# Gene-environment interactions relevant to estrogen and risk of breast cancer: can gene-environment interactions be detected only among candidate SNPs from genome-wide association studies?

Park, J.; Choi, J.Y.; Choi, J.; Chung, S.; Song, N.; Park, S.K.; ... ; NBCS Collaborators

cancers

MDPI

*Article*

# Gene-Environment Interactions Relevant to Estrogen and Risk of Breast Cancer: Can Gene-Environment Interactions Be Detected Only among Candidate SNPs from Genome-Wide Association Studies?

**JooYong Park [1,2], Ji-Yeob Choi [1,2,3,4,*], Jaesung Choi [3], Seokang Chung [1], Nan Song [5], Sue K. Park [1,4,6], Wonshik Han [4,7], Dong-Young Noh [4,7], Sei-Hyun Ahn [8], Jong Won Lee [8], Mi Kyung Kim [9], Sun Ha Jee [10], Wanqing Wen [11], Manjeet K. Bolla [12], Qin Wang [12], Joe Dennis [12], Kyriaki Michailidou [12,13,14], Mitul Shah [15], Don M. Conroy [15], Patricia A. Harrington [15], Rebecca Mayes [15], Kamila Czene [16], Per Hall [16,17], Lauren R. Teras [18], Alpa V. Patel [19], Fergus J. Couch [19], Janet E. Olson [20], Elinor J. Sawyer [21], Rebecca Roylance [22], Stig E. Bojesen [23,24,25], Henrik Flyger [26], Diether Lambrechts [27,28], Adinda Baten [29], Keitaro Matsuo [30,31], Hidemi Ito [31], Pascal Guénel [32], Thérèse Truong [32], Renske Keeman [33], Marjanka K. Schmidt [33,34], Anna H. Wu [35], Chiu-Chen Tseng [35], Angela Cox [36], Simon S. Cross [37], kConFab Investigators [38,39,†], Irene L. Andrulis [40,41], John L. Hopper [42], Melissa C. Southey [43,44,45], Pei-Ei Wu [46], Chen-Yang Shen [47,48], Peter A. Fasching [49,50], Arif B. Ekici [51], Kenneth Muir [52], Artitaya Lophatananon [52], Hermann Brenner [53,54,55], Volker Arndt [53], Michael E. Jones [56], Anthony J. Swerdlow [56,57], Reiner Hoppe [58,59], Yon-Dschun Ko [60], Mikael Hartman [61,62,63], Jingmei Li [64], Arto Mannermaa [65,66,67], Jaana M. Hartikainen [65,66], Javier Benitez [68,69], Anna González-Neira [69], Christopher A. Haiman [35], Thilo Dörk [70], Natalia V. Bogdanova [70,71,72], Soo Hwang Teo [73,74], Nur Aishah Mohd Taib [75], Olivia Fletcher [76], Nichola Johnson [76], Mervi Grip [77], Robert Winqvist [78,79], Carl Blomqvist [80,81], Heli Nevanlinna [82], Annika Lindblom [83,84], Camilla Wendt [85], Vessela N. Kristensen [86,87,†], NBCS Collaborators [86,87,88,89,90,91,92,93,94,95,96,97,†], Rob A. E. M. Tollenaar [98], Bernadette A. M. Heemskerk-Gerritsen [99], Paolo Radice [100], Bernardo Bonanni [101], Ute Hamann [102], Mehdi Manoochehri [102], James V. Lacey [103,104], Maria Elena Martinez [105,106], Alison M. Dunning [15], Paul D. P. Pharoah [12,15], Douglas F. Easton [12,15], Keun-Young Yoo [6] and Daehee Kang [1,4,6]**

check for updates

**Citation:** Park, J.; Choi, J.-Y.; Choi, J.; Chung, S.; Song, N.; Park, S.K.; Han, W.; Noh, D.-Y.; Ahn, S.-H.; Lee, J.W.; et al. Gene-Environment Interactions Relevant to Estrogen and Risk of Breast Cancer: Can Gene-Environment Interactions Be Detected Only among Candidate SNPs from Genome-Wide Association Studies?. *Cancers* **2021**, *13*, 2370. https://doi.org/10.3390/cancers13102370

Academic Editor: Samuel Cos

Received: 9 April 2021
Accepted: 30 April 2021
Published: 14 May 2021



1 Department of Biomedical Sciences, Seoul National University Graduate School, Seoul 03080, Korea; terran2712@snu.ac.kr (J.P.); chungsk@neca.re.kr (S.C.); suepark@snu.ac.kr (S.K.P.); dhkang@snu.ac.kr (D.K.)
2 BK21plus Biomedical Science Project, Seoul National University College of Medicine, Seoul 03080, Korea
3 Institute of Health Policy and Management, Seoul National University Medical Research Center, Seoul 03080, Korea; cksjwjsja@snu.ac.kr
4 Cancer Research Institute, Seoul National University, Seoul 03080, Korea; hanw@snu.ac.kr (W.H.); dynoh@snu.ac.kr (D.-Y.N.)
5 College of Pharmacy, Chungbuk National University, Cheongju-si 28160, Korea; nan.song@chungbuk.ac.kr
6 Department of Preventive Medicine, Seoul National University College of Medicine, Seoul 03080, Korea; kyyoo@snu.ac.kr
7 Department of Surgery, Seoul National University College of Medicine, Seoul 03080, Korea
8 Department of Surgery, Medicine and ASAN Medical Center, University of Ulsan College, Seoul 05505, Korea; ahnsh@amc.seoul.kr (S.-H.A.); leejw@amc.seoul.kr (J.W.L.)
9 Division of Cancer Epidemiology and Management, National Cancer Center, Goyang-si 10408, Korea; alrud@ncc.re.kr
10 Department of Epidemiology and Health Promotion, Institute for Health Promotion, Graduate School of Public Health, Yonsei University, Seoul 03722, Korea; jsunha@yuhs.ac
11 Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, USA; wanqing.wen@vanderbilt.edu
12 Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB2 0SR, UK; mkh39@medschl.cam.ac.uk (M.K.B.); qw232@medschl.cam.ac.uk (Q.W.); jgd29@cam.ac.uk (J.D.); kyriakimi@cing.ac.cy (K.M.); pp10001@medschl.cam.ac.uk (P.D.P.P.); dfe20@medschl.cam.ac.uk (D.F.E.)
13 Biostatistics Unit, The Cyprus Institute of Neurology & Genetics, Nicosia 2371, Cyprus
14 Cyprus School of Molecular Medicine, The Cyprus Institute of Neurology & Genetics, Nicosia 23462, Cyprus
15 Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge CB1 8RN, UK; ms483@medschl.cam.ac.uk (M.S.); dmc40@medschl.cam.ac.uk (D.M.C.); pah25@medschl.cam.ac.uk (P.A.H.); rm480@medschl.cam.ac.uk (R.M.); amd24@medschl.cam.ac.uk (A.M.D.)

[16] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 65 Stockholm, Sweden; kamila.czene@ki.se (K.C.); Per.Hall@ki.se (P.H.)
[17] Department of Oncology, Södersjukhuset, 118 83 Stockholm, Sweden
[18] Department of Population Science, American Cancer Society, Atlanta, GA 30303, USA; lauren.teras@cancer.org
[19] Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA; alpa.patel@cancer.org (A.V.P.); couch.fergus@mayo.edu (F.J.C.)
[20] Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA; olsonj@mayo.edu
[21] School of Cancer & Pharmaceutical Sciences, Comprehensive Cancer Centre, Guy's Campus, King's College London, London SE1 9RT, UK; elinor.sawyer@kcl.ac.uk
[22] Department of Oncology, UCLH Foundation Trust, London NW1 2PG, UK; r.roylance@ucl.ac.uk
[23] Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, 2730 Herlev, Denmark; stig.egil.bojesen@regionh.dk
[24] Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, 2730 Herlev, Denmark
[25] Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark
[26] Department of Breast Surgery, Herlev and Gentofte Hospital, Copenhagen University Hospital, 2730 Herlev, Denmark; henrik.flyger@regionh.dk
[27] VIB Center for Cancer Biology, 3001 Leuve, Belgium; diether.lambrechts@kuleuven.vib.be
[28] Laboratory for Translational Genetics, Department of Human Genetics, University of Leuven, 3000 Leuven, Belgium
[29] Department of Radiotherapy Oncology, KU Leuven—University of Leuven, University Hospitals Leuven, 3000 Leuven, Belgium; Adinda.Baten@uzleuven.be
[30] Division of Cancer Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya 464-8681, Japan; kmatsuo@aichi-cc.jp
[31] Division of Cancer Epidemiology, Nagoya University Graduate School of Medicine, Nagoya 466-8550, Japan; hidemi@aichi-cc.jp
[32] Center for Research in Epidemiology and Population Health (CESP), Team Exposome and Heredity, INSERM, University Paris-Saclay, 94805 Villejuif, France; pascal.guenel@inserm.fr (P.G.); therese.truong@inserm.fr (T.T.)
[33] Division of Molecular Pathology, The Netherlands Cancer Institute—Antoni van Leeuwenhoek Hospital, 1066 CX Amsterdam, The Netherlands; r.keeman@nki.nl (R.K.); mk.schmidt@nki.nl (M.K.S.)
[34] Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute—Antoni van Leeuwenhoek Hospital, 1066 CX Amsterdam, The Netherlands
[35] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; Anna.Wu@med.usc.edu (A.H.W.); Chiu-Chen.Tseng@med.usc.edu (C.-C.T.); Christopher.Haiman@med.usc.edu (C.A.H.)
[36] Sheffield Institute for Nucleic Acids (SInFoNiA), Department of Oncology and Metabolism, University of Sheffield, Sheffield S10 2TN, UK; a.cox@shef.ac.uk
[37] Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield S10 2TN, UK; s.s.cross@sheffield.ac.uk
[38] Peter MacCallum Cancer Center, Melbourne, VIC 3000, Australia; heather.thorne@petermac.org
[39] Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, VIC 3000, Australia
[40] Fred A, Litwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada; andrulis@lunenfeld.ca
[41] Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada
[42] Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC 3010, Australia; j.hopper@unimelb.edu.au
[43] Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC 3168, Australia; melissa.southey@monash.edu
[44] Department of Clinical Pathology, The University of Melbourne, Melbourne, VIC 3010, Australia
[45] Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC 3004, Australia
[46] Taiwan Biobank, Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan; peiei@gate.sinica.edu.tw
[47] Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan; bmcys@ibms.sinica.edu.tw
[48] School of Public Health, China Medical University, Taichung 404, Taiwan
[49] Department of Medicine Division of Hematology and Oncology, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA; peter.fasching@fau.de
[50] Department of Gynecology and Obstetrics, Comprehensive Cancer Center ER-EMN, University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg, 91054 Erlangen, Germany
[51] Institute of Human Genetics, Comprehensive Cancer Center Erlangen-EMN, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, 91054 Erlangen, Germany; Arif.Ekici@uk-erlangen.de

