Author/s:
Barrdahl, M;Canzian, F;Gaudet, MM;Gapstur, SM;Trichopoulou, A;Tsilidis, K;van Gils, CH;Borgquist, S;Weiderpass, E;Khaw, KT;Giles, GG;Milne, RL;Le Marchand, L;Haiman, C;Lindström, S;Kraft, P;Hunter, DJ;Ziegler, R;Chanock, SJ;Yang, XR;Buring, JE;Lee, IM;Kaaks, R;Campa, D

Title:
A comprehensive analysis of polymorphic variants in steroid hormone and insulin-like growth factor-1 metabolism and risk of in situ breast cancer: Results from the Breast and Prostate Cancer Cohort Consortium

Date:
2018-03-15

Citation:
Barrdahl, M., Canzian, F., Gaudet, M. M., Gapstur, S. M., Trichopoulou, A., Tsilidis, K., van Gils, C. H., Borgquist, S., Weiderpass, E., Khaw, K. T., Giles, G. G., Milne, R. L., Le Marchand, L., Haiman, C., Lindström, S., Kraft, P., Hunter, D. J., Ziegler, R., Chanock, S. J. ,... Campa, D. (2018). A comprehensive analysis of polymorphic variants in steroid hormone and insulin-like growth factor-1 metabolism and risk of in situ breast cancer: Results from the Breast and Prostate Cancer Cohort Consortium. International Journal of Cancer, 142 (6), pp.1182-1188. https://doi.org/10.1002/ijc.31145.

A comprehensive analysis of polymorphic variants in steroid hormone and IGF-1 metabolism and risk of in situ breast cancer: results from the Breast and Prostate Cancer Cohort (BPC3) Consortium

Myrto Barrdahl [1], Federico Canzian [2], Mia M. Gaudet [3], Susan M. Gapstur [3], Antonia Trichopoulou [4], Kostas Tsilidis [5,6], Carla H. van Gils [7], Signe Borgquist [8,9], Elisabete Weiderpass [10-13], Kay-Tee Khaw [14], Graham G. Giles [15-17], Roger L. Milne [15,16], Loic Le Marchand [18], Christopher Haiman [19], Sara Lindström [20], Peter Kraft [21], David J. Hunter [21], Regina Ziegler [22], Stephen J. Chanock [22,23], Xiaohong R Yang [22], Julie E. Buring [24,25], I-Min Lee [24,25], Rudolf Kaaks [1], Daniele Campa [26]

[1] Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

[2] Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany

[3] Epidemiology Research Program, American Cancer Society, Atlanta GA-30303, USA

[4] Hellenic Health Foundation, Alexandroupoleos 23, Athens 11527

[5] Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece

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

[7] Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, STR 6.131, PO Box 85500, 3508GA Utrecht, the Netherlands

[8] Clinical Trial Unit, Skåne University Hospital, Lund, Sweden

[9] Division of Oncology and Pathology, Clinical Sciences, Lund, Lund University, Sweden

[10] Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, The Arctic University of Norway, Tromsø, Norway

[11] Department of Research, Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway

[12] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[13] Genetic Epidemiology Group, Folkhälsan Research Center, Helsinki, Finland

[14] Department of Public Health and Primary Care, School of Clinical Medicine, University of Cambridge Addenbrooke's Hospital, Hills Rd UK-CB2 0SP, UK

[15] Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, VIC-3004, Australia

[16] Centre for Epidemiology and Biostatistics, School of Population and Global Health, The University of Melbourne, VIC-3010, Australia

[17] Faculty of Medicine, Monash University, Melbourne VIC-3800, Australia

[18] Cancer Research Center of Hawaii, University of Hawaii, Honolulu HI-96813, USA

[19] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles CA-90033, USA

[20] Department of Epidemiology, University of Washington; Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle WA-98109, USA

[21] Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston MA-02115, USA

[22] Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9000 Rockville Pike Bethesda MD-20892, USA

[23] Core Genotyping Facility Frederick National Laboratory for Cancer Research, Gaithersburg MD-21701, USA

[24] Divisions of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston MA-02215, USA

[25] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston MA-02215, USA

[26] Department of Biology, University of Pisa, Via Derna 1, 56126 Pisa, Italy

Correspondence to: Rudolf Kaaks

Division of Cancer Epidemiology

German Cancer Research Center (DKFZ)

