Breast Cancer Research

Genetic variation of ESR1 and its co-activator PPARGC1B is synergistic in augmenting the risk of estrogen receptor positive breast cancer

Breast Cancer Research 2011, **13**:R10 doi:10.1186/bcr2817

Yuqing Li (liy1@gis.a-star.edu.sg)
Yi Li (liy3@gis.a-star.edu.sg)
Sara Wedren (sara.wedren@ki.se)
Guoliang Li (ligl@gis.a-star.edu.sg)
Tze Howe Charn (charnh@gis.a-star.edu.sg)
Desai Kartiki Vasant (desaikv@gis.a-star.edu.sg)
Carine Bonnard (carine.bonnard@bmsi.a-star.edu.sg)
Kamila Czene (Kamila.Czene@ki.se)
Keith Humphreys (Keith.Humphreys@ki.se)
Hatef Darabi (hatef.darabi@ki.se)
Kristjana Einarsdttir (keinarsdottir@meddent.uwa.edu.au)
Tuomas Heikkinen (tuomas.heikkinen@helsinki.fi)
Kristiina Aittomaki (kristiina.aittomaki@hus.fi)
Carl Blomqvist (carl.blomqvist@helsinki.fi)
Kee Seng Chia (ephcks@nus.edu.sg)
Heli Nevanlinna (Heli.Nevanlinna@hus.fi)
Per Hall (Per.Hall@ki.se)
Edison T Liu (liue@gis.a-star.edu.sg)
Jianjun Liu (liuj3@gis.a-star.edu.sg)

ISSN 1465-5411

Article type Research article

Submission date 29 September 2010

Acceptance date 26 January 2011

Publication date 26 January 2011

Article URL http://breast-cancer-research.com/content/13/1/R10

Breast Cancer Research

Genetic variation of *ESR1* and its co-activator *PPARGC1B* is synergistic in augmenting the risk of estrogen receptor positive breast cancer

Yuqing Li[1], Yi Li[1], Sara Wedrén[2], Guoliang Li[3], Tze Howe Charn[4], Desai Kartiki Vasant[4], Carine Bonnard[5], Kamila Czene [6], Keith Humphreys[6], Hatef Darabi[6], Kristjana Einarsdóttir[7], Tuomas Heikkinen[8], Kristiina Aittomäki[9], Carl Blomqvist[10], Kee Seng Chia[11], Heli Nevanlinna[8], Per Hall[6], Edison T Liu[4] and Jianjun Liu[1#]

[1]Human Genetics, Genome Institute of Singapore, 60 Biopolis Street, Singapore 138672, Singapore

[2]Institute for Environmental Medicine, Karolinska Institutet, Box 210, Stockholm, 17177, Sweden

[3]Computational & Mathematical Biology, Genome Institute of Singapore, 60 Biopolis Street, Singapore 138672, Singapore

[4]Cancer Biology, Genome Institute of Singapore, 60 Biopolis Street, Singapore 138672, Singapore

[5]Laboratory of Human Embryology, Institute of Medical Biology of Singapore, 8A Biomedical Grove, Singapore 138648, Singapore

[6]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, Stockholm, 17177, Sweden

[7]Centre for Health Services Research, School of Population Health, University of Western Australia, 35 Stirling Highway, Perth, Western Australia 6009, Australia

[8]Departments of Obstetrics and Gynecology, Helsinki University Central Hospital, P.O. Box 700, Helsinki, 00029 HUS, Finland

[9]Clinical Genetics, Helsinki University Central Hospital, P.O. Box 700, Helsinki, 00029 HUS, Finland

[10]Oncology , Helsinki University Central Hospital, P.O. Box 700, Helsinki, 00029 HUS, Finland

[11]Department of Epidemiology and Public Health, Yong Loo Lin School of Medicine, National University of Singapore, MD3, 16 Medical Drive, Singapore 117597, Singapore

To whom correspondence should be addressed: Jianjun Liu, liuj3@gis.a-star.edu.sg

ABSTRACT

Introduction: Given the role of estrogen in breast carcinogenesis and the modification of estrogen receptor (ER) activity by its biochemical cofactors, we hypothesize that genetic variation within ER cofactor genes alters cellular response to estrogen exposure and consequently modifies the risk for ER positive breast cancer.

Methods: We genotyped 790 tagging single nucleotide polymorphisms (SNPs) within 60 ER cofactor genes in1257 cases and 1464 controls from Sweden and 2215 cases and 1265 controls from Finland and tested their associations with either ER+ or ER- breast cancer.

Results: Seven SNPs showed consistent association with ER+ breast cancer in the two independent samples, and six of them were located within *PPARGC1B*, encoding an ER co-activator, with the strongest association at rs741581 (OR=1.41, P=4.84E-05) that survived Bonferroni correction for multiple testing in the combined ER+ breast cancer sample ($P_{corrected}$=0.03). Moreover, we also observed significant synergistic interaction ($P_{interaction}$ =0.008) between the genetic polymorphisms within *PPARGC1B* and *ESR1* in ER+ breast cancer. By contrast, no consistent association was observed in ER- breast cancer. Furthermore, we found that administration of estrogen in the MCF-7 cell line induced *PPARGC1B* expression and enhanced occupancies of ER and RNA polymerase II within the region of SNP association, suggesting the up-regulation of *PPARGC1B* expression by *ESR1* activation.

Conclusions: Our study revealed that DNA polymorphisms of *PPARGC1B*, coding a bona fide ER co-activator, are associated with ER positive breast cancer risk. The feed-forward transcriptional regulatory loop between *PPARGC1B* and *ESR1* further augments their protein interaction, which provides a plausible mechanistic explanation for the synergistic genetic interaction between *PPARGC1B* and *ESR1* in ER+ breast cancer. Our study also highlights that biochemically and genomically informed candidate gene studies can enhance the discovery of interactive disease susceptibility genes.

INTRODUCTION

It is known that the risk of breast cancer is related to lifetime exposure to estrogen [1, 2]. Estrogen stimulates cell proliferation and increases the frequency of spontaneous mutations, leading to a malignant phenotype [3]. Breast cells respond to estrogen via estrogen receptors (ERs) through a defined biochemical process: upon ligand binding, ERs undergo a conformational change that facilitates receptor dimerization, DNA binding, recruitment of ER cofactors, and modulation of target gene expression [4-6]. Endocrine therapy provides strong evidence that attenuation of ER (*ESR1*) activity can reduce breast cancer risk [7], and women with ER positive tumor would be the most likely to benefit from these treatments [7, 8]. The genetic studies of *ESR1*, however, have had contradictory results. Only recently, through a very large genetic association study has there been demonstrated a small but significant association of polymorphisms within *ESR1* with the risk of breast cancer [9-11]. Two plausible explanations for the inconsistent results might be due to the small sample sizes and thus limited statistical power of these studies, or that the risk was not evaluated by stratifying breast cancer patients based on tumor ER status. However, there is at least one further possibility: ER cofactors can either enhance transcriptional activity of ER as co-activators or inhibit the activity as co-repressors. The genetic variants within ER cofactors have not been systematically investigated in term of association with breast cancer risk, although some coding variants within individual genes, such as *NCOA3* and *CCND1*, have been investigated [12-15].

Given the modification of ER activity by its cofactors through their physical and functional interactions [16], the cofactor proteins which bind to ER may be as important as the receptor itself in mediating transcriptional response to estrogen exposure [17]. We therefore hypothesized that genetic variation within ER cofactor genes may alter cellular response to estrogen exposure and consequently, alone with or by interacting with genetic variations within *ESR1*, modify breast cancer risk in an ER status-dependent fashion. To assess this hypothesis, we investigated the association of common genetic variation, using a tagging SNP approach, within 60 cofactor genes in two large case-control samples of breast cancer from Sweden and Finland and investigated their interaction with genetic variation within *ESR1* in term of influencing the risk of hormone-driven breast cancer.