[52] Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester M13 9PL, UK; kenneth.muir@manchester.ac.uk (K.M.); artitaya.lophatananon@manchester.ac.uk (A.L.)
[53] Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; h.brenner@dkfz.de (H.B.); v.arndt@dkfz.de (V.A.)
[54] Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), 69120 Heidelberg, Germany
[55] German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
[56] Division of Genetics and Epidemiology, The Institute of Cancer Research, London SM2 5NG, UK; michael.jones@icr.ac.uk (M.E.J.); Anthony.Swerdlow@icr.ac.uk (A.J.S.)
[57] Division of Breast Cancer Research, The Institute of Cancer Research, London SW7 3RP, UK
[58] Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, 70376 Stuttgart, Germany; reiner.hoppe@ikp-stuttgart.de
[59] University of Tübingen, 72074 Tübingen, Germany
[60] Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, 53177 Bonn, Germany; Yon-Dschun.Ko@bn.johanniter-kliniken.de
[61] Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore 117549, Singapore; ephbamh@nus.edu.sg
[62] Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore 119228, Singapore
[63] Department of Surgery, National University Health System, Singapore 119228, Singapore
[64] Human Genetics Division, Genome Institute of Singapore, Singapore 138672, Singapore; lijm1@gis.a-star.edu.sg
[65] Translational Cancer Research Area, University of Eastern Finland, 70210 Kuopio, Finland; arto.mannermaa@uef.fi (A.M.); jaana.hartikainen@uef.fi (J.M.H.)
[66] Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, 70210 Kuopio, Finland
[67] Biobank of Eastern Finland, Kuopio University Hospital, 70210 Kuopio, Finland
[68] Biomedical Network on Rare Diseases (CIBERER), 28029 Madrid, Spain; jbenitez@cnio.es
[69] Human Cancer Genetics Programme, Spanish National Cancer Research Centre (CNIO), 28029 Madrid, Spain; agonzalez@cnio.es
[70] Gynaecology Research Unit, Hannover Medical School, 30625 Hannover, Germany; doerk.thilo@mh-hannover.de (T.D.); bogdanova.natalia@mh-hannover.de (N.V.B.)
[71] Department of Radiation Oncology, Hannover Medical School, 30625 Hannover, Germany
[72] NN Alexandrov Research Institute of Oncology and Medical Radiology, 223040 Minsk, Belarus
[73] Breast Cancer Research Programme, Cancer Research Malaysia, Subang Jaya 47500, Malaysia; soohwang.teo@cancerresearch.my
[74] Department of Surgery, Faculty of Medicine, University of Malaya, Kuala Lumpur 50603, Malaysia
[75] Breast Cancer Research Unit, University Malaya Cancer Research Institute, Faculty of Medicine, University of Malaya, Kuala Lumpur 50603, Malaysia; naisha@um.edu.my
[76] The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London SW7 3RP, UK; Olivia.fletcher@icr.ac.uk (O.F.); Nichola.johnson@icr.ac.uk (N.J.)
[77] Department of Surgery, Oulu University Hospital, University of Oulu, 90220 Oulu, Finland; mervi.grip@ppshp.fi
[78] Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit, Biocenter Oulu, University of Oulu, 90570 Oulu, Finland; robert.winqvist@oulu.fi
[79] Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre Oulu, Oulu 90570, Finland
[80] Department of Oncology, Helsinki University Hospital, University of Helsinki, 00290 Helsinki, Finland; carl.blomqvist@helsinki.fi
[81] Department of Oncology, Örebro University Hospital, 70185 Örebro, Sweden
[82] Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, 00290 Helsinki, Finland; Heli.Nevanlinna@hus.fi
[83] Department of Molecular Medicine and Surgery, Karolinska Institutet, 171 76 Stockholm, Sweden; Annika.Lindblom@ki.se
[84] Department of Clinical Genetics, Karolinska University Hospital, 171 76 Stockholm, Sweden
[85] Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet, 118 83 Stockholm, Sweden; camilla.wendt@sll.se
[86] Department of Medical Genetics, Oslo University Hospital and University of Oslo, 0450 Oslo, Norway; v.n.kristensen@medisin.uio.no (V.N.K.); greal@ous-hf.no (NBCS Collaborators)
[87] Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, 0372 Oslo, Norway
[88] Department of Research, Vestre Viken Hospital, 3004 Drammen, Norway
[89] Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, 0450 Oslo, Norway

[90] Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, 0450 Oslo, Norway
[91] Section for Breast- and Endocrine Surgery, Department of Cancer, Division of Surgery, Cancer and Transplantation Medicine, Oslo University Hospital-Ullevål, 0450 Oslo, Norway
[92] Department of Radiology and Nuclear Medicine, Oslo University Hospital, 0450 Oslo, Norway
[93] Department of Pathology at Akershus University Hospital, 1478 Lørenskog, Norway
[94] Department of Oncology, Division of Surgery and Cancer and Transplantation Medicine, University Hospital-Radiumhospitalet, 0405 Oslo, Norway
[95] National Advisory Unit on Late Effects after Cancer Treatment, Department of Oncology, Oslo University Hospital, 0405 Oslo, Norway
[96] Department of Oncology, Akershus University Hospital, 1478 Lørenskog, Norway
[97] Oslo Breast Cancer Research Consortium, Oslo University Hospital, 0405 Oslo, Norway
[98] Department of Surgery, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands; r.a.e.m.tollenaar@lumc.nl
[99] Department of Medical Oncology, Erasmus MC Cancer Institute, 3015 CN Rotterdam, The Netherlands; b.heemskerk-gerritsen@erasmusmc.nl
[100] Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Research, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), 20133 Milan, Italy; Paolo.Radice@istitutotumori.mi.it
[101] Division of Cancer Prevention and Genetics, IEO, European Institute of Oncology IRCCS, 20141 Milan, Italy; bernardo.bonanni@ieo.it
[102] Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; u.hamann@dkfz-heidelberg.de (U.H.); mehdi.manoochehri@igb.fraunhofer.de (M.M.)
[103] Department of Computational and Quantitative Medicine, City of Hope, Duarte, CA 91010, USA; jlacey@coh.org
[104] City of Hope Comprehensive Cancer Center, City of Hope, Duarte, CA 91010, USA
[105] Moores Cancer Center, University of California San Diego, La Jolla, CA 92037, USA; e8martinez@health.ucsd.edu
[106] Herbert Wertheim School of Public Health and Longevity Science, University of California San Diego, La Jolla, CA 92161, USA
* Correspondence: miso77@snu.ac.kr; Tel.: +82-2-740-8922; Fax: +82-2-3673-2167
† Membership of the kConFab Investigators and NBCS Collaborators are provided in the Acknowledgments section of the manuscript.

**Simple Summary:** Breast cancer is the most common cancer in females worldwide. To date, many gene-environment interaction (GxE) studies have been conducted to better understand how genetic factors combine with environmental factors to influence risk. However, previous studies have not found or found only a few interactions by using SNPs which were discovered from genome-wide association studies and have been conducted, for the most part, within European populations. In this study, we focused on estrogen-related lifestyle factors that have been identified for breast cancer, including several well-established reproductive factors that are mediated by hormonal mechanisms. We aimed to examine whether there are any gene and environmental factor interactions related to estrogen exposure or metabolism using a candidate approach in Korean women. We found two interactions in this study, although they were not replicated in the independent large consortium data. These findings suggest specificity in Koreans for breast cancer risk.

**Abstract:** In this study we aim to examine gene-environment interactions (GxEs) between genes involved with estrogen metabolism and environmental factors related to estrogen exposure. GxE analyses were conducted with 1970 Korean breast cancer cases and 2052 controls in the case-control study, the Seoul Breast Cancer Study (SEBCS). A total of 11,555 SNPs from the 137 candidate genes were included in the GxE analyses with eight established environmental factors. A replication test was conducted by using an independent population from the Breast Cancer Association Consortium (BCAC), with 62,485 Europeans and 9047 Asians. The GxE tests were performed by using two-step methods in GxEScan software. Two interactions were found in the SEBCS. The first interaction was shown between rs13035764 of NCOA1 and age at menarche in the GE|2df model ($p$-2df = $1.2 \times 10^{-3}$). The age at menarche before 14 years old was associated with the high risk of breast cancer, and the risk was higher when subjects had homozygous minor allele G. The second GxE was shown between rs851998 near ESR1 and height in the GE|2df model ($p$-2df = $1.1 \times 10^{-4}$). Height taller than 160 cm was associated with a high risk of breast cancer, and the risk increased when the minor allele was

added. The findings were not replicated in the BCAC. These results would suggest specificity in Koreans for breast cancer risk.

**Keywords:** breast cancer; estrogen; gene-environment interaction

## 1. Introduction

Breast cancer is the most common cancer in females worldwide, with an estimated 2.3 million incident cases globally in 2020. Moreover, breast cancer ranks as the most common cancer for women in 159 countries, and the most common cause of cancer deaths in 110 countries [1]. In Korea, the incidence rate of breast cancer, as well as breast cancer mortality, has been persistently increasing [2].

Numerous epidemiological studies have been performed to identify risk factors for breast cancer and many non-genetic factors, referred to as environmental factors, have been established [3–5]. Furthermore, many genome-wide association studies (GWAS) have been conducted, which have provided significant opportunities to discover the potential effects of common genetic factors on complex diseases [6]. To date, more than 200 common susceptibility loci for breast cancer have been identified through GWAS [7–10]. Although *BRCA1* and *BRCA2* are the most well-known risk genes associated with about a 20-fold increased breast cancer risk, they only account for 20% of familial breast cancer due to the low frequency of mutations [11,12]. Even if all the common genetic factors are taken together, they are estimated to explain only about 30% of the familial risk [7]. Consequently, further research is needed to identify the missing heritability [13], and to better understand how genetic factors combine with environmental factors to influence risk [14].

Currently, a few alternatives have been proposed, and one of them is a gene-environment interaction (GxE) analysis. By accounting for interactions between genetic and environmental factors, better estimates of the population-attributable risk or effects in specific subgroups exposed by certain environments can be obtained [15,16]. To date, many GxE studies for breast cancer risk have been conducted, but few studies have shown significant results after accounting for multiple testing [14,17–24]. Most of these studies used the standard case-control analysis of GxE interaction based on logistic regression and added interaction terms and tended to have poor statistical power. Some methods have been developed and proposed [25], and two-step approaches have been reported to have greater power than other approaches [13]. However, previous studies have not found, or have only found a few interactions by using SNPs which were discovered from GWAS and have been conducted, for the most part, within European populations [14,17–24].

In this study we focused on estrogen-related lifestyle factors that have been identified for breast cancer, including several well-established reproductive factors that are mediated by hormonal mechanisms. Endogenous estrogen levels are likely controlled by genetic factors of the estrogen metabolism pathway and, in turn, play an integral part as determinants of breast cancer risk [26–28]. Thus, we aimed to examine whether there are any interactions between genes and environmental factors related to estrogen exposure or metabolism by using a candidate approach in Korean women.

## 2. Results

Table 1 shows the distribution of demographic and reproductive factors and their association with breast cancer risk in the Seoul Breast Cancer Study (SEBCS) population. Controls were generally older than cases, and most of the environmental factors were significantly different between cases and controls. Taller height, lower BMI, younger age at menarche, older age at first full term pregnancy (FFTP), no breastfeeding experience and shorter duration of breastfeeding, and longer duration of estrogen exposure before FFTP (EEBF) were associated with increased risk for breast cancer among the eight candidate environmental factors.

**Table 1.** Distributions of demographic and reproductive factors and associations with the risk of breast cancer, Seoul Breast Cancer Study.

| Demographic and Reproductive Factors | Cases *N* = 1970 | | Controls *N* = 2052 | | OR [c] | 95% CI |
|---|---|---|---|---|---|---|
| | Mean ± SD | % | Mean ± SD | % | | |
| Age | 49.0 ± 8.26 | | 51.4 ± 7.75 | | 0.99 | 0.98–1.01 |
| Family history. yes | | 4.8 | | 1.8 | 2.64 | 1.77–3.94 |
| Height | 157.9 ± 5.04 | | 155.9 ± 5.00 | | 1.07 | 1.06–1.08 |
| <160 cm | | 59.6 | | 75.7 | 1.00 | reference |
| ≥160 cm | | 40.4 | | 24.3 | 1.92 | 1.67–2.20 |
| BMI (kg/m$^2$) | 23.2 ± 2.94 | | 23.7 ± 2.96 | | 0.97 | 0.95–1.00 |
| <25 | | 76.8 | | 71.2 | 1.00 | reference |
| ≥25 | | 23.2 | | 28.9 | 0.81 | 0.69–0.93 |
| Menarche age | 14.8 ± 1.65 | | 15.2 ± 1.77 | | 0.90 | 0.86–0.93 |
| ≥14 years | | 78.9 | | 85.0 | 1.00 | reference |
| <14 years | | 21.1 | | 15.0 | 1.27 | 1.07–1.50 |
| Ever pregnancy | | 93.5 | | 92.8 | 0.26 | 0.14–0.46 |
| Age at FFTP [a] | 25.9 ±3.48 | | 25.1 ± 3.36 | | 1.05 | 1.03–1.07 |
| <27 years | | 55.6 | | 66.4 | 1.00 | reference |
| Nulliparity or ≥27 years | | 44.4 | | 33.6 | 1.49 | 1.30–1.70 |
| No. of children | 2.2 ± 0.85 | | 2.3 ± 0.91 | | 1.06 | 0.97–1.16 |
| ≥2 | | 85.2 | | 88.2 | 1.00 | reference |
| <2 | | 14.8 | | 11.8 | 1.11 | 0.91–1.36 |
| Never breastfeeding | | 22.3 | | 14.5 | 1.31 [d] | 1.10–1.57 |
| Breastfeeding duration (months) | 18.0 ± 22.59 | | 24.0 ± 21.53 | | 0.99 [d] | 0.99–1.00 |
| Breastfeeding duration per child (months) | 7.3 ± 8.01 | | 9.9 ± 7.27 | | 0.97 [d] | 0.96–0.98 |
| ≥2.5 month | | 45.1 | | 68.3 | 1.00 | reference |
| <2.5 month | | 54.9 | | 31.7 | 3.81 [d] | 3.16–4.59 |
| Duration of EEBF [b] | 12.5 ± 6.78 | | 11.7 ± 8.00 | | 1.04 [e] | 1.02–1.05 |
| <13 years | | 61.1 | | 76.2 | 1.00 | reference |
| ≥13 years | | 38.9 | | 23.8 | 1.75 [e] | 1.51–2.03 |
| Postmenopausal women | | 38.2 | | 54.8 | 0.60 | 0.50–0.72 |
| Age at menopause | 48.5 ± 5.40 | | 49.2 ± 4.55 | | 0.96 | 0.94–0.98 |

[a] FFTP first full-term pregnancy (years); [b] EEBF estrogen exposure before the first full-term pregnancy (years); [c] ORs analyzed by logistic regression adjusting for age, family history, categorical FFTP, age at menarche, and menopausal status; [d] ORs analyzed by logistic regression adjusting for age, family history, categorical FFTP, age at menarche, menopausal status, and the number of children; [e] ORs analyzed by logistic regression adjusting for age, family history, ever pregnancy, and menopausal status.