Im Neuenheimer Feld 280

D-69120 Heidelberg, Germany

Tel. +49-6221-42 2219

Fax +49-6221-422203

E-mail r.kaaks@dkfz.de

ABSTRACT

We assessed the association between 1,414 single nucleotide polymorphisms (SNPs) in genes involved in synthesis and metabolism of steroid hormones and IGF-1, and risk of breast cancer in situ (BCIS), with the aim of determining whether any of these were disease specific. This was done using 1,062 BCIS cases and 10,126 controls as well as 6,113 invasive breast cancer cases from the Breast and Prostate Cancer Cohort Consortium (BPC3). Three SNPs showed at least one nominally significant association in homozygous minor versus homozygous major models. *ACVR2A*-rs2382112 (OR_{hom}=3.05, 95%CI = 1.72-5.44, P_{hom}= 1.47×10^{-4}), *MAST2*-rs12124649 (OR_{hom}=1.73, 95% CI =1.18-2.54, P_{hom}= 5.24×10^{-3}), and *INSR*-rs10500204 (OR_{hom}= 1.96, 95% CI =1.44-2.67, P_{hom}=1.68×10^{-5}) were associated with increased risk of BCIS; however only the latter association was significant after correcting for multiple testing. Furthermore, *INSR*-rs10500204 was more strongly associated with risk of BCIS than invasive disease in case-only analyses using the homozygous minor versus homozygous major model (OR_{hom}=1.78, 95% CI =1.30-2.44, P_{hom}= 3.23×10^{-4}).

The SNP *INSR*-rs10500204 is located in an intron of the *INSR* gene and is likely to affect binding of the promyelocytic leukemia (PML) protein. The *PML* gene is known as a tumor suppressor and growth regulator in cancer. However, it is not clear on what pathway the A-allele of rs10500204 could operate to influence the binding of the protein. Hence, functional studies are warranted to investigate this further.

Novelty and impact: The present study provides a first indication of a possible association between the A-allele of *INSR*-rs10500204 and BCIS risk. However, further epidemiologic studies are needed to confirm this finding, and functional studies are required to verify the hypothesized biological mechanism.

Keywords: SNP, breast cancer in situ, BPC3, genetic epidemiology

BACKGROUND

Breast carcinoma in situ (BCIS) is a non-invasive breast cancer and a non-obligate precursor of invasive breast cancer (BC). Around 20% of all diagnosed breast tumors are non-invasive, and the most common in situ histological subtype, ductal carcinoma in situ (DCIS) corresponds to roughly 80% of BCIS diagnoses in the U.S [1, 2]. While BCIS and invasive breast cancer share epidemiological and, in part also genetic risk factors, it is important to identify any disease specific genetic risk factors, since only a fraction of existing BCIS tumors will progress to the invasive stage [3, 4].

Even though a growing number of common, low-penetrance susceptibility *loci* for invasive breast cancer are being identified [5-12], comparatively few attempts to detect SNPs that are specific for breast carcinoma in situ (BCIS) have been made so far and of these several were underpowered [4, 13-16]. A previous study conducted within the Million Women Study identified 2p-rs4666451 as being more strongly associated with BCIS risk than with invasive disease [13]. Additionally, a study within the Breast Cancer Association Consortium (BCAC) found that the SNP 5p12-rs10941679 was more strongly associated with DCIS and lower grade tumors [15]. We also reported the polymorphic variant *CDKN2BAS*-rs1011970 as a potential BCIS specific variant in a study conducted within the Breast and Prostate Cancer Cohort Consortium (BPC3) [4]. More recently, the *CCND1*-rs75915166 and *CCND1*-rs554219 SNPs were found to be associated with low and intermediate grade DCIS risk within the BCAC study [16], but these associations were also observed for invasive disease [7, 9, 10].

In a previous effort, we investigated 1,414 SNPs in 37 steroid hormone metabolism genes and 24 IGF-I pathway genes in relation to invasive BC susceptibility [17]. While only modest associations between the selected SNPs and BC were detected, there is convincing epidemiologic and molecular evidence that the concentrations of endogenous steroid hormones and insulin-like growth factor play a crucial role in modulating the risk of developing breast cancer [18-20], and it is possible that genetic variants in those pathways could influence early and non-invasive as well as more advanced stages of disease development. The purpose of the present study was to evaluate these 1,414 SNPs in relation to BCIS risk, using up to 1,062 BCIS cases and up to 10,126 controls, as well as 6,113 invasive BC cases in a two phase association study within the BPC3.

METHODS

Study population

The Breast and Prostate Cancer Cohort Consortium (BPC3) has been described extensively elsewhere [21]. Briefly, it consists of case-control studies nested within large, prospective cohorts in Europe, Australia and the United States that have both DNA samples and extensive questionnaire information collected at baseline. For the present study, cases were women who had been diagnosed with BCIS after enrolment in one of the BPC3 cohorts. Controls were healthy women selected from each cohort. In addition, 6,113 incident, invasive BC cases from the BPC3 cohorts were also used to determine whether the variants were specific BCIS alleles.

Relevant institutional review boards from each cohort approved the project and informed consent was obtained from all participants.

SNP selection and genotyping

In a first phase of the study we analyzed 1,414 SNPs involved in steroid hormone and IGF-1 metabolism using 624 BCIS cases and 8,135 controls.

In the second phase of the study, we aimed to replicate the best associations (p-value<0.001) using an additional set of 438 BCIS cases and 2,044 controls.