MATERIALS AND METHODS

Study population

The Swedish sample was from a population-based case-control study which has been described in detail previously [18]. Briefly, 1322 cases were Swedish-born women diagnosed with incident primary invasive breast cancer between October 1993 and March 1995 and contributed blood samples. All cases were postmenopausal and between 50 and 74 years of age at diagnosis. All the cases were identified through the six regional cancer registries in Sweden. The controls (n=1524) were randomly selected from the Swedish Registry of Total Population with no previous breast cancer and were frequency-matched for age with the cases. Questionnaires were used to collect risk factor information.

The Finnish sample was from a hospital-based case-control study, in which the cases consisted of two series of unselected breast cancer patients and additional familial patients diagnosed at the Helsinki University Central Hospital. The first set of cases were 884 patients collected in 1997-1998 and 2000 covering 79% of all newly diagnosed breast cancer cases during those periods [19, 20]. The second set of cases, consisting of 986 newly diagnosed breast cancer patients, were collected in 2001 – 2004 and covered 87% of all such patients during that period [21]. An additional 538 familial breast cancer cases were also collected at the same hospital as previously described [22, 23]. Women with a prior diagnosis of breast cancer in-situ were excluded, leaving 2215 invasive breast cancer cases for analysis. Healthy female population controls (n=1287) were collected from the same geographical regions of Finland as the cases.

Information on reproductive and hormonal risk factors was available for the Swedish sample and showed expected association patterns with breast cancer [24-26]. Such information was not available for the Finnish controls.

Hormone receptor status information was retrieved from medical records of all participating cases and was available for both the Swedish and Finnish cases.

Approval for the study was obtained from the Institutional Review Boards in Sweden, Finland and the National University of Singapore. All subjects provided written informed consent.

DNA Isolation

DNA was extracted from 4 ml of whole blood using the QIAamp DNA Blood Maxi Kit (Qiagen) according to manufacturer's instructions.

Candidate Gene and Tagging SNP Selection

In this study, key words of "ER cofactor", "ER coactivator" and "ER corepressor" were used in a literature search to identify ER cofactor genes. Boolean searching ("AND" "OR") was used to narrow or broaden the search in Pubmed. Using this method, 60 ER-cofactor genes were identified as candidate genes. Tagging SNPs within the 60 candidate genes were selected based on the HapMap CEU data (Rel #22/phase II Apr07, on NCBI B36 assembly, dbSNP b126). In brief, for each gene, all common SNP with a minor allele frequency (MAF) over than 0.05 within the gene and 5kb surrounding region were first identified from the HapMap database. Tagging SNPs were then selected in Haploview version 4.1 [27] using a pair-wise SNP tagging approach with $r^2 > 0.8$ used as the criterion for selection. A total of 806 tagging SNPs were selected within the 60 ER co-factor genes.

Genotyping

Illumina's GoldenGate assay was used for genotyping SNPs, following manufacturers' instructions. In brief, all 806 tagging SNPs were subjected to genotyping assay design, out of which 790 SNPs were successfully designed and subjected to genotyping analysis. DNA samples were randomly assigned to the plates carrying positive and negative controls, and all genotyping results were generated and checked by laboratory staff unaware of the case-control status. SNPs with a call rate < 96% (81 SNPs failed in the Swedish sample and 42 SNPs failed in the Finnish sample) and minor allele frequency (MAF) < 1% (18 SNPs in the Swedish sample and 40 SNPs in the Finnish sample) were excluded from further analysis. Deviation of genotype frequencies

from those expected under Hardy-Weinberg Equilibrium (HWE) were assessed in the control subjects. SNPs with P value of HWE less than 7.4E-5 (0.05/675) (6 SNPs failed in the Swedish sample and 15 SNPs failed in the Finnish sample) were excluded. In total, 685 SNPs from the Swedish sample and 693 SNPs from the Finnish sample were used for statistical analysis, and 675 shared SNPs between the Swedish and Finnish samples were used for analysis in the combined sample.

Genotyping was duplicated in two percent of samples (in both Swedish and Finnish samples) and there was concordance in >99% of the duplicated samples, suggesting high genotyping accuracy. With r^2 over than 0.8, the average coverage of common variation (MAF>5%) within the 60 candidate genes was 91%. Out of these, 51 genes had coverage over 80% (Table S1 in Additional file 1).

RT-qPCR analysis

MCF-7 cells were cultured in DMEM (Invitrogen) medium with 10% FBS (Invitrogen). Prior to hormone treatment, cells were maintained in phenol-red free DMEM F-12 containing 5% charcoal stripped serum for 72 hours for hormone depletion. Cells were treated with 10nM 17-β Estradiol (E2) (Sigma) for a period of 0 or 3 hours. Cells were harvested and total RNA and RT-qPCR analysis was carried out as described previously [28]. DMSO (Merck)/vehicle treated cells were used as controls for the same time course. Real-time PCR analysis was performed in the ABI Prism 7700 sequence detection system using SYBR Green from ABI. Primers were designed using the online Primer 3 program. All experiments were repeated at least twice. We used two sets of

primers for identifying different isoforms of PPARGC1B. The oligonucleotide sequences were as follows: PPARGC1B_1 isoform (NM_001172699.1), forward: 5'GAAGAGGAAGAAGGGGAGGA3' reverse: 5'CTCTGGTAGGGGCAGTGGT3'; and PPARGC1B_2 isoform (NM_133263.3), forward: 5'CCTGAAGATGACGTGGGTCT3', reverse 5'CCTTCCTTCTGGGTGTCAGA 3'. Beta-actin specific primers (forward: 5'TCCCTGGAGAAGAGCTACGA3', reverse 5' AGGAAGGAAGGCTGGAAGAG3') were used as an internal control to normalize the amounts of reverse transcribed product used in the PCR reaction. Ct values obtained for PPARGC1B isoforms were normalized to Beta-ACTIN Ct values. The normalized Ct (ΔCt) values were then used to calculate the difference ($\Delta\Delta$Ct) between E2 and DMSO treated samples. Fold change of PPARGC1B was calculated as $2^{-\Delta\Delta Ct}$.

Statistical Analysis

To measure the magnitude of association between SNPs and breast cancer risk, per-allele ORs (assuming a log-additive model) and 95 percent confidence intervals (CI) were estimated using logistic regression. As the controls were younger than cases in the Finnish samples, age at diagnosis / enrollment (as a continuous variable) was included in the regression models in the Finnish analysis for OR adjustment. The Cochran-Armitage trend test was used to calculate p-values in the Swedish and Finnish sample sets, separately in subtypes, and in cases overall. Inverse variance weighting was used in a meta-analysis for two independent datasets. Individual OR was obtained from age-unajusted in Swedish and age-adjusted in Finnish analysis. To evaluate differences in

ORs between studies, a test of homogeneity was carried out for each individual SNP analysis (data not shown).

To determine the model of inheritance, associations between SNPs within the PPARGC1B gene and ER positive breast cancer risk were estimated by assuming dominant, recessive and additive models in the two sample sets. We then performed these analyses with meta-analysis using inverse variance weighting approach. Individual OR from two independent studies followed-up age-unadjusted in Swedish and age-adjusted in Finnish.

Forward stepwise logistic regression was used to explore whether the associations at the six SNPs were independent to each other. The selection criterion was P value less than 0.2. The analysis was performed in ER+ breast cancer risk in the two sample sets separately as well as in the combined ER+ sample data set. To account for different MAFs in the two populations, a binary indicator variable for study was included in the regression models as well as age, in the combined data regression analysis.

Pair-wise interaction analysis was performed under a dominant mode of inheritance using logistic regression and likelihood ratio tests. To maximize the statistical power, we pooled sample sets from Swedish and Finnish. Age and study were included in the model as covariables. The full model included an interaction term between the two interacting variables for the risk of breast cancer. In this multivariate logistic regression analysis, each coefficient provided an estimate of the log odds ratio whilst adjusting for all other variables included in the model. Likelihood ratio tests, comparing models with and without the interaction term, were used to generate P-values.

All analyses were performed using STATA version 8.0. Linkage Disequilibrium (LD) calculation was performed in the Haploview version 4.1. All statistical tests were two-sided.