We conducted the replication using both a European population and an Asian population in the Breast Cancer Association Consortium (BCAC) (Table S3), and their distribution of demographic and other factors are shown in Tables S4 and S5, as well as associations with breast cancer. The associations with the risk of breast cancer were heterogeneous by ethnicity in height, age at menarche, age at FFTP, and menopause status.

We found two interactions through the GxE analysis (Table 2). High LDs were observed among the three SNPs (i.e., rs13035764, rs11125629, and rs11688818) in NCOA1 on chromosome 2.

**Table 2.** Top two SNPs from the gene and environment interaction analysis between 11,555 SNPs involved in estrogen metabolism and 8 environmental factors related to estrogen, Seoul Breast Cancer Study.

| **SNP** | **Chromosome** | **Position_b36** | **Gene** | **Minor/Major_Allele** | **MAF** | | |
|---|---|---|---|---|---|---|---|
| rs13035764 | 2 | 24571432 | *NCOA1* | G/C | 0.2971 | | |
| | **N(Cases/Controls)** | **OR [a] (95%CI)** | **Analysis methods** | **Environment** | ***p*-2df** | **Threshold** | ***p*-GxE** |
| | 3925(1949/1976) | 0.84 (0.76–0.93) | Weighted GE\|2df | Age at menarche | $1.2 \times 10^{-3}$ | 0.005 [b] | 0.0524 |
| **SNP** | **Chromosome** | **Position_b36** | **Gene** | **Minor/Major_Allele** | **MAF** | | |
| rs851998 | 6 | 152025031 | *ESR1* | A/G | 0.4426 | | |
| | **N(Cases/Controls)** | **OR [a] (95%CI)** | **Analysis methods** | **Environment** | ***p*-2df** | **Threshold** | ***p*-GxE** |
| | 3940(1955/1985) | 0.82 (0.75–0.90) | Subset GE\|2df | height | $6.8 \times 10^{-5}$ | 0.000110 [c] | 0.0471 |

[a] risk of SNP for breast cancer; [b] threshold of bin 1 in the weighted model; [c] 0.05 divided by 456 SNPs which were tested in step 2.

These three SNPs, i.e., rs13035764, rs11125629, and rs11688818, showed interactions with age at menarche by the weighted GE|2df method ($p_{2df}$ = 0.0012, 0.0018, 0.0027 respectively, significant threshold $p$ = 0.005, bin1) (Figure S1). Age at menarche before 14 years old (<25% quartile) was associated with increased risk of breast cancer (OR = 1.27, 95% CI 1.07–1.50); this association was much stronger for subjects who had a homozygous minor allele G ($OR_{\text{age at menarche < 14 years | rs13035764\_GG}}$ = 2.85, 95% CI 1.43–5.67). When we examined the associations of age at menarche with breast cancer risk in 44 studies, eight studies (seven European and one Asian) showed consistent effects with the SEBCS (Figure S3), however, interaction with rs13035764 was not observed in the replication set from either theEuropean or Asian populations (Table 3 and Table S6).

Figure S2 shows the results of the interaction between height and three SNPs (rs851998, rs851971, and rs851967) located near the ESR1 gene in chromosome 6 by the subset GE|2df method ($p_{2df}$ = 0.000068, 0.000094 and 0.0001, respectively, the significance threshold was 0.000110 because the number of SNPs tested in step 2 was 456). Height taller than 160 cm (≥75% quartile) was associated with increased risk of breast cancer (OR = 1.92, 95% CI 1.67–2.20). The association of height with breast cancer was stronger when the minor allele was considered ($OR_{\text{height} \geq 160 \text{ cm | rs851998\_GG}}$ = 1.53, 95% CI 1.20–1.95; $OR_{\text{height} \geq 160 \text{ cm | rs851998\_GA}}$ = 2.03, 95% CI 1.65–2.49; and $OR_{\text{height} \geq 160 \text{ cm | rs851998\_AA}}$ = 2.39, 95% CI 1.74–3.29) (Table 3). This association was observed among Asian studies of the BCAC. The risk of breast cancer associated with greater height was enhanced among those who had minor alleles ($OR_{\text{height} \geq 160 \text{ cm | rs851998\_AA}}$ = 1.38, 95% CI 1.04–1.82) but the interaction test was not significant ($p$-GxE = 0.4642) (Table 3). Although there were ten studies (nine European and one Asian) that showed the consistent effects of height with SEBCS (Figure S4), a significant interaction with rs851998 was not found, even in the combined population, which showed a consistent effect with the SEBCS (Table S7).

The associations between each environmental factor and the risk of breast cancer by genotype were not different according to the estrogen receptor (ER) or progesterone receptor (PR) status in the SEBCS population (Table S8).

**Table 3.** Association between an environmental factor and breast cancer by genotypes.

| Study | Exposure | Category | N(Cases/Controls) | OR | 95% CI | *p*-GxE |
|---|---|---|---|---|---|---|
| SEBCS | Age at menarche <14 years old [a] | Effect of E in overall | 4022(1970/2052) | 1.27 | 1.07–1.50 | 0.0524 |
| | | E \| rs13035764_CC | 1963(1007/956) | 1.19 | 0.95–1.51 | |
| | | E \| rs13035764_CG | 1706(810/896) | 1.17 | 0.89–1.53 | |
| | | E \| rs13035764_GG | 337(147/190) | 2.85 | 1.43–5.67 | |
| | Height ≥ 160 cm [b] | Effect of E in overall | 4022(1970/2052) | 1.92 | 1.67–2.20 | 0.0471 |
| | | E \| rs851998_GG | 1281(677/604) | 1.53 | 1.20–1.95 | |
| | | E \| rs851998_GA | 1921(938/983) | 2.03 | 1.65–2.49 | |
| | | E \| rs851998_AA | 819(355/464) | 2.39 | 1.74–3.29 | |
| BCAC-European | Age at menarche <12 years old [a] | Effect of E in overall | 62,485(36,894/25,591) | 1.03 | 0.99–1.08 | 0.5314 |
| | | E \| rs13035764_CC | 9953(5921/4032) | 1.02 | 0.91–1.14 | |
| | | E \| rs13035764_CG | 34,124(20,140/13,984) | 1.03 | 0.97–1.09 | |
| | | E \| rs13035764_GG | 18,408(10,833/7575) | 1.05 | 0.97–1.14 | |
| | Height ≥ 169 cm [b] | Effect of E in overall | 56,363(31,994/24,369) | 0.97 | 0.93–1.01 | 0.2425 |
| | | E \| rs851998_GG | 27,224(15,513/11,711) | 0.93 | 0.88–0.99 | |
| | | E \| rs851998_GA | 23,843(13,459/10,384) | 0.99 | 0.93–1.05 | |
| | | E \| rs851998_AA | 5296(3022/2274) | 1.03 | 0.90–1.17 | |
| BCAC-Asian | Age at menarche <12 years old [a] | Effect of E in overall | 9047(4392/4655) | 0.97 | 0.82–1.14 | 0.7817 |
| | | E \| rs13035764_CC | 4867(2407/2460) | 0.95 | 0.75–1.21 | |
| | | E \| rs13035764_CG | 3504(1666/1838) | 0.99 | 0.76–1.28 | |
| | | E \| rs13035764_GG | 676(319/357) | 1.04 | 0.57–1.91 | |
| | Height ≥ 160 cm [b] | Effect of E in overall | 5789(2879/2910) | 1.22 | 1.08–1.39 | 0.4642 |
| | | E \| rs851998_GG | 1751(899/852) | 1.20 | 0.95–1.52 | |
| | | E \| rs851998_GA | 2846(1422/1424) | 1.19 | 0.99–1.42 | |
| | | E \| rs851998_AA | 1192(558/634) | 1.38 | 1.04–1.82 | |

ORs analyzed by logistic regression adjusting for age, family history, categorical FFTP, age at menarche, and menopausal status. [a] age at menarche <25% quartile; [b] height ≥75% quartile.

## 3. Discussion

We performed GxE analyses between 11,555 SNPs in and near the promoters of the 136 genes involved with estrogen metabolism and eight environmental factors related to estrogen exposure by using GxEScan. Various results from the two-step methods were obtained, and two interactions involving NCOA1 with age at menarche and ESR1 with height were found.

The minor allele of rs13035764 in the NCOA1 was associated with an enhanced risk of breast cancer for younger age at menarche in the SEBCS. Our study suggested that women whose menstruation was late, possibly had an increased protective effect with this variant. NCOA1 is known as the steroid receptor coactivator 1 (SRC-1) encoding a protein that acts as a transcriptional coactivator for steroid and nuclear hormone receptors. Including SRC-1, SRC coactivators have important roles in development, growth, reproduction, and even in cancer [29]. It has been shown that SRC-1 proteins were overexpressed from 19% to 29% of human breast cancers, and this overexpression has been associated with large and high-grade tumors [30–32]. Although there was no epidemiological study that examined the association between SRC-1 and age at menarche directly, previous reports have indicated the regulatory role of SRC-1 during a specific phase in the menstrual cycle [33,34]. Thus, our findings could suggest that the risk for breast cancer decreases in the late menarche group as they were exposed not only to the hormone but also to the SRC-1 effects later.

We also found that the minor allele of rs851998, which is located near the ESR1 gene, was associated with an increased risk of breast cancer for females who were taller than 160 cm. Previously, numerous studies have suggested the association between the ESR1

gene and breast cancer [35–38], and have reported the role of ESR1 mutations for breast cancer [39–41]. In addition, adult height has been linked with breast cancer risk from many epidemiological studies [42–46]. As a key factor for the development of female sex organs and sex characteristics, estrogen also plays a role in height as a part of the growth process [47,48]. Although which part of estrogen metabolism and how it acts on the growth in vivo is unclear [49], many studies have supported that because estrogen synthesis and secretion are increased during puberty, these factors could lead to epiphyseal fusion, termination of linear growth, and determination of final height [50–53]. With several reports suggesting the association between ESR1 and height [54–56], we think that the ESR1 gene and height possibly have an association with breast cancer risk, since significant interaction results were found in our results.

This study was conducted by the candidate gene approach with a hypothesis for statistical efficiency and biological plausibility. There have been a considerable number of criticisms [57–59]. The main concerns have included inadequate analytic procedures during the selection of genetic or environmental factors and using statistical methods with small sample size and lower power, as well as publication bias. However, an opposite commentary has refuted that the candidate GxE approach remains to be the most commonly studied despite its limitations, and the goal of a candidate gene-environment interaction study is distinct from that of a GWAS-based interaction study [60]. GxE based on GWAS focuses on gene-environmental correlations, while candidate GxE with a hypothesis attempts to discover causes. Moreover, GxE based on GWAS mainly aims to overcome missing heritability and to examine the modified effect of genetic factors by environmental factors. There are other approaches to discover novel genetic factors, which have interactions with established environmental factors from genome-wide ranges; however, these approaches could not report significant interactions, possibly due to the burden of multiple testing [20,21]. Meanwhile, the candidate GxE approach, which aimed to detect genetic factors that have interactions with established environmental factors from candidate genes and to examine the modified effects of environmental factors by genotype, could decrease the statistical burden and interpret easily due to biological plausibility.

