.21	*1.35 (1.10-1.68); 0.005*	0.97 (0.80–1.17); 0.74	1.11 (0.91–1.35); 0.30	0.92 (0.80–1.07); 0.28	*1.30 (1.05-1.61); 0.02*
TXNRD1[a]	rs75436938	0.79 (0.61–1.03); 0.09	1.09 (0.76–1.58); 0.63	*0.62 (0.44-0.88); 0.008*	0.97 (0.68–1.38); 0.84	0.91 (0.70–1.18); 0.49	0.77 (0.54–1.11); 0.17
TXNRD2[a]	rs3788314	0.91 (0.79–1.06); 0.23	0.95 (0.77–1.18); 0.65	0.95 (0.78–1.17); 0.62	1.05 (0.87–1.27); 0.63	1.05 (0.90–1.22); 0.56	*0.76 (0.62-0.93); 0.009*
TXNRD3[a]	rs9637365	0.97 (0.84–1.12); 0.68	0.87 (0.72–1.06); 0.18	1.03 (0.85–1.24); 0.77	*0.71 (0.58-1.86); 0.001*	*0.85 (0.74-0.98); 0.02*	*1.23 (1.00-1.52); 0.05*
Pathway 2 = Antioxidants and Redox							
GPX5	rs451774	*0.79 (0.67-0.93); 0.005*	1.11 (0.89–1.39); 0.34	0.85 (0.70–1.04); 0.12	1.00 (0.82–1.21); 0.96	*0.82 (0.71-0.95); 0.009*	*0.89 (0.71-1.11); 0.30*
RPS6KB1	rs7217337	0.90 (0.70–1.16); 0.41	0.77 (0.56–1.07); 0.13	1.00 (0.71–1.40); 0.98	*0.60 (0.44-0.83); 0.002*	1.00 (0.77–1.28); 0.98	0.75 (0.52–1.09); 0.14
TXN	rs4135212	1.02 (0.83–1.26); 0.84	0.96 (0.72–1.29); 0.80	0.99 (0.75–1.31); 0.93	1.11 (0.85–1.45); 0.44	1.17 (0.94–1.44); 0.16	*0.57 (0.41-0.80); 0.001*
TXN2	rs11089790	0.92 (0.73–1.14); 0.43	*0.61 (0.43-0.86); 0.005*	0.82 (0.61–1.11); 0.20	0.81 (0.60–1.09); 0.16	1.02 (0.82–1.28); 0.84	*0.65 (0.46-0.92); 0.02*

Associations with breast cancer risk (conditional logistic regression).
Results given as Odds Ratio (95 % Confidence Interval); *P*-value (those in italics are at ≤ 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.
>=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 equal to or above their respective medians.
[a] Selenoprotein gene.
[b] Se biosynthesis gene.

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 between the interaction of rs1050450 (*GPX1)* and rs1799725 (now = rs4880) in *SOD2* with increased BC risk [66]. Méplan et al. (2013) reported in The Danish Diet, Cancer, and Health cohort (building on their previous 2006 study) [45] that rs1050450 was associated with an increased risk of non-ductal BC only, and as described above, that there may be an interaction effect between the variant and GPX1 activity impacting BC development [13]. Subsequently, the large, international BPC3 study, which included approximately 400 ER- BC case-control pairs included in this study, found no significant interaction between rs1050450 (*GPX1)* and BC risk in postmenopausal women, or in interaction with rs4880 in *SOD2* (although a lower prostate cancer risk in rs1050450 carriers was observed) [42]. There was also no association between rs1050450 and BC risk in the current analysis (although rs4880 was just at the significance threshold; $P = 0.05$) or, as investigated here for the first time, in interaction with the three measured biomarkers of Se status. Finally, in a large UK study, limited evidence was found to link common variation in *SOD1, SOD2, GPX1, GPX4, GSR, TXNRD1*, and *TXN2* with BC development, except for the missense variant A66S (rs5748469) and tagSNP rs756661 in *TXNRD2* [67].