RESULTS

Study Subjects

Two independent case-control samples of breast cancer from Sweden and Finland were investigated in this study, whose characteristics are summarized in Table 1. The cases and controls of the Swedish sample were frequency-matched on age, whereas the Finnish controls were younger than the Finnish cases ($P<0.0001$). In the Swedish sample, there were significant differences between the cases and controls in terms of age at first birth ($P=0.0002$), age at menopause ($P=0.0001$), HRT use ($P=0.017$), and parity ($P=0.0001$), which is consistent with the well-established role of these reproductive factors in breast cancer development. The reproductive factor information was not available for the Finnish controls. In both the Swedish and Finnish cases, there were similar percentages of ER+ (81.9% vs. 80.9%) and ER- (18.1% vs. 19.1%) cases.

SNP Association Analysis

First, single SNP association analyses were performed using trend tests in the Swedish and Finnish samples separately by stratifying the cases into ER+ and ER- groups, with 685 SNPs being tested in the Swedish sample and 693 SNPs being tested in the Finnish sample. 48 SNPs (7.00%) in the Swedish sample and 50 SNPs (7.28%) in the Finnish sample showed association with ER+ breast cancer risk with a nominal P

value less than 0.05. Seven SNPs showed consistent association between the two independent samples (Table S2 in Additional file 1), and six of them were located within the *PPARGC1B* gene. In contrast, 21 and 50 SNPs showed association with ER- breast cancer with a nominal P value of 0.05 or less, in the Swedish and Finnish samples respectively, but no SNPs show consistent associations between the two independent samples.

We then analyzed SNP associations in the combined Swedish and Finnish samples. In general, SNPs showed stronger evidence of association with ER+ breast cancer than ER- breast cancer (Table 2 & Table S3 in Additional file 1). The most significant association was identified at rs741581 within the second intron of *PPARGC1B* (OR=1.41, P=4.84E-05) in ER+ breast cancer, which survived the Bonferroni correction for multiple testing ($P_{corrected}$=0.03). rs741581 was one of the seven SNPs that showed consistent associations between the Swedish and Finnish samples.

We also evaluated the SNP association with overall breast cancer risk and found 55 SNPs (8.03%) from the Swedish samples and 61 SNPs (8.80%) from the Finnish samples to show association with overall breast cancer risk with a nominal P value of 0.05 or less . Only two SNPs, however, showed consistent association between the two independent samples (Table S2 in Additional file 1), and none of the associations survived Bonferroni correction for multiple testing in the combined samples (smallest $P_{correct}$=0.198).

Genotype Association Analysis of *PPARGC1B* in ER+ Breast Cancer

PPARGC1B is located on 5q33.1 and encodes for peroxisome proliferative activated receptor gamma coactivator beta (PGC-1β), a bona fide co-activator of ERα. In

this study, 40 tagging SNPs within *PPARGC1B* were successfully genotyped in both the Swedish and Finnish samples, which could capture 80% of common variants (131 out of 162 SNPs) within *PPARGC1B* with a minimal r^2 value of 0.8 (mean r^2 value=0.95, according to the HapMap CEU data).

To have a better understanding of the association within *PPARGC1B*, we performed genotype-based association analysis by assuming dominant, recessive and additive model of inheritance. We found that the top 3 SNPs yielding the most significant association evidence in the dominant model compared with other models (Table S5 in Additional file 1). Under the dominant model, the same 6 SNPs (as for the trend tests) of the 40 SNPs within *PPARGC1B* showed consistent association with ER+ breast cancer between the Swedish and Finnish samples (Table 3). The strength of the association (ORs) at the 6 SNPs was stronger in ER+ breast cancer than in overall or ER- ones, with the strongest association identified at rs741581 (P=1.9E-02 in the Swedish samples, P=6.1E-05 in the Finnish samples, and P=1.8E-05 in the combined samples).

The six SNPs showing consistent association with ER+ breast cancer were located within two regions of high LD (Figure 1 B&C), suggesting that the associations at those SNPs may not be completely independent. We therefore performed a forward stepwise logistic regression (cutoff P value=0.20) and revealed two independent associations with ER+ breast cancer at rs741581 (P=0.031) and rs6895698 (P=0.014) in the combined sample. Under the dominant model, we found that rs741581, rs6895698, age and study sample were four independent variables associated with ER+ breast cancer risk. Similarly, the same stepwise analysis of ER+ breast cancer in the two individual samples also revealed two independent associations at rs741581 (P=0.172) and rs2340621

(P=0.036) in the Swedish sample and rs741581 (P=0.023) and rs6895698 (P=0.053) in the Finnish sample. Notably, rs6895698 and rs2340621 lies within the same LD block and are highly correlated (r^2=0.72, according to HapMap CEU data).

Association of *ESR1* variation with ER+ Breast Cancer

Our previous study suggested an association between *ESR1* polymorphisms and breast cancer risk [9, 29] in the same Swedish sample. The association was within a region flanked by rs988328 to rs3020318 and was manifested by three haplotypes. Given that a large association study of *ESR1* by the Breast Cancer Association Consortium (BCAC) also revealed a significant association within the same region under a dominant model [10], we searched for SNPs that were in high LD (r^2 >0.5) with the three haplotypes, but were not genotyped in our previous study. We did this by performing a genotype-based association analysis under a dominant model of inheritance. Using the haplotype information from the HapMap CEU data, we identified three common SNPs that were in high LD (r^2=0.89) with one of the three haplotypes (TAG18~21) [9], while no SNPs were found with r^2 >0.5 for the other two haplotypes, based on the HapMap CEU data (Rel #22/phase II Apr07, on NCBI B36 assembly, dbSNP b126). Given that the three SNPs were in perfect LD (r^2 = 1), we genotyped one of the three SNPs, rs7761846, in our Swedish and Finnish samples. As expected, rs7761846 showed association with ER+ breast cancer (OR=1.28, P=0.014) in the combined sample. The two independent Swedish and Finnish samples also revealed consistent association, although the association in the Finnish sample did not reach statistical significance (Table 4).

Genetic Interaction between the Polymorphisms of *PPARGC1B* (rs6895698, rs2340621 and rs741581) and *ESR1* (rs7761846)

Given the known modification of ER activity by PPARGC1B in cellular response to estrogen exposure, we investigated the genetic interaction between rs741581, rs2340621 and rs6895698 within *PPARGC1B* and rs7761846 within *ESR1* in terms of modulating ER+ breast cancer risk. The analysis in the combined sample identified a significant synergistic interaction between rs2340621 (representing *PPARGC1B*) and rs7761846 (representing *ESR1)*, $P_{interaction}$=0.008 (Table 5). Women carrying both *PPARGC1B* (rs2340621) and *ESR1* (rs7761846) risk genotypes (*GA/AA* & *CT/CC*) had a much higher risk for breast cancer than non-carriers (*GG&TT*) (OR=1.94, P=2.03E-06). Similar patterns of genetic interaction were also observed between the remaining SNPs rs741581 (*PPARGC1B*) and rs7761846 (*ESR1*) as well as rs6895698 (*PPARGC1B*) and rs7761846 (*ESR1*), although these interactions did not achieve statistical significance, likely due to the low minor allele frequencies of rs741581 and rs6895698. However, the significant genetic interactions could not be detected in overall or ER- breast cancer (Table S6 & S7 in Additional file 2).