There are several limitations to this study. First, we dichotomized the continuous variables when the environmental factors were included in the GxE analysis. Although this approach could have some statistical problems, such as losing information, underestimating the extent of variation, and concealing nonlinearity, binary variables make interpretation easy with possible mechanistic conclusions [61] and facilitate testing for multi-factor interactions [62]. Second, our statistically significant results were obtained from the EG | 2df method rather than from the EDGxE, which demonstrated the best power and efficiency. Furthermore, the 2df test is called the joint gene and GxE test, in which the hypothesis is $\beta g = \beta gxe = 0$. Thus, a polymorphism can be detected even when there is no interaction, but it has a marginal effect. However, our results from the GE | 2df showed that the effect of environmental factors differed according to the number of minor alleles (Table 3). Thus, they might not be false-positive results. Third, the interactions found from the SEBCS were not replicated in the independent large consortium data, even in the replication test using each ethnicity population and combined population which showed consistent effects with the SEBCS. To date, numerous studies have performed gene-environment interaction tests for breast cancer risk, however, the results have rarely been replicated [63–65]. Most of the studies have been conducted in the European population and all replicated results have also been found in the European descent study population. This could be explained by racial differences impacting environmental factors and also genetic backgrounds [66–70] (as shown the Figures S3 and S4 and Tables S6 and S7). A larger study population or consortium data including Asian descent are demanded to replicate the findings successfully. Fourth, the sample size of this study was not large relative to other consortium studies; however, two-step methods generally require a smaller sample size to achieve enough power for the GxE test as compared with the standard GxE test, and a previous study used only 2382 subjects and found novel genes by using GxEScan [13]. Lastly, breast cancer was considered

as a whole entity rather than subtypes of breast cancers such as pre-menopausal and post-menopausal breast cancer due to the limitation of sample size. Further studies with a larger sample size would find that differences in the GxE depend on menopause status.

Although breast cancer is the most common cancer in women worldwide, the incidence or mortality of breast cancer differs from country to country [71]. According to GLOBCAN 2018 and the Korean Breast Cancer Society, the incidence rate of most of these countries is decreasing, while it is still increasing in Korea. The mortality of breast cancer is also dramatically increasing [72,73]. Korean females have shown a decreasing trend of age at menarche and an increasing trend of height [74–77]. These aspects possibly explain why the incidence rate of breast cancer is increasing in Korea, and therefore there is a need for research on breast cancer specific to Korean women.

## 4. Methods

### *4.1. Study Population of the Gene–Environment Interaction (GxE) Study*

The Seoul Breast Cancer Study (SEBCS) subjects were used in this study. As a multicenter case-control study, 4040 women with histologically confirmed breast cancer were recruited between 2001 and 2007 from the Seoul National University Hospital and the Asan Medical Center. During the same period, 1818 non-cancer patients aged 32 to 75 years were enrolled as hospital-based controls in the same hospital, and 2052 healthy women aged 39 to 71 years who were participants in a community health screening program from 2002 to 2007 were also included in the control group [78].

A questionnaire survey was administered by trained interviewers using in-person interviews for all subjects. Demographic information, family history, behavioral and dietary habits, and reproductive information were collected by this survey. Among the participants, we excluded those who had previous histories of cancer, hysterectomy, or oophorectomy, and only included those who were born from 1930 to 1969, following the same criteria as a previous study [78]. From these exclusion criteria, 3332 cases and 3620 controls were selected with questionnaire data. All subjects in the SEBCS had written informed consent, and the study design was approved by the Committee on Human Research of the Seoul National University Hospital (IRB no. H-0503-144-004).

### *4.2. Study Population and Data Used in the Replication*

A large dataset pooled from the Breast Cancer Association Consortium (BCAC) was used for the replication. Participants who were not of European or Asian descent and had missing data regarding sex, age, and genotype information of SNPs were excluded. A replication dataset for interaction between age at menarche and an SNP that had shown interaction with age at menarche from the SEBCS was made after excluding subjects who had missing data on age at menarche. Likewise, a replication dataset for interaction between height and an SNP that had shown an interaction with height from the SEBCS was created after excluding missing data in height. Missing values of covariates including family history, age at first full term pregnancy, and menopause status were categorized separately. In total, 62,485 women of European descent (36,894 cases and 25,591 controls in 37 studies) and 9047 women of Asian descent (4392 cases and 4655 controls in 7 studies) were included in the replication dataset to examine the interaction between age at menarche and an SNP in NCOA1, and 56,363 women of European descent (31,994 cases and 24,369 controls in 29 studies) and 5789 women of Asian descent (2879 cases and 2910 controls in 4 studies) were included in a replication dataset to assess the interaction between height and an SNP in ESR1 (Table S3).

### *4.3. GWAS Data*

The SEBCS also performed GWAS by using the Affymetrix Genome-Wide Human SNP Array 6.0 chip (Affymetrix, Inc., Santa Clara, CA, USA). A total of 4394 women (2342 cases and 2052 controls) were scanned, and we compiled a final dataset that included 555,117 SNPs after quality control exclusions. Detailed descriptions of genotyping and

quality control procedures can be found in the previous literature [79,80]. To conduct gene-environment interaction analyses, we included as eligible subjects a total 4022 women (1970 cases and 2052 controls) who had both questionnaire and GWAS data.

### 4.4. Candidate Gene Selection

We used the search term "estrogen metabolism" in several databases to identify candidate genes involved with estrogen metabolism. Ninety-one genes were identified using the Ensembl genome browser 84 (www.ensembl.org/ (accessed on 30 May 2016)), and 89 genes were selected except for two genes located on the X chromosome. Twenty-eight genes were searched using AmiGO (http://amigo1.geneontology.org/ (accessed on 30 May 2016)), but all genes overlapped with the previous selection. Twenty-nine genes were searched using AmiGO2 (http://amigo.geneontology.org/amigo (accessed on 30 May 2016)), and two genes, which did not overlap with the previous selection, were added. We also searched the literature to find more genes related to estrogen metabolism. Thirty-seven genes were added among 199 SNPs, which showed significant results from the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis [81], and nine genes were further added from previous studies [82–94]. A total of 137 candidate genes were on our list.

### 4.5. SNPs Extraction and Imputation

The location and range of candidate genes data were searched using Ensembl genome browser 84 and the National Center for Biotechnology Information (NCBI) website (GRCh38 /hg38 ver.). To cover the promoter regions of most candidate genes, we included SNPs from candidate genes that were within 20 Kb from the start and the end positions of each range [95]. Of the 555,117 SNPs mentioned in the above GWAS data, 2482 SNPs were extracted from 136 genes and one gene had an empty extraction result. After the Hardy–Weinberg equilibrium (HWE) test, 2472 SNPs were selected with HWE > 0.001.

We also performed an SNP imputation based on HapMap 2.0 (release 22/hg18) with the same range of candidate genes. In total, 9750 SNPs with quality scores > 0.3 were obtained by imputation, and 9083 SNPs with minor allele frequency (MAF) > 0.01 were included in the GxE analysis. In total, 11,555 SNPs were tested in the GxE analysis.

### 4.6. Environmental Factors Related to Estrogen Exposure

We selected seven variables related to estrogen exposure in this analysis including body mass index (BMI), height, age at menarche, age at first full term pregnancy (FFTP), number of children, breastfeeding, and total duration of breastfeeding. All exposures were continuous variables, except for breastfeeding which was collected as yes/no, and most of the continuous variables had missing values. We performed an imputation for missing data to the median value of each variable in the stratified birth year groups (1930–1950, 1950–1959, and 1960–1969), and therefore the effect of birth cohort could be considered without reducing power with the sample size (Table 1 and Table S1).

To assess the adjusted effect of breastfeeding duration, breastfeeding duration per child was calculated by dividing the total duration of breastfeeding by the number of children. We also made a new variable to examine the combined effects of age at menarche and age at FFTP. The duration of estrogen exposure before FFTP (EEBF) was calculated by subtracting the age at menarche from the age at FFTP for parous women. For nulliparous women, EEBF was calculated by subtracting the age at menarche from the age at interview for pre-menopause women, or from the age at menopause for post-menopause women.

To reduce computational burden and to facilitate interpretation, all continuous variables were dichotomized by quartile of risk direction. This resulted in eight environmental factors in the GxE analysis that were categorized as follows: age at menarche ($\geq$14 years vs. <14 years, 25% quartile); age at FFTP (<27 years vs. nulliparity or $\geq$27 years, 75% quartile); height (<160 cm vs. $\geq$160 cm. 75% quartile); BMI (<25 kg/m$^2$ vs. $\geq$25 kg/m$^2$ Korea criteria of obesity); number of children ($\geq$2 vs. <2, 25% quartile); breastfeeding (often vs. never);

breastfeeding duration per child (≥2.5 month vs. <2.5 month, 25% quartile); and EEBF (<13 years vs. ≥13 years, 75% quartile). Likewise, continuous variables from the BCAC were dichotomized by quartile of risk direction in the European and Asian populations, respectively, as follows: age at menarche (≥12 years vs. <12 years, 25% quartile in both European and Asian populations); height (<169 cm vs. ≥169 cm, 75% quartile in European population and <160 cm vs. ≥160 cm, 75% quartile in Asian population); and age at FFTP (<28 years vs. nulliparity or ≥28 years, 75% quartile in both European and Asian populations).

The estrogen receptor (ER) or progesterone receptor (PR) status was collected by the medical and pathological records and was determined when the tumor cells showed 10% or more positive nuclear staining from an immunohistochemistry assay [78].

### 4.7. Statistical Analysis

To estimate the effect of each environmental factor for breast cancer with odds ratios (OR) and 95% confidence intervals (CI), age, family history, parity, age at FFTP, age at menarche, and menopausal status were commonly included in the model. For variables regarding parity such as the number of children, breastfeeding history, and breastfeeding duration per child, further adjustments were performed by adding the number of children to the model. As an exception, age, family history, parity, and menopausal status were included in the model when the analysis was performed for EEBF. These models were consistently used in the GxE analysis, and in the logistic regression to estimate genetic risk factors, which showed significant interactions in the GxE analysis.

GxEScan (ver. Beta 0.4.0; http://biostats.usc.edu/software (accessed on 1 June 2016)) was mainly used in this study to conduct the GxE analysis. This software was developed by Gauderman et al. [13] and was designed to conduct GxE analysis in case-control study designs using the traditional GxE test and also a few two-step methods. Especially, the two-step methods have generally been proposed to provide greater statistical power than the traditional case-control GxE test while preserving the type I error rate. Four two-step methods were provided by GxEScan, i.e., DG | GxE, GE | GxE, EDGxE, and GE | 2df and we mainly observed GxE results from these methods. There were two hypotheses testing approaches in the two-step methods, 'subset testing' and 'weighted testing'. Detailed descriptions of these hypotheses can be found elsewhere [13,96] and both were used in this study.

As an initial step, we carried out GxE analyses by birth year groups (1930–1950, 1950–1959, and 1960–1969) because each birth year group had a different distribution of environmental factors (Table S2). The results of GxE analyses of each birth year group were inconsistent and were not statistically significant, possibly due to the modest sample size. For these reasons, we performed a meta-analysis of the three sets of results obtained from each birth year group by using METAL software [97], and then compared our results with the GxE result from the overall pooled data. There was a negligible difference between the two results; we presented the result from the pooled data analysis.

PLINK (ver. 1.07; http://pngu.mgh.harvard.edu/purcell/plink/ (accessed on 7 March 2016)) was used to examine GWAS data and for extracting candidate SNPs. Overall management of questionnaire data and statistical analyses were performed by SAS (ver. 9.4, SAS Institute Inc., Cary, NC, USA).

Replication analyses were performed in the two datasets for gene x height and for gene x age at menarche, separately. Interactions were tested by ethnicity first, and then in each study that showed the consistent effects of exposures with SEBCS. SAS (ver. 9.4; SAS Institute Inc., Cary, NC, USA) was used to test replication.

## 5. Conclusions

Through the GxE test between 136 candidate genes involved in estrogen metabolism and eight environmental factors related to estrogen exposure, we found two interactions in middle-aged Korean females. Although the results of our study were not replicated,

they do suggest specificity in Koreans for breast cancer risk. The explainable interaction might be a better method to interpret and apply the results than deep medicine or artificial intelligence [98]. We expect to make a step forward in predicting and preventing breast cancer, as more interactions are to be investigated.

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10.3390/cancers13102370/s1, Figure S1: Result plot of GxE between 11,555 SNPs and age at menarche for breast cancer risk by using the weighted GE | 2df method, Figure S2: Result plot of GxE between 11,555 SNPs and height for breast cancer risk by using the GE | 2df method, Figure S3: Forest plot presenting risk of age at menarche for breast cancer by study in BCAC, Figure S4: Forest plot presenting risk of height for breast cancer by study in BCAC, Table S1: Distributions of demographic and reproductive factors and associations with the risk of breast cancer, Seoul Breast Cancer Study (pre-imputation), Table S2: Distributions of demographic and reproductive factors in cases and controls by birth year groups, Seoul Breast Cancer Study (imputed + binary), Table S3: Studies including BCAC replication data set, Table S4: Distributions of demographic and reproductive factors and associations with the risk of breast cancer by ethnicity, Breast Cancer Association Consortium replication dataset for age at menarche x NCOA1, Table S5: Distributions of demographic and reproductive factors and associations with the risk of breast cancer by ethnicity, Breast Cancer Association Consortium replication dataset for height x ESR1, Table S6: Association between age at menarche and the risk of breast cancer in studies from BCAC that showed consistence with SeBCS, Table S7: Association between height and the risk of breast cancer in studies from BCAC that showed consistence with SeBCS, Table S8: Association between environmental factor and breast cancer by genotypes according to ER/PR status, Seoul Breast Cancer Study.