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 genes, *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

primarily to cancer risk through its antioxidant functions. Here, it harboured two SNPs with nominal *P*-values ≤0.01 for association with PR+ BC risk. While rs451774 in *GPX5* and rs974334 in *GPX6* were the only SNPs in the GPX family associated with overall BC risk, all members (*GPXs* 1–7, of which 1–4 & 6 are selenoprotein genes) harboured at least one SNP nominally significantly associated ($P \leq 0.05$) with least one BC subtype (Supplementary Table S11). This, along with the significant associations with BC risk in premenopausal women, further suggests that these genes may play a key role in BC development (although noting again that GPX6 has not been found to be expressed in breast tumours) [61]. In particular, *GPX2* harboured three SNPs with nominal *P*-values ≤0.01 for association with PR+ BC, while *GPX7* contained two SNPs with nominal *P*-values ≤0.01 for association with HER2- BC. An association of rs7529595 in *GPX7* with increased CRC risk was also observed in our 2019 EPIC study [35]. These observations are compatible with the role of GPXs in BC [58].

Among the other pathway 2 genes, *GSR* encodes glutathione reductase, a key enzyme in redox homeostasis and cellular oxidative stress defense. It has been implicated in cancer initiation and progression due to its antioxidant functions [71]. In a Danish cohort study of 703 BC case-control pairs the functional *GSR* rs1002149 variant affecting enzyme activity was not itself significantly associated with BC risk, but was associated with increased BC risk in interaction with higher alcohol consumption [72]. In this study, only rs8190996 in *GSR* was nominally associated with an increased BC risk ($P = 0.03$; *P* for rs1002149 = 0.08) but there were four SNPs with nominal *P*-values ≤0.01 for association with HER2- BC ($P = 0.0008$ for an increased risk with rs1002149), and two 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 TXRN 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 rs3805435 in *GPX3* and rs2695234 in *SOD3*) were significant when assessed at Se levels both above and below the median; all 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 [8,16].

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 subsites. 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

Conceptualization, D.J.H., L.S. M.J., C.M., V.F., and L.D.; Data curation, D.J.H., L.S., C.B., A.M., and L.D.; Formal analysis, C.B., A.M., L. D., L.S., Q.S., K.M., C.M., and D.J.H.; Funding acquisition, D.J.H., M.J., L.S., V.F., and C.M. Methodology, D.J.H., M.J., L.S., V.F., C.M., C.B., and L.D.; Project administration, D.J.H., L.S., M.J., M.M., and L.D.; Resources, D.J.H., M.J., L.S., L.D., E.W., and all other EPIC co-authors (A. O., A.T., K.O., M.S., T.H.N., G.S., K.S.O., F.R., S.G., D.P., G.M., R.T., F·P., P.A., S.M.C.Y., A.A., M.J.S., E.A., M.S., A.A. (Anne Andersson), A.P.C., R. T. (Ruth Travis), A.K.H.); Supervision, D.J.H., L.S., M.J., and L.D.; Writing—original draft, D.J.H., M.M., M.J., C.M., L.S., V.F., and L.D.; Writing—review and editing, all other EPIC co-authors (E.W., A.O., A.T., K.O., M.S., T.H.N., G.S., K.S.O., F.R., S.G., D.P., G.M., R.T., F.P., P.A., S. M.C.Y., A.A., M.J.S., E.A., M.S., A.A. (Anne Andersson), A.P.C., R.T.

(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, Scho 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'Education 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-Rehbruecke (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/M012190/1 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

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, Ca - Cancer J. Clin. 71 (2021) 209–249.
[2] S. Wu, W. Zhu, P. Thompson, Y.A. Hannun, Evaluating intrinsic and non-intrinsic cancer risk factors, Nat. Commun. 9 (2018) 3490.