Transcriptional Regulation of *PPARGC1B* by *ERα*

To understand the molecular mechanism underlying the observed genetic interaction, we investigated whether there was any transcriptional "cross-talk" between the two genes beyond the known ligand dependent, co-activating interaction of the

PPARGC1B protein with ERα [30, 31], using the ER responsive MCF7 breast cancer cell line. First, we examined the expression of *PPARGC1B* in MCF7 and noted a two-fold induction of *PPARGC1B* expression by ER activation after estradiol administration (Figure S1 in Additional file 3). As a marker of transcriptional activity, ChIP-seq analysis in the same MCF7 cell line identified a significant peak of RNA polymerase II (RNAPII) occupancy close to the transcriptional start site (TSS) of *PPARGC1B* within the LD region of SNP association, and the RNAPII occupancy was further enhanced by estradiol treatment. This confirms the transcriptional responsiveness of PPARGC1B to estradiol. Moreover, the ChIP-seq analysis also identified 5 ER binding sites in and around *PPARGC1B* (one site located approximately 50 kb 5' of the TSS, one in the second intron of the gene within 13 kb of the associated SNP rs741581, and the other three binding sites approximately 10, 31 and 57 kb 3' of the polyadenylation signal sequence) and within the LD region of significant association with ER+ breast cancer (Figure 1F&H). Interestingly, the sites showing highest of ER occupancy were seen at two locations, one ~13kb from the significant SNP rs741581 and the second within 31 kb 3' of the polyadenylation signal sequence. We have recently described the identification of all ER binding site interactions in the human genome [32, 33] and had defined that genes engaged in chromatin loop formation by a transcription factor were definitively regulated by the factor. Our data indicated that all the ER binding sites around *PPARGC1B* were engaged in chromatin loop formation centered on the *PPARGC1B* gene (Figure 1I), which indicates that ERα directly regulates *PPARGC1B.* Taken together, these data strongly indicate that *PPARGC1B* expression could be directly regulated by ERα and, when coupled with the known enhancement of ERα transcriptional activity by

PPARGC1B protein at the site of binding, suggests a feed-forward regulatory loop between the two genes that augments ER signaling when the two factors are present.

Discussion

To our knowledge, this is the first comprehensive association analysis of common variation within ER cofactor genes in breast cancer where 36 ER co-activators and 24 ER co-repressors were investigated. The utilization of two independent case-control samples of northern European origin allowed us to identify an association based on not only the overall significance in the large combined sample, but also on the consistency of the SNP association between the two individual samples. We found significant associations between *PPARGC1B* polymorphisms and risk for ER+ breast cancer, and importantly, revealed a synergistic effect between the genetic polymorphisms within *PPARGC1B* and *ESR1*.

Genetic association studies of ER cofactor genes have been limited so far. Burwinkel, et al. reported a significant association of coding variants Q586H and T960T of *NCOA3* with familial breast cancer risk and further suggested that familial breast cancer patients may condense the rare allele's contribution to the protective effect of breast cancer [12]. Whilst two studies [34, 35] have reported an association of the variant Pro241Pro in *CCND1* with breast cancer risk, other studies have reported negative results for this variant [14, 36, 37]. In particular, Wirtenberger, et al. investigated the coding variant Ala203Pro of *PPARGC1B* and found it to be associated with familial breast

cancer susceptibility [38]. In our study, we did not observe significant association between polymorphisms in *NCOA3* and *CCND1* with breast cancer risk. However, the Ala203Pro (rs7732671) variant of *PPARGC1B* is 10kb away and not correlated with *PPARGC1B* SNP, rs741581 (r^2<0.05 in HapMap CEU data) and thus would not have been detected by our tagging SNP approach. Nevertheless, both Wirtenberger et al.'s and our study support the association of genetic variation of *PPARGC1B* with particular subtypes of breast cancer.

Importantly, the association of *PPARGC1B* as well as its synergistic interaction with *ESR1* was only observed in breast cancer patients with ER+ tumor, as would be expected according to the biochemical mechanism of interaction. There is growing evidence that the impact of genetic risk factors on breast cancer varies by hormone receptor status. For example, recent studies by the Breast Cancer Association Consortium have led to the discovery of novel breast cancer susceptibility loci in *FGFR2, TNRC9*, 8q24, 2q35, and 5p12 that showed stronger association with ER+ disease than ER- disease [39-42] with FGFR being also a direct target of ER. These data suggest that the risk of ER+ tumors which has been shown to be driven by reproductive factors in epidemiologic studies have also a genomic basis based on the constituents of the ER gene regulatory network [43, 44]. Thus, in our study, although the sample sizes of two ER+ data sets were smaller compared with the two overall data sets, the number of overlapping SNPs between the Swedish and Finnish studies was larger than that observed in the overall breast cancer analysis. Recently, we also demonstrated that genetic variation of the estrogen metabolism pathway, particularly the ones involved in the production of estrogen through androgen conversion, also influences the risk of the

development of estrogen-sensitive breast cancer [45]. As with this study, the effect size of the metabolism gene polymorphisms are relatively small but taken together with PPARGC1B and FGFR, show that the estrogen receptor signaling axis that engages both upstream and downstream components may have, in the composite, a significant role in the genesis of the most common form of breast cancer.

The genetic interaction between *PPARGC1B* and *ESR1* is biologically plausible. The *PPARGC1B* protein, peroxisome proliferate activated receptor gamma coactivator beta (PGC-1β), is a bona fide ER coactivator [31] that physically interacts with ERα and plays a role in "amplifying" ER signaling all of which provides a convincing biological mechanism for the observed genetic interaction between the two genes. Furthermore, our series of transcriptional regulation analyses in the MCF7 ER+ breast cancer cell line have demonstrated that *PPARGC1B* expression can be induced by estrogen treatment, and this transcriptional response of *PPARGC1B* is likely mediated by five functional ER binding sites around *PPARGC1B* that are all engaged in interlocking chromatin loops highly indicative of an ER regulated gene [32]. Thus, *PPARGC1B* may be involved in a feed forward control mechanism with ERα such that ER induction (for example, by estradiol treatment) heightens the expression of a co-activator *PPARGC1B* of ER, which in turn increase ER action at the DNA binding site. The feed-forward looping mechanism will therefore further augment the protein interaction between *PPARGC1B* and *ESR1*. This putative “amplification” effect, if confirmed, is another mechanistic model for epistatic interactions between genetic loci and maybe one reason for the strength of its signal in the association study as compared to the other ER co-factors studied.

There are some limitations in our study. Coverage of common variation is not sufficient (less than 80%) for some genes (Table S1 in Additional file 1), so that some associations may have been missed. In addition, our tagging SNP selection provides a rather limited coverage of 5kb surrounding sequences of the candidate genes, which may have contributed to some associations of regulatory SNPs being undetected, such as the one reported within ESR1 [11]. The number of overlapping SNPs between the two data sets is small for both ER+ and overall breast cancer analyses. The limited overlapping could be due to ethnic heterogeneity between the two population samples and their moderate sample sizes. On the one hand, the ethnic heterogeneity may patially explain the low overlapping SNPs between two datasets; on the other hand, the current sample size is not large enough to capture the moderate effect of associated SNPs. Thus, some of the top SNPs for each individual sample set are probably false positive, which cause the small overlap between the numbers of significant SNPs in both datasets. Besides that, the sample size limitation in ER- patients also could lead the non-significant results in ER- analysis, as we observed that some associations in ER- are in the same direction with ER+ analysis. As it is known that ER cofactors work as a multi-component protein complex, but due to sample size limitation, we are unable to detect interaction among three or more genes simultaneously. It is also worth noting that the contribution of genetic variation to cancer risk is based on both their prevalence and penetrance, and thus the relative importance of individual SNPs may vary from population to population. Hence, further confirmation of our findings in other populations is warranted.

Conclusions

Our study has revealed an association of genetic variation within *PPARGC1B* with the risk of ER+ breast cancer. Consistent with the known interaction of *PPARGC1B* and ER at the molecular level, where *PPARGC1B* modulates ER activity and thus ER signaling, our study revealed a synergistic effect between genetic variation within the *PPARGC1B* and *ESR1* genes. *PPARGC1B* has been shown to alter responses to the selective ER modulator, tamoxifen [30]. Kressler et al also demonstrated that *PPARGC1B* indirectly co-activates tamoxifen-bound ERα which cooperates with *NCOA1* to enable tamoxifen agonism in kidney and osteosarcoma cell lines. Lastly, the synergism demonstrated in this study also suggests that disrupting the interaction between an ER co-activator, such as *PPARGC1B*, and ERα, or blocking their mutual activation, may represent a sensitive and "leveraged" strategy for cancer prevention [7]. Our study therefore provides new biological insight into the genetic basis of the more common estrogen receptor positive breast cancer and highlights that biochemically and genomically informed candidate gene study can enhance the discovery of interactive disease susceptibility genes.