The genotyping for the first phase was conducted using TaqMan assays (Applied Biosystems, Foster City, CA, USA) and a custom Illumina Golden Gate array. More detailed information on genotyping and quality control is given elsewhere [17]. The genotyping of the second phase was performed using TaqMan as specified by the producer. Cases and controls were mixed in 384 well plates in order to genotype approximately the same number in each PCR run. Laboratory personnel were blinded to whether the samples were from cases or controls. To ensure the quality of the genotyping procedure, duplicate samples (~8%) were also included. All SNPs were tested for Hardy-Weinberg equilibrium (HWE) using the control samples.

Statistical analyses

The association between SNPs and BCIS risk was investigated using an unconditional logistic regression and a co-dominant model of inheritance. In order to investigate the possible differential association of SNP alleles with invasive and non-invasive breast disease, we also carried out case-only analyses comparing BCIS cases with a reference group of invasive cases to assess whether the associations were specific for BCIS or shared with

invasive BC. In addition, since DCIS is the most common form of BCIS, and information on histological subtype was available for 31% of the cases, we investigated associations between the SNP alleles and risk of DCIS. All analyses were adjusted for age at recruitment and cohort study. Since the SNPs for this study were selected on basis of an analysis comprising a total of 1,414 SNPs, we took all of these into account when adjusting the p-value threshold of statistical significance for multiple testing. To account for residual linkage disequilibrium (LD) between the SNPs we calculated the effective number of independent variants, M_{eff}, using the SNP Spectral Decomposition approach (*simpleM* method) (13). All statistical tests were two sided and all statistical analyses were performed with SAS version 9.2.

In-silico analyses

In order to assess any possible functional relevance of the identified SNP associations, a number of bioinformatics tools were used. These included the RegulomeDB (http://regulome.stanford.edu/) [22], which was used to investigate possible regulatory effects of the identified alleles or of other alleles in the surrounding region, and HaploReg v4.1 [23] which facilitated the identification of other SNPs in high LD with the variants in our study. We also used GTEX [24] to identify eQTLs.

RESULTS

Data filtering and quality control

Quality control results for the genotyping of the first study phase have been reported elsewhere [17]. Three SNPs were selected for genotyping in phase 2, because of the association found with risk of BCIS ($p<0.001$). None of these SNPs were out of HWE in the controls. The quality control analysis showed a concordance rate of 100% between duplicate samples for each of the SNP tested. The mean SNP call rate was 99.36%, with the lowest (98.55%) observed for rs10500204 and the highest (99.88%) for rs238112.

The variants in these analyses were already investigated in relation to invasive BC, but no significant associations were detected [17]. However, any association with BCIS specific risk remains unclear for these SNPs [17].

SNPs main results

The adjusted threshold for statistical significance was set to $p=0.05/1,065=4.6\times10^{-5}$, since the M_{eff} was found to be 1,065.

In the first phase of this study we identified three SNPs (*ACVR2A*-rs2382112, *INSR*-rs10500204 and *MAST2*-rs12124649) that were associated with increased risk of developing BCIS ($p<0.001$). All three associations were found when comparing homozygotes for the minor allele with those who were homozygous for the common allele in the co-dominant model. The lowest p-value was observed for the carriers of the major allele (A) of *INSR*-rs10500204 with an $OR_{hom}=2.08$, 95% CI =1.39-3.10, $P_{hom}=3.52\times10^{-4}$). The other two signals were *ACVR2A*-rs2382112 with an $OR_{hom}=3.48$, 95% CI = 1.69-7.18, $P_{hom}= 7.22\times10^{-4}$; and *MAST2*-rs12124649 with an $OR_{hom}=2.23$, 95% CI =1.40-3.57, $P_{hom}= 7.92\times10^{-4}$. The results for all the 1,414 polymorphic variants of phase one are reported in Supplementary table 1 and the results of the three SNPs with $p<0.001$ are reported in table 2. In phase two these three variants were genotyped in an additional set of cases and controls from the BPC3 cohort. When considering only this second set of individuals, the associations with BCIS risk were nominally significant for *INSR*-rs10500204 with an $OR_{hom}=1.68$, 95% CI =1.03-2.74, $P_{hom}=0.037$. When analyzing all the individuals together strong associations between all three SNPs and increased risk of developing BCIS were observed. In particular, for *INSR*-rs10500204 the association was significant after correcting for multiple testing ($OR_{hom}= 1.96$, 95% CI =1.44-2.67, $P_{hom}=1.68\times10^{-5}$), whereas for the other two SNPs the associations were consistent with the results in phase one but did not reach statistical significance after Bonferroni's correction. The results of all stages of the analysis are shown in table 2.

Case-only analyses

In the case-only analyses (BCIS *vs*. invasive BC) none of the associations were significant accounting for multiple testing. The result with the lowest p-value was observed for *INSR*-rs10500204 ($OR_{hom}=1.78$, 95% CI =1.30-2.44, $P_{hom}= 3.23\times10^{-4}$) indicating a stronger association with BCIS (Table 3).