[3] S. Łukasiewicz, M. Czeczelewski, A. Forma, J. Baj, R. Sitarz, A. Stanisławek, Breast cancer-epidemiology, risk factors, classification, prognostic markers, and current treatment strategies-an updated review, Cancers 13 (2021).
[4] Q. Xie, M.L. Chen, Y. Qin, Q.Y. Zhang, H.X. Xu, Y. Zhou, M.T. Mi, J.D. Zhu, Isoflavone consumption and risk of breast cancer: a dose-response meta-analysis of observational studies, Asia Pac. J. Clin. Nutr. 22 (2013) 118–127.
[5] T. Norat, D. Aune, D. Chan, D. Romaguera, Fruits and vegetables: updating the epidemiologic evidence for the WCRF/AICR lifestyle recommendations for cancer prevention, Cancer Treat Res. 159 (2014) 35–50.
[6] P.K.S. Mahalingaiah, K.P. Singh, Chronic oxidative stress increases growth and tumorigenic potential of MCF-7 breast cancer cells, PLoS One 9 (2014), e87371.
[7] A.R. Nourazarian, P. Kangari, A. Salmaninejad, Roles of oxidative stress in the development and progression of breast cancer, Asian Pac. J. Cancer Prev. APJCP 15 (2014) 4745–4751.
[8] J.D. Lee, Q. Cai, X.O. Shu, S.J. Nechuta, The role of biomarkers of oxidative stress in breast cancer risk and prognosis: a systematic review of the epidemiologic literature, J. Wom. Health 26 (2017) 467–482, 2002.
[9] S.M. Hamdy, A.K. Latif, E.A. Drees, S.M. Soliman, Prevention of rat breast cancer by genistin and selenium, Toxicol. Ind. Health 28 (2012) 746–757.
[10] X. Cai, C. Wang, W. Yu, W. Fan, S. Wang, N. Shen, P. Wu, X. Li, F. Wang, Selenium exposure and cancer risk: an updated meta-analysis and meta-regression, Sci. Rep. 6 (2016) 1–18.
[11] N. Babaknejad, F. Sayehmiri, K. Sayehmiri, P. Rahimifar, S. Bahrami, A. Delpesheh, F. Hemati, S. Alizadeh, The relationship between selenium levels and breast cancer: a systematic review and meta-analysis, Biol. Trace Elem. Res. 159 (2014) 1–7.
[12] K. Demircan, Y. Bengtsson, Q. Sun, A. Brange, J. Vallon-Christersson, E. Rijntjes, M. Malmberg, L.H. Saal, L. Rydén, Å. Borg, J. Manjer, L. Schomburg, Serum selenium, selenoprotein P and glutathione peroxidase 3 as predictors of mortality and recurrence following breast cancer diagnosis: a multicentre cohort study, Redox Biol. 47 (2021), 102145.
[13] C. Méplan, L.O. Dragsted, G. Ravn-Haren, A. Tjønneland, U. Vogel, J. Hesketh, Association between polymorphisms in glutathione peroxidase and selenoprotein P genes, glutathione peroxidase activity, HRT use and breast cancer risk, PLoS One 8 (2013), e73316.
[14] V.M. Labunskyy, D.L. Hatfield, V.N. Gladyshev, Selenoproteins: molecular pathways and physiological roles, Physiol. Rev. 94 (2014) 739–777.
[15] Y.J. Hu, A.M. Diamond, Role of glutathione peroxidase 1 in breast cancer: loss of heterozygosity and allelic differences in the response to selenium, Cancer Res. 63 (2003) 3347–3351.
[16] Y. Zhang, Y.J. Roh, S.J. Han, I. Park, H.M. Lee, Y.S. Ok, B.C. Lee, S.R. Lee, Role of selenoproteins in redox regulation of signaling and the antioxidant system: a review, Antioxidants 9 (2020).