**Author Contributions:** Conceptualization, J.P. and J.-Y.C.; Methodology, J.P. and J.-Y.C.; Formal analysis, J.P., J.C., S.C. and N.C; Investigation, J.P.; Data curation, management and provision, J.-Y.C., S.K.P., W.H., D.-Y.N., S.-H.A., J.W.L., M.K.K., S.H.J., W.W., M.K.B., Q.W., J.D., K.M. (Kyriaki Michailidou), M.S., D.M.C., P.A.H., R.M., K.C., P.H., L.R.T., A.V.P., F.J.C., J.E.O., E.J.S., R.R., S.E.B., H.F., D.L., A.B., K.M. (Keitaro Matsuo), H.I., P.G., T.T., R.K., M.K.S., A.H.W., C.-C.T., A.C. and S.S.C.; kConFab Investigators, I.L.A., J.L.H., M.C.S., P.-E.W., C.-Y.S., P.A.F., A.B.E., K.M. (Kenneth Muir), A.L. (Artitaya Lophatananon), H.B., V.A., M.E.J., A.J.S., R.H., Y.-D.K., M.H., J.L., A.M., J.M.H., J.B., A.G.-N., C.A.H., T.D., N.V.B., S.H.T., N.A.M.T., O.F., N.J., M.G., R.W., C.B., H.N., A.L. (Annika Lindblom), C.W. and V.N.K.; NBCS Collaborators, R.A.E.M.T., B.A.M.H.-G., P.R., B.B., U.H., M.M., J.V.L., M.E.M., A.M.D., P.D.P.P., D.F.E., K.-Y.Y. and D.K.; Validation, S.H.J., W.W. and J.P.; Visualization, J.P.; Writing—original draft preparation, J.P.; Writing—review and editing, all authors; Supervision, J.-Y.C.; Funding acquisition, J.-Y.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the Seoul National University Hospital (2021) and the Seoul National University Hospital Research Fund (0320150110 (2015–1126)). BCAC is funded by Cancer Research UK (C1287/A16563), the European Union's Horizon 2020 Research and Innovation Programme (grant numbers 634935 and 633784 for BRIDGES and B-CAST, respectively), and by the European Community's Seventh Framework Programme under grant agreement number 223175 (grant number HEALTH-F2-2009-223175) (COGS). The EU Horizon 2020 Research and Innovation Programme funding source had no role in study design, data collection, data analysis, data interpretation, or writing of the report. Genotyping of the OncoArray was funded by the NIH grant U19 CA148065, and Cancer UK grant C1287/A16563 and the PERSPECTIVE project supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research (grant GPH-129344), and the Ministère de l'Économie, Science et Innovation du Québec through Genome Québec and the PSRSIIRI-701 grant, and the Quebec Breast Cancer Foundation. Funding for the iCOGS infrastructure came from the European Community's Seventh Framework Programme under grant agreement no. 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, and C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065, and 1U19 CA148112 the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, and Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund. The DRIVE Consortium was funded by U19 CA148065. The Australian Breast Cancer Family Study (ABCFS) was supported by

grant UM1 CA164920 from the National Cancer Institute (USA). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government or the BCFR. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia), and the Victorian Breast Cancer Research Consortium. J.L.H. is a National Health and Medical Research Council (NHMRC) Senior Principal Research Fellow. M.C.S. is a NHMRC Senior Research Fellow. The ABCS study was supported by the Dutch Cancer Society (grants NKI 2007-3839 and 2009 4363). The ACP study is funded by the Breast Cancer Research Trust, UK. K.M. and A.L. are supported by the NIHR Manchester Biomedical Research Centre, the Allan Turing Institute under the EPSRC grant EP/N510129/1. The work of the BBCC was partly funded by ELAN-Fond of the University Hospital of Erlangen. The BBCS is funded by Cancer Research UK and Breast Cancer Now and acknowledges NHS funding to the NIHR Biomedical Research Centre, and the National Cancer Research Network (NCRN). The CECILE study was supported by the Fondation de France, Institut National du Cancer (INCa), the Ligue Nationale contre le Cancer, Agence Nationale de Sécurité Sanitaire, de l'Alimentation, de l'Environnement et du Travail (ANSES), and the Agence Nationale de la Recherche (ANR). The CGPS was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council, and Herlev and Gentofte Hospital. The CNIO-BCS was supported by the Instituto de Salud Carlos III, the Red Temática de Investigación Cooperativa en Cáncer, and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923 and PI12/00070). The American Cancer Society funds the creation, maintenance, and updating of the CPS-II cohort. The California Teachers Study and the research reported in this publication were supported by the National Cancer Institute of the National Institutes of Health under award numbers U01-CA199277, P30-CA033572, P30-CA023100, UM1-CA164917, and R01-CA077398. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. The collection of cancer incidence data used in the California Teachers Study was supported by the California Department of Public Health pursuant to California Health and Safety Code Section 103885; Centers for Disease Control and Prevention's National Program of Cancer Registries, under cooperative agreement 5NU58DP006344; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201800032I awarded to the University of California, San Francisco; contract HHSN261201800015I awarded to the University of Southern California; and contract HHSN261201800009I awarded to the Public Health Institute. The opinions, findings, and conclusions expressed herein are those of the author(s) and do not necessarily reflect the official views of the State of California, Department of Public Health, the National Cancer Institute, the National Institutes of Health, the Centers for Disease Control and Prevention or their Contractors and Subcontractors, or the Regents of the University of California or any of its programs. The ESTHER study was supported by a grant from the Baden Württemberg Ministry of Science, Research and Arts. Additional cases were recruited in the context of the VERDI study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe). The GENICA was funded by the Federal Ministry of Education and Research (BMBF) Germany grants 01KW9975/5, 01KW9976/8, 01KW9977/0, and 01KW0114, the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany. GLACIER was supported by Breast Cancer Now, CRUK and Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London. The HEBCS was financially supported by the Helsinki University Hospital Research Fund, the Finnish Cancer Society, and the Sigrid Juselius Foundation. The HERPACC was supported by MEXT Kakenhi (no. 170150181 and 26253041) from the Ministry of Education, Science, Sports, Culture and Technology of Japan, by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from Ministry Health, Labour, and Welfare of Japan, by Health and Labour Sciences Research Grants for Research on Applying Health Technology from the Ministry Health of Labour and Welfare of Japan, by National Cancer Center Research and Development Fund, and "Practical Research for Innovative Cancer Control (15ck0106177h0001 and 20ck0106553)" from Japan Agency for Medical Research and development, AMED, and Cancer Bio Bank Aichi. The HMBCS was supported by a grant from the German Research Foundation (Do 761/10-1). ICICLE was supported by Breast Cancer Now,

CRUK, and Biomedical Research Centre at Guy's and St Thomas' N.H.S. Foundation Trust and King's College London. Financial support for KARBAC was provided through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Cancer Society, The Gustav V Jubilee foundation, and Bert von Kantzows foundation. The KARMA study was supported by Märit and Hans Rausings Initiative Against Breast Cancer. The KBCP was financially supported by the special Government Funding (VTR) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, and by the strategic funding of the University of Eastern Finland. kConFab is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia, and the Cancer Foundation of Western Australia. Financial support for the AOCS was provided by the United States Army Medical Research and Materiel Command (DAMD17-01-1-0729), Cancer Council Victoria, Queensland Cancer Fund, Cancer Council New South Wales, Cancer Council South Australia, The Cancer Foundation of Western Australia, Cancer Council Tasmania, and the National Health and Medical Research Council of Australia (NHMRC) (400413, 400281, and 199600). G.C.T. and P.W. are supported by the NHMRC. R.B was a Cancer Institute NSW Clinical Research Fellow. Clinical Follow Up Study has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia, and the National Institute of Health (USA). LAABC is supported by grants (1RB-0287, 3PB-0102, 5PB-0018, 10PB-0098) from the California Breast Cancer Research Program. Incident breast cancer cases were collected by the USC Cancer Surveillance Program (CSP) which is supported under subcontract by the California Department of Health. The CSP is also part of the National Cancer Institute's Division of Cancer Prevention and Control Surveillance, Epidemiology, and End Results Program, under contract number N01CN25403. LMBC is supported by the 'Stichting tegen Kanker'. D.L. is supported by the FWO. The MARIE study was supported by the Deutsche Krebshilfe e.V. (70-2892-BR I, 106332, 108253, 108419, 110826, and 110828), the Hamburg Cancer Society, the German Cancer Research Center (DKFZ), and the Federal Ministry of Education and Research (BMBF) Germany (01KH0402). MBCSG is supported by grants from the Italian Association for Cancer Research (AIRC). The MCBCS was supported by the NIH grants R35CA253187, R01CA192393, R01CA116167, and R01CA176785 an NIH Specialized Program of Research Excellence (SPORE) in Breast Cancer (P50CA116201), and the Breast Cancer Research Foundation. The Melbourne Collaborative Cohort Study (MCCS) cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further augmented by Australian National Health and Medical Research Council grants 209057, 396414, and 1074383 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database. The MEC was supported by NIH grants CA63464, CA54281, CA098758, CA132839, and CA164973. MYBRCA is funded by research grants from the Malaysian Ministry of Higher Education (UM.C/HlR/MOHE/06) and Cancer Research Malaysia. MYMAMMO is supported by research grants from Yayasan Sime Darby LPGA Tournament and Malaysian Ministry of Higher Education (RP046B-15HTM). The NBCS has received funding from the K.G. Jebsen Centre for Breast Cancer Research, the Research Council of Norway grant 193387/V50 (to A-L Børresen-Dale and V.N. Kristensen) and grant 193387/H10 (to A-L Børresen-Dale and V.N. Kristensen), South Eastern Norway Health Authority (grant 39346 to A-L Børresen-Dale), and the Norwegian Cancer Society (to A-L Børresen-Dale and V.N. Kristensen). The NBHS was supported by NIH grant R01CA100374. Biological sample preparation was conducted by the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The OBCS was supported by research grants from the Finnish Cancer Foundation, the Academy of Finland (grant number 250083, 122715, and Center of Excellence grant number 251314), the Finnish Cancer Foundation, the Sigrid Juselius Foundation, the University of Oulu, the University of Oulu Support Foundation, and the special Governmental EVO funds for Oulu University Hospital-based research activities. The Northern California Breast Cancer Family Registry (NC-BCFR) and Ontario Familial Breast Cancer Registry (OFBCR) were supported by grants U01CA164920 and U01CA167551 from the U.S. National Cancer Institute of the National Institutes of Health. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR) or the Colon Cancer Family Registry (CCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government or the BCFR or CCFR. The ORIGO study was supported by the Dutch Cancer Society (RUL 1997–1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16). The

PBCS was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The SASBAC study was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR.); the U.S. National Institute of Health (NIH); and the Susan G. Komen Breast Cancer Foundation. The SBCGS was supported primarily by NIH grants R01CA64277, R01CA148667, UMCA182910, and R37CA70867. Biological sample preparation was conducted by the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The scientific development and funding of this project were, in part, supported by the Genetic Associations and Mechanisms in Oncology (GAME-ON) Network U19 CA148065. The SBCS was supported by Sheffield Experimental Cancer Medicine Centre and Breast Cancer Now Tissue Bank. SEARCH is funded by Cancer Research UK (C490/A10124, C490/A16561] and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. The University of Cambridge has received salary support for PDPP from the NHS in the East of England through the Clinical Academic Reserve. SECS was supported by the BRL (Basic Research Laboratory) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2012-0000347). SGBCC is funded by the National Research Foundation Singapore, NUS start-up grant, National University Cancer Institute Singapore (NCIS) Centre Grant, Breast Cancer Prevention Programme, Asian Breast Cancer Research Fund, and the NMRC Clinician Scientist Award (SI Category). Additional controls were recruited by the Singapore Consortium of Cohort Studies-Multi-ethnic cohort (SCCS-MEC), which was funded by the Biomedical Research Council, grant number 05/1/21/19/425. SKKDKFZS is supported by the DF.Z. The TWBCS is supported by the Taiwan Biobank project of the Institute of Biomedical Sciences, Academia Sinica, Taiwan. The UKBGS is funded by Breast Cancer Now and the Institute of Cancer Research (ICR), London. The ICR acknowledges NHS funding to the NIHR Biomedical Research Centre. UKBGS thanks Breast Cancer Now and the Institute of Cancer Research for support and funding of the Generations Study, and the study participants, study staff, and the doctors, nurses and other health care providers and health information sources who have contributed to the study. We acknowledge NHS funding to the Royal Marsden/ICR NH.R Biomedical Research Centre. We acknowledge funding to the Manchester NIHR Biomedical Research Centre (IS-BRC-1215-20007).