Ductal carcinoma in situ

Finally, the risk association with DCIS was investigated for all three SNPs showing a significant association with risk of DCIS in phase one, but none of the associations reached nominal statistical significance (see Table 4).

In silico analyses

The *INSR*-rs10500204 SNP, which showed a significant association with BCIS risk, was investigated further in terms of possible functional relevance. According to information

derived from the RegulomeDB (http://regulome.stanford.edu/) [22], rs10500204 is likely to affect binding (RegulomeDB score=2b) of the following proteins: promyelocytic leukemia (PML), DNA-binding protein Ikaros (IKZF1), Nuclear factor of activated T-cells, cytoplasmic 1 (NFATC1), and Max-interacting protein 1 (MXI1). The binding was predicted by the software in lymphoblastic cell lines. Additionally, using HaploReg v4.1 [23], 35 SNPs in moderate to high LD ($R^2 > 0.6$) with rs10500204 were identified. Out of these 35 SNPs, HaploReg provided information on functionality for only four (rs3815902, rs2059806, rs3786681, rs6510956), and none of them appeared to be functionally related to BCIS. In addition, for all of these 35 SNPs the RegulomeDB indicated either that they were less likely to affect binding or that minimal binding evidence was available. Using data from GTEX [24], we did not observe any significant eQTL results that were relevant for the present analyses.

DISCUSSION

The aim of this study was to identify novel BCIS risk variants. This was done, analysing 1,414 common SNPs in 1,062 BCIS cases and 10,126 controls, as well as 6,113 invasive breast cancer cases from the Breast and Prostate Cancer Cohort Consortium (BPC3).

The most interesting results of the present study was the significant association between *INSR*-rs10500204 and risk of BCIS, where two copies of the common (A)-allele were associated with an almost two fold increased risk, compared with two copies of the rare C-allele ($P_{hom}=1.68\times10^{-5}$).

The insulin receptor (*INSR*) gene codes for the insulin receptor which binds insulin and insulin-like growth factors. Mutations in this gene have been involved in insulin resistance as well as in several types of obesity related invasive cancers, such as colorectal, pancreatic, liver and breast cancer [25, 26]. In addition, plasma levels of IGF-1 and IGFBP3 have been suggested to increase the risk of premenopausal DCIS [27]. It is therefore reasonable to hypothesise that SNPs located in regions related to these growth factors could influence BCIS risk.

The SNP rs10500204 is located in an intron of the *INSR* gene and according to information found in the RegulomeDB, it likely affects binding of the promyelocytic leukemia (PML) protein. Mutations in the *PML* gene have been related to various insulin resistance syndromes [28, 29] and the PML protein has been associated with tumor suppression in patients of acute PML [30] and is also known to be a tumor suppressor and growth regulator in cancer [31]. Loss of PML gene expression has been associated with progression of primary breast tumors to lymph node metastases [32] and PML expression in breast cancer has also been

associated with reduced time until recurrence [31]. Taken together, this information suggests a protective effect of the *PML* gene on tumorigenesis, and one might hypothesize that the A-allele of rs10500204 could decrease the effect of the *PML* gene while the minor allele (C) enhances the binding of PML to the *INSR* gene. However, further investigation is needed in order to confirm this and to identify the exact biological mechanism through which the alleles operate.

The present study focuses on selected SNPs in genes related to steroid hormone metabolism and IGF-1 – pathways that are known to be involved in breast carcinogenesis. Thus, our findings should be interpreted against the background of a greater than average prior probability of finding associations with disease risk, compared with random SNP variants in the overall genome. The two stage design with the subsequent genotyping allowed a cost-effective increase in statistical power for the potentially interesting associations discovered in the first stage, through extension of the phase-one study. Overall, the present study is one of the larger efforts made to investigate BCIS specific associations of SNPs in regions related to metabolism of steroid hormones and IGF-1.

While there are genetic risk factors which are specific for either lobular or ductal breast cancer [14, 16] a limitation of the present study was that the information on histological subtype of the tumors was relatively sparse and so the power to detect DCIS specific associations was limited.

Taken together, the present study provides a first indication of a possible association between the A-allele of *INSR*-rs10500204 and BCIS risk. While this finding could be plausible, further epidemiologic studies are needed to confirm this finding, and functional studies are required to verify the hypothesized biological mechanism.

ACKNOWLEDGEMENTS

This work was supported by the Australian National Health and Medical Research Council (NHMRC, grant 1088405). MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by NHMRC grants 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. Cases were ascertained through the Victorian Cancer Registry (VCR) and the Australian Cancer Database (Australian Institute of Health and Welfare).

The CPS-II cohort was initiated and is maintained by the American Cancer Society (Atlanta, GA, USA).

EPIC Greece was supported by the Hellenic Health foundation.