Abbreviations

ER, estrogen receptor; *ESR1*, estrogen receptor 1 ; *NCOA3*, nuclear receptor coactivator 3; *CCND1*, cyclin D1; *PPARGC1B*, peroxisome proliferative activated receptor gamma coactivator beta; HapMap, Haplotype Map; HRT, hormone replacement treatment; ChIP-seq, chromatin immunoprecipitation-sequencing; *FGFR2*, fibroblast growth factor receptor 2; *TNRC9*, TOX high mobility group box family member 3; *NCOA1*, nuclear receptor coactivator 1.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YQ, SW,KE, HN, KS, PH, ET and JJ initiated and designed the study; CB, KC and HN provided the study material and patients information; YQ, CB, TH, KA and SW collected and organized the data; YQ, YL, GL, TH, DK, JJ, KH and HD performed data analysis and interpreted results; YQ, SW, YL, DK, KH, PH, ET and JJ drafted the manuscript; CB, HD, KH, KE, HN, PH, ET and JJ performed critical review and revised the manuscript. All authors read and approved the manuscript.

Acknowledgements

We are grateful to all the women who took the time and effort to participate in this study. We would also like to thank Heng Khai Koon and Ong Eng Hu Jason for genotyping, Frans Verhoeff for processing the genotyping data and the Finnish Cancer Registry for providing cancer data.

This work was supported by funding from the Agency for Science, Technology and Research of Singapore (A_STAR), Nordic Cancer Union (both for Sweden and Finland) and the Märit and Hans Rausing's Initiative Against Breast Cancer. In addition, the Finnish study has been financially supported by the Helsinki University Central Hospital Research Fund, Academy of Finland (132473), the Finnish Cancer Society, and the Sigrid Juselius Foundation.

References

1. Key T, Appleby P, Barnes I, Reeves G: **Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies.** *J Natl Cancer Inst* 2002, **94:**606-616.

2. Missmer SA, Eliassen AH, Barbieri RL, Hankinson SE: **Endogenous estrogen, androgen, and progesterone concentrations and breast cancer risk among postmenopausal women.** *J Natl Cancer Inst* 2004, **96:**1856-1865.

3. Dickson RB, Stancel GM: **Estrogen receptor-mediated processes in normal and cancer cells.** *J Natl Cancer Inst Monogr* 2000**:**135-145.

4. Gao X, Loggie BW, Nawaz Z: **The roles of sex steroid receptor coregulators in cancer.** *Mol Cancer* 2002, **1:**7.

5. Smith CL, O'Malley BW: **Coregulator function: a key to understanding tissue specificity of selective receptor modulators.** *Endocr Rev* 2004, **25:**45-71.

6. Tremblay GB, Giguere V: **Coregulators of estrogen receptor action.** *Crit Rev Eukaryot Gene Expr* 2002, **12:**1-22.

7. Osborne CK, Schiff R: **Estrogen-receptor biology: continuing progress and therapeutic implications.** *J Clin Oncol* 2005, **23:**1616-1622.

8. Johnston SR: **New strategies in estrogen receptor-positive breast cancer.** *Clin Cancer Res* 2010, **16:**1979-1987.

9. Einarsdottir K, Darabi H, Li Y, Low YL, Li YQ, Bonnard C, Sjolander A, Czene K, Wedren S, Liu ET, Hall P, Humphreys K, Liu J: **ESR1 and EGF genetic variation in relation to breast cancer risk and survival.** *Breast Cancer Res* 2008, **10:**R15.

10. Dunning AM, Healey CS, Baynes C, Maia AT, Scollen S, Vega A, Rodriguez R, Barbosa-Morais NL, Ponder BA, Low YL, Bingham S, Haiman CA, Le Marchand L, Broeks A, Schmidt MK, Hopper J, Southey M, Beckmann MW, Fasching PA, Peto J, Johnson N, Bojesen SE, Nordestgaard B, Milne RL, Benitez J, Hamann U, Ko Y, Schmutzler RK, Burwinkel B, Schurmann P, et al.: **Association of ESR1 gene tagging SNPs with breast cancer risk.** *Hum Mol Genet* 2009, **18:**1131-1139.

11. Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, Wen W, Levy S, Deming SL, Haines JL, Gu K, Fair AM, Cai Q, Lu W, Shu XO: **Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1.** *Nat Genet* 2009, **41:**324-328.

12. Burwinkel B, Wirtenberger M, Klaes R, Schmutzler RK, Grzybowska E, Forsti A, Frank B, Bermejo JL, Bugert P, Wappenschmidt B, Butkiewicz D, Pamula J, Pekala W, Zientek H, Mielzynska D, Siwinska E, Bartram CR, Hemminki K: **Association of NCOA3 polymorphisms with breast cancer risk.** *Clin Cancer Res* 2005, **11:**2169-2174.

13. Ceschi M, Sun CL, Van Den Berg D, Koh WP, Yu MC, Probst-Hensch N: **The effect of cyclin D1 (CCND1) G870A-polymorphism on breast cancer risk is modified by oxidative stress among Chinese women in Singapore.** *Carcinogenesis* 2005, **26:**1457-1464.

14. Haiman CA, Garcia RR, Hsu C, Xia L, Ha H, Sheng X, Le Marchand L, Kolonel LN, Henderson BE, Stallcup MR, Greene GL, Press MF: **Screening and association testing of common coding variation in steroid hormone receptor co-activator and co-repressor genes in relation to breast cancer risk: the Multiethnic Cohort.** *BMC Cancer* 2009, **9:**43.

15. Montgomery KG, Chang JH, Gertig DM, Dite GS, McCredie MR, Giles GG, Southey MC, Hopper JL, Campbell IG: **The AIB1 glutamine repeat polymorphism is not associated with risk of breast cancer before age 40 years in Australian women.** *Breast Cancer Res* 2005, **7:**R353-356.

16. Hall JM, McDonnell DP: **Coregulators in nuclear estrogen receptor action: from concept to therapeutic targeting.** *Mol Interv* 2005, **5:**343-357.

17. McKenna NJ, Lanz RB, O'Malley BW: **Nuclear receptor coregulators: cellular and molecular biology.** *Endocr Rev* 1999, **20:**321-344.

18. Wedren S, Lovmar L, Humphreys K, Magnusson C, Melhus H, Syvanen AC, Kindmark A, Landegren U, Fermer ML, Stiger F, Persson I, Baron J, Weiderpass E: **Oestrogen receptor alpha gene haplotype and postmenopausal breast cancer risk: a case control study.** *Breast Cancer Res* 2004, **6:**R437-449.

19. Kilpivaara O, Bartkova J, Eerola H, Syrjakoski K, Vahteristo P, Lukas J, Blomqvist C, Holli K, Heikkila P, Sauter G, Kallioniemi OP, Bartek J, Nevanlinna H: **Correlation of CHEK2 protein expression and c.1100delC mutation status with tumor characteristics among unselected breast cancer patients.** *Int J Cancer* 2005, **113:**575-580.

20. Syrjakoski K, Vahteristo P, Eerola H, Tamminen A, Kivinummi K, Sarantaus L, Holli K, Blomqvist C, Kallioniemi OP, Kainu T, Nevanlinna H: **Population-based study of BRCA1 and BRCA2 mutations in 1035 unselected Finnish breast cancer patients.** *J Natl Cancer Inst* 2000, **92:**1529-1531.

21. Fagerholm R, Hofstetter B, Tommiska J, Aaltonen K, Vrtel R, Syrjakoski K, Kallioniemi A, Kilpivaara O, Mannermaa A, Kosma VM, Uusitupa M, Eskelinen M, Kataja V, Aittomaki K, von Smitten K, Heikkila P, Lukas J, Holli K, Bartkova J, Blomqvist C, Bartek J, Nevanlinna H: **NAD(P)H:quinone oxidoreductase 1 NQO1*2 genotype (P187S) is a strong prognostic and predictive factor in breast cancer.** *Nat Genet* 2008, **40:**844-853.