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and was approved by the Committee on Human Research of the Seoul National University Hospital (IRB no. H-0503-144-004).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data are available on request due to restraints imposed by the ethics committees of individual. To request the data from the Breast Cancer Association Consortium (BCAC), readers are instructed to submit a concept proposal, which will be reviewed by the BCAC Data Access Coordination Committee (DACC).

**Acknowledgments:** We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians, and administrative staff who have enabled this work to be carried out. The COGS study would not have been possible without the contributions of the following: Per Hall (COGS), Douglas F. Easton, Paul Pharoah, Kyriaki Michailidou, Manjeet K. Bolla, Qin Wang (BCAC); Andrew Berchuck (OCAC); Rosalind A. Eeles, Douglas F. Easton, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL); Georgia Chenevix-Trench, Antonis Antoniou, Lesley McGuffog, Fergus Couch and Ken Offit (CIMBA); Joe Dennis, Alison M. Dunning, Andrew Lee, Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory; Javier Benitez, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit; Jacques Simard and Daniel C. Tessier, Francois Bacot, Daniel Vincent, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre; Stig E. Bojesen, Sune F. Nielsen, Borge G. Nordestgaard, and the staff of the Copenhagen DNA laboratory; and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility. ABCFS thank Maggie Angelakos, Judi Maskiell, Gillian Dite. ABS. thanks the Blood bank Sanquin, The Netherlands. The ACP study wishes to thank the participants in the Thai Breast Cancer study. Special thanks also go to the Thai Ministry of Public Health (MOPH), doctors, and nurses who helped with the data collection process. Finally, the study would like to thank Prat Boonyawongviroj, the former Permanent Secretary of MOPH and Pornthep Siriwanarungsan, the former Department Director-General of Disease Control who have supported the study throughout. BBCS thanks Eileen Williams, Elaine Ryder-Mills, and Kara Sargus. CGPS thanks staff and participants of the Copenhagen

General Population Study for the excellent technical assistance, Dorthe Uldall Andersen, Maria Birna Arnadottir, Anne Bank, and Dorthe Kjeldgård Hansen. CNIO-BCS thanks Guillermo Pita, Charo Alonso, Nuria Álvarez, Pilar Zamora, Primitiva Menendez, and the Human Genotyping CEGEN Unit (CNIO). Investigators from the CPS-II cohort thank the participants and Study Management Group for their invaluable contributions to this research. They also acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, as well as cancer registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program. The authors would like to thank the California Teachers Study Steering Committee that is responsible for the formation and maintenance of the study within which this research was conducted. A full list of California Teachers Study team members is available at https://www.calteachersstudy.org/team. ESTHER thanks Hartwig Ziegler, Sonja Wolf, Volker Hermann, Christa Stegmaier, Katja Butterbach. The GENICA Network; Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany (Hiltrud Brauch, Wing-Yee Lo, and R.H.); Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany (Y-D. K. and Christian Baisch); Institute of Pathology, University of Bonn, Germany (Hans-Peter Fischer); Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany (Ute Hamann); Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany (Thomas Brüning, Beate Pesch, Sylvia Rabstein, Anne Lotz); and the Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany (Volker Harth). GLACIER thanks Kelly Kohut, Patricia Gorman, Maria Troy. HEBCS thanks Sofia Khan, Johanna Kiiski, Carl Blomqvist, Kristiina Aittomäki, Rainer Fagerholm, Kirsimari Aaltonen, Karl von Smitten, and Irja Erkkilä. HMBCS thanks Natalia Antonenkova, Peter Schürmann, Peter Hillemanns, Hans Christiansen, and Johann H. Karstens. ICICLE thanks Kelly Kohut, Michele Caneppele, and Maria Troy. KARMA and SASBAC thank the Swedish Medical Research Counsel. KBCP thanks Eija Myöhänen. kConFab/AOCS wish to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study for their contributions to this resource, and the many families who contribute to kConFab. LAABC thanks all the study participants and the entire data collection team, especially Annie Fung and June Yashiki. LMBC thanks Gilian Peuteman, Thomas Van Brussel, EvyVanderheyden, and Kathleen Corthouts. MARIE thanks Petra Seibold, Dieter Flesch-Janys, Judith Heinz, Nadia Obi, Alina Vrieling, Sabine Behrens, Ursula Eilber, Muhabbet Celik, Til Olchers, and Stefan Nickels. MBCSG (Milan Breast Cancer Study Group): Paolo Peterlongo, Siranoush Manoukian, Bernard Peissel, Jacopo Azzollini, Erica Rosina, Daniela Zaffaroni, Irene Feroce, Mariarosaria Calvello, Aliana Guerrieri Gonzaga, Monica Marabelli, Davide Bondavalli, and the personnel of the Cogentech Cancer Genetic Test Laboratory. The MCCS was made possible by the contribution of many people, including the original investigators, the teams that recruited the participants and continue working on follow-up, and the many thousands of Melbourne residents who continue to participate in the study. MYBRCA thanks study participants and research staff (particularly Patsy Ng, Nurhidayu Hassan, Yoon Sook-Yee, Daphne Lee, Lee Sheau Yee, Phuah Sze Yee, and Norhashimah Hassan) for their contributions and commitment to this study. The following are NBCS collaborators: Kristine K. Sahlberg, Lars Ottestad, Rolf Kåresen, Ellen Schlichting, Marit Muri Holmen, Toril Sauer, Vilde Haakensen, Olav Engebråten, Bjørn Naume, Alexander Fosså, Cecile E. Kiserud, Kristin V. Reinertsen, Åslaug Helland, Margit Riis, Jürgen Geisler, OSBREAC and Grethe I. Grenaker Alnæs. NBHS and SBCGS thank study participants and research staff for their contributions and commitment to the studies. OBCS thanks Arja Jukkola-Vuorinen, Mervi Grip, Saila Kauppila, Meeri Otsukka, Leena Keskitalo, and Kari Mononen for their contributions to this study. The OFBCR thanks Teresa Selander, Nayana Weerasooriya, and Steve Gallinger. ORIGO thanks E. Krol-Warmerdam, and J. Blom for patient accrual, administering questionnaires, and managing clinical information. The LUMC survival data were retrieved from the Leiden hospital-based cancer registry system (ONCDOC) with the help of J. Molenaar. PBCS thanks Louise Brinton, Mark Sherman, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao, and Michael Stagner. SBCS thanks Sue Higham, Helen Cramp, Dan Connley, Ian Brock, Sabapathy Balasubramanian, and Malcolm W.R. Reed. We thank the SEARCH and EPIC teams. SGBCC thanks the participants and all research coordinators for their excellent help with recruitment, data, and sample collection. SKKDKFZS. thanks all study participants, clinicians, family doctors, researchers and technicians for their contributions and commitment to this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References


1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**. [CrossRef]
2. Kang, S.Y.; Kim, Y.S.; Kim, Z.; Kim, H.Y.; Kim, H.J.; Park, S.; Bae, S.Y.; Yoon, K.H.; Lee, S.B.; Lee, S.K.; et al. Breast Cancer Statistics in Korea in 2017: Data from a Breast Cancer Registry. *J. Breast Cancer* **2020**, *23*, 115–128. [CrossRef] [PubMed]
3. Kluttig, A.; Schmidt-Pokrzywniak, A. Established and Suspected Risk Factors in Breast Cancer Aetiology. *Breast Care* **2009**, *4*, 82–87. [CrossRef] [PubMed]
4. Key, T.J.; Verkasalo, P.K.; Banks, E. Epidemiology of breast cancer. *Lancet Oncol.* **2001**, *2*, 133–140. [CrossRef]
5. Rojas, K.; Stuckey, A. Breast Cancer Epidemiology and Risk Factors. *Clin. Obstet. Gynecol.* **2016**, *59*, 651–672. [CrossRef] [PubMed]
6. Hindorff, L.A.; Sethupathy, P.; Junkins, H.A.; Ramos, E.M.; Mehta, J.P.; Collins, F.S.; Manolio, T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9362–9367. [CrossRef]
7. Michailidou, K.; The Breast and Ovarian Cancer Susceptibility Collaboration; Hall, P.; Gonzalez-Neira, A.; Ghoussaini, M.; Dennis, J.; Milne, R.L.; Schmidt, M.K.; Chang-Claude, J.; Bojesen, S.E.; et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat. Genet.* **2013**, *45*, 353–361. [CrossRef]
8. Michailidou, K.; Beesley, J.; Lindstrom, S.; Canisius, S.; Dennis, J.; Lush, M.J.; Maranian, M.J.; Bolla, M.K.; Wang, Q.; Shah, M.; et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat. Genet.* **2015**, *47*, 373–380. [CrossRef]
9. Maas, P.; Barrdahl, M.; Joshi, A.D.; Auer, P.L.; Gaudet, M.M.; Milne, R.L.; Schumacher, F.R.; Anderson, W.F.; Check, D.; Chattopadhyay, S.; et al. Breast Cancer Risk from Modifiable and Nonmodifiable Risk Factors among White Women in the United States. *JAMA Oncol.* **2016**, *2*, 1295–1302. [CrossRef]
10. Han, M.-R.; Long, J.; Choi, J.-Y.; Low, S.-K.; Kweon, S.-S.; Zheng, Y.; Cai, Q.; Shi, J.; Guo, X.; Matsuo, K.; et al. Genome-wide association study in East Asians identifies two novel breast cancer susceptibility loci. *Hum. Mol. Genet.* **2016**, *25*, 3361–3371. [CrossRef] [PubMed]
11. Mavaddat, N.; Antoniou, A.C.; Easton, D.F.; Garcia-Closas, M. Genetic susceptibility to breast cancer. *Mol. Oncol.* **2010**, *4*, 174–191. [CrossRef]
12. Stratton, M.R.; Rahman, N. The emerging landscape of breast cancer susceptibility. *Nat. Genet.* **2007**, *40*, 17–22. [CrossRef]
13. Gauderman, W.J.; Zhang, P.; Morrison, J.L.; Lewinger, J.P. Finding novel genes by testing G × E interactions in a genome-wide association study. *Genet. Epidemiol.* **2013**, *37*, 603–613. [CrossRef]
14. Nickels, S.; Truong, T.; Hein, R.; Stevens, K.; Buck, K.; Behrens, S.; Eilber, U.; Schmidt, M.; Häberle, L.; Vrieling, A.; et al. Evidence of Gene-Environment Interactions between Common Breast Cancer Susceptibility Loci and Established Environmental Risk Factors. *PLoS Genet.* **2013**, *9*, e1003284. [CrossRef] [PubMed]
15. Hunter, D.J. Gene-environment interactions in human diseases. *Nat. Rev. Genet.* **2005**, *6*, 287–298. [CrossRef] [PubMed]
16. Thomas, D. Gene-environment-wide association studies: Emerging approaches. *Nat. Rev. Genet.* **2010**, *11*, 259–272. [CrossRef] [PubMed]
17. Milne, R.L.; Network, G.; Gaudet, M.M.; Spurdle, A.B.; Fasching, P.A.; Couch, F.J.; Benítez, J.; Pérez, J.I.A.; Zamora, M.P.; Malats, N.; et al. Assessing interactions between the associations of common genetic susceptibility variants, reproductive history and body mass index with breast cancer risk in the breast cancer association consortium: A combined case-control study. *Breast Cancer Res.* **2010**, *12*, R110. [CrossRef]
18. Travis, R.C.; Reeves, G.K.; Green, J.; Bull, D.; Tipper, S.J.; Baker, K.; Beral, V.; Peto, R.; Bell, J.; Zelenika, D.; et al. Gene-environment interactions in 7610 women with breast cancer: Prospective evidence from the Million Women Study. *Lancet* **2010**, *375*, 2143–2151. [CrossRef]
19. Campa, D.; Kaaks, R.; Le Marchand, L.; Haiman, C.A.; Travis, R.C.; Berg, C.D.; Buring, J.E.; Chanock, S.J.; Diver, W.R.; Dostal, L.; et al. Interactions between Genetic Variants and Breast Cancer Risk Factors in the Breast and Prostate Cancer Cohort Consortium. *J. Natl. Cancer Inst.* **2011**, *103*, 1252–1263. [CrossRef] [PubMed]
20. Hein, R.; Network, T.G.; Flesch-Janys, D.; Dahmen, N.; Beckmann, L.; Lindström, S.; Schoof, N.; Czene, K.; Mittelstraß, K.; Illig, T.; et al. A genome-wide association study to identify genetic susceptibility loci that modify ductal and lobular postmenopausal breast cancer risk associated with menopausal hormone therapy use: A two-stage design with replication. *Breast Cancer Res. Treat.* **2013**, *138*, 529–542. [CrossRef]
21. Rudolph, A.; Hein, R.; Lindström, S.; Beckmann, L.; Behrens, S.; Liu, J.; Aschard, H.; Bolla, M.K.; Wang, J.; Truong, T.; et al. Genetic modifiers of menopausal hormone replacement therapy and breast cancer risk: A genome-wide interaction study. *Endocr. Relat. Cancer* **2013**, *20*, 875–887. [CrossRef]
22. Barrdahl, M.; Canzian, F.; Joshi, A.D.; Travis, R.C.; Chang-Claude, J.; Auer, P.L.; Gapstur, S.M.; Gaudet, M.; Diver, W.R.; Henderson, B.E.; et al. Post-GWAS gene-environment interplay in breast cancer: Results from the Breast and Prostate Cancer Cohort Consortium and a meta-analysis on 79,000 women. *Hum. Mol. Genet.* **2014**, *23*, 5260–5270. [CrossRef]
23. Schoeps, A.; Rudolph, A.; Seibold, P.; Dunning, A.M.; Milne, R.L.; Bojesen, S.E.; Swerdlow, A.; Andrulis, I.; Brenner, H.; Behrens, S.; et al. Identification of New Genetic Susceptibility Loci for Breast Cancer through Consideration of Gene-Environment Interactions. *Genet. Epidemiol.* **2014**, *38*, 84–93. [CrossRef]