Conflict of interest: The authors declare that they have no conflict of interest.

REFERENCES

1. Leonard GD, Swain SM. Ductal carcinoma in situ, complexities and challenges. *J Natl Cancer Inst* 2004;**96**: 906-20.

2. Ward EM, DeSantis CE, Lin CC, Kramer JL, Jemal A, Kohler B, Brawley OW, Gansler T. Cancer statistics: Breast cancer in situ. *CA: a cancer journal for clinicians* 2015;**65**: 481-95.

3. Trentham-Dietz A, Newcomb PA, Storer BE, Remington PL. Risk factors for carcinoma in situ of the breast. *Cancer Epidemiol Biomarkers Prev* 2000;**9**: 697-703.

4. Campa D, Barrdahl M, Gaudet MM, Black A, Chanock SJ, Diver WR, Gapstur SM, Haiman C, Hankinson S, Hazra A, Henderson B, Hoover RN, et al. Genetic risk variants associated with in situ breast cancer. *Breast Cancer Res* 2015;**17**: 82.

5. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, Maranian MJ, Bolla MK, Wang Q, Shah M, Perkins BJ, Czene K, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet* 2015;**47**: 373-80.

6. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* 2013;**45**: 353-61, 61e1-2.

7. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;**447**: 1087-93.

8. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007;**39**: 870-4.

9. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2007;**39**: 865-9.

10. Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, Jonsson GF, Jakobsdottir M, Bergthorsson JT, Gudmundsson J, Aben KK, Strobbe LJ, Swinkels DW, et al. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2008;**40**: 703-6.

11. Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, Millikan RC, Wang X, Ademuyiwa F, Ahmed S, Ambrosone CB, Baglietto L, Balleine R, et al. A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. *Nat Genet* 2011;**43**: 1210-4.

12. Siddiq A, Couch F, Chen G, Lindström S, Eccles D, Millikan R, Michailidou K, Stram D, Beckmann L, Rhie S, Ambrosone C, Aittomäki K, et al. A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum Mol Genet* 2012;**21**: 5373-84.

13. Reeves GK, Travis RC, Green J, Bull D, Tipper S, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M, Million Women Study C. Incidence of breast cancer and its subtypes in relation to individual and multiple low-penetrance genetic susceptibility loci. *J Am Med Ass* 2010;**304**: 426-34.

14. Sawyer E, Roylance R, Petridis C, Brook MN, Nowinski S, Papouli E, Fletcher O, Pinder S, Hanby A, Kohut K, Gorman P, Caneppele M, et al. Genetic predisposition to in situ and invasive lobular carcinoma of the breast. *PLoS Genet* 2014;**10**: e1004285.

15. Milne RL, Goode EL, Garcia-Closas M, Couch FJ, Severi G, Hein R, Fredericksen Z, Malats N, Zamora MP, Arias Perez JI, Benitez J, Dork T, et al. Confirmation of 5p12 as a susceptibility locus for progesterone-receptor-positive, lower grade breast cancer. *Cancer Epidemiol Biomarkers Prev* 2011;**20**: 2222-31.

16. Petridis C, Brook MN, Shah V, Kohut K, Gorman P, Caneppele M, Levi D, Papouli E, Orr N, Cox A, Cross SS, Dos-Santos-Silva I, et al. Genetic predisposition to ductal carcinoma in situ of the breast. *Breast Cancer Res* 2016;**18**: 22.

17. Canzian F, Cox DG, Setiawan VW, Stram DO, Ziegler RG, Dossus L, Beckmann L, Blanche H, Barricarte A, Berg CD, Bingham S, Buring J, et al. Comprehensive analysis of common genetic variation in 61 genes related to steroid hormone and insulin-like growth factor-I metabolism and breast cancer risk in the NCI breast and prostate cancer cohort consortium. *Hum Mol Genet* 2010;**19**: 3873-84.

18. Pike MC, Spicer DV, Dahmoush L, Press MF. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. *Epidemiologic reviews* 1993;**15**: 17-35.

19. Khandwala HM, McCutcheon IE, Flyvbjerg A, Friend KE. The effects of insulin-like growth factors on tumorigenesis and neoplastic growth. *Endocr Rev* 2000;**21**: 215-44.

20. Endogenous H, Breast Cancer Collaborative G, Key TJ, Appleby PN, Reeves GK, Roddam AW. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. *The Lancet Oncology* 2010;**11**: 530-42.

21. Hunter DJ, Riboli E, Haiman CA, Albanes D, Altshuler D, Chanock SJ, Haynes RB, Henderson BE, Kaaks R, Stram DO, Thomas G, Thun MJ, et al. A candidate gene approach to searching for low-penetrance breast and prostate cancer genes. *Nature reviews Cancer* 2005;**5**: 977-85.

22. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. *Genome research* 2012;**22**: 1790-7.