[17] F.G. Ottaviano, S.-S. Tang, D.E. Handy, J. Loscalzo, Regulation of the extracellular antioxidant selenoprotein plasma glutathione peroxidase (GPx-3) in mammalian cells, Mol. Cell. Biochem. 327 (2009) 111–126.
[18] G.F. Combs Jr., Biomarkers of selenium status, Nutrients 7 (2015) 2209–2236.
[19] M.P. Rayman, M.P. Rayman, The argument for increasing selenium intake, Proc. Nutr. Soc. 61 (2002) 203–215.
[20] M.P. Rayman, Selenium and human health, Lancet 379 (2012) 1256–1268.
[21] C. Méplan, J. Hesketh, Selenium and cancer: a story that should not be forgotten-insights from genomics, Cancer Treat Res. 159 (2014) 145–166.
[22] C. Méplan, Selenium and chronic diseases: a nutritional genomics perspective, Nutrients 7 (2015) 3621–3651.
[23] M. Sandsveden, Y. Bengtsson, O. Melander, A.H. Rosendahl, J. Manjer, Genetic variation interacts with selenium exposure regarding breast cancer risk: assessing dietary intake, serum levels and genetically elevated selenium levels, Nutrients 14 (2022).
[24] M. Sandsveden, J. Manjer, Selenium and breast cancer risk: a prospective nested case-control study on serum selenium levels, smoking habits and overweight, Int. J. Cancer 141 (2017) 1741–1750.
[25] D. Guo, M. Hendryx, X. Liang, J.E. Manson, K. He, M.Z. Vitolins, Y. Li, J. Luo, Association between selenium intake and breast cancer risk: results from the Women's Health Initiative, Breast Cancer Res. Treat. 183 (2020) 217–226.
[26] Y. Bengtsson, M. Sandsveden, J. Manjer, Risk of breast cancer in relation to dietary intake of selenium and serum selenium as a marker of dietary intake: a prospective cohort study within the Malmö Diet and Cancer Study, Cancer Causes Control 32 (2021) 815–826.
[27] L. Schomburg, D.J. Hughes, The missing link? The potential role of selenium in the development of liver cancer and significance for the general population, Expet Rev. Gastroenterol. Hepatol. 11 (2017) 707–709.
[28] D.J. Hughes, V. Fedirko, M. Jenab, L. Schomburg, C. Méplan, H. Freisling, H. B. Bueno-de-Mesquita, S. Hybsier, N.P. Becker, M. Czuban, A. Tjønneland, M. Outzen, et al., Selenium status is associated with colorectal cancer risk in the European prospective investigation of cancer and nutrition cohort, Int. J. Cancer 136 (2015) 1149–1161.
[29] E. Riboli, R. Kaaks, The EPIC project: rationale and study design. European prospective investigation into cancer and nutrition, Int. J. Epidemiol. 26 (Suppl 1) (1997) S6–S14.

[30] E. Riboli, K.J. Hunt, N. Slimani, P. Ferrari, T. Norat, M. Fahey, U.R. Charrondière, B. Hémon, C. Casagrande, J. Vignat, K. Overvad, A. Tjønneland, et al., European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection, Publ. Health Nutr. 5 (2002) 1113–1124.
[31] E. Caboux, C. Lallemand, G. Ferro, B. Hémon, M. Mendy, C. Biessy, M. Sims, N. Wareham, A. Britten, A. Boland, Sources of pre-analytical variations in yield of DNA extracted from blood samples: analysis of 50,000 DNA samples in EPIC, PLoS One 7 (2012), e39821.
[32] D.J. Hughes, T. Duarte-Salles, S. Hybsier, A. Trichopoulou, M. Stepien, K. Aleksandrova, K. Overvad, A. Tjønneland, A. Olsen, A. Affret, G. Fagherazzi, M. C. Boutron-Ruault, et al., Prediagnostic selenium status and hepatobiliary cancer risk in the European Prospective Investigation into Cancer and Nutrition cohort, Am. J. Clin. Nutr. 104 (2016) 406–414.