24. Rudolph, A.; Milne, R.L.; Truong, T.; Knight, J.A.; Seibold, P.; Flesch-Janys, D.; Behrens, S.; Eilber, U.; Bolla, M.K.; Wang, Q.; et al. Investigation of gene-environment interactions between 47 newly identified breast cancer susceptibility loci and environmental risk factors. *Int. J. Cancer* **2015**, *136*, E685–E696. [CrossRef] [PubMed]
25. Hutter, C.M.; Mechanic, L.E.; Chatterjee, N.; Kraft, P.; Gillanders, E.M.; On Behalf of the NCI Gene-Environment Think Tank. Gene-Environment Interactions in Cancer Epidemiology: A National Cancer Institute Think Tank Report. *Genet. Epidemiol.* **2013**, *37*, 643–657. [CrossRef]
26. Travis, R.C.; Key, T.J. Oestrogen exposure and breast cancer risk. *Breast Cancer Res.* **2003**, *5*, 239–247. [CrossRef] [PubMed]
27. Cohen, A.; Burgos-Aceves, M.A.; Smith, Y. Estrogen repression of microRNA as a potential cause of cancer. *Biomed. Pharmacother.* **2016**, *78*, 234–238. [CrossRef]
28. Cohen, A.; Burgos-Aceves, M.A.; Kahan, T.; Smith, Y. Estrogen Repression of MicroRNAs Is Associated with High Guanine Content in the Terminal Loop Sequences of Their Precursors. *Biomedicines* **2017**, *5*, 47. [CrossRef]
29. Xu, J.; Wu, R.-C.; O'Malley, B.W. Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. *Nat. Rev. Cancer* **2009**, *9*, 615–630. [CrossRef]
30. Walsh, C.A.; Qin, L.; Tien, J.C.-Y.; Young, L.S.; Xu, J. The Function of Steroid Receptor Coactivator-1 in Normal Tissues and Cancer. *Int. J. Biol. Sci.* **2012**, *8*, 470–485. [CrossRef] [PubMed]
31. Fleming, F.J.; Hill, A.D.K.; McDermott, E.W.; O'Higgins, N.J.; Young, L.S. Differential Recruitment of Coregulator Proteins Steroid Receptor Coactivator-1 and Silencing Mediator for Retinoid and Thyroid Receptors to the Estrogen Receptor-Estrogen Response Element by β-Estradiol and 4-Hydroxytamoxifen in Human Breast Cancer. *J. Clin. Endocrinol. Metab.* **2004**, *89*, 375–383. [CrossRef]
32. Fleming, F.J.; Myers, E.R.; Kelly, G.M.; Crotty, T.B.; McDermott, E.W.; O'Higgins, N.J.; Hill, A.D.K.; Young, L.S. Expression of SRC-1, AIB1, and PEA3 in HER2 mediated endocrine resistant breast cancer; a predictive role for SRC-1. *J. Clin. Pathol.* **2004**, *57*, 1069–1074. [CrossRef] [PubMed]
33. Shiozawa, T.; Shih, H.-C.; Miyamoto, T.; Feng, Y.-Z.; Uchikawa, J.; Itoh, K.; Konishi, I. Cyclic Changes in the Expression of Steroid Receptor Coactivators and Corepressors in the Normal Human Endometrium. *J. Clin. Endocrinol. Metab.* **2003**, *88*, 871–878. [CrossRef]
34. Wieser, F.; Schneeberger, C.; Hudelist, G.; Singer, C.; Kurz, C.; Nagele, F.; Gruber, C.; Huber, J.C.; Tschugguel, W. Endometrial nuclear receptor co-factors SRC-1 and N-CoR are increased in human endometrium during menstruation. *Mol. Hum. Reprod.* **2002**, *8*, 644–650. [CrossRef]
35. Chen, L.; Kang, H.; Jin, G.-J.; Chen, X.; Zhang, Q.-Y.; Lao, W.-T.; Li, R. The association between a novel polymorphism (rs1062577) in ESR1 and breast cancer susceptibility in the Han Chinese women. *Gynecol. Endocrinol.* **2016**, *32*, 553–556. [CrossRef]
36. Li, N.; Dong, J.; Hu, Z.; Shen, H.; Dai, M. Potentially functional polymorphisms in ESR1 and breast cancer risk: A meta-analysis. *Breast Cancer Res. Treat.* **2009**, *121*, 177–184. [CrossRef] [PubMed]
37. Zheng, W.; Long, J.; Gao, Y.-T.; Li, C.; Zheng, Y.; Xiang, Y.-B.; Wen, W.; Levy, S.; Deming, S.L.; Haines, J.L.; et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat. Genet.* **2009**, *41*, 324–328. [CrossRef] [PubMed]
38. Dunning, A.M.; Healey, C.S.; Baynes, C.; Maia, A.-T.; Scollen, S.; Vega, A.; Rodríguez, R.; Barbosa-Morais, N.L.; Ponder, B.A.; Low, Y.-L.; et al. Association of ESR1 gene tagging SNPs with breast cancer risk. *Hum. Mol. Genet.* **2009**, *18*, 1131–1139. [CrossRef] [PubMed]
39. Robinson, D.R.; Wu, Y.-M.; Vats, P.; Su, F.; Lonigro, R.J.; Cao, X.; Kalyana-Sundaram, S.; Wang, R.; Ning, Y.; Hodges, L.; et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat. Genet.* **2013**, *45*, 1446–1451. [CrossRef] [PubMed]
40. Jeselsohn, R.; Buchwalter, G.; De Angelis, C.; Brown, M.; Schiff, R. ESR1 mutations—A mechanism for acquired endocrine resistance in breast cancer. *Nat. Rev. Clin. Oncol.* **2015**, *12*, 573–583. [CrossRef] [PubMed]
41. Alluri, P.G.; Speers, C.; Chinnaiyan, A.M. Estrogen receptor mutations and their role in breast cancer progression. *Breast Cancer Res.* **2014**, *16*, 494. [CrossRef]
42. Zhang, B.; Shu, X.-O.; Delahanty, R.J.; Zeng, C.; Michailidou, K.; Bolla, M.K.; Wang, Q.; Dennis, J.; Wen, W.; Long, J.; et al. Height and Breast Cancer Risk: Evidence from Prospective Studies and Mendelian Randomization. *J. Natl. Cancer Inst.* **2015**, *107*, 15–32. [CrossRef]
43. Friedenreich, C.M. Review of anthropometric factors and breast cancer risk. *Eur. J. Cancer Prev.* **2001**, *10*, 15–32. [CrossRef]
44. Okasha, M.; McCarron, P.; Gunnell, D.; Smith, G.D. Exposures in Childhood, Adolescence and Early Adulthood and Breast Cancer Risk: A Systematic Review of the Literature. *Breast Cancer Res. Treat.* **2003**, *78*, 223–276. [CrossRef]
45. Gunnell, D.; Okasha, M.; Smith, G.D.; Oliver, S.; Sandhu, J.; Holly, J. Height, leg length, and cancer risk: A systematic review. *Epidemiol. Rev.* **2001**, *23*, 313–342. [CrossRef] [PubMed]
46. Ahlgren, M.; Melbye, M.; Wohlfahrt, J.; Sørensen, T.I.A. Growth Patterns and the Risk of Breast Cancer in Women. *N. Engl. J. Med.* **2004**, *351*, 1619–1626. [CrossRef] [PubMed]
47. Simm, P.J.; Bajpai, A.; Russo, V.C.; Werther, G.A. Estrogens and growth. *Pediatr. Endocrinol. Rev.* **2008**, *6*, 32–41. [PubMed]
48. Khosla, S. Oestrogen, bones and men: When testosterone just isn't enough. *Clin. Endocrinol.* **2002**, *56*, 291–293. [CrossRef] [PubMed]
49. Rochira, V.; Kara, E.; Carani, C. The Endocrine Role of Estrogens on Human Male Skeleton. *Int. J. Endocrinol.* **2015**, *2015*, 165215. [CrossRef]
50. Carter, S.L. The Genetic Basis of Human Height: The Role of Estrogen. Ph.D. Thesis, Queensland University of Technology, Brisbane, QLD, Australia, 2008.