23. Cechova H, Lassuthova P, Novakova L, Belickova M, Stemberkova R, Jencik J, Stankova M, Hrabakova P, Pegova K, Zizkova H, Cermak J. Monitoring of methylation changes in 9p21 region in patients with myelodysplastic syndromes and acute myeloid leukemia. *Neoplasma* 2012;**59**: 168-74.

24. Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobank* 2015;**13**: 307-8.

25. Cowey S, Hardy RW. The metabolic syndrome: A high-risk state for cancer? *Am J Pathol* 2006;**169**: 1505-22.

26. Parekh N, Guffanti G, Lin Y, Ochs-Balcom HM, Makarem N, Hayes R. Insulin receptor variants and obesity-related cancers in the Framingham Heart Study. *Cancer causes & control : CCC* 2015;**26**: 1189-95.

27. Bohlke K, Cramer DW, Trichopoulos D, Mantzoros CS. Insulin-like growth factor-I in relation to premenopausal ductal carcinoma in situ of the breast. *Epidemiology* 1998;**9**: 570-3.

28. Choi JH, Kang M, Kim JH, Cho J, Kim GH, Yoo HW. Identification and Functional Characterization of Two Novel Nonsense Mutations in the beta-Subunit of INSR That Cause Severe Insulin Resistance Syndrome. *Horm Res Paediatr* 2015;**84**: 73-8.

29. Bathi RJ, Parveen S, Mutalik S, Rao R. Rabson-Mendenhall syndrome: two case reports and a brief review of the literature. *Odontology* 2010;**98**: 89-96.

30. Rego EM, Wang ZG, Peruzzi D, He LZ, Cordon-Cardo C, Pandolfi PP. Role of promyelocytic leukemia (PML) protein in tumor suppression. *J Exp Med* 2001;**193**: 521-29.

31. Carracedo A, Weiss D, Leliaert AK, Bhasin M, de Boer VC, Laurent G, Adams AC, Sundvall M, Song SJ, Ito K, Finley LS, Egia A, et al. A metabolic prosurvival role for PML in breast cancer. *The Journal of clinical investigation* 2012;**122**: 3088-100.

32. Gurrieri C, Capodieci P, Bernardi R, Scaglioni PP, Nafa K, Rush LJ, Verbel DA, Cordon-Cardo C, Pandolfi PP. Loss of the tumor suppressor PML in human cancers of multiple histologic origins. *J Natl Cancer Inst* 2004;**96**: 269-79.

Table 1. Description of the study population

	CPS2		EPIC		MEC		NHS		PLCO		ALL	
	control	**case**	**control**	**case**	**control**	**case**	**control**	**case**	**control**	**case**	**control**	**case**
No.	489	103	4344	543	1934	13	1656	201	1087	202	9510	1062
Ethnicity												
White	482	101	4344	543	428	6	1555	189	978	178	7787	1017
Hispanic	1	.	.	.	383	1	5	1	13	2	402	4
African American	4	1	.	.	424	2	10	2	43	14	481	19
Asian	2	1	.	.	414	4	9	1	44	7	469	13
Hawaian	.	.	.	.	285	.	.	.	3	1	288	1
Other	.	.	.	.	.	.	77	8	6	.	83	8
Subtype												
Ductal	.	.	.	329	.	.	.	.	.	.	.	329
Lobular	.	.	.	29	.	.	.	.	.	.	.	29
Other	.	103	.	185	.	13	.	201	.	202	.	704
Age at diagnosis, mean (sd)	. (.)	69.58 (6.26)	. (.)	56.01 (7.91)	. (.)	64.40 (8.20)	. (.)	61.23 (7.56)	. (.)	66.00 (5.59)	. (.)	63.34 (8.11)
BMI, mean (sd)	25.49 (4.53)	24.26 (4.26)	25.88 (4.46)	25.78 (4.52)	26.95 (5.95)	27.71 (5.82)	25.55 (4.62)	25.53 (4.72)	27.01 (5.35)	27.91 (5.28)	26.15 (4.98)	25.88 (4.65)
Menopausal status												
premenopausal	.	.	712	42	317	3	299	52	.	.	1328	97
postmenopausal	489	103	1365	58	1584	10	1216	118	1078	201	5732	490
perimenopausal	.	.	221	5	33	.	141	31	9	1	404	37
menop unknown	.	.	2046	438	.	.	.	.	.	.	2046	438
Family history												
no relatives diagnosed with BC	396	80	.	.	1716	10	1429	161	907	161	4448	412
1 relative diagnosed	64	17	.	.	192	3	208	32	161	35	625	87
2 or more relatives diagnosed	9	6	.	.	23	.	19	8	14	6	65	20
Family history unknown	20	.	4344	543	3	.	.	.	5	.	4372	543