[33] L. Flohé, W.A. Günzler, Assays of glutathione peroxidase, Methods Enzymol. 105 (1984) 114–121.
[34] J. Hoeflich, B. Hollenbach, T. Behrends, A. Hoeg, H. Stosnach, L. Schomburg, The choice of biomarkers determines the selenium status in young German vegans and vegetarians, Br. J. Nutr. 104 (2010) 1601–1604.
[35] V. Fedirko, M. Jenab, C. Méplan, J.S. Jones, W. Zhu, L. Schomburg, A. Siddiq, S. Hybsier, K. Overvad, A. Tjønneland, Association of selenoprotein and selenium pathway genotypes with risk of colorectal cancer and interaction with selenium status, Nutrients 11 (2019) 935.
[36] D.M. Evans, G. Zhu, V. Dy, A.C. Heath, P.A. Madden, J.P. Kemp, G. McMahon, B. St Pourcain, N.J. Timpson, J. Golding, Genome-wide association study identifies loci affecting blood copper, selenium and zinc, Hum. Mol. Genet. 22 (2013) 3998–4006.
[37] J. Gong, L. Hsu, T. Harrison, I.B. King, S. Stürup, X. Song, D. Duggan, Y. Liu, C. Hutter, S.J. Chanock, C.B. Eaton, J.R. Marshall, et al., Genome-wide association study of serum selenium concentrations, Nutrients 5 (2013) 1706–1718.
[38] M.C. Cornelis, M. Fornage, M. Foy, P. Xun, V.N. Gladyshev, S. Morris, D. I. Chasman, F.B. Hu, E.B. Rimm, P. Kraft, Genome-wide association study of selenium concentrations, Hum. Mol. Genet. 24 (2015) 1469–1477.
[39] J. Mao, J.J. Vanderlelie, A.V. Perkins, C.W. Redman, K.R. Ahmadi, M.P. Rayman, Genetic polymorphisms that affect selenium status and response to selenium supplementation in United Kingdom pregnant women, Am. J. Clin. Nutr. 103 (2016) 100–106.
[40] A. Siddiq, F.J. Couch, G.K. Chen, S. Lindström, D. Eccles, R.C. Millikan, K. Michailidou, D.O. Stram, L. Beckmann, S.K. Rhie, C.B. Ambrosone, K. Aittomäki, et al., A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11, Hum. Mol. Genet. 21 (2012) 5373–5384.
[41] G. Fehringer, P. Kraft, P.D. Pharoah, R.A. Eeles, N. Chatterjee, F.R. Schumacher, J. M. Schildkraut, S. Lindström, P. Brennan, H. Bickeböller, R.S. Houlston, M. T. Landi, et al., Cross-cancer genome-wide analysis of lung, ovary, breast, prostate, and colorectal cancer reveals novel pleiotropic associations, Cancer Res. 76 (2016) 5103–5114.
[42] S. Blein, S. Berndt, A.D. Joshi, D. Campa, R.G. Ziegler, E. Riboli, D.G. Cox, Factors associated with oxidative stress and cancer risk in the breast and prostate cancer cohort Consortium, Free Radic. Res. 48 (2014) 380–386.
[43] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. Roy. Stat. Soc. B 57 (1995) 289–300.
[44] R. Hurst, C.N. Armah, J.R. Dainty, D.J. Hart, B. Teucher, A.J. Goldson, M. R. Broadley, A.K. Motley, S.J. Fairweather-Tait, Establishing optimal selenium status: results of a randomized, double-blind, placebo-controlled trial, Am. J. Clin. Nutr. 91 (2010) 923–931.