51. Grumbach, M.M.; Auchus, R.J. Estrogen: Consequences and Implications of Human Mutations in Synthesis and Action[1]. *J. Clin. Endocrinol. Metab.* **1999**, *84*, 4677–4694. [CrossRef] [PubMed]
52. Emons, J.; Chagin, A.S.; Sävendahl, L.; Karperien, M.; Wit, J.M. Mechanisms of growth plate maturation and epiphyseal fusion. *Horm. Res. Paediatr.* **2011**, *75*, 383–391. [CrossRef]
53. Chagin, A.; Sävendahl, L. Estrogens and growth: Review. *Pediatr. Endocrinol. Rev.* **2007**, *4*, 329–334. [PubMed]
54. Dahlgren, A.; Lundmark, P.; Axelsson, T.; Lind, L.; Syvänen, A.-C. Association of the Estrogen Receptor 1 (ESR1) Gene with Body Height in Adult Males from Two Swedish Population Cohorts. *PLoS ONE* **2008**, *3*, e1807. [CrossRef]
55. Schuit, S.C.E.; Van Meurs, J.B.J.; Bergink, A.P.; Van Der Klift, M.; Fang, Y.; Leusink, G.; Hofman, A.; Van Leeuwen, J.P.T.M.; Uitterlinden, A.G.; Pols, H.A.P. Height in Pre- and Postmenopausal Women Is Influenced by Estrogen Receptor $\alpha$ Gene Polymorphisms. *J. Clin. Endocrinol. Metab.* **2004**, *89*, 303–309. [CrossRef] [PubMed]
56. Lehrer, S.; Rabin, J.; Stone, J.; Berkowitz, G. Association of an Estrogen Receptor Variant with Increased Height in Women. *Horm. Metab. Res.* **1994**, *26*, 486–488. [CrossRef] [PubMed]
57. Dick, D.M.; Agrawal, A.; Keller, M.C.; Adkins, A.; Aliev, F.; Monroe, S.; Hewitt, J.K.; Kendler, K.S.; Sher, K.J. Candidate Gene-Environment Interaction Research. *Perspect. Psychol. Sci.* **2015**, *10*, 37–59. [CrossRef]
58. Munafò, M.R. Understanding the candidate gene × environment interaction debate: Epistemological or evidential divide? *Int. J. Epidemiol.* **2015**, *44*, 1130–1132. [CrossRef]
59. Border, R.; Keller, M.C. Commentary: Fundamental problems with candidate gene-by-environment interaction studies—reflections on Moore and Thoemmes (2016). *J. Child Psychol. Psychiatry* **2017**, *58*, 328–330. [CrossRef] [PubMed]
60. Moore, S.R. Commentary: What is the case for candidate gene approaches in the era of high-throughput genomics? A response to Border and Keller (2017). *J. Child Psychol. Psychiatry* **2017**, *58*, 331–334. [CrossRef]
61. Joshi, A.D.; Lindström, S.; Hüsing, A.; Barrdahl, M.; VanderWeele, T.J.; Campa, D.; Canzian, F.; Gaudet, M.M.; Figueroa, J.D.; Baglietto, L.; et al. Additive interactions between susceptibility single-nucleotide polymorphisms identified in genome-wide association studies and breast cancer risk factors in the Breast and Prostate Cancer Cohort Consortium. *Am. J. Epidemiol.* **2014**, *180*, 1018–1027. [CrossRef]
62. Usset, J.L.; Raghavan, R.; Tyrer, J.P.; McGuire, V.; Sieh, W.; Webb, P.; Chang-Claude, J.; Rudolph, A.; Anton-Culver, H.; Berchuck, A.; et al. Assessment of Multifactor Gene-Environment Interactions and Ovarian Cancer Risk: Candidate Genes, Obesity, and Hormone-Related Risk Factors. *Cancer Epidemiol. Biomark. Prev.* **2016**, *25*, 780–790. [CrossRef] [PubMed]
63. Rudolph, A.; Chang-Claude, J.; Schmidt, M.K. Gene-environment interaction and risk of breast cancer. *Br. J. Cancer* **2016**, *114*, 125–133. [CrossRef] [PubMed]
64. Barrdahl, M.; Rudolph, A.; Hopper, J.L.; Southey, M.C.; Broeks, A.; Fasching, P.A.; Beckmann, M.W.; Gago-Dominguez, M.; Castelao, J.E.; Guénel, P.; et al. Gene-environment interactions involving functional variants: Results from the Breast Cancer Association Consortium. *Int. J. Cancer* **2017**, *141*, 1830–1840. [CrossRef] [PubMed]
65. Kapoor, P.M.; Lindström, S.; Behrens, S.; Wang, X.; Michailidou, K.; Bolla, M.K.; Wang, Q.; Dennis, J.; Dunning, A.M.; Pharoah, P.D.P.; et al. Assessment of interactions between 205 breast cancer susceptibility loci and 13 established risk factors in relation to breast cancer risk, in the Breast Cancer Association Consortium. *Int. J. Epidemiol.* **2020**, *49*, 216–232. [CrossRef] [PubMed]
66. Pinheiro, S.P.; Holmes, M.D.; Pollak, M.N.; Barbieri, R.L.; Hankinson, S.E. Racial Differences in Premenopausal Endogenous Hormones. *Cancer Epidemiol. Biomark. Prev.* **2005**, *14*, 2147–2153. [CrossRef]
67. Setiawan, V.W.; Pike, M.C.; Kolonel, L.N.; Nomura, A.M.; Goodman, M.T.; Henderson, B.E. Racial/Ethnic Differences in Endometrial Cancer Risk: The Multiethnic Cohort Study. *Am. J. Epidemiol.* **2006**, *165*, 262–270. [CrossRef]
68. Dunn, B.K.; Agurs-Collins, T.; Browne, D.; Lubet, R.; Johnson, K.A. Health disparities in breast cancer: Biology meets socioeconomic status. *Breast Cancer Res. Treat.* **2010**, *121*, 281–292. [CrossRef]
69. Sexton, K.R.; Franzini, L.; Day, R.S.; Brewster, A.; Vernon, S.W.; Bondy, M.L. A review of body size and breast cancer risk in Hispanic and African American women. *Cancer* **2011**, *117*, 5271–5281. [CrossRef] [PubMed]
70. Reding, K.W.; Chen, C.; Lowe, K.; Doody, D.R.; Carlson, C.S.; Chen, C.T.; Houck, J.; Weiss, L.K.; Marchbanks, P.A.; Bernstein, L.; et al. Estrogen-related genes and their contribution to racial differences in breast cancer risk. *Cancer Causes Control* **2012**, *23*, 671–681. [CrossRef]
71. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D.M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer. J. Int. Du Cancer* **2015**, *136*, E359–E386. [CrossRef]
72. Korean Breast Cancer Society. *Breast Cancer Facts & Figures 2019*; Korean Breast Cancer Society: Seoul, Korea, 2019.
73. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J. Clin.* **2018**, *68*, 394–424. [CrossRef]
74. Cho, G.J.; Park, H.T.; Shin, J.H.; Hur, J.Y.; Kim, Y.T.; Kim, S.H.; Lee, K.W.; Kim, T. Age at menarche in a Korean population: Secular trends and influencing factors. *Eur. J. Pediatrics* **2010**, *169*, 89–94. [CrossRef]
75. Lee, M.-H.; Kim, S.H.; Oh, M.; Lee, K.-W.; Park, M.-J. Age at menarche in Korean adolescents: Trends and influencing factors. *Reprod. Health* **2016**, *13*, 121. [CrossRef]
76. Kim, J.Y.; Oh, I.H.; Lee, E.Y.; Choi, K.S.; Choe, B.K.; Yoon, T.Y.; Lee, C.G.; Moon, J.S.; Shin, S.H.; Choi, J.M. Anthropometric changes in children and adolescents from 1965 to 2005 in Korea. *Am. J. Phys. Anthropol.* **2008**, *136*, 230–236. [CrossRef]

77. Moon, J.S. Secular trends of body sizes in Korean children and adolescents: From 1965 to 2010. *Korean J. Pediatrics* **2011**, *54*, 436–442. [CrossRef] [PubMed]
78. Chung, S.; Park, S.K.; Sung, H.; Song, N.; Han, W.; Noh, D.Y.; Ahn, S.H.; Yoo, K.Y.; Choi, J.Y.; Kang, D. Association between chronological change of reproductive factors and breast cancer risk defined by hormone receptor status: Results from the Seoul Breast Cancer Study. *Breast Cancer Res. Treat.* **2013**, *140*, 557–565. [CrossRef]
79. Kim, H.C.; Lee, J.Y.; Sung, H.; Choi, J.Y.; Park, S.K.; Lee, K.M.; Kim, Y.J.; Go, M.J.; Li, L.; Cho, Y.S.; et al. A genome-wide association study identifies a breast cancer risk variant in ERBB4 at 2q34: Results from the Seoul Breast Cancer Study. *Breast Cancer Res.* **2012**, *14*, R56. [CrossRef] [PubMed]
80. Cai, Q.; Zhang, B.; Sung, H.; Low, S.K.; Kweon, S.S.; Lu, W.; Shi, J.; Long, J.; Wen, W.; Choi, J.Y.; et al. Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1. *Nat. Genet.* **2014**, *46*, 886–890. [CrossRef]
81. Li, J.; Humphreys, K.; Heikkinen, T.; Aittomaki, K.; Blomqvist, C.; Pharoah, P.D.; Dunning, A.M.; Ahmed, S.; Hooning, M.J.; Martens, J.W.; et al. A combined analysis of genome-wide association studies in breast cancer. *Breast Cancer Res. Treat.* **2011**, *126*, 717–727. [CrossRef]
82. Haiman, C.A.; Hankinson, S.E.; De Vivo, I.; Guillemette, C.; Ishibe, N.; Hunter, D.J.; Byrne, C. Polymorphisms in steroid hormone pathway genes and mammographic density. *Breast Cancer Res. Treat.* **2003**, *77*, 27–36. [CrossRef] [PubMed]
83. Shen, Y.; Li, D.K.; Wu, J.; Zhang, Z.; Gao, E. Joint effects of the CYP1A1 MspI, ERalpha PvuII, and ERalpha XbaI polymorphisms on the risk of breast cancer: Results from a population-based case-control study in Shanghai, China. *Cancer Epidemiol. Biomark. Prev.* **2006**, *15*, 342–347. [CrossRef]
84. Torresan, C.; Oliveira, M.M.; Torrezan, G.T.; de Oliveira, S.F.; Abuazar, C.S.; Losi-Guembarovski, R.; Lima, R.S.; Urban, C.A.; Cavalli, I.J.; Ribeiro, E.M. Genetic polymorphisms in oestrogen metabolic pathway and breast cancer: A positive association with combined CYP/GST genotypes. *Clin. Exp. Med.* **2008**, *8*, 65–71. [CrossRef] [PubMed]
85. Antognelli, C.; Del Buono, C.; Ludovini, V.; Gori, S.; Talesa, V.N.; Crino, L.; Barberini, F.; Rulli, A. CYP17, GSTP1, PON1 and GLO1 gene polymorphisms as risk factors for breast cancer: An Italian case-control study. *BMC Cancer* **2009**, *9*, 115. [CrossRef]
86. Reding, K.W.; Weiss, N.S.; Chen, C.; Li, C.I.; Carlson, C.S.; Wilkerson, H.W.; Farin, F.M.; Thummel, K.E.; Daling, J.R.; Malone, K.E. Genetic polymorphisms in the catechol estrogen metabolism pathway and breast cancer risk. *Cancer Epidemiol. Biomark. Prev.* **2009**, *18*, 1461–1467. [CrossRef] [PubMed]
87. Cerne, J.Z.; Novakovic, S.; Frkovic-Grazio, S.; Pohar-Perme, M.; Stegel, V.; Gersak, K. Estrogen metabolism genotypes, use of long-term hormone replacement therapy and risk of postmenopausal breast cancer. *Oncol. Rep.* **2011**, *26*, 479–485. [CrossRef]
88. Cerne, J.Z.; Pohar-Perme, M.; Novakovic, S.; Frkovic-Grazio, S.; Stegel, V.; Gersak, K. Combined effect of CYP1B1, COMT, GSTP1, and MnSOD genotypes and risk of postmenopausal breast cancer. *J. Gynecol. Oncol.* **2011**, *22*, 110–119. [CrossRef]
89. dos Santos, R.A.; Teixeira, A.C.; Mayorano, M.B.; Carrara, H.H.; de Andrade, J.; Takahashi, C.S. Variability in estrogen-metabolizing genes and their association with genomic instability in untreated breast cancer patients and healthy women. *J. Biomed. Biotechnol.* **2011**, *2011*, 571784. [CrossRef] [PubMed]
90. Higginbotham, K.S.; Breyer, J.P.; Bradley, K.M.; Schuyler, P.A.; Plummer, W.D., Jr.; Freudenthal, M.E.; Trentham-Dietz, A.; Newcomb, P.A.; Sanders, M.E.; Page, D.L.; et al. A multistage association study identifies a breast cancer genetic locus at NCOA7. *Cancer Res.* **2011**, *71*, 3881–3888. [CrossRef]
91. Lee, E.; Schumacher, F.; Lewinger, J.P.; Neuhausen, S.L.; Anton-Culver, H.; Horn-Ross, P.L.; Henderson, K.D.; Ziogas, A.; Van Den Berg, D.; Bernstein, L.; et al. The association of polymorphisms in hormone metabolism pathway genes, menopausal hormone therapy, and breast cancer risk: A nested case-control study in the California Teachers Study cohort. *Breast Cancer Res.* **2011**, *13*, R37. [CrossRef]
92. Johnson, N.; Walker, K.; Gibson, L.J.; Orr, N.; Folkerd, E.; Haynes, B.; Palles, C.; Coupland, B.; Schoemaker, M.; Jones, M.; et al. CYP3A variation, premenopausal estrone levels, and breast cancer risk. *J. Natl. Cancer Inst.* **2012**, *104*, 657–669. [CrossRef]
93. Martinez-Ramirez, O.C.; Perez-Morales, R.; Castro, C.; Flores-Diaz, A.; Soto-Cruz, K.E.; Astorga-Ramos, A.; Gonsebatt, M.E.; Casas, L.; Valdes-Flores, M.; Rubio, J. Polymorphisms of catechol estrogens metabolism pathway genes and breast cancer risk in Mexican women. *Breast (Edinb. Scotl.)* **2013**, *22*, 335–343. [CrossRef]
94. Johnson, N.; Dudbridge, F.; Orr, N.; Gibson, L.; Jones, M.E.; Schoemaker, M.J.; Folkerd, E.J.; Haynes, B.P.; Hopper, J.L.; Southey, M.C.; et al. Genetic variation at CYP3A is associated with age at menarche and breast cancer risk: A case-control study. *Breast Cancer Res.* **2014**, *16*, R51. [CrossRef] [PubMed]
95. Smith, A.V.; Thomas, D.J.; Munro, H.M.; Abecasis, G.R. Sequence features in regions of weak and strong linkage disequilibrium. *Genome Res.* **2005**, *15*, 1519–1534. [CrossRef]
96. Hsu, L.; Jiao, S.; Dai, J.Y.; Hutter, C.; Peters, U.; Kooperberg, C. Powerful cocktail methods for detecting genome-wide gene-environment interaction. *Genet. Epidemiol.* **2012**, *36*, 183–194. [CrossRef] [PubMed]
97. Willer, C.J.; Li, Y.; Abecasis, G.R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **2010**, *26*, 2190–2191. [CrossRef] [PubMed]
98. Matuchansky, C. Deep medicine, artificial intelligence, and the practising clinician. *Lancet* **2019**, *394*, 736. [CrossRef]