Table 2. Risk associations between SNP alleles and breast cancer in situ

SNP	Cases			Controls			Heterozygous / hom major*		Hom minor / hom major*		Study phase
	AA	Aa	aa	AA	Aa	aa	OR_{het} (95% CI)	P_{het}	OR_{hom} (95% CI)	P_{hom}	
ACVR2A-rs2382112	535	78	11	7018	998	38	0.82 (0.64,1.06)	1.34×10^{-1}	3.48 (1.69,7.18)	7.22×10^{-4}	Phase I
INSR-rs10500204	352	244	28	628	3057	4365	1.91 (1.27,2.87)	1.82×10^{-3}	2.08 (1.39,3.10)	3.52×10^{-4}	Phase I
MAST2-rs12124649	467	123	24	6330	1554	125	0.92 (0.74,1.14)	4.48×10^{-1}	2.23 (1.40,3.57)	7.92×10^{-4}	Phase I
ACVR2A-rs2382112	375	57	6	1716	318	10	0.82 (0.61,1.11)	2.00×10^{-1}	2.75 (0.99,7.60)	5.19×10^{-2}	Phase II
INSR-rs10500204	232	183	20	1029	828	149	1.65 (1.01,2.70)	4.76×10^{-2}	1.68 (1.03,2.74)	3.73×10^{-2}	Phase II
MAST2-rs12124649	339	89	10	1531	470	38	0.86 (0.66,1.10)	2.30×10^{-1}	1.19 (0.59,2.41)	6.32×10^{-1}	Phase II
ACVR2A-rs2382112	910	135	17	8734	1316	48	0.85 (0.70,1.03)	1.02×10^{-1}	3.05 (1.72,5.44)	1.47×10^{-4}	all
INSR-rs10500204	584	427	48	5394	3885	777	1.86 (1.36,2.54)	9.30×10^{-5}	1.96 (1.44,2.67)	1.68×10^{-5}	all
MAST2-rs12124649	806	212	34	7861	2024	163	0.90 (0.76,1.05)	1.80×10^{-1}	1.73 (1.18,2.54)	5.24×10^{-3}	all

*for rs10500204 all ORs indicate the effect of the major allele

Table 3. Case-only analyses, comparing BCIS and invasive cases

SNP	BCIS Cases			Invasive Cases			Heterozygous / hom major		Hom minor / hom major	
	AA	Aa	aa	AA	Aa	aa	OR_{het} (95% CI)	P_{het}	OR_{hom} (95% CI)	P_{hom}
rs2382112	910	135	17	5294	775	28	0.86 (0.70,1.05)	1.28×10^{-1}	2.84 (1.55,5.21)	7.55×10^{-4}
rs10500204	584	427	48	3260	2411	414	1.58 (1.15,2.17)	4.89×10^{-3}	1.78 (1.30,2.44)	3.23×10^{-4}
rs12124649	806	212	34	4840	1156	89	0.92 (0.78,1.09)	3.40×10^{-1}	1.86 (1.24,2.79)	2.62×10^{-3}

Table 4. Risk of ductal carcinoma in situ (DCIS)

SNP	DCIS Cases			Controls			Heterozygous / hom major		Hom minor / hom major	
	AA	Aa	aa	AA	Aa	aa	OR_{het} (95% CI)	P_{het}	OR_{hom} (95% CI)	P_{hom}
rs2382112	285	41	3	8734	1316	48	0.92 (0.77,1.11)	4.04×10^{-1}	1.84 (0.90,3.74)	9.27×10^{-2}
rs10500204	175	138	13	5394	3885	777	1.27 (0.97,1.65)	7.98×10^{-2}	1.29 (1.00,1.68)	5.12×10^{-2}
rs12124649	251	69	9	7861	2024	163	0.99 (0.85,1.15)	8.78×10^{-1}	1.19 (0.78,1.83)	4.18×10^{-1}

Table 1. Description of the study population

	CPS2		EPIC		MEC		NHS		PLCO		ALL	
	control	**case**	**control**	**case**	**control**	**case**	**control**	**case**	**control**	**case**	**control**	**case**
No.	489	103	4344	543	1934	13	1656	201	1087	202	9510	1062
Ethnicity												
White	482	101	4344	543	428	6	1555	189	978	178	7787	1017
Hispanic	1	.	.	.	383	1	5	1	13	2	402	4
African American	4	1	.	.	424	2	10	2	43	14	481	19
Asian	2	1	.	.	414	4	9	1	44	7	469	13
Hawaian	.	.	.	.	285	.	.	.	3	1	288	1
Other	.	.	.	.	.	.	77	8	6	.	83	8
Subtype												
Ductal	.	.	.	329	.	.	.	.	.	.	.	329
Lobular	.	.	.	29	.	.	.	.	.	.	.	29
Other	.	103	.	185	.	13	.	201	.	202	.	704
Age at diagnosis, mean (sd)	. (.)	69.58 (6.26)	. (.)	56.01 (7.91)	. (.)	64.40 (8.20)	. (.)	61.23 (7.56)	. (.)	66.00 (5.59)	. (.)	63.34 (8.11)
BMI, mean (sd)	25.49 (4.53)	24.26 (4.26)	25.88 (4.46)	25.78 (4.52)	26.95 (5.95)	27.71 (5.82)	25.55 (4.62)	25.53 (4.72)	27.01 (5.35)	27.91 (5.28)	26.15 (4.98)	25.88 (4.65)
Menopausal status												
premenopausal	.	.	712	42	317	3	299	52	.	.	1328	97
postmenopausal	489	103	1365	58	1584	10	1216	118	1078	201	5732	490
perimenopausal	.	.	221	5	33	.	141	31	9	1	404	37
menop unknown	.	.	2046	438	.	.	.	.	.	.	2046	438
Family history												
no relatives diagnosed with BC	396	80	.	.	1716	10	1429	161	907	161	4448	412
1 relative diagnosed	64	17	.	.	192	3	208	32	161	35	625	87
2 or more relatives diagnosed	9	6	.	.	23	.	19	8	14	6	65	20
Family history unknown	20	.	4344	543	3	.	.	.	5	.	4372	543