[45] G. Ravn-Haren, A. Olsen, A. Tjønneland, L.O. Dragsted, B.A. Nexø, H. Wallin, K. Overvad, O. Raaschou-Nielsen, U. Vogel, Associations between GPX1 Pro198Leu polymorphism, erythrocyte GPX activity, alcohol consumption and breast cancer risk in a prospective cohort study, Carcinogenesis 27 (2006) 820–825.
[46] H.R. Harris, L. Bergkvist, A. Wolk, Selenium intake and breast cancer mortality in a cohort of Swedish women, Breast Cancer Res. Treat. 134 (2012) 1269–1277.
[47] J. Lubinski, W. Marciniak, M. Muszynska, T. Huzarski, J. Gronwald, C. Cybulski, A. Jakubowska, T. Debniak, M. Falco, J. Kladny, J. Kotsopoulos, P. Sun, et al., Serum selenium levels predict survival after breast cancer, Breast Cancer Res. Treat. 167 (2018) 591–598.
[48] M. Sandsveden, E. Nilsson, S. Borgquist, A.H. Rosendahl, J. Manjer, Prediagnostic serum selenium levels in relation to breast cancer survival and tumor characteristics, Int. J. Cancer 147 (2020) 2424–2436.
[49] P. van 't Veer, R.P. van der Wielen, F.J. Kok, R.J. Hermus, F. Sturmans, Selenium in diet, blood, and toenails in relation to breast cancer: a case-control study, Am. J. Epidemiol. 131 (1990) 987–994.
[50] P.A. van den Brandt, R.A. Goldbohm, P. van't Veer, P. Bode, E. Dorant, R. J. Hermus, F. Sturmans, Toenail selenium levels and the risk of breast cancer, Am. J. Epidemiol. 140 (1994) 20–26.
[51] K. Charalabopoulos, A. Kotsalos, A. Batistatou, A. Charalabopoulos, P. Vezyraki, D. Peschos, V. Kalfakakou, A. Evangelou, Selenium in serum and neoplastic tissue in breast cancer: correlation with CEA, Br. J. Cancer 95 (2006) 674–676.
[52] M. Vinceti, T. Filippini, C. Del Giovane, G. Dennert, M. Zwahlen, M. Brinkman, M. P. Zeegers, M. Horneber, R. D'Amico, C.M. Crespi, Selenium for preventing cancer, Cochrane Database Syst. Rev. 1 (2018) Cd005195.
[53] W. Lou, B. Ding, S. Wang, P. Fu, Overexpression of GPX3, a potential biomarker for diagnosis and prognosis of breast cancer, inhibits progression of breast cancer cells in vitro, Cancer Cell Int. 20 (2020) 378.
[54] L. Lundholm, M. Putnik, M. Otsuki, S. Andersson, C. Ohlsson, J.A. Gustafsson, K. Dahlman-Wright, Effects of estrogen on gene expression profiles in mouse hypothalamus and white adipose tissue: target genes include glutathione peroxidase 3 and cell death-inducing DNA fragmentation factor, alpha-subunit-like effector A, J. Endocrinol. 196 (2008) 547–557.
[55] A.T. Reddy, S.P. Lakshmi, A. Banno, R.C. Reddy, Role of GPx3 in PPARγ-induced protection against COPD-associated oxidative stress, Free Radic. Biol. Med. 126 (2018) 350–357.
[56] R. Brigelius-Flohé, Tissue-specific functions of individual glutathione peroxidases, Free Radic. Biol. Med. 27 (1999) 951–965.
[57] C. Rock, P.J. Moos, Selenoprotein P protects cells from lipid hydroperoxides generated by 15-LOX-1, Prostaglandins Leukot. Essent. Fatty Acids 83 (2010) 203–210.
[58] W. Lou, B. Ding, S. Wang, P. Fu, Overexpression of GPX3, a potential biomarker for diagnosis and prognosis of breast cancer, inhibits progression of breast cancer cells in vitro, Cancer Cell Int. 20 (2020) 378.