22. Vahteristo P, Bartkova J, Eerola H, Syrjakoski K, Ojala S, Kilpivaara O, Tamminen A, Kononen J, Aittomaki K, Heikkila P, Holli K, Blomqvist C, Bartek J, Kallioniemi OP, Nevanlinna H: **A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer.** *Am J Hum Genet* 2002, **71:**432-438.

23. Eerola H, Blomqvist C, Pukkala E, Pyrhonen S, Nevanlinna H: **Familial breast cancer in southern Finland: how prevalent are breast cancer families and can we trust the family history reported by patients?** *Eur J Cancer* 2000, **36:**1143-1148.

24. Magnusson C, Baron JA, Correia N, Bergstrom R, Adami HO, Persson I: **Breast-cancer risk following long-term oestrogen- and oestrogen-progestin-replacement therapy.** *Int J Cancer* 1999, **81:**339-344.

25. Magnusson C, Colditz G, Rosner B, Bergstrom R, Persson I: **Association of family history and other risk factors with breast cancer risk (Sweden).** *Cancer Causes Control* 1998, **9:**259-267.

26. Magnusson CM, Persson IR, Baron JA, Ekbom A, Bergstrom R, Adami HO: **The role of reproductive factors and use of oral contraceptives in the aetiology of breast cancer in women aged 50 to 74 years.** *Int J Cancer* 1999, **80:**231-236.

27. Barrett JC, Fry B, Maller J, Daly MJ: **Haploview: analysis and visualization of LD and haplotype maps.** *Bioinformatics* 2005, **21:**263-265.

28. Liu J, Kartiki Vasant Desai, Li Y, Shakeela Banu, Yew Kok Lee, Dianbo Qu, Tuomas Heikkinen, Kirsimari Aaltonen, Taru A. Muranen , Tasneem Shabbir Kajiji, Carine Bonnard, Kristiina Aittoma¨ki, Karl von Smitten, Carl Blomqvist, John L. Hopper, Southey MC, Brauch H, Consortium TG, Chenevix-Trench G, Beesley J, Spurdle AB, Chen X, Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer AOCSG, Czene K, Hall P, Nevanlinna H, Liu ET: **Germ-line variation at a functional p53 binding site increases susceptibility to breast cancer development** *The HUGO Journal* 2009, **Volume 3:**31-40.

29. Li Y, Sung WK, Liu JJ: **Association mapping via regularized regression analysis of single-nucleotide-polymorphism haplotypes in variable-sized sliding windows.** *Am J Hum Genet* 2007, **80:**705-715.

30. Kressler D, Hock MB, Kralli A: **Coactivators PGC-1beta and SRC-1 interact functionally to promote the agonist activity of the selective estrogen receptor modulator tamoxifen.** *J Biol Chem* 2007, **282:**26897-26907.

31. Kressler D, Schreiber SN, Knutti D, Kralli A: **The PGC-1-related protein PERC is a selective coactivator of estrogen receptor alpha.** *J Biol Chem* 2002, **277:**13918-13925.

32. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, et al.: **An oestrogen-receptor-alpha-bound human chromatin interactome.** *Nature* 2009, **462:**58-64.

33. Li G, Fullwood MJ, Xu H, Mulawadi FH, Velkov S, Vega V, Ariyaratne PN, Mohamed YB, Ooi HS, Tennakoon C, Wei CL, Ruan Y, Sung WK: **ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing.** *Genome Biol* 2010, **11:**R22.

34. Onay UV, Aaltonen K, Briollais L, Knight JA, Pabalan N, Kilpivaara O, Andrulis IL, Blomqvist C, Nevanlinna H, Ozcelik H: **Combined effect of CCND1 and COMT polymorphisms and increased breast cancer risk.** *BMC Cancer* 2008, **8:**6.

35. Yu CP, Yu JC, Sun CA, Tzao C, Ho JY, Yen AM: **Tumor susceptibility and prognosis of breast cancer associated with the G870A polymorphism of CCND1.** *Breast Cancer Res Treat* 2008, **107:**95-102.

36. Driver KE, Song H, Lesueur F, Ahmed S, Barbosa-Morais NL, Tyrer JP, Ponder BA, Easton DF, Pharoah PD, Dunning AM: **Association of single-nucleotide polymorphisms in the cell cycle genes with breast cancer in the British population.** *Carcinogenesis* 2008, **29:**333-341.

37. Krippl P, Langsenlehner U, Renner W, Yazdani-Biuki B, Wolf G, Wascher TC, Paulweber B, Weitzer W, Leithner A, Samonigg H: **The 870G>A polymorphism of the cyclin D1 gene is not associated with breast cancer.** *Breast Cancer Res Treat* 2003, **82:**165-168.

38. Wirtenberger M, Tchatchou S, Hemminki K, Schmutzhard J, Sutter C, Schmutzler RK, Meindl A, Wappenschmidt B, Kiechle M, Arnold N, Weber BH, Niederacher D, Bartram CR, Burwinkel B: **Associations of genetic variants in the estrogen receptor coactivators PPARGC1A, PPARGC1B and EP300 with familial breast cancer.** *Carcinogenesis* 2006, **27:**2201-2208.

39. Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, Jonsson GF, Jakobsdottir M, Bergthorsson JT, Gudmundsson J, Aben KK, Strobbe LJ, Swinkels DW, van Engelenburg KC, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Saez B, Lambea J, Godino J, Polo E, Tres A,

Picelli S, Rantala J, Margolin S, Jonsson T, Sigurdsson H, Jonsdottir T, Hrafnkelsson J, et al.: **Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2008, **40:**703-706.

40. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, Albers-Akkers MT, Swinkels DW, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Godino J, Garcia-Prats MD, Polo E, Tres A, Mouy M, Saemundsdottir J, Backman VM, Gudmundsson L, Kristjansson K, Bergthorsson JT, Kostic J, et al.: **Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2007, **39:**865-869.

41. Milne RL, Benitez J, Nevanlinna H, Heikkinen T, Aittomaki K, Blomqvist C, Arias JI, Zamora MP, Burwinkel B, Bartram CR, Meindl A, Schmutzler RK, Cox A, Brock I, Elliott G, Reed MW, Southey MC, Smith L, Spurdle AB, Hopper JL, Couch FJ, Olson JE, Wang X, Fredericksen Z, Schurmann P, Bremer M, Hillemanns P, Dork T, Devilee P, van Asperen CJ, et al.: **Risk of estrogen receptor-positive and -negative breast cancer and single-nucleotide polymorphism 2q35-rs13387042.** *J Natl Cancer Inst* 2009, **101:**1012-1018.

42. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, Bojesen SE, Nordestgaard BG, Axelsson CK, Arias JI, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Zamora P, Brauch H, Justenhoven C, Hamann U, Ko YD, Bruening T, Haas S, Dork T, Schurmann P, Hillemanns P, Bogdanova N, Bremer M, Karstens JH, Fagerholm R, Aaltonen K, Aittomaki K, et al.: **Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics.** *PLoS Genet* 2008, **4:**e1000054.

43. Althuis MD, Fergenbaum JH, Garcia-Closas M, Brinton LA, Madigan MP, Sherman ME: **Etiology of hormone receptor-defined breast cancer: a systematic review of the literature.** *Cancer Epidemiol Biomarkers Prev* 2004, **13:**1558-1568.

44. Garcia-Closas M, Chanock S: **Genetic susceptibility loci for breast cancer by estrogen receptor status.** *Clin Cancer Res* 2008, **14:**8000-8009.

45. Low YL, Li Y, Humphreys K, Thalamuthu A, Darabi H, Wedren S, Bonnard C, Czene K, Iles MM, Heikkinen T, Aittomaki K, Blomqvist C, Nevanlinna H, Hall P, Liu ET, Liu J: **Multi-variant pathway association analysis reveals the importance of genetic determinants of estrogen metabolism in breast and endometrial cancer susceptibility.** *PLoS Genet* 2010, **6:**e1001012.