Table 2. Risk associations between SNP alleles and breast cancer in situ

SNP	Cases			Controls			Heterozygous / hom major*		Hom minor / hom major*		Study phase
	AA	Aa	aa	AA	Aa	aa	OR_{het} (95% CI)	P_{het}	OR_{hom} (95% CI)	P_{hom}	
ACVR2A-rs2382112	535	78	11	7018	998	38	0.82 (0.64,1.06)	1.34×10^{-1}	3.48 (1.69,7.18)	7.22×10^{-4}	Phase I
INSR-rs10500204	352	244	28	628	3057	4365	1.91 (1.27,2.87)	1.82×10^{-3}	2.08 (1.39,3.10)	3.52×10^{-4}	Phase I
MAST2-rs12124649	467	123	24	6330	1554	125	0.92 (0.74,1.14)	4.48×10^{-1}	2.23 (1.40,3.57)	7.92×10^{-4}	Phase I
ACVR2A-rs2382112	375	57	6	1716	318	10	0.82 (0.61,1.11)	2.00×10^{-1}	2.75 (0.99,7.60)	5.19×10^{-2}	Phase II
INSR-rs10500204	232	183	20	1029	828	149	1.65 (1.01,2.70)	4.76×10^{-2}	1.68 (1.03,2.74)	3.73×10^{-2}	Phase II
MAST2-rs12124649	339	89	10	1531	470	38	0.86 (0.66,1.10)	2.30×10^{-1}	1.19 (0.59,2.41)	6.32×10^{-1}	Phase II
ACVR2A-rs2382112	910	135	17	8734	1316	48	0.85 (0.70,1.03)	1.02×10^{-1}	3.05 (1.72,5.44)	1.47×10^{-4}	all
INSR-rs10500204	584	427	48	5394	3885	777	1.86 (1.36,2.54)	9.30×10^{-5}	1.96 (1.44,2.67)	1.68×10^{-5}	all
MAST2-rs12124649	806	212	34	7861	2024	163	0.90 (0.76,1.05)	1.80×10^{-1}	1.73 (1.18,2.54)	5.24×10^{-3}	all

*for rs10500204 all ORs indicate the effect of the major allele

Table 3. Case-only analyses, comparing BCIS and invasive cases

SNP	BCIS Cases			Invasive Cases			Heterozygous / hom major		Hom minor / hom major	
	AA	Aa	aa	AA	Aa	aa	OR_{het} (95% CI)	P_{het}	OR_{hom} (95% CI)	P_{hom}
rs2382112	910	135	17	5294	775	28	0.86 (0.70,1.05)	1.28×10^{-1}	2.84 (1.55,5.21)	7.55×10^{-4}
rs10500204	584	427	48	3260	2411	414	1.58 (1.15,2.17)	4.89×10^{-3}	1.78 (1.30,2.44)	3.23×10^{-4}
rs12124649	806	212	34	4840	1156	89	0.92 (0.78,1.09)	3.40×10^{-1}	1.86 (1.24,2.79)	2.62×10^{-3}

Table 4. Risk of ductal carcinoma in situ (DCIS)

SNP	DCIS Cases			Controls			Heterozygous / hom major		Hom minor / hom major	
	AA	Aa	aa	AA	Aa	aa	OR_{het} (95% CI)	P_{het}	OR_{hom} (95% CI)	P_{hom}
rs2382112	285	41	3	8734	1316	48	0.92 (0.77,1.11)	4.04×10^{-1}	1.84 (0.90,3.74)	9.27×10^{-2}
rs10500204	175	138	13	5394	3885	777	1.27 (0.97,1.65)	7.98×10^{-2}	1.29 (1.00,1.68)	5.12×10^{-2}
rs12124649	251	69	9	7861	2024	163	0.99 (0.85,1.15)	8.78×10^{-1}	1.19 (0.78,1.83)	4.18×10^{-1}