[59] M.M. Mohamed, S. Sabet, D.-F. Peng, M.A. Nouh, M. El-Shinawi, W. El-Rifai, Promoter hypermethylation and suppression of glutathione peroxidase 3 are associated with inflammatory breast carcinogenesis, Oxid. Med. Cell. Longev. 2014 (2014), 787195.
[60] P. Saelee, T. Pongtheerat, T. Sophonnithiprasert, Reduced expression of GPX3 in breast cancer patients in correlation with clinical significance, Global Med. Genetics 7 (2020) 87–91.
[61] K. Demircan, Y. Bengtsson, T.S. Chillon, J. Vallon-Christersson, Q. Sun, C. Larsson, M. Malmberg, L.H. Saal, L. Rydén, Å. Borg, J. Manjer, L. Schomburg, Matched analysis of circulating selenium with the breast cancer selenotranscriptome: a multicentre prospective study, J. Transl. Med. 21 (1) (2023) 658.
[62] M.W. Pitts, P.R. Hoffmann, Endoplasmic reticulum-resident selenoproteins as regulators of calcium signaling and homeostasis, Cell Calcium 70 (2018) 76–86.
[63] E. Holzerova, K. Danhauser, T.B. Haack, L.S. Kremer, M. Melcher, I. Ingold, S. Kobayashi, C. Terrile, P. Wolf, J. Schaper, E. Mayatepek, F. Baertling, et al., Human thioredoxin 2 deficiency impairs mitochondrial redox homeostasis and causes early-onset neurodegeneration, Brain 139 (2015) 346–354.
[64] P. Seibold, R. Hein, P. Schmezer, P. Hall, J. Liu, N. Dahmen, D. Flesch-Janys, O. Popanda, J. Chang-Claude, Polymorphisms in oxidative stress-related genes and postmenopausal breast cancer risk, Int. J. Cancer 129 (2011) 1467–1476.
[65] J. Lu, A. Holmgren, Thioredoxin system in cell death progression, Antioxidants Redox Signal. 17 (2012) 1738–1747.
[66] D.G. Cox, R.M. Tamimi, D.J. Hunter, Gene x Gene interaction between MnSOD and GPX-1 and breast cancer risk: a nested case-control study, BMC Cancer 6 (2006) 217.
[67] A. Cebrian, P.D. Pharoah, S. Ahmed, P.L. Smith, C. Luccarini, R. Luben, K. Redman, H. Munday, D.F. Easton, A.M. Dunning, B.A. Ponder, Tagging single-nucleotide polymorphisms in antioxidant defense enzymes and susceptibility to breast cancer, Cancer Res. 66 (2006) 1225–1233.
[68] A.J. Pellatt, R.K. Wolff, E.M. John, G. Torres-Mejia, L.M. Hines, K.B. Baumgartner, A.R. Giuliano, A. Lundgreen, M.L. Slattery, SEPP1 influences breast cancer risk among women with greater native american ancestry: the breast cancer health disparities study, PLoS One 8 (2013), e80554.
[69] D.T. Lincoln, E.M. Ali Emadi, K.F. Tonissen, F.M. Clarke, The thioredoxin-thioredoxin reductase system: over-expression in human cancer, Anticancer Res. 23 (2003) 2425–2433.
[70] M. Reeves, P. Hoffmann, The human selenoproteome: recent insights into functions and regulation, Cell. Mol. Life Sci. 66 (2009) 2457–2478.
[71] N. Couto, J. Wood, J. Barber, The role of glutathione reductase and related enzymes on cellular redox homoeostasis network, Free Radic. Biol. Med. 95 (2016) 27–42.
[72] T.I. Kopp, U. Vogel, L.O. Dragsted, A. Tjonneland, G. Ravn-Haren, Association between single nucleotide polymorphisms in the antioxidant genes CAT, GR and SOD1, erythrocyte enzyme activities, dietary and life style factors and breast cancer risk in a Danish, prospective cohort study, Oncotarget 8 (2017) 62984–62997.