Figure 1. Transcriptional regulation of PPARGC1B by ERα in MCF7 cell line. From top to bottom, A. Schematic diagram of the genes from UCSC database; B. Map positions of six significant SNPs within PPARGC1B; C. Recombination rate surrounding PPARGC1B from HapMAP CEU database; D. MCF7 input DNA density for CHIP-seq analysis; E. CHIP-seq RNA PolII occupancy density without E2 treatment; F. CHIP-seq ER occupancy without E2 treatment; G. CHIP-seq RNA PolII occupancy density with E2 treatment; H. CHIP-seq ER occupancy with E2 treatment; I. ER interaction loop detected by ChIA-PET (from Fullwood et.al. 2009).

Table 1. Selected characteristics of the cases and controls in the Swedish and the Finnish samples

Entire Study	Swedish			Finnish		
Characteristic	Number (case/ctrl)	Mean (case/ctrl)	P value*	Number (case/ctrl)	Mean (case/ctrl)	P value*
Age (years)	1257 /1464	63.11 / 63.12	0.96	2214 / 1265	56.07 / 40.88	<1.00E-4
Age at first birth	1072/1321	25.50 / 24.76	2.00E-4	1185 / -	26.43 /-	/
Age at menopause	1247 / 1460	50.6 / 49.97	1.00E-4	1341 / -	50.34 /-	/
BMI (Recent)	1250 / 1443	25.71 / 25.67	0.81	1525 /-	25.03 /-	/
Case only	Number	Percentage		Number	Percentage	
All cases	1257	/		2215	/	
ER positive	684	81.92		1709	80.92	
ER negative	151	18.08		403	19.08	
Controls	1464	/		1265	/	

BMI: Body Mass Index; ER positive, estrogen receptor positive; ER negative, estrogen receptor negative.

*:Two-sided t-test was used for P value estimation.

Table 2. Twenty-five most significant SNPs associated with ER positive breast cancer in Swedish and Finnish samples

Gene	SNP	Position	P-value†	Adjusted P-value‡	OR (95%CI) †
PPARGC1B	rs741581*§	chr5:149182978	4.84E-05	0.033	1.414 (1.197,1.672)
PPARGC1B	rs1012543*§	chr5:149157138	9.98E-05	0.067	1.223 (1.105,1.353)
PPARGC1B	rs6895698§	chr5:149120455	2.73E-04	0.184	1.225 (1.098,1.366)
CARM1	rs1529711	chr19:10884434	4.26E-04	0.288	1.229 (1.096,1.378)
RBM23	rs7469*	chr14:22440037	1.15E-03	0.778	1.248 (1.092,1.427)
PPARGC1B	rs4705365*§	chr5:149093146	1.78E-03	–	1.193 (1.068,1.333)
NCOR2	rs10846670	chr12:123456184	2.67E-03	–	0.872 (0.798,0.954)
RBM23	rs3811187*	chr14:22439134	2.96E-03	–	1.158 (1.051,1.275)
PELP1	rs4790674	chr17:4529772	3.23E-03	–	1.171 (1.054,1.3)
PPARGC1B	rs2340621§	chr5:149122509	4.00E-03	–	1.149 (1.045,1.262)
CCND1	rs649392*§	chr11:69173974	4.43E-03	–	0.877 (0.801,0.96)
PPARGC1B	rs10036538§	chr5:149135781	5.51E-03	–	1.156 (1.043,1.28)
PELP1	rs7214635	chr17:4547769	5.67E-03	–	1.166 (1.046,1.3)
PPARGC1B	rs4705382	chr5:149161559	6.32E-03	–	0.872 (0.791,0.962)
NEDD4	rs11071224*	chr15:53902817	6.32E-03	–	0.764 (0.63,0.927)
MED13	rs4968469*	chr17:57491867	6.90E-03	–	1.145 (1.038,1.263)
NCOR2	rs12321007	chr12:123449054	7.44E-03	–	1.137 (1.035,1.25)
MED13	rs9889324*	chr17:57481404	8.89E-03	–	1.14 (1.033,1.258)
NCOR2	rs1794973	chr12:123391545	9.45E-03	–	0.889 (0.813,0.972)
PPARGC1B	rs1422429	chr5:149146627	1.11E-02	–	1.124 (1.027,1.231)
NCOR2	rs10846666	chr12:123450306	1.17E-02	–	0.867 (0.776,0.969)
NCOA1	rs17046513	chr2:24817999	1.26E-02	–	1.284 (1.055,1.563)
NCOA1	rs17046462	chr2:24759054	1.37E-02	–	1.285 (1.053,1.568)
NCOR2	rs10846667	chr12:123450377	1.40E-02	–	0.895 (0.819,0.978)
SNW1	rs3759728*	chr14:77299912	1.49E-02	–	0.855 (0.753,0.97)

ER positive, estrogen receptor positive; SNP: single nucleotide polymorphism rs id; chr: chromosome; OR, odds ratio; 95% CI, 95% confidence interval.

*:SNP belongs to top 25 most significant SNPs associated with ER negative breast cancer.

§:SNP was significantly associated with ER positive breast cancer in both Swedish and Finnish dataset.

†: P-value and OR were obtained from meta-analysis based on inverse variance method for two independent datasets. Individual OR was obtained from age-unadjusted in Swedish and age-adjusted in Finnish.

‡: P-value was adjusted by Bonferroni correction (n=675); "–", adjusted P-value >1.

Table 3. Six overlapping SNPs in PPARGC1B associated with ER positive breast cancer in Swedish and Finnish samples

			Swedish		Finnish	
SNP ID	Allele[§]	Subtype	MAF[†]	OR‡ (95%CI)	MAF[†]	OR‡ (95%CI)
rs4705365	G/A	ER+		1.26 (1.05,1.52)		1.26 (1.05,1.52)
		ER-		1.14 (0.81,1.61)		1.17 (0.9,1.52)
		All cases	0.21	1.14 (0.97,1.33)	0.17	1.18 (0.99,1.4)
rs6895698	G/A	ER+		1.27 (1.06,1.53)		1.39 (1.15,1.67)
		ER-		1.05 (0.74,1.48)		1.14 (0.88,1.49)
		All cases	0.22	1.12 (0.96,1.3)	0.17	1.25 (1.05,1.49)
rs2340621	G/A	ER+		1.3 (1.08,1.57)		1.22 (1.03,1.46)
		ER-		0.86 (0.61,1.2)		1.05 (0.82,1.34)
		All cases	0.31	1.12 (0.96,1.31)	0.32	1.14 (0.97,1.34)
rs10036538	C/G	ER+		1.19 (0.99,1.42)		1.2 (1,1.43)
		ER-		0.91 (0.65,1.28)		1.05 (0.81,1.35)
		All cases	0.26	1.03 (0.89,1.2)	0.22	1.11 (0.94,1.31)
rs1012543	A/G	ER+		1.26 (1.05,1.51)		1.26 (1.06,1.5)
		ER-		1.08 (0.77,1.51)		1.08 (0.84,1.38)
		All cases	0.26	1.11 (0.95,1.29)	0.23	1.18 (1,1.39)
rs741581	G/A	ER+		1.32 (1.05,1.67)		1.76 (1.33,2.31)
		ER-		0.81 (0.49,1.32)		1.21 (0.81,1.82)
		All cases	0.08	1.12 (0.92,1.37)	0.05	1.53 (1.18,1.98)

ER+, estrogen receptor positive; ER-, estrogen receptor negative; OR, odds ratio; 95% CI, 95% confidence interval.

§Major allele / Minor allele.

† Minor Allele Frequency (MAF) was from control samples only.

‡ Odds Ratios were performed on dominant model based, which were age-unadjusted in Swedish and age-adjusted in Finnish.

Table 4. Association analysis of rs7761846 within ESR1 under dominant model in ER+ cases analysis

	Control	Case	OR (95%CI)	Pvalue
Swedish	1442	675	1.43 (1.10, 1.86)	0.007
Finnish	1246	1669	1.10 (0.81, 1.48)	0.55
Combined*	2688	2344	1.28 (1.05, 1.56)	0.014

ER+, estrogen receptor positive; OR, odds ratio; 95% CI, 95% confidence interval.
*: Combined analysis was performed on logistic regression adjusted by age and study.