[73] C. Méplan, F. Nicol, B.T. Burtle, L.K. Crosley, J.R. Arthur, J.C. Mathers, J. E. Hesketh, Relative abundance of selenoprotein P isoforms in human plasma depends on genotype, se intake, and cancer status, Antioxidants Redox Signal. 11 (2009) 2631–2640.
[74] C. Méplan, L.K. Crosley, F. Nicol, G.J. Beckett, A.F. Howie, K.E. Hill, G. Horgan, J. C. Mathers, J.R. Arthur, J.E. Hesketh, Genetic polymorphisms in the human selenoprotein P gene determine the response of selenoprotein markers to selenium supplementation in a gender-specific manner (the SELGEN study), Faseb. J. 21 (2007) 3063–3074.
[75] N. Karunasinghe, D.Y. Han, S. Zhu, J. Yu, K. Lange, H. Duan, R. Medhora, N. Singh, J. Kan, W. Alzaher, B. Chen, S. Ko, 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. 7 (2012) 179–190.
[76] A.L. Van Dyke, M.L. Cote, A.S. Wenzlaff, J. Abrams, S. Land, P. Iyer, A.G. Schwartz, Chromosome 5p region SNPs are associated with risk of NSCLC among women, J. Cancer Epidemiol. 2009 (2009), 242151.
[77] D.N. Ekoue, S. Zaichick, K. Valyi-Nagy, M. Picklo, C. Lacher, K. Hoskins, M. A. Warso, M.G. Bonini, A.M. Diamond, 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. 39 (2017) 227–233.
[78] S. Mohammaddoust, Z. Salehi, H. Saeidi Saedi, SEPP1 and SEP15 gene polymorphisms and susceptibility to breast cancer, Br. J. Biomed. Sci. 75 (2018) 36–39.
[79] M. Udler, A.T. Maia, A. Cebrian, C. Brown, D. Greenberg, M. Shah, C. Caldas, A. Dunning, D. Easton, B. Ponder, P. Pharoah, Common germline genetic variation in antioxidant defense genes and survival after diagnosis of breast cancer, J. Clin. Oncol. 25 (2007) 3015–3023.

[80] M.S. Geybels, C.M. Hutter, E.M. Kwon, E.A. Ostrander, R. Fu, Z. Feng, J.L. Stanford, U. Peters, Variation in selenoenzyme genes and prostate cancer risk and survival, Prostate 73 (2013) 734–742.
[81] M.L. Cooper, H.-O. Adami, H. Grönberg, F. Wiklund, F.R. Green, M.P. Rayman, Interaction between SNPs in selenoprotein P and mitochondrial superoxide dismutase determines prostate cancer risk, Cancer Res. 68 (2008), 10171.
[82] A. Steinbrecher, C. Méplan, J. Hesketh, L. Schomburg, T. Endermann, E. Jansen, B. Åkesson, S. Rohrmann, J. Linseisen, Effects of selenium status and polymorphisms in selenoprotein genes on prostate cancer risk in a prospective study of European men, Cancer Epidemiol. Prevention Biomarkers 19 (2010) 2958–2968.
[83] S. Peng, B. Lü, W. Ruan, Y. Zhu, H. Sheng, M. Lai, Genetic polymorphisms and breast cancer risk: evidence from meta-analyses, pooled analyses, and genome-wide association studies, Breast Cancer Res. Treat. 127 (2011) 309–324.
[84] E. Jablonska, J. Gromadzinska, B. Peplonska, W. Fendler, E. Reszka, M.B. Krol, E. Wieczorek, A. Bukowska, P. Gresner, M. Galicki, O. Zambrano Quispe, Z. Morawiec, et al., Lipid peroxidation and glutathione peroxidase activity relationship in breast cancer depends on functional polymorphism of GPX1, BMC Cancer 15 (2015) 657.