Table 5. Analysis of pair-wise interaction between SNPs within PPARGC1B and ESR1 on the ER+ breast cancer in the combined Swedish and Finnish samples under dominant model

ESR1 (rs7761846)	**PPARGC1B**					
Genotype	**GG (rs2340621)**			**GA/AA (rs2340621)**		
	case (%)	ctrl (%)	OR(95%CI)	case(%)	ctrl(%)	OR(95%CI)
TT	916 (39)	1121 (42)	1	1160 (50)	1280 (48)	1.18 (1.03, 1.34)
CT/CC	106 (5)	151 (6)	0.95 (0.71, 1.29)	161 (7)	134 (5)	1.94 (1.47, 2.55)
Interaction P value*	**0.008**					
	GG (rs6895698)			**GA/AA (rs6895698)**		
	case (%)	ctrl (%)	OR(95%CI)	case(%)	ctrl(%)	OR(95%CI)
TT	1284 (55)	1566 (58)	1	791 (34)	837 (31)	1.28 (1.11, 1.47)
CT/CC	160 (7)	184 (7)	1.21 (0.93, 1.55)	106 (5)	101 (4)	1.77 (1.30, 2.42)
Interaction P value*	0.506					
	GG (rs741581)			**GA/AA (rs741581)**		
	case (%)	ctrl (%)	OR(95%CI)	case(%)	ctrl(%)	OR(95%CI)
TT	1741 (74)	2076 (77)	1	335 (14)	326 (12)	1.41 (1.17, 1.70)
CT/CC	225 (10)	251 (9)	1.25 (1.01, 1.55)	42 (2)	34 (1)	2.18 (1.32, 3.59)
Interaction P value*	0.459					

SNP: single nucleotide polymorphism rs id ; ER+, estrogen receptor positive; OR, odds ratio; 95% CI, 95% confidence interval.
*: Analysis was performed on combined data set, in which study and age were regarded as covariables.

Additional files

Additional file 1: Table S1 to Table S5. Table S1 presents coverage evaluation of common variant in 60 ER cofactor genes. Table S2 presents ORs and P values of the consistent SNPs between the Swedish and Finnish samples from the analyses of ER+ and overall breast cancer. Table S3 presents the twenty-five most significant SNPs in ER negative association analysis in Swedish and Finnish samples. Table S4 presents the twenty-five most significant SNPs in overall association analysis in Swedish and Finnish samples. Table S5 presents the comparison of P value among additive, dominant and recessive models in the analysis of ER positive breast cancer in PPARGC1B in combined Swedish and Finnish samples.

Additional file 2: Table S6 to Table S7. Table S6 presents analysis of pair-wise interaction effect between SNPs within PPARGC1B and ESR1 on the overall breast cancer in the combined Swedish and Finnish samples. Table S7 presents the analysis of pair-wise interaction effect between SNPs within PPARGC1B and ESR1 on the ER negative breast cancer in the combined Swedish and Finnish samples.

Additional file 3: Figure S1. Figure S1 presents relative expression of PPARGC1B gene in MCF7 cells 3 hrs post E2 treatment.

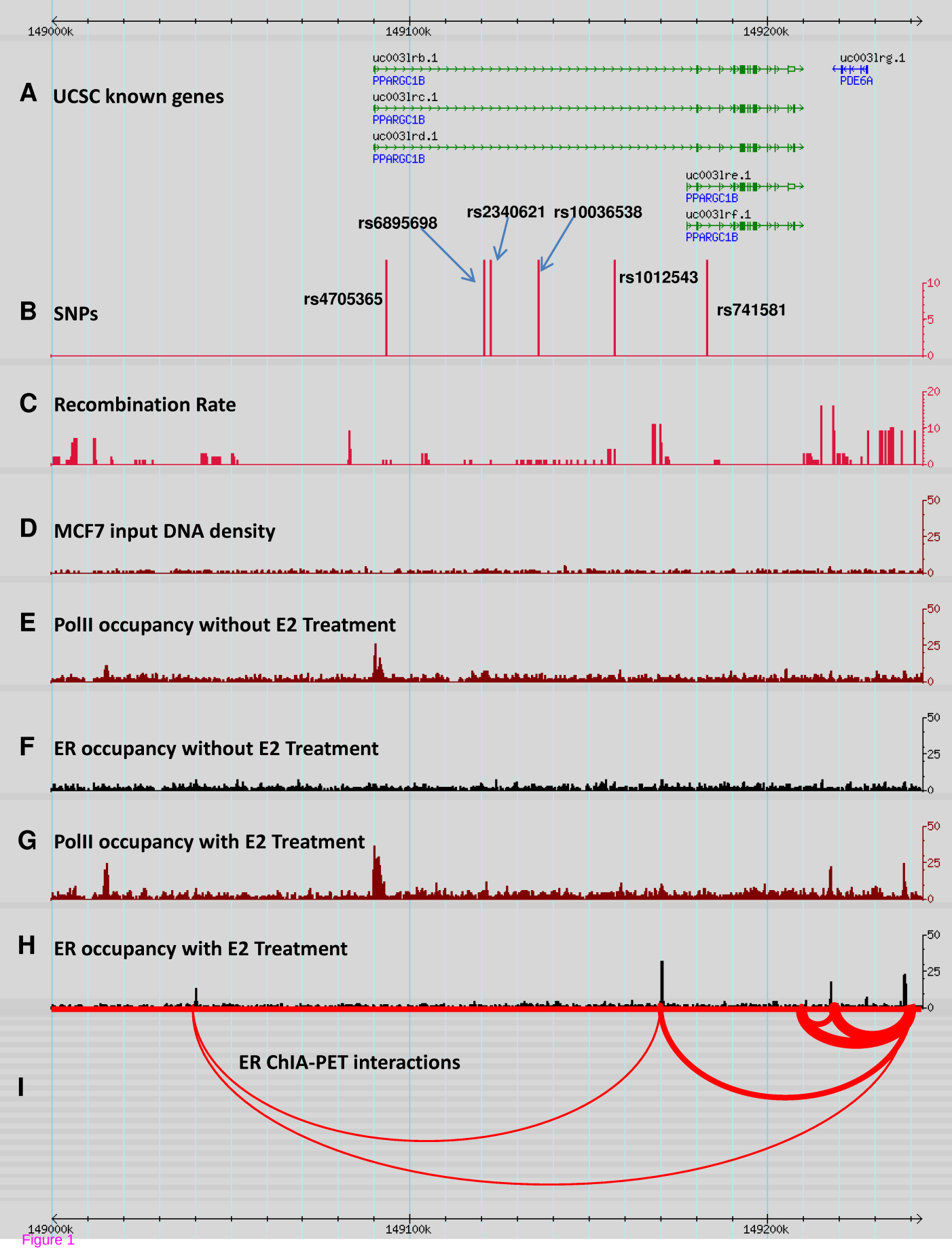
149000k
149100k
149200k
A
UCSC known genes
uc003lrb.1
PPARGC1B
uc003lrc.1
PPARGC1B
uc003lrd.1
PPARGC1B
uc003lrg.1
PDE6A
uc003lre.1
PPARGC1B
uc003lrf.1
PPARGC1B
rs6895698
rs2340621
rs10036538
rs1012543
rs4705365
rs741581
B
SNPs
10
5
0
C
Recombination Rate
20
10
0
D
MCF7 input DNA density
50
25
0
E
PolII occupancy without E2 Treatment
50
25
0
F
ER occupancy without E2 Treatment
50
25
0
G
PolII occupancy with E2 Treatment
50
25
0
H
ER occupancy with E2 Treatment
50
25
0
ER ChIA-PET interactions
I
149000k
149100k
149200k

Figure 1

Additional files provided with this submission:

Additional file 1: Additional data 1.doc, 239K
http://breast-cancer-research.com/imedia/9043167885111484/supp1.doc
Additional file 2: Additional data 2.doc, 55K
http://breast-cancer-research.com/imedia/8370953551114599/supp2.doc
Additional file 3: Additional data 3.doc, 34K
http://breast-cancer-research.com/imedia/1968431254511145/supp3